

Julian Nida-Rümelin

## Vernunft und Freiheit

Da ich überzeugt bin, dass menschliche Freiheit angemessen nur über die Rolle von Gründen geklärt werden kann, gibt es einen unlösbaren Zusammenhang zwischen Rationalitätstheorie und Freiheitskonzeption.

Der erste Teil meines Vortrags wird daher die Frage „Was ist Vernunft?“ zu klären versuchen. Dabei werden wir uns nicht auf die praktische Dimension beschränken können. Theoretische und praktische Vernunft, die Rationalität des Handelns und des Urteilens sind unauflöslich miteinander verbunden. In einem zweiten Teil ist dann zu klären, was Freiheit mit Vernunft zu tun hat. Und schließlich möchte ich im dritten Teil Überlegungen zur metaphysischen Dimension des Verhältnisses von Vernunft und Freiheit präsentieren.

Dieser Text kann die detaillierteren Ausführungen in *Strukturelle Rationalität* (Reclam 2001) und *Über menschliche Freiheit* (Reclam 2005) natürlich nicht ersetzen, wohl aber präzisieren und um neue Aspekte ergänzen. Ich hoffe, dass er dazu beiträgt das Gesamtbild der Argumentation deutlich zu machen.

## **Inhalt**

I. Was ist Vernunft? .....	3
II. Was ist Freiheit? .....	11
III. Metaphysische Aspekte.....	17
Literatur.....	25

## I. Was ist Vernunft?

§1 Insbesondere in der Kantischen Tradition des philosophischen Denkens ist die Unterscheidung zwischen **Rationalität und Vernunft** üblich und setzt sich bis in die Gegenwart fort. So war John Rawls ursprünglich darauf bedacht, die Theorie der Gerechtigkeit als einen Zweig der rationalen Entscheidungstheorie zu präsentieren – und in der Tat sind es die rationalen Entscheidungen eigeninteressierter und wechselseitig desinteressierter Individuen unter dem sogenannten Schleier des Nichtwissens, die die *fairness* der Prinzipien und der auf diesen Prinzipien beruhenden institutionellen Grundstrukturen sichern – um aber dann doch *rationality* von *reason* zu unterscheiden und die Theorie in ihrer Endfassung als Konzeption öffentlichen Vernunftgebrauchs (*public reasoning*) zu präsentieren. Mir scheint die Unterscheidung zwischen Rationalität und Vernunft wenig sinnvoll zu sein. Damit sind bestimmte Formulierungen terminologisch unzulässig, etwa: „Dies ist zwar rational, aber unvernünftig“. „Rational“ meint in Formulierungen dieser Art in der Regel „im eigenen Interesse“ und dann wäre es in der Tat eine offene Frage, ob es immer vernünftig ist, im eigenen Interesse zu handeln. Aber dieses *Argument der offenen Frage* lässt sich auch hinsichtlich der Rationalität stellen: „Dies mag zwar in deinem eigenen Interesse sein, aber ist es auch rational?“ ist nicht trivialerweise unsinnig. Rationalität mit Eigeninteresse zu identifizieren oder es ausschließlich über Eigeninteresse zu charakterisieren wäre eine begründungsbedürftige inhaltliche Stellungnahme zur Theorie praktischer Rationalität. Rationalitätsaussagen sind *normativ*. Das Eigeninteresse betreffende Feststellungen sind *deskriptiv*. Rationalitätsbehauptungen können daher nicht dasselbe bedeuten wie das Eigeninteresse betreffende Behauptungen. Wer behauptet, eine Handlung sei genau dann rational, wenn sie das Eigeninteresse des Akteurs optimiere, stellt eine normative Behauptung auf, er vertritt eine normative Theorie, die praktische Rationalität als Optimierung des Eigeninteresses kriterial bestimmt.

§2 Vernunft (oder Rationalität) kann nur über **Gründe** bestimmt werden: Eine Handlung A ist rationaler als eine Handlung B, wenn bessere Gründe für A, als für B sprechen. Eine Handlung ist rational (oder vernünftig), wenn sie wohlbegründet ist, wenn gute Gründe für sie sprechen. Damit kann das Argument der offenen Frage angewandt auf das Prädikat *rational* (oder vernünftig) bezüglich jeder kriterialen Theorie auf die Form gebracht werden: Diese Handlung A optimiert das Eigeninteresse der Person P, aber sprechen auch die besseren Gründe für A? So mag es im **Eigeninteresse** eines jungen Mannes sein, der wenig Neigung und Begabung für Erwerbstätigkeit mitbringt, seine Erbtante umbringen zu lassen, aber die gewichtigeren Gründe sprechen dafür, dieses Vorhaben nicht auszuführen. Daher ist die

Handlung unvernünftig. Wenn eine Person meint, eine Handlung sei in ihrem eigenen Interesse, so ist dies lediglich ein *prima facie* Grund, der für diese Handlung spricht. Dieser Grund muss aber gegen andere Gründe, die möglicherweise gegen diese Handlung sprechen, abgewogen werden. Diese Abwägung ist uns lebensweltlich bestens vertraut und es ist keine Gesellschaft menschlicher Wesen vorstellbar, die ohne Abwägungen dieser Art auskommt.<sup>1</sup>

§3 Es ist die Theorie (der Rationalität oder der Ethik)<sup>2</sup>, die sich gegenüber dem Komplex guter Gründe bewähren muss. Die ethische Theorie validiert nicht erst, was gute Gründe sind. Wenn man unter **Rationalismus** philosophische und wissenschaftliche Auffassungen zusammenfasst, die lebensweltlich etablierte Gründe, seien es theoretische Gründe für Überzeugungen oder praktische Gründe für Handlungen, durch philosophische oder wissenschaftliche Prinzipien ersetzen will, dann kann die von mir vertretene Konzeption nicht unter „Rationalismus“ subsumiert werden. Als Alternative zum Rationalismus gilt in der Philosophie der **Intuitionismus**. Lange Zeit war es im angelsächsischen Sprachraum sogar üblich, zwischen einer wissenschaftlichen Ethik des utilitaristischen Typs einerseits und *intuitionistischen* Ethiken deontologischen Typs andererseits zu unterscheiden.<sup>3</sup> Wenn man unter „Intuitionismus“ Auffassungen zusammenfasst, die einen bestimmten Bereich normativer Überzeugungen als unmittelbar, d.h. ohne jedes Raisonement einsichtig und deswegen einer Kritik grundsätzlich entzogen ansehen, dann gibt es zwischen Rationalismus und Intuitionismus ein Drittes. Es ist dann möglich, sowohl gegen die rationalistische Ersetzung lebensweltlich verankerter Gründe als auch gegen die intuitionistische Verabsolutierung bestimmter normativer Überzeugungen Stellung zu nehmen. Gegen rationalistische Begründungsprogramme spricht, dass die Prinzipien, auf denen die deduktiven Schlüsse rationalistischer Theorien aufbauen, selbst einer Begründung bedürfen und gegen den Intuitionismus, dass es keine isolierten (normativen) Überzeugungen gibt, die nicht in bestimmten Situationen mit anderen (normativen) Überzeugungen in Konflikt geraten können. Typischerweise sind intuitionistische Theorien pluralistisch, d.h. sie beharren auf der unmittelbaren Einsichtigkeit (oder Erfahrbarkeit) intrinsischer Qualitäten und optieren daher

---

<sup>1</sup> In den 50er und 60er Jahren galt der sogenannte *good reasons approach* als eine theoretische Alternative sowohl zum Utilitarismus wie zum Kantianismus. Insbesondere in der Variante von Steven Toulmin *An Examination of the Place of Reason in Ethics* 1950 werden die lebensweltlich etablierten praktischen Gründe gegenüber ethischen Prinzipien aufgewertet. Die späteren Schriften Toulmins, insbesondere *Cosmopolis* 1990 zeigen, dass Toulmins Anti-Rationalismus ohne antimoderne und antiaufklärerische Implikationen auskommt.

<sup>2</sup> Auf diese Differenzierung kommen wir noch zu sprechen.

<sup>3</sup> Diese Gegenüberstellung prägt auch die ethische Taxonomie Henry Sidgwick's in *Methods of Ethics* 1874

für einen erkenntnistheoretischen Fundamentalismus (im Sinne von *foundationalism*).<sup>4</sup> Das stärkste Argument gegen intuitionistische Konzeptionen ist die Nicht-Isolierbarkeit einzelner (normativer) Überzeugungen. Jede normative wie deskriptive Überzeugung hängt mit jeder anderen zusammen: In bestimmten Äußerungs-Situationen werden diese simultan relevant, geraten in Konflikt oder ergänzen sich, die eine erscheint als ein Spezialfall der anderen (die eine scheint auf die andere reduzierbar zu sein), beide lassen sich unter eine dritte (allgemeinere) normative oder deskriptive Überzeugung (eine Regel, eine Gesetzmäßigkeit) subsumieren etc. Dabei ist es weder so, dass das Allgemeine jeweils durch das Konkretere begründet wird, noch umgekehrt, dass das Konkretere durch das Allgemeinere begründet wird. Beide Begründungsformen kommen vor und keine von diesen führt zu nicht mehr bezweifelbaren Urteilen. Der pluralistische und im Extremfall partikularistische Intuitionismus optiert für die erste Interpretation, der Rationalismus für die zweite. Der hartnäckige Fortbestand beider konkurrierender Theorie-Optionen lässt sich gut damit erklären, dass beide Begründungsformen lebensweltlich gut etabliert sind.

