Elemente der Philosophie des Geistes und der Wissenschaftstheorie (2)[1]

Reduzierbarkeit und freier Wille

U. Mortensen

[1] KognitionFreierWillRekonst.pdf, 18. 01. 2026, letzte Korrektur 12. 09. 2026

Inhaltsverzeichnis

1 Zur Philosophie des Freien Willens **4**

1.1 Freier Wille, Moral, und die Wissenschaft 4

1.1.1 Determinismus und Kausalität 6

1.1.2 Physik, Kausalität, und freier Wille 9

1.1.3 Libets Experimente, die Rolle des Bereitschaftspotentials . 18

1.2 Einige philosophische Ansätze 24

1.2.1 David Hume 24

1.2.2 Kant, Schopenhauer, Nietzsche 25

1.2.2.1 Immanuel Kant 25

1.2.2.2 Arthur Schopenhauer 36

1.2.2.3 Friedrich Nietzsche 37

1.2.3 Martin Heisenbergs Befunde 41

1.2.4 Ursachen, Gründe, Wittgenstein 43

1.2.5 Externalismus, Enaktivismus, und freier Wille 46

1.2.5.1 Frederik I. Dretske 48

1.2.5.2 Jürgen Habermas 49

1.2.5.3 Julian Nida-Rümelin 67

1.2.5.4 Alva Noë 72

1.2.6 Max Plancks Betrachtungen 75

1.2.7 Seth Lloyds rekursive Argumentation 80

1.2.8 Freier Wille und moralische Verantwortung 82

1.2.9 Galen Strawsons Folgerungen 85

1.2.10 Albert Einsteins Bekenntnis 86

1.2.11 Moritz Schlicks Argumentationen 86

1.3 Philosophie und Empirie 89

1.3.1 Allgemeine Betrachtungen 89

1.3.2 Der mereologische Fehlschluss 91

1.3.3 Entscheidungsmodelle 95

1.3.4 Philosophie versus Empirie 96

2 Anhang **109**

2.1 Turingmaschinen 109

2.2 Turings Halte-Problem 110
2.3 LLoyds Beweis 112

Literatur **114**

Index **126**

Die Wissenschaft denkt nicht.[2]
Martin Heidegger (1951)

1 Zur Philosophie des Freien Willens

1.1 Freier Wille, Moral, und die Wissenschaft

Die Frage, ob der menschliche Wille frei ist oder nicht, wird spätestens seit Parmenides (520/515 – 460/455 v.Chr.) kontrovers diskutiert (O'Connor et al. (2022)). Die Frage ist offenbar weniger ein Thema empirischer Forschung als ein Thema philosophischer Analyse und insofern kein klassisches Thema der Wissenschaftstheorie. Gleichwohl ist die These der Willensfreiheit auch von wissenschaftstheoretischer Bedeutung, weil u.a. die Rolle von Definitionen und Annahmen deutlich wird, und nach Ansicht einiger Philosophen, aber auch einiger Wissenschaftler, die Existenz eines freien Willens sogar Voraussetzung für die Wissenschaft (Zeilinger (2010)) ist. Aber auch Neurowissenschaftler vertreten die These, der Wille sei frei, und begründen diese Sicht mit einem Hinweis auf neurobiologische Prozesse (verg. Zitat Z 69, Seite 98).

Umgangssprachlich scheint die Bedeutung des Ausdrucks 'freier Wille' klar zu sein, – der Wille eines Menschen ist "frei", wenn der Mensch nicht "unfrei" entscheiden muß, d.h. wenn keinerlei Zwänge die Entscheidung zwischen den jeweils möglichen Alternativen bestimmen. Allerdings ergeben sich schon hier verschiedene Möglichkeiten, den Ausdruck "Zwang" zu interpretieren. Philosophen wie David Hume betrachten den Willen als frei, weil sie unter 'Zwang' äußere Bedingungen verstehen, die Menschen sich so entscheiden lassen, wie es ihrem Willen *nicht* entspricht; mit der Kausalität hat Hume schon deswegen kein Problem, weil er sie für ein illusionäres Konstrukt hält[3]. Andere Philosophen wie etwa Arthur Schopenhauer und Friedrich Nietzsche sehen auch die Rolle "innerer" Bedingungen, die den Willen einer Person beeinflussen oder gar determinieren und kommen zu dem Schluß, der Wille sei nicht frei. Wieder andere Philosophen wie Immanuel Kant gehen davon aus, dass natürliche Prozesse wegen der angenommenen allgemeinen Gültigkeit des Kausalprinzips determinierte Prozesse seien, gedankliche Prozesse dagegen dem Kausalprinzip nicht unterlägen. Kant folgert die Freiheit der Gedanken aus der Fähigkeit, moralisch handeln zu können, und muß dann, um die Annahme der Freiheit der Gedanken beibehalten zu können, postulieren, dass es eine Zweite Welt, die Welt der Noumena (der Gedanken) geben muß, in der kognitive Prozesse ablaufen, die das Treffen von Entscheidungen ausmachen. Diese Annahme hat ihren Preis: sie fordert eine Antwort auf die Frage, wie denn