§4 Diese erkenntnistheoretische Position ist **kohärentistisch**: Sie bringt gegen rationalistische und intuitionistische Auffassungen die Verbundenheit normativer wie deskriptiver Überzeugungen in Stellung. Diese kohärentistische Position darf nicht selbst in die rationalistische Falle tappen. Und sie darf auch nicht so verstanden werden, als beruhe sie auf einer systematischen Theorie der Kohärenz. Dieser Kohärentismus rechtfertigt sich nicht erst durch eine Theorie, auch nicht durch eine Theorie der Kohärenz. Er rechtfertigt sich in alltäglichen Modi des Begründens. Jedes Begründen hat ein Ende. Am Grund allen Begründens steht die praktizierte **Lebensform** als Ganzes. Diese Beobachtung, dass das Begründen irgendwann mal ein Ende hat, ist uns allen aus der Lebenswelt, aber auch aus philosophischen Disputen vertraut. Dieses Ende lässt sich in einem Prinzip oder einem System von Axiomen einer Theorie nicht erfassen. Dies schließt keineswegs aus, dass sich das Gesamt der Begründungsrelationen auf wenige Prinzipien und Inferenzregeln reduzieren lässt. Diese Prinzipien und Inferenzregeln hätten dann aber keinen letzt-begründenden, keinen fundamentalen Status. Sie bewährten sich an den lebensweltlich und disziplinär etablierten Begründungsformen.<sup>5</sup>

---

<sup>4</sup> Der *locus classicus* einer solchen intuitionistischen und zugleich im erkenntnistheoretischen Sinne fundamentalistischen Theorie ist die Wertkonzeption, die George Edward Moore in *Principia Ethica* 1903 entwickelt.

<sup>5</sup> Ich habe in diesem Zusammenhang in Anlehnung an Wittgenstein von „Begründungsspielen“ gesprochen (Vgl. *Demokratie und Wahrheit* 2006, Kap. 3). Dies kann aber das Missverständnis nahe legen, dass es sich lediglich um Spiele, d.h. um so, aber auch anders konzipierbare regelgeleitete Interaktionsformen handele (vgl. V.

§5 Ich habe in diesem Zusammenhang den Begriff der **strukturellen Rationalität** eingeführt. Damit setze ich mich – bei aller Übereinstimmung in der erkenntnistheoretischen Grundhaltung, also in dem, was ich *Wittgensteinsche Perspektive* nenne – von den irrationalistischen Tendenzen, wenn nicht des späten Wittgenstein selbst, so zumindest eines Gutteils seiner Anhänger ab. Die Begründungsspiele stehen nicht unabhängig voneinander. Sie bilden ein komplexes **Netzwerk**, das die alltägliche Praxis, die Interaktionen, Sprachhandlungen und Meinungsäußerungen trägt. Die Akteure dieser Praxis müssen über die Zeit und über ihre Teilnahme an unterschiedlichen Begründungsspielen hinweg erkennbar bleiben, eine **personale Identität** aufweisen. Ihr (öffentliches) Verhalten und ihre (öffentlichen) Meinungsäußerungen müssen einen Sinn ergeben, die orientierenden theoretischen wie praktischen Gründe müssen nachvollziehbar sein, dies schließt eine lediglich punktuelle, auf ein spezifisches Begründungsspiel bezogene Begründung aus. Gründe legen in diese lebensweltliche Praxis **Strukturen** – Strukturen, die sich über Raum und Zeit, über unterschiedliche soziale, kulturelle, temporale, biographische Kontexte hinweg durchhalten. Diese Strukturen sind nicht einfach gegeben, da die Gründe der kritischen Revision zugänglich sind. Das Gesamt dieser Struktur, und damit das Gesamt der epistemischen und konativen Einstellungen propositionaler und nicht-propositionaler Art steht unter der ständigen Prüfung der Stimmigkeit, der **Kohärenz**. Da es ständig Spannungen zwischen verschiedenen Strängen und Knoten dieses Netzwerkes gibt, findet eine kontinuierliche Revision statt. Dieser Prozess ist nicht arational, sondern Ergebnis von praktischen wie theoretischen Deliberationen. Wären die Begründungsspiele lediglich lokal und punktuell, so gäbe es eine lediglich lokale oder punktuelle, aber keine übergreifende Rationalität und das **Projekt der Aufklärung**, also der Kritik von Lebens- und Gesellschaftsformen, würde in der Tat in der Weise kollabieren, wie es manche Wittgensteinianer und postmoderne Theoretiker postulieren. Der Terminus *Strukturelle Rationalität*<sup>6</sup> soll auf diese größeren Zusammenhänge, diese strukturellen Merkmale in erster Linie unserer lebensweltlichen Praxis, aber dann auch unserer lebensweltlichen Theorie hinweisen.

§6 Es scheint keinen größeren Gegensatz als den zwischen dem auf Frank P. Ramsey und Rudolf Carnap zurückgehenden Programm einer Bayesianischen Rationalitäts- und Erkenntnistheorie einerseits und der Wittgensteinschen Perspektive in dem oben erläuterten

---

Gerhardt: *Partizipation. Das Prinzip der Politik* 2007) und daher spreche ich hier statt von „Begründungsspielen“ von „Begründungsformen“.

<sup>6</sup> *Strukturelle Rationalität. Ein philosophischer Essay über praktische Vernunft* 2001.

Sinne andererseits zu geben. In der Tat scheinen hier zwei unvereinbare Paradigmen miteinander verknüpft zu werden: Eine in der Tradition des logischen Empirismus stehende Theorie rationaler Entscheidungen und rationaler Meinungsänderungen (*rational belief dynamics*) und ein an Sprechakt-Theorie und Wittgensteinscher Sprachspiel-Theorie orientiertes pluralistisches Verständnis von Lebensformen. Abgesehen davon, dass wir es hier mit unterschiedlichen philosophischen „Schulen“ zu tun haben, deren Anhänger wechselseitig wenig Respekt füreinander aufbringen, kann ich aber kein Argument dafür erkennen, dass sich diese beiden Perspektiven nicht integrieren ließen. Gefordert ist lediglich, dass der **kohärentistische Charakter der Axiome der Wahrscheinlichkeits- und Entscheidungstheorie** ernst genommen wird und diesen Axiomen keine konsequentialistische Gründe gebende Lesart untergeschoben wird, wie das in der entscheidungstheoretischen Literatur und erst recht in den ökonomischen Anwendungen überwiegend erfolgt. Andererseits muss gegen einen falsch verstandenen Pluralismus von Begründungsspielen die Einheit des (rationalen) Akteurs über unterschiedliche Kontexte hinweg geltend gemacht werden. Diese Einheit verlangt nach Kohärenz – konativer wie epistemischer. Minimalbedingungen von Kohärenz werden in der Bayesianischen Entscheidungs- und Wahrscheinlichkeitstheorie expliziert. Da die lebensweltlich etablierten praktischen Gründe in den seltensten Fällen Handeln lediglich unter Bezugnahme auf seine Konsequenzen rechtfertigen, sind die Strukturen begründeter Praxis nur in sehr seltenen Fällen konsequentialistisch, zumeist hingegen deontologisch zu interpretieren. Diese **deontologische Imprägnierung** unserer lebensweltlichen Praxis verletzt jedoch nicht die minimalen Kohärenzbedingungen der rationalen Entscheidungstheorie. Optimierung ist in den seltensten Fällen das Ziel des rationalen Akteurs. Da die Präferenzen des rationalen Akteurs, auch wenn er deontologischen Gründen folgt, kohärent sind, lassen sich diese Präferenzen in Gestalt einer subjektiven Wertfunktion repräsentieren.<sup>7</sup>

§7 Diese konsequent kohärentistische Perspektive erlaubt eine **spieltheoretische Lesart von struktureller Rationalität**. Diese war sogar der ursprüngliche Taufpate bei der Namensgebung. Ich möchte dies an einem besonders durchsichtigen Beispiel, dem vieldiskutierten *prisoner's dilemma* erläutern. Das Dilemma besteht hier darin, dass ein spieltheoretischer Gleichgewichtspunkt in dominanten Strategien pareto-ineffizient ist, d.h. dass es eine andere Strategien-Kombination gibt, die alle Akteure besser stellt, als in diesem

---

<sup>7</sup> Näher ausgeführt habe ich das in *Why Rational Deontological Action Optimizes Subjective Value*, in: *Protosociology* 21: S. 182-193.

Gleichgewichtspunkt. Diese andere Strategiekombination ist jedoch kein Gleichgewichtspunkt, d.h. es gibt unter der Voraussetzung, dass die anderen Beteiligten bei ihrer Strategie bleiben, jeweils individuell eine Möglichkeit, sich durch Abweichung zu verbessern. Dies gilt für alle an der Interaktion Beteiligten. Die entscheidungstheoretische Orthodoxie quittiert Interaktionssituationen dieses Typs gewissermaßen mit einem Achselzucken: Rationale Akteure wählen dominante Strategien. Die Tatsache, dass die Kombination dominanter Strategien in Interaktionssituationen dieses Typs pareto-ineffiziente Auszahlungen hat, ist bedauerlich, ändert aber nichts an den Rationalitäts-Kriterien. Der empirische Befund, dass ein hoher Prozentsatz von Akteuren in solchen Situationen die dominierte kooperative Strategie wählt, also eine Strategie, die unter der Bedingung, dass auch die anderen Interaktionsbeteiligten ihre jeweils dominierten Strategien wählen, ein pareto-effizientes Ergebnis hat, wird als Zeichen verbreiteter Irrationalität interpretiert. Weniger orthodoxe Entscheidungstheoretiker versuchen, diese Anomalie des realen Entscheidungsverhaltens dadurch einen Anschein des Rationalen zu geben, dass sie es in den sozialen Kontext iterierter Interaktionen stellen<sup>8</sup> was aber die Problematik nicht befriedigend löst. Eine *strukturelle* Sicht auf das *prisoner's dilemma* ergibt jedoch schlagartig ein neues Bild. Betrachten wir die Gründe, die Akteure für ihre Entscheidungen anführen werden. Diejenigen, die sich für die dominante Strategie entscheiden, werden darauf hinweisen, dass sie eine möglichst hohe Auszahlung erreichen wollen und da sie nicht wissen, wie sich der andere Akteur verhält, auf „Nummer sicher“ gehen wollten (dies jedenfalls die Auskunft von Studenten mit denen ich das *prisoner's dilemma*-Spiel durchgeführt habe). Diejenigen, die die kooperative Strategie wählen (also die dominierte) werden z.B. sagen, dass sie die Erwartung hatten, dass auch der andere kooperiert und dass sie ihren Teil zur kooperativen Lösung beitragen wollten. Der orthodoxe Spieltheoretiker interpretiert dies als einen Übergang vom *prisoner's dilemma*-Spiel zum *assurance game*. Er behauptet, dass die Präferenzen nun eben andere seien und damit das *prisoner's dilemma* nicht fortbesteht. Die strukturelle Sichtweise ist eine andere. Sie unterscheidet zwischen den Handlungszielen, die – wie wir annehmen können – weiterhin auf eine möglichst hohe Auszahlung gerichtet sind (zumal wenn es sich um Geldbeträge zwischen zufällig zusammengewürfelten Spielern handelt), fügt jedoch als Bestimmungselement individueller Rationalität die *Struktur der Interaktion* hinzu. Demnach wäre die Information über die unterschiedlichen Auszahlungen an den einzelnen Akteur unvollständig. Dieser sollte zusätzlich die **Struktur der Interaktion** kennen, um

---

<sup>8</sup> Vgl. Axelrod's *Evolution of Cooperation*

gegebenenfalls *kooperativ* handeln zu können, d.h. sein Handeln an dem Motiv der Kooperation ausrichten zu können. Die rationale Entscheidung bedarf einer dichteren Information als lediglich der individuell zu erwartenden Auszahlungen. Dies ist kein Bruch mit dem methodologischen Individualismus, da es mir um die adäquate Beschreibung individueller Handlungsmotive geht. Viele Handlungsmotive beziehen sich eben nicht lediglich auf die zu erwartenden Folgen einer Handlung, sondern auch auf die (strukturelle) Rolle der eigenen Handlung im Kontext der Handlungen anderer. Die Übersetzbarkeit von strukturellen Merkmalen in diesem Sinne in Auszahlungsfunktionen ist ein Vorurteil orthodoxer Entscheidungstheoretiker. Die Konzeption struktureller Rationalität bricht mit diesem Vorurteil, ohne auch nur eines der Axiome der Entscheidungs- und Wahrscheinlichkeitstheorie aufgeben zu müssen. Die Repräsentation der kohärenten Präferenzen einer rationalen Person durch eine quantitative Wertfunktion, also das, was in der Regel als *utility function* bezeichnet wird, ist dann wieder im ursprünglichen Sinne, nämlich strikt kohärentistisch, nicht konsequentialistisch zu verstehen. Präferenzen, sofern sie die Axiome der Entscheidungstheorie erfüllen, lassen sich durch eine reellwertige, bis auf positiv-lineare Transformationen eindeutige Funktion repräsentieren. Die Präferenzen, wie auch immer motiviert, werden **repräsentiert**, es wird nicht die Rationalität von Präferenzen durch ein Kriterium der Optimierung von Handlungskonsequenzen festgestellt. Die Entscheidungstheorie ist gegenüber den Motiven der Akteure *inhaltlich neutral*. Sie ist nicht nur in dem Sinne formal, dass ihre Axiome und Theoreme sich in einer formalen Sprache repräsentieren lassen, sondern auch in dem Sinne, dass sie über die inhaltlichen Bestimmungen der Handlungsmotive nichts aussagt. Sie ist mit allen Typen von Gründen kompatibel, sofern diese in der praktischen Deliberation zu kohärenten Präferenz-Relationen führen.

§8 Das Konzept der strukturellen Rationalität postuliert eine **spezifische Freiheit**, nämlich die des rationalen Akteurs, Interaktionsstrukturen insgesamt und nicht lediglich individuelle Auszahlungsfunktionen seiner Entscheidung zu Grunde zu legen. Ein Akteur, der lediglich Auszahlungsfunktionen optimiert, wäre berechenbar (außer im Falle von Indifferenz). Diese spezifische Freiheit macht den Akteur unberechenbar. Um es paradox zuzuspitzen: Die Möglichkeit, im Einzelfall nicht lediglich individuelle Auszahlungsfunktionen zu optimieren, kann Bedingung des Handlungserfolgs werden, selbst wenn man den Handlungserfolg in Auszahlungen misst. Der konsequente Optimierer seiner individuellen Auszahlungsfunktion

ist z.B. erpressbar.<sup>9</sup> Diese von der Konzeption Struktureller Rationalität postulierte Freiheit geht über die der etablierten spieltheoretischen Kriterien hinaus. Es ist dabei wichtig zu sehen, dass beim Übergang von Spielen gegen die Natur zu Spielen gegen andere rationale Akteure eine wesentliche begriffliche Veränderung stattfindet. In Spielen gegen die Natur postuliert die rationale Entscheidungstheorie lediglich die Optimierung der subjektiven Wertfunktion, die kohärente Präferenzrelationen repräsentiert. Die Gewichtung möglicher Ergebnisse einer Strategiewahl mit Wahrscheinlichkeiten (*natural luck*) wird jedoch in Interaktionssituationen nicht fortgesetzt. Würde es fortgesetzt, so sollte es in jeder Interaktionssituation mindestens eine empfohlene Lösung, also mindestens eine rationale Strategie für jeden einzelnen Teilnehmer geben (so wie in Spielen gegen die Natur). In der spieltheoretischen Analyse wird *implicit* angenommen, dass die Akteure sich nicht wechselseitig Wahrscheinlichkeiten zuschreiben und dann ihre je individuelle Bewertungsfunktion relativ zu diesen subjektiven Wahrscheinlichkeiten optimieren; vielmehr werden nur solche Rationalitätsempfehlungen zugelassen, die sich simultan an alle Interaktionsbeteiligten richten können. Dies ist ja die Rechtfertigung dafür, nur Gleichgewichtspunkte als Lösungen von Spielen zuzulassen. Für die meisten Personalisten (oder Subjektivisten) unter den Wahrscheinlichkeitstheoretikern hat jedes Ereignis aus der Perspektive der handelnden Person eine (subjektive) Wahrscheinlichkeit, die durch den maximalen Wettquotienten eruiert werden kann. Insofern müsste auch das Ereignis, dass ein bestimmter Spieler eine bestimmte Strategie wählt, eine subjektive Wahrscheinlichkeit aus der Sicht jedes einzelnen Interaktionsbeteiligten haben. Diese subjektiven Wahrscheinlichkeitszuschreibungen könnten neben den Kolmogorov-Axiomen weiteren Restriktionen unterworfen sein, die sich auch aus dem Spiel-Format ergeben könnten. Diesen Weg ist die Spieltheorie bis heute nicht gegangen. Ich denke, aus gutem Grund. Denn Personen lassen sich nicht in dieser Weise kalkulieren, sie rekurren nicht nur auf erwartete Strategien Anderer, sondern auch auf die Erwartungen Anderer bezüglich eigener Erwartungen etc., ein iterativer Prozess, der grundsätzlich unbegrenzt ist.

---

<sup>9</sup> Die auf wechselseitiger nuklearer Abschreckung beruhende relative Stabilität des Kalten Krieges beruhte auf der wechselseitig unterstellten ‚Irrationalität‘, nach einem wie auch immer gearteten Angriff, selbst mit den überlegenen konventionellen Panzerarmeen des Warschauer Paktes an der deutsch-tschechischen Grenze, bereit zu sein, durch den Einsatz von Nuklearwaffen den eigenen Untergang in Kauf zu nehmen. Diese ‚Irrationalitätsannahme‘, wurde in den Zeiten der Reagan-Administration durch den Aufbau von *Theater Nuclear Forces* (taktische im Gegensatz zu strategischen Kernwaffen) auf dem potentiellen europäischen Kampffeld zurückgenommen oder wenigstens abgeschwächt, während die französische Nuklear-Strategie bis heute an ihr festhält.