[2] Heidegger, M. (1951) Im Vortrag "Was heißt Denken?". Gemeint ist damit nach Heidegger "Man kann nicht mit den Methoden der Physik sagen, was die Physik ist. Sondern was die Physik ist, kann ich nur denken." Roth (2004) zitiert den Münsteraner Philosophen Joachim Ritter, der der Ansicht war, dass Naturwissenschaftler die natürliche Welt zwar erforschen können, die Interpretation der Daten aber den Philosophen überlassen müssen.

[3] In seinem Werk Treatise on Human Nature (1739), Erstes Buch, Dritter Teil – Knowledge and Probability

gedankliche Prozesse initiiert werden, wenn sie eben keine Kausalstruktur haben. Kant kennt die Frage und gibt eine merkwürdige Antwort: die gedankliche Dynamik (das ist nicht sein Ausdruck, es soll hier damit nur ausgedrückt werden, dass Gedanken ja auseinander folgen können) ergäbe sich aus der Spontaneität der Vernunft[4]. Der Begriff der Vernunft schwebt hier wie *deus ex machina* aus dem philosophischen Himmel herab[5], und inwiefern erklärt der Ausdruck 'Spontaneität' die gedankliche Aktivität? In der modernen Theorie der Systeme, sofern sie äquivalent der Theorie der Differentialgleichungen ist, werden ebenfalls Wechselwirkungen verschiedener Variablen beschrieben, ohne dass der Begriff der Kausalität verwendet wird, aber daran hat Kant sicherlich nicht gedacht. Sein Begriff der Spontaneität hat etwas Unerklärtes, Vages, Apodiktisches, das zumindest bei einer Reihe von Lesern Unzufriedenheit erzeugt. Nietzsche hielt die Idee der *causa suis*, die sich selbst erzeugende Ursache schlicht und überzeugend für eine Art "logischer Notzucht", worauf weiter unten noch eingegangen wird. Der Biologe Martin Heisenberg argumentiert allerdings, die biologische Möglichkeit der *causa suis* nachgewiesen zu haben; auch hierauf wird im Folgenden näher eingegangen. Neuere Philosophen versuchen, die Kantsche dualistische Metaphysik der akausalen Noumena zu vermeiden, indem sie einen "epistemischen Dualismus" (Habermas) oder "Common-sense-Dualismus" (Nida-Rümelin) einführen und argumentieren, dass Kognitionen gar nichts mit der Tätigkeit des Gehirns zu tun haben, sondern irgendwie außerhalb des Kopfes stattfinden, und dass Entscheidungen nicht auf Ursachen, sondern auf Gründen beruhen; eine noch radikalere Variante dieses Ansatzes liefert der Philosoph Alva Noë, der u.a. die Verleihung von Nobelpreisen für Medizin[6] als verfehlt betrachtet, denn die Wahrnehmung finde gar nicht im Kopf statt. Die Argumente dieser Philosophen werden kurz vorgestellt, ebenso wie einige Philosophen, die der Hypothese des freien Willens skeptisch gegenüber stehen.

So kommt man auf die Frage, ob die als Hypothese aufgefasste Behauptung, der Wille sei frei, nicht empirisch entschieden werden kann. Kant hatte die Frage schon negativ beantwortet: der Begriff des freien Willens sei ein transzendentales Konstrukt ohne Ort in der Empirie. Auch diese Position wird diskutiert werden; der Begriff des freien Willens löst sich gewissermaßen auf, je detaillierter er analysiert wird. Zu dieser Schlußfolgerung ist auch Max Planck gekommen, nachdem er über Jahrzehnte hinweg versucht hat, die These des freien Willens begründen zu können, und in der gleichen Weise argumentierte der Physiker und Philosoph Moritz Schlick (1930) schon vor Planck. In der empirischen Psychologie findet man denn auch keine Versuche, die Freiheit des Willens nachzuweisen, dafür aber Untersuchungen zum Ablauf von Willensbildungen bzw. Entscheidungen. Zum Abschluß werden einige neurobiologische Befunde zur Entstehung des Bewußtseins, die eng mit dem bewußten Willen einhergeht, vorgestellt.