Die Unberechenbarkeit rationaler Akteure ist Voraussetzung der Rationalitätskriterien der Spieltheorie.<sup>10</sup>

## II. Was ist Freiheit?

§9 Unsere alltäglichen Intuitionen menschliche Freiheit betreffend sind besonders stark hinsichtlich einer subjektiven Erfahrung ausgeprägt, die ich als diejenige der **Willkürfreiheit** bezeichnen möchte. Jemanden, der behauptet mein Handeln sei determiniert, glaube ich damit widerlegen zu können, dass ich ihn um eine Prognose bitte, was ich in der nächsten Minute tun werde – etwa sitzen bleiben oder aufstehen – und dann jede seiner Prognosen widerlege. Oder kurz: Ich *entscheide*, ob ich mich jetzt erhebe, oder noch eine Minute länger sitze. Für solche Situationen in denen wir willkürlich so oder auch anders entscheiden können, ist typisch, dass das Abwägen von Gründen keine Rolle spielt. Es ist egal, ob ich sitzen bleibe oder aufstehe und gerade hier scheint sich die menschliche Freiheit besonders eindringlich zu manifestieren. Dieses Argument lässt sich natürlich dadurch unterlaufen, dass man die geäußerte Prognose als kausale Ursache der jeweiligen Zuwiderhandlung annimmt, was die Iteration allerdings nur auf die nächste Stufe hebt. Diese Vermutung der kausalen Determination ließe sich wieder durch die entgegengesetzte Handlung ‚widerlegen‘. Dieses Phänomen der Unberechenbarkeit menschlichen Handelns, das in Willkür-Situationen besonders augenscheinlich wird, sollte in der philosophischen Analyse ernst genommen werden. Es lässt sich, wie mir scheint, nicht adäquat in Gestalt eines Zufallsoperators analysieren, der jeweils bei Indifferenz der subjektiven Bewertungen des Akteurs zur Anwendung käme, etwa in Gestalt des Leibniz-Kriteriums: Wenn es keinen Grund gibt, eine von zwei Optionen für wahrscheinlicher zu halten, dann halte beide für gleich wahrscheinlich. Der willkürlich Handelnde erscheint von außen wie von einem Zufallsoperator gesteuert. Die Innenperspektive, die Perspektive der ersten Person ist eine andere. Es ist egal, was ich tue, aber was ich tue, ist unter meiner Kontrolle. Es ist kein Zufallsoperator, der festlegt, was ich jeweils tue. Für Willkürfreiheit spielen Gründe keine Rolle. Die Kontrolle des eigenen Handelns minimiert sich auf die freie (durch Gründe unterbestimmte) Entscheidung, also die Ausbildung einer handlungsvorausgehenden Absicht, die durch die Handlung selbst erfüllt wird.

---

<sup>10</sup> Das Newcombsche Paradoxon ergibt sich aus der Annahme, dass das Entscheidungsverhalten rationaler Akteure grundsätzlich vorhersehbar ist. Entsprechend löst sich das Newcombsche Paradoxon auf, sobald man diese Annahme fallen lässt. Kurz: Die Leugnung menschlicher Freiheit führt in eine schwer auflösbare Paradoxie. Vgl. Campbell und Sowden 1985.

§10 Entscheidungen kann man als Abschluss einer **Deliberation** charakterisieren. Zwischen Entscheidung und erfüllender Handlung findet in der Regel keine Abwägung mehr statt. Sollte sie dennoch stattfinden, dann wird die Entscheidung aufgehoben, dispensiert und gegebenenfalls eine neue Entscheidung getroffen. Entscheidungen bilden einen spezifischen Typus von Intentionen (Absichten) aus. Zwischen diesen Absichten und den sie erfüllenden Handlungen kann eine unterschiedlich große Zeitspanne liegen. Im einen Extrem beendet eine Entscheidung weitere Abwägungen für Jahre und die dann erfüllende Handlung wird aus der Ich-Perspektive als lange zuvor festgelegt wahrgenommen. Im anderen Extrem folgt die Handlung unmittelbar der Entscheidung, ja *a limine* ist zwischen Entscheidung und Handlung die Zeitspanne nicht mehr groß genug, um die Entscheidung auf Dispens setzen zu können. Es gibt eine neurophysiologisch bestimmte untere Grenze dieser Zeitspanne. Die berühmten gewordenen Libet-Experimente spielen sich in diesem Bereich ab. Es handelt sich zweifellos um Willkür-Entscheidungen, da es den Versuchsteilnehmern egal ist, zu welchem Zeitpunkt sie ihren Finger bewegen und die Ausführung der Entscheidung folgt jeweils unmittelbar, d.h. an der unteren Grenze des (neuro-)physiologisch Möglichen. Ob die in den Versuchen reportierten Zeitpunkte der Entscheidung verlässlich sind, ist in der neurophysiologischen Literatur umstritten. Die starke Varianz der Angaben und die Tatsache, dass die reportierten Zeitpunkte der Entscheidung sich wiederum auf Wahrnehmungen eines Zeigerstandes beziehen, lassen jedenfalls Zweifel an der Verlässlichkeit dieser Angaben aufkommen. Eine Abweichung in der Größenordnung einer Zehntelsekunde würde ausreichen, um dem Libet-Experiment jede Relevanz für die Willensfreiheits-Debatte zu nehmen.<sup>11</sup>

§11 In *Über menschliche Freiheit* habe ich die Willkürfreiheit vollständig aus der Betrachtung ausgeschlossen und die Rolle handlungsleitender Gründe in den Mittelpunkt der Analyse gestellt.<sup>12</sup> Der Titel gibt einen Hinweis darauf, warum: Es ging mir um das Spezifikum *menschlicher* Freiheit. Ich glaube nicht, dass nicht-menschliche Tiere allesamt als (algorithmische) Maschinen funktionieren, die einem genetisch verankerten Programm folgend auf sensorische Stimuli – ggf. modifiziert durch die Geschichte vorausgegangener

---

<sup>11</sup> Die Tatsache, dass außer dem Libet-Experiment und seiner Nachfolger keine weiteren experimentellen Anordnungen entwickelt werden konnten, die die These einiger (weniger) Neurophysiologen stützen, dass sich die Existenz menschlicher Willensfreiheit (genauer sollte man sagen Handlungsfreiheit) durch empirische Daten heute widerlegen lasse, macht stutzig. In letzter Zeit werden auch innerhalb der neurophysiologischen Forschung die Vorbehalte gegen die übliche Interpretation der Libet-Befunde deutlicher artikuliert. (Vgl. zu den Libet-Experimenten Kane 2002, insbesondere die Beiträge von Benjamin Libet und Henrik Walter.)

<sup>12</sup> Allerdings kann man auch die existenziellen Entscheidungen, deren zentrale Rolle für die Lebensgestaltung ich im letzten Kapitel von *Strukturelle Rationalität* analysiert habe, als eine Form von Willkürfreiheit interpretieren - zumindest hat diese existenzielle Freiheit etwas mit der oben benannten Willkürfreiheit gemeinsam: sie ist nicht – primär – von Gründen bestimmt.

sensorischer Stimuli – in jeweils eindeutig festgelegter Weise reagieren. Selbst wenn diese Programme mit Zufallsoperatoren angereichert werden, scheinen mir solche mechanistischen Interpretationen des Tier-Verhaltens unzureichend zu sein. Zumindest höher entwickelte Säugetiere scheinen ihr Verhalten zumindest teilweise zu kontrollieren und in jedem Fall über so etwas wie Willkürfreiheit zu verfügen. Das Abwägen von Gründen und die langfristige Strukturierung des eigenen Lebens über akzeptierte Handlungs-Gründe ist Tieren jedoch fremd. Diese zentrale Rolle der Gründe – Gründe, die dem Leben eine Struktur geben<sup>13</sup> und die die zwischenmenschlichen Interaktionen generell und unsere Verständnispraxis speziell tragen – ist eine, wenn nicht *die conditio humana*, sie macht das Besondere der menschlichen Existenz aus. Sich von Gründen affizieren zu lassen, Gründe zu nehmen und Gründe zu geben, scheinen mir zentrale Charakteristika des Humanen<sup>14</sup> zu sein. Entsprechend lässt sich die Position eines **theoretischen Humanismus** dadurch charakterisieren, dass sie dem Gründe-Geben und Gründe-Nehmen einen zentralen Stellenwert in der Analyse menschlicher Praxis beimisst. Dieser humanistische Ansatz stellt sich gegen naturalistische, deren Analyse auf die Eskamotierung von Gründen gerichtet ist. In der vollendeten (wissenschaftlichen) Erklärung menschlichen Verhaltens dürfen nach naturalistischer Auffassung Gründe keine Rolle spielen. Gründe spielen demnach keine eigenständige, sondern bestenfalls eine abgeleitete Rolle, in manchen Varianten des Naturalismus ist diese Rolle in deterministische Kausalzusammenhänge der Neurophysiologie übersetzbar. Unter ‚Naturalismus‘ werden unterschiedliche metaphysische Positionen und Forschungsprogramme subsumiert. Eine Gemeinsamkeit ist jedoch die Überzeugung, dass in der naturwissenschaftlichen Beschreibungsform alles Wesentliche über die Welt einschließlich der Rolle des Menschen in ihr, ausgesagt werden kann. Der Humanismus, wie ich ihn verstehe, hält dagegen, dass menschliches Handeln ohne eine Bezugnahme auf Gründe nicht vollständig beschreibbar und erklärbar ist und dass die Rede von Gründen sich nicht ohne Verlust in naturwissenschaftliche Beschreibungsformen (weder der Physik, noch der Biologie, noch der Neurophysiologie) übersetzen lässt. Diese These der Unübersetzbarkeit darf nicht im Sinne eines Aspekt-Dualismus einer einheitlichen Welt gelesen werden. Die Beschreibung menschlichen Handelns, menschlichen Lebens, menschlicher Verständigungspraxis unter Rekurs auf Gründe beinhaltet mehr als die Verwendung einer zweiten Sprachform oder den Wechsel auf eine andere linguistische Ebene. Auch der weiche Naturalismus, wie er kürzlich von Jürgen

---

<sup>13</sup> vgl. *Strukturelle Rationalität* Kap. 4.

<sup>14</sup> Vgl. *Über menschliche Freiheit* Kap. 1

Habermas vertreten wurde,<sup>15</sup> ist mit der Grammatik unserer Alltagskommunikation nicht in Einklang zu bringen. Diese Grammatik hat u. a. die ontologische Präsupposition, dass **Gründe eine kausale Rolle** für unser Verhalten spielen. Wenn Gründe Ursachen sind oder genauer formuliert, wenn das Akzeptieren von Gründen und die vorausgehende Deliberation kausal relevant sind, also einen Einfluss auf den Weltverlauf haben, dann wäre der Weltverlauf anders, wenn anstelle dieser *andere* Gründe vom Akteur akzeptiert worden wären. Damit A Ursache von B ist, muss A für B unverzichtbar sein. Diese Unverzichtbarkeit muss auf hinreichende Bedingungen von B relativiert werden. Es gibt also einen Komplex von Bedingungen, die zusammen hinreichend für B sind und in dem A notwendig ist. Ursachen sind jeweils für sich genommen in der Regel keine notwendigen Bedingungen für das Ereignis, das sie verursachen. Es gibt andere Konstellationen von Bedingungen, die für B hinreichend sind und in denen A nicht vorkommt.<sup>16</sup> Gründe werden kausal relevant nur darüber, dass sie von Akteuren akzeptiert werden. Die kausale Wirksamkeit dieses Akzeptierens von Gründen erfolgt durch Handlungen. Handlungen sind – nicht nur, aber auch – raum-zeitliche Vorgänge, in der Regel Körperbewegungen. Dass Körperbewegungen eine kausale Rolle im Weltverlauf spielen, kann als unstrittig gelten. Strittig ist jedoch die Verbindung zwischen Körperbewegungen und Gründen (dem Akzeptieren von Gründen). Das Akzeptieren von Gründen äußert sich in der Ausbildung bestimmter konativer Einstellungen, nicht umgekehrt.<sup>17</sup> Im Sinne des üblichen (naturalistischen) Kausalitätsverständnisses<sup>18</sup> sind Körperbewegungen, oder allgemeiner der physikalische Aspekt unseres Handelns für den Weltverlauf kausal relevant, sie stellen (natürliche) Ursachen für weitere (natürliche) Ereignisse dar.

**§12** Kann es eine kausale Verbindung zwischen dem nicht-natürlichen Ereignis des Akzeptierens eines Grundes und dem natürlichen Ereignis einer Körperbewegung (im Sinne eines physikalisch-biologischen Prozesses) geben? Die Antwort muss je nach Kausalitätsbegriff, je nach dem Verständnis von Kausalität unterschiedlich ausfallen. Im Sinne einer **Regularitäts-Theorie der Kausalität**, die noch bis vor wenigen Jahren die analytische Philosophie dominierte,<sup>19</sup> fällt es schwer, einen solchen Kausalzusammenhang zu rekonstruieren. Es wäre erforderlich, eine Sukzessions-Regularität zwischen dem Akzeptieren

---

<sup>15</sup> dieser weiche Naturalismus ist als ein Versöhnungsversuch von „epistemischem Dualismus“ und „ontologischem Monismus“ zu verstehen, vgl. Habermas 2006.

<sup>16</sup> Vgl. Mackie 1974.

<sup>17</sup> Die entgegengesetzte Position vertritt Bernard Williams 1981.

<sup>18</sup> Das z.B. W.V.O.Quine, nicht aber Donald Davidson zu Grunde legt.

<sup>19</sup> Vgl. Stegmüller 1983, Kap. VII.

eines Grundes und bestimmten Körperbewegungen zu beobachten. Mit anderen Worten: Es wäre ein Gesetz erforderlich, das es erlaubte, aus dem nicht-natürlichen Ereignis des Akzeptierens eines Grundes das natürliche Ereignis einer raum-zeitlichen Körperbewegung abzuleiten. Solche Gesetzmäßigkeiten können sich nicht auf *token*, also einzelne Vorkommnisse des Akzeptierens eines Grundes und einzelne Vorkommnisse einer Körperbewegung, beziehen, sondern nur auf *types*. Es scheint aber völlig aussichtslos zu sein, Gründe eines bestimmten Typs mit Körperbewegungen in einen gesetzesartigen Zusammenhang zu bringen.<sup>20</sup> Günstiger steht es um eine kausale Beziehung zwischen Gründen und Körperbewegungen im Rahmen einer **probabilistischen (und epistemischen) Theorie der Kausalität**. Eine bestimmte Körperbewegung erscheint ohne vorausgehende Intentionen des Akteurs unverständlich, ihre Apriori-Wahrscheinlichkeit ist 0 oder sehr niedrig. Bei Kenntnis eines akzeptierten Handlungsgrundes steigt die (epistemische) Wahrscheinlichkeit für dieses raum-zeitliche Ereignis. Dabei ist natürlich wesentlich, dass das natürliche Ereignis der Körperbewegung als Bestandteil einer Handlung interpretiert wird. Handlungen aber sind duale Ereignisse, sie haben einen äußeren und einen inneren Aspekt, eine physische und eine psychische Komponente. Bloße Körperbewegungen sind keine Handlungen. Körperbewegungen werden zu Handlungen oder können als äußere Form einer Handlung interpretiert werden, wenn sie durch geeignete Intentionen hervorgerufen und begleitet sind. Muss dann nicht auch die Verbindung zwischen diesen (handlungs-)konstitutiven Intentionalitäten kausal sein? Reintegriert die Annahme einer kausal bestimmten Intentionalitätsdynamik aber dann nicht doch Gründe in eine naturalistische Gesamtperspektive? Mir scheint folgende Sichtweise am plausibelsten: Die Deliberation selbst, **das Abwägen von Gründen ist kein kausaler Prozess**. Diese Deliberation wird jedoch in der natürlichen Welt kausal wirksam dadurch, dass diese Deliberation zum Abschluss gebracht wird. Dieser Abschluss trägt in der Regel den Namen „Entscheidung“. Zwischen Entscheidung und erfüllender Handlung besteht in vielen Fällen erneut die Möglichkeit des intentionalen Eingriffs, der dann die Entscheidung aufhebt (dispensiert, obsolet macht). Unter Normalbedingungen heißt „sich entscheiden“ jedoch, die

---

<sup>20</sup> Wer dennoch Gründe für Ursachen hält, wie etwa Donald Davidson, wird in der Regel dem sog. anomalen Monismus zugerechnet, eine Terminologie, die nur dann Sinn macht, wenn der „Monismus“ durch Kausalitätsbeziehungen konstituiert wird. Die Welten des Mentalen unter Einschluss der Gründe und Präferenzen und die Welten des Physischen sind dann über Kausalitätsbeziehungen gewissermaßen vereint. Wer jedoch für den Bereich des Physischen eine Regularitätstheorie der Kausalität vertritt oder stärkere, realistische Kausalitätsbegriffe zugrunde legt, wie etwa Mario Bunge oder Karl Popper, ist dann im zweiten Schritt in einem erneuten Dualismus gefangen, da erst einsichtig gemacht werden müsste, was es rechtfertigt, mit zwei so unterschiedlichen Kausalitätsbegriffen zu operieren.

Deliberation abzuschließen, keine weitere Abwägung von Gründen vorzunehmen. Im Rückblick sprechen wir im Falle eines Dispenses dann auch davon, dass ich zwar „dachte, mich schon entschieden zu haben, dann aber feststellte, dass ich mich doch noch nicht entschieden habe“ oder ähnlich. Aber auch dann, wenn die Entscheidung ultimativ ist, wenn sie nicht mehr dispensiert wird, kommt erneut Intentionalität in Gestalt der Kontrolle desjenigen Verhaltens zustande, das diese Entscheidung als vorausgehende Absicht, realisiert. Grob beschrieben ist in dieser letzten Form von Intentionalität (vor der Handlungsausführung) ein Suchprozess eingeschaltet, der die geeignete Form der Ausführung bestimmt. Dieser Suchprozess ist selbst wieder von Gründen gesteuert; Gründe aber, die nicht mehr darauf gerichtet sind, die Entscheidung selbst zu rechtfertigen oder zu überprüfen. Die „Gerichtetheit“ bezieht sich auf die geeignete Mittelwahl, betrifft also eine ganz andere Form von Normativität als diejenige, die zur Entscheidung geführt hat. Am Ende resultiert dieser komplexe, hier nur ganz grob beschriebene intentionale Prozess in einem konkreten, raumzeitlich lokalisierten und mit den Mitteln der Naturwissenschaft beschreibbaren **Verhalten**.