[4] "Der Verstand ist ein Vermögen der Spontaneität des Denkens". Kritik der reinen Vernunft, A51/B75)

[5] Anhänger der Philosophie Kants sehen das nicht so!

[6] speziell die Verleihung des Nobelpreises 1981 an David Hubel, Thorsten Wiesel und Roger Sperry für ihre Arbeiten zur Informationsverarbeitung im visuellen System.

1.1.1 Determinismus und Kausalität

Der klassischen Vostellung über die Rolle des Kausalitätsprinzip entsprechend entsprechend wird unter einem Naturgesetz eine deterministische Aussage verstanden. Ein einfaches Beispiel ist Galileis Gesetz des freien Falls: Wird ein Körper zum Zeitpunkt $t = 0$ fallen gelassen, so hat er nach t Sekunden die Strecke $s(t) = \frac{1}{2}gt^2$ Meter durchfallen (wenn er bis dahin nicht durch ein Hindernis gestoppt wurde; $g \approx 9.81[m/sec^2]$ ist die Gravitationskonstante, und da die Masse m des Körpers nicht in die Formel eingeht, spielt die Masse keine Rolle, d.h. verschieden schwere Körper fallen gleich schnell. Eine Nebenbedingung ist, dass der Körper reibungsfrei fällt. Dies bedeutet, dass das Gesetz nur streng gilt, wenn der Körper im Vakuum fällt. Gibt man die Fallhöhe s vor, so folgt, dass der Körper $t = \sqrt{2s/g}$ Sekunden fällt. Die Fallzeit ist eine deterministische Funktion von s, Abweichungen vom vorhergesagte Wert t gelten als Messfehler, z.B. Fehler, die sich beim Ablesen der Uhr ergeben.

Die Frage ist, ob psychologische Prozesse – und damit Entscheidungsprozesse – durch Naturgesetze beschreibbar sind. Gemeint sind dabei *strikte Gesetze*. Sie beschreiben Zusammenhänge zwischen Ereignissen, die überall und für jede Zeit gelten, d.h. unter identischen Bedingungen wird dieselbe Ereignisfolge beobachtet; gelegentlich wird auch gefordert, dass es sich bei strikten Gesetzen auch um determministische Gesetze handelt, aber die Deterministik ist keine notwendige Bedingung für Striktheit; so gibt es in der Quantenmechanik auch strikte Wahrscheinlichkeitsaussagen, etwa die Bornsche Regel, derzufolge das Quadrat des Betrags der Wellenfunktion die Wahrscheinlichkeitsdichte dafür angibt, dass ein Teilchen an einem bestimmten Ort mit einem bestimmten Messwert angetroffen wird. Srikte Gesetze werden von *Ceteris-paribus-Gesetzen* (c.p.-Gesetze) unterschieden, d.h. von Gesetzen, die *unter sonst gleichen Bedingungen* gelten. c.p.-Gesetze beschreiben den Effekt (falls es einen Effekt gibt) einer oder mehrerer experimentelle Variable auf bestimmte abhängige Größen, wenn alle anderen möglichen Einflußgrößen konstant gehalten werden. Mit c.p.-Gesetzen hat man es insbesondere bei der Untersuchung von komplexen Systemen zu tun, wie etwa in der Psychologie.

Man kann strikte Gesetze wie folgt charakerisieren (Gadenne (2004), p. 124): Gegeben seien Ereignisse vom Typ A und B. Ein Gesetz ist ein striktes, deterministisches Gesetz $G(A \rightarrow B)$, wenn ein Ereignis vom Typ A stets zu einem Ereignis vom Typ B führt. Dabei kann A aus Teilereignissen $A_1, A_2, \ldots$ zusammengesetzt sein. Solche Gesetze heißen auch *Sukzessionsgesetze*, im Unterschied zu *Zustandsgesetzen*, denen zufolge bestimmte Variablenausprägungen stets miteinander einhergehen, ohne dass die Ausprägung der einen die der anderen nach sich zieht oder bewirkt. *Kausalgesetze* sind stets Sukzessionsgesetze: wenn