**§13** Unsere Freiheitsintuitionen haben nur dann ein *fundamentum in re*, wenn die (praktischen) Gründe, die unsere handlungsrelevante Intentionalität prägen, kausal relevant sind. Wenn das Ergebnis der Deliberation vor aller Deliberation immer schon festläge, etwa durch einen mit naturwissenschaftlichen Mitteln beschriebenen Weltzustand und die wirkenden Naturgesetze allein, dann müssten unsere Freiheitsintuitionen als – vielleicht – nützliche Illusionen gelten. Dies ist jedenfalls die These der **Freiheit als naturalistischer Unterbestimmtheit von Gründen**. Da wir – auffälligerweise – viel eher geneigt sind, theoretischen Gründen, also Gründen, etwas zu glauben, von etwas überzeugt zu sein, Freiheit im Sinne naturalistischer Unterbestimmtheit zuzugestehen, ist die weitgehend analoge Rolle theoretischer und praktischer Gründe eine Stützung dieser Freiheitstheorie.<sup>21</sup> Die Vorstellung, theoretische Gründe nach dem Muster physikalischer und deterministischer Kausalität zu erfassen, würde voraussetzen, dass es einen algorithmischen Prozess gibt, der dieses Wirken abbildet.<sup>22</sup> Da aber schon die Theoreme der Prädikatenlogik erster Stufe algorithmisch nicht beweisbar sind, ist diese Vorstellung unplausibel. Zudem geriete sie in einen Konflikt mit der Annahme, dass menschliches Wissen für den weiteren Weltverlauf, etwa über die Anwendung von wissenschaftlichen Ergebnissen in Technik, Wirtschaft und Gesellschaft, kausale Effekte hat.

---

<sup>21</sup> Vgl. „Zur Reichweite theoretischer Vernunft in der Ethik“ in *Ethische Essays*.

<sup>22</sup> Vgl. „Über menschliche Freiheit“, in Liessmann 2007, S. 30 ff.

### III. Metaphysische Aspekte

§14 Atomistische und fundamentalistische Interpretationen lebensweltlicher, wie auch wissenschaftlicher Überzeugungssysteme sind mit der tatsächlichen Begründungspraxis – der lebensweltlichen wie der wissenschaftlichen – nicht in Einklang zu bringen. Epistemologien, die diese Merkmale aufweisen, bleiben entweder abstrakte Gedankenspiele oder greifen gewaltsam in die etablierte Begründungspraxis ein. Das ursprüngliche Programm der *ideal language philosophy*, die logizistische Rekonstruktion der Mathematik und die Axiomatisierung naturwissenschaftlicher Theorien kann heute als gescheitert gelten, womit natürlich keineswegs gesagt sein soll, dass formale Sprachen in der Philosophie nichts zu suchen hätten. Weit offenkundiger ist die Inadäquatheit atomistischer und fundamentalistischer Epistemologien für lebensweltliche Überzeugungssysteme, etablierte Verständigungspraxis und Alltagssprache. Die Alternative ist **kohärentistisch** und **holistisch**: Jede einzelne Überzeugung hat eine unendliche Vielfalt von Präsuppositionen und die Begründungsrelationen sind zu komplex, um eine Einteilung in eine Klasse begründender und eine Klasse begründeter Überzeugungen plausibel erscheinen zu lassen. Wenn im Folgenden einige metaphysische Aspekte des Verhältnisses von Vernunft und Freiheit angesprochen werden, so ist dies vor dem Hintergrund einer kohärentistischen und holistischen epistemologischen Perspektive zu sehen und d.h. unter anderem, dass es keine strikte Trennung zwischen metaphysischen und nicht-metaphysischen oder empirischen Aspekten geben kann. Auch das transzendente Programm, wie es etwa bei Peter Strawson als „deskriptive Metaphysik“ entwickelt wurde, hat sich von den fundamentalistischen Ursprüngen der abendländischen Metaphysik-Tradition noch nicht konsequent genug abgelöst, wie mir scheint. Dennoch ist der Ausdruck „deskriptive Metaphysik“ hilfreich. Die geteilte Verständigungspraxis wird zum Ausgangspunkt der philosophischen Analyse. Der Philosoph tritt nicht von außen an die etablierten Formen des Begründens heran, um diese zu validieren (nach einem Maßstab, der dann letztlich im Dunkeln bleiben muss); vielmehr beschränkt er sich auf die Klärung der Rolle, die bestimmte, besonders zentral erscheinende Überzeugungen für diese Verständigungspraxis *de facto* haben. Da diese Verständigungspraxis jedoch Inkohärenzen aufweist, kann diese Form der zurückhaltenden philosophischen Analyse dennoch normative Kraft entfalten. Manche Überzeugung, die revidierbar erschien, kann sich als zu zentral erweisen, um mit Gründen bezweifelt werden zu

können und andere Überzeugungen können sich angesichts ihrer Unvereinbarkeit mit jenen als unplausibel herausstellen.<sup>23</sup> Wenn man im Rahmen einer so verstandenen kohärentistischen Epistemologie dennoch von metaphysischen Aspekten spricht, dann ist das **gradualistisch** zu verstehen. Metaphysische Aspekte betreffen besonders abstrakte Annahmen, vor allem solche, die erst in der philosophischen Analyse als Präsuppositionen oder Verallgemeinerungen konkreterer Überzeugungen zu Tage treten.

**§15** Die Charakterisierung menschlicher Freiheit als naturalistische Unterbestimmtheit unserer Deliberationen hat einen solchen, in diesem schwachen Sinne metaphysischen Charakter. **Ob diese Form einer menschlichen Freiheit real ist, lässt sich nicht durch einen empirischen Test feststellen.** Auch die Gegenthese, dass Menschen in diesem Sinne nicht frei seien, dass diese Form menschlicher Freiheit nicht existiere, kann nicht für sich genommen einem empirischen Test unterzogen werden. Die These ist lediglich in dem Maße plausibel, als sie sich gut in das Gesamt unserer Sprach- und Interaktionspraxis einbetten lässt. Die These ist ihrer Gegenthese überlegen, wenn sie sich besser als diese einbetten lässt. Die Entscheidung zwischen diesen beiden Thesen ist unmöglich, wenn sich beide gleichermaßen einbetten lassen oder (eine ganz andere Möglichkeit) wenn unsere Sprach- und Interaktionspraxis so starke Inkohärenzen aufwies, dass beide Optionen bestehen: die Revision dieser Sprach- und Interaktionspraxis auf der Grundlage der These mit dem Ergebnis, dass diese ihre innere Widersprüchlichkeit überwindet und zugleich ihre Revision auf der Grundlage der Gegenthese ebenfalls mit dem Resultat, dass sie ihre innere Widersprüchlichkeit überwindet. Ich glaube nicht, dass es sich so verhält, obwohl die nun in unserem Kulturkreis seit über 2000 Jahren anhaltende Auseinandersetzung um Freiheit und Determinismus eine solche Aporie nahe zu legen scheint. Ich glaube, wie die vorstehenden Argumente deutlich zu machen versuchen, dass es eine Sichtweise gibt, die unser Selbstbild mit unserem wissenschaftlichen Weltbild so verbindet, dass es nicht zu einer derart fundamentalen inneren Widersprüchlichkeit kommt. Ich glaube mit anderen Worten nicht, dass wir zu einer dieser beiden Revisionen gezwungen sind und uns dezisionistisch für den einen oder anderen – jeweils metaphysisch motivierten – Umbau unserer Überzeugungssysteme entschließen müssen. Aber ich gebe zu, dass allein die Möglichkeit einer solchen epistemischen Situation beunruhigend ist. Wer, wie ich, für einen Ontologie-

---

<sup>23</sup> *Das Projekt der Aufklärung* wurde oft als Distanzierung von unseren lebensweltlichen Überzeugungen missverstanden, ja in extremen Varianten, als Neu-Konstruktion *ab ovo*, beginnend bei nicht mehr bezweifelbaren Axiomen und endend in den empirischen Einzel-Anwendungen, die durch Deduktion gewonnen werden. Dieses Missverständnis reicht von Decartes über Leibniz und Spinoza bis zu Husserl, Lorenzen und Hare.

freien, „unaufgeregten“ Realismus plädiert, wer epistemische Wahrheitsdefinitionen für inadäquat hält, muss auf die Konvergenz unserer Begründungsspiele hoffen, muss hoffen, dass die Vielfalt der Begründungstypen in ihrer jeweiligen Praxis-Verankerung nicht zu divergenten, jeweils lediglich intern kohärenten Überzeugungssystemen führen. Die Forderung der Inklusion als einer diskursethischen Norm<sup>24</sup>, dieses Prinzip des Aufklärungsprojektes erweist sich damit als unverzichtbar für einen nicht-metaphysischen (ontologiefreien) Realismus<sup>25</sup>.

**§16** Im Sinne einer so verstandenen „deskriptiven Metaphysik“: Was genau präsupponiert unsere alltägliche Verständigungspraxis? Wenn sie, wie Strawson meint, Freiheit präsupponiert, dann stellt sich die Anschlussfrage, welches Verständnis von Freiheit hier genau präsupponiert wird. Kommt diese Freiheit ohne alternative Möglichkeiten aus und ist damit Determinismus-verträglich? Die Strawson'sche Analyse<sup>26</sup> hat eine, wie wir es nennen können, *sentimentalistische Einseitigkeit*: Sie fokussiert auf moralische Gefühle und reaktive Einstellungen und blendet die Gründe, die diese moralischen Gefühle und reaktiven Einstellungen leiten, aus der Analyse aus. Dadurch entsteht die problematische Gegenüberstellung einer subjektiven, von Zusammengehörigkeitsgefühlen (*attachment*) geprägten und einer objektiven, den anderen als bloßes Objekt der Manipulation ansehenden, Perspektive.<sup>27</sup> Ich habe gegen diesen Sentimentalismus geltend gemacht, dass die Verbindung von Freiheit und moralischer Missbilligung über Gründe hergestellt wird.<sup>28</sup> Genauer: dass es die wechselseitige Zuschreibung der Fähigkeit, Gründe abzuwägen und nach Gründen zu handeln, ist, die uns als moralische Akteure erscheinen lässt, die daher moralische Missbilligung und andere moralische Gefühle und reaktive Einstellungen rechtfertigt. Der Austausch von Gründen, wie er in unserer Lebenswelt etabliert ist, setzt jedoch die **Existenz objektiver normativer Kriterien** voraus: Es gibt eben gute und schlechte Gründe und ob es ein guter Grund oder ein schlechter Grund ist, entscheidet sich nicht an den Gefühlslagen, sondern letztlich an den besseren Argumenten. Dass diese Argumente selbst von Gefühlslagen nicht völlig unabhängig sind, dass Interessen, Empfindlichkeiten, insbesondere in die für die menschliche Existenz so wesentliche Selbstachtung eingehen (müssen), liegt auf

---

<sup>24</sup> Vgl. Habermas 1996.

<sup>25</sup> Vgl. Hilary Putnam 2004, der unterdessen – möglicherweise auch unter dem Einfluss der John-Dewey-Lektüre – nach vielen Wendungen und Irrfahrten nun genau diese Position einzunehmen scheint. Habermas votiert dagegen für eine epistemische Wahrheitsdefinition im Bereich des Normativen und für einen „metaphysischen“ Realismus im Bereich des Empirischen (vgl. Habermas 1999, insbesondere Kapitel 6).

<sup>26</sup> Vgl. Strawson 1962 (Wiederabdruck Strawson 1974).

<sup>27</sup> Dieser Sentimentalismus verwundert, wenn man bedenkt, dass Strawson einen Gutteil seiner Philosophie in engem Anschluss an Kantische Erkenntnistheorie entwickelte.

<sup>28</sup> *Über menschliche Freiheit*, Kap. 1.

der Hand.<sup>29</sup> Damit rücken die praktischen und die theoretischen Deliberationen, die unser Handeln und unsere Meinungsbildung anleiten, ins Zentrum der Analyse, während die den Austausch von Gründen begleitenden Gefühlslagen die Analyse komplettieren. Mit dieser notwendigen Verschiebung ist aber der Weg frei für eine präzisere Bestimmung der spezifischen Freiheit, die unsere lebensweltliche Verständigungspraxis präsupponiert. Die Freiheit und Verantwortung, die wir uns wechselseitig zuschreiben, ist Ausdruck der wechselseitigen Anerkennung als Akteure, die ihre je eigenen Gründe haben, die für dieses oder jenes Handeln, für diese oder jene Überzeugung sprechen und denen wir zumuten können, dass sie sich so weit von ihren Augenblicksneigungen, ihrer Sozialisation, ihrer genetischen und epigenetischen Prägung distanzieren können, dass sie jedenfalls den überwiegenden Teil ihrer Überzeugungen und Handlungen am Ergebnis der Abwägung von Gründen – theoretischen, wie praktischen – ausrichten. Mit dieser Auskunft ist die spezifische Freiheit, die wir in unserer Verständigungs- und Interaktionspraxis voraussetzen, schon näher charakterisiert: Sie verlangt, dass das Ergebnis unserer Deliberationen vor aller Deliberation nicht immer schon feststeht, dass unsere **Deliberationen** ausschlaggebend sind – man könnte auch sagen **kausal relevant** – für das, was wir tun und was wir glauben. Verantwortung tragen Menschen für ihre Handlungen, aber auch für ihre Überzeugungen und denjenigen Teil ihrer Gefühle, die einen kognitiven Gehalt aufweisen.

§17 Es ist ein unverzichtbarer Teil unseres lebensweltlichen Orientierungswissens, dass Menschen Gründe für ihre Handlungen und ihre Überzeugungen haben und auch für manche ihrer Gefühle. Grenzen wir dieses Phänomen zunächst etwas näher ein, um die Integration von Gründen in unser Weltbild vorzubereiten. Manche Menschen haben eine Spinnen-Phobie, d.h. sie haben Angst vor Spinnen, laufen weg, wenn sie Spinnen sehen, wenden den Blick ab, wenn sie Photographien von Spinnen sehen, obwohl sie wissen, dass von Spinnen in unseren Breiten keine Gefahr ausgeht. An diesem Beispiel lässt sich das Verhältnis von Gründen sehr schön illustrieren. Die Person weiß, dass sie für dieses besondere Gefühl der Spinnenangst keinen Grund hat, dass von Spinnen *de facto* keine Gefahr ausgeht. Dennoch kann sie dieses Gefühl nicht unterdrücken. Nennen wir Gefühle dieser Art, die einen kognitiven Gehalt haben, sich von Gründen jedoch nicht beeinflussen lassen, pathologische. Gefühle dieser Art *erleidet* man, sie lassen sich nicht kontrollieren (Personen, die mit solchen pathologischen Gefühlen konfrontiert sind, erleben dies auch als eine Art Ohnmacht oder Hilflosigkeit). Was begründet ist und was verursacht ist, klafft auseinander. Die Erfahrung des Kontrollverlustes

---

<sup>29</sup> *Über menschliche Freiheit*, Kap. 5.

hängt damit zusammen, dass sich das kognitiv Eingesehene (von Spinnen geht keine Gefahr aus) in den Gefühlslagen nicht adäquat niederschlägt. Es gibt Strategien, um in solchen Fällen die Kontrolle zurück zu gewinnen, während des Essens Fotos von Spinnen zu betrachten soll im Laufe der Zeit helfen. Was geschieht hier? Die Person weiß um die Irrationalität ihres Gefühls, eines Gefühls, das einen kognitiven Gehalt hat (es gibt auch andere Formen der Irrationalität von Gefühlen) und sie wählt eine Verhaltensstrategie, die den kausalen Nexus zwischen sensorischen Stimuli und Gefühlslagen (hier dem sensorischen Stimulus in Spinnengestalt zum Gefühl der Phobie) unterbricht, um den Gründen die gewünschte Wirkung auf Gefühlslagen erst zu verschaffen. Man könnte auch sagen, diese **Einsicht wird kausal wirksam**. Wir haben es also mit einer Konkurrenz zwischen der kausalen Wirksamkeit eines nicht von Gründen gesteuerten Prozesses (der Prozess, der zwischen den sensorischen Stimuli der Spinnengestalt und dem Gefühl der Angst abläuft) und dem kausalen Prozess, der zwischen der Einsicht (Ungefährlichkeit der Spinnen) und dem Gefühl (entspannte Reaktion auf Spinnenwahrnehmung) abläuft, zu tun. In dieser Interpretation wird – hoffentlich – deutlich, dass eine Zwei-Aspekte-Metaphysik nicht befriedigen kann. Der kausale Prozess, der von der Wahrnehmung einer Spinne zur pathologischen Spinnenangst führt, steht im Konflikt zu den Gründen, die gegen Spinnenphobie sprechen. Es ist nicht zutreffend, dass (naturwissenschaftlich grundsätzlich erfassbare) Kausalität prinzipiell nicht in Konflikt mit Gründen kommen kann, da diese einer anderen linguistischen Ebene angehören. Es ist nicht zutreffend, dass die intentionalistische Beschreibung menschlichen Handelns mit der kausalistischen schon deswegen nicht in Konflikt kommen kann, weil die ontologischen Präsuppositionen beider Beschreibungen voneinander unabhängig sind. Dieses einfache Beispiel zeigt, dass diese ontologische Unabhängigkeit nicht besteht. Wenn der Kausalitätsbegriff nicht an nomologische Erklärungsformen der Naturwissenschaften gebunden wird, kann dieser Konflikt auch folgendermaßen zugespitzt werden. Spielen Gründe (genauer: Deliberations-Ergebnisse) eine kausale Rolle? Oder kürzer und unter Umgehung des Kausalitätsbegriffes: Sind Gründe wirksam, haben Gründe eine Wirkung auf das was wir glauben, das was wir tun und das was wir fühlen? Wir können annehmen, dass die kausale Wirksamkeit von Gründen über neurophysiologische Prozesse vermittelt ist. Auch konkrete Deliberationen, das konkrete Abwägen von Gründen, realisiert sich in neurophysiologischen Prozessen, wie wir annehmen können. Aber die Regeln guten Begründens ergeben sich nicht aus naturwissenschaftlichen, auch nicht aus psychologischen Gesetzmäßigkeiten. Es gibt keinen Grund, hinter die Psychologismus-Kritik Freges und Husserls zurückzufallen. Wenn logische Inferenzregeln nicht als psychologische

Gesetzmäßigkeiten interpretiert werden können, so *a fortiori* nicht als neurophysiologische. Die Welt 3, von der Popper in seinen späteren Schriften redet, die Welt der logischen Inferenzregeln, des Wissens, der Begründungen, der Theorien lässt sich weder auf die naturalistische Welt, bestehend aus natürlichen Tatsachen, die grundsätzlich mit begrifflichen Mitteln der Naturwissenschaft beschreibbar und erklärbar sind, noch auf die Welt der subjektiven Gegebenheiten, der mentalen Zustände, der Dispositionen, der Wünsche und Überzeugungen reduzieren. Die Theorie des naturalistischen Fehlschlusses, die *naturalistic fallacy* von George Edward Moore, deren Begründung in *Principia Ethica* unzureichend ist, die aber bis heute vom ganz überwiegenden Teil der Ethiker und Philosophen akzeptiert ist, markiert nur einen Spezialfall dieser These der Nicht-Reduzierbarkeit. Die Inadäquatheit einer epistemischen Definition deskriptiver oder normativer Wahrheit scheint mir das entscheidende Kriterium dieser Nicht-Reduzierbarkeit des Logischen (im weitesten Sinne), also nicht nur Inferenzregeln, sondern auch Propositionen umfassend, auf das Mentale zu sein, so wie die Nicht-Definierbarkeit von Qualia durch physische (oder ‚natürliche‘) Prädikate ein umfassendes Kriterium für die Nicht-Reduzierbarkeit des Mentalen auf das Physische markiert. Wenn wir etwa davon sprechen, dass wir diese Gründe eingesehen haben oder dass eine Person zwar glaubte einen Grund zu haben, tatsächlich aber keinen guten Grund hatte, etwas zu tun, dann ordnen wir Gründe (in diesem Sprachgebrauch) der logischen Sphäre zu. In diesem Sinne kann man auch von ‚objektiven‘ oder ‚guten‘ Gründen sprechen.

**§18** Es kann gute Gründe geben, etwas zu tun oder zu unterlassen, ohne dass es auch nur eine Person gibt, die diese Gründe einsieht. Die Ermordung einer unschuldigen Person kann auch dann ungerechtfertigt sein, keinen guten Grund für sich haben, wenn es niemanden gibt, der dies erkennt. Alle diese Behauptungen sind auf den alltäglichen Sprachgebrauch gestützt. Sie rekurrieren nicht auf eine davon unabhängige Metaphysik oder Ontologie. Ja, wir sollten der Versuchung widerstehen, diese Befunde einer deskriptiven Metaphysik zu ontologisieren. Der deskriptive und normative Objektivismus, der in unsere alltägliche Verständigungs- und Interaktionspraxis eingelassen ist, lässt eine Reduktion der (objektiven) Gründe auf (subjektive) Motive, wie es der Hume’schen Tradition in der praktischen Philosophie entspricht, nicht zu. Dieser **Objektivismus** ist unaufgebbar, wenn man an der etablierten Verständigungspraxis festhält und diese nicht *in toto* verwerfen will. Eine solche Verwerfung wäre nicht möglich; derjenige der dies versucht, würde sich außerhalb der Verständigungsgemeinschaft stellen und keine Chancen auf Gehör haben, außer vielleicht in sektenartigen Klüngeln von „Halbirren“, um den Wittgenstein’schen Ausdruck zu

gebrauchen.<sup>30</sup> Die Isolierung, d.h. die Abkoppelung von lebensweltlichen Verständigungsprozessen, ist aber dann die Voraussetzung. Gute Philosophie und gute Wissenschaft sind jedoch mit der lebensweltlichen Verständigungspraxis kompatibel, im besten Falle verbunden durch Regeln der lebensweltlichen Anwendung und komplementär im Sinne eines erweiterten lebensweltlichen Orientierungswissens.<sup>31</sup> Epistemische Revolutionen gibt es in der Wissenschaft,<sup>32</sup> sie ziehen in der Regel keine epistemischen Revolutionen unseres lebensweltlichen Orientierungswissens nach sich. Die besonders unter Naturwissenschaftlern verbreitete Vorstellung, dass irriige Alltagsvorstellungen von der Wissenschaft revidiert werden, gilt nur in marginalen Bereichen lebensweltlichen Urteilens. Selbst der Übergang vom geo- zum heliozentrischen Weltbild erfordert keine Revisionen lebensweltlichen Orientierungswissens. Die metaphysischen, religiösen und generell weltanschaulichen Folgen, die wissenschaftliche Revolutionen nach sich ziehen können, stehen dagegen auf einem ganz anderen Blatt. In der Tat führen besonders erfolgreiche naturwissenschaftliche Forschungsprogramme meist einen weltanschaulichen Überschuss mit sich, der unzulässige Verallgemeinerungen auf Bereiche, für die diese wissenschaftlichen Theorien nicht gedacht sind, beinhaltet. Dies gilt für den Sozial-Darwinismus des 19. und 20. Jahrhunderts, für die vermeintlichen anti-ethischen Implikationen der Soziobiologie und für die Fundamentalkritik der Freiheit und Verantwortung, die manche aus zeitgenössischen neurophysiologischen Befunden glauben ableiten zu können. Es ist eine irreführende Vorstellung zu meinen, dass die Naturwissenschaften unser lebensweltliches Orientierungswissen erst validieren. Dies gilt weder in historischer, noch in systematischer Betrachtungsweise. In historischer nicht, weil unser lebensweltliches deskriptives wie empirisches Orientierungswissen eine erstaunliche Invarianz über die Zeiten aufweist, ohne damit obsolet zu werden. Die sog. *folk psychology*, also die wechselseitige Zuschreibung von mentalen Eigenschaften, Dispositionen und Intentionen trägt das alltägliche Beziehungs- und Interaktionsgefüge und es ist weder durch die Entwicklung der Psychologie zu einer empirischen Wissenschaft seit der vorletzten Jahrhundertwende, noch durch die jüngste Entwicklung der Neurophysiologie in ihrer Orientierungsleistung erschüttert. Diese Feststellung ist hier deswegen wichtig, weil das gesamte Projekt einer im oben erläuterten

---

<sup>30</sup> Dass die eine oder andere Philosophenschule im Laufe der abendländischen Geistesgeschichte diesem Charakteristikum entsprochen haben mag, will ich gar nicht bestreiten. Ich bestreite aber, dass die Diskurse, die im geschützten Bereich eines Oberseminars oder einer Akademie möglich sind, diesem entsprechen.

<sup>31</sup> Vgl. „Lebensweltliches Orientierungswissen“ in Nida-Rümelin 2002.

<sup>32</sup> Thomas S. Kuhn hat diese in eindringlicher Form beschrieben, ohne jedoch die wissenschaftstheoretischen Implikationen zureichend zu erfassen. Die irrationalistischen Interpretationen vieler seiner Anhänger sind jedenfalls nicht zwingend.

Sinne *deskriptiven* Metaphysik in Frage gestellt wäre, träfe diese Feststellung nicht zu. Wenn unser lebensweltliches Orientierungswissen als Ruinenfeld längst vergangener wissenschaftlicher Theorien mit Anbauten aus neueren wissenschaftlichen Theorien gelten müsste, könnte auch die Analyse der begrifflichen, normativen und deskriptiven Präsuppositionen unserer lebensweltlichen Verständigungs- und Interaktionspraxis bestenfalls die Fundamente dieses Ruinenfeldes und seiner Anbauten freilegen. Warum sollte man ein systematisches Interesse an diesem Unterfangen haben, wenn doch alle Hoffnung auf die Aufklärungsleistung der empirischen Einzelwissenschaften gerichtet ist? Das Projekt einer deskriptiven Metaphysik setzt – *implicit* – die Zurückweisung des Szientismus voraus.

## Literatur

- Axelrod, R. (1984), The Evolution of Cooperation. New York.
- Campbell, R. und L. Sowden (1985), Hrsg., Paradoxes of Rationality and Cooperation: Prisoner's Dilemma and Newcomb's Problem. Vancouver.
- Habermas, J. (1996), Die Einbeziehung des Anderen. Studien zur politischen Theorie. Frankfurt am Main.
- Habermas, J. (1999), Wahrheit und Rechtfertigung. Philosophische Aufsätze. Frankfurt am Main.
- Habermas, J. (2006), *Das Sprachspiel verantwortlicher Urheberschaft und das Problem der Willensfreiheit: Wie lässt sich der epistemische Dualismus mit einem ontologischen Monismus versöhnen?*, in: Deutsche Zeitschrift für Philosophie 54: S. 669-707.
- Kane, R. (2002), Hrsg., The Oxford Handbook of Free Will. Oxford.
- Liessmann, K. P. (2007). Hrsg., Die Freiheit des Denkens. Wien
- Mackie, J. L. (1974), The Cement of the Universe. A Study of Causation. Oxford.
- Moore, G. E. (1903), Principia Ethica. Cambridge.
- Nida-Rümelin, J. (2001), Strukturelle Rationalität. Ein philosophischer Essay über praktische Vernunft. Stuttgart.
- Nida-Rümelin, J. (2002), Ethische Essays, herausgegeben von J. Nida-Rümelin. Frankfurt (Main): S. 96-112.
- Nida-Rümelin, J. (2005), Über menschliche Freiheit. Stuttgart.
- Nida-Rümelin, J. (2006), Demokratie und Wahrheit. München
- Putnam, H. (2004), Ethics Without Ontology. Cambridge, Mass. ; London.
- Sidgwick, H. (1874), The Methods of Ethics., London
- Stegmüller, W. (1983), Probleme und Resultate der Wissenschaftstheorie und analytischen Philosophie, Bd. I: Erklärung, Begründung, Kausalität. Berlin.
- Strawson, P. F. (1962), *Freedom and Resentment*, in: Proceedings of the British Academy 98: S. 187-211.
- Strawson, P. F. (1974), *Freedom and Resentment*, in: Freedom and Resentment, herausgegeben von P. F. Strawson. London: S. 1-25.
- Toulmin, S. (1986), The Place of Reason in Ethics. Chicago.

Toulmin, S. E. (1990), Cosmopolis: The Hidden Agenda of Modernity. New York.

Williams, B. (1981), *Internal and External Reasons*, in: Moral Luck: Philosophical Papers 1973-1980, herausgegeben von B. Williams. Cambridge: S. 101-113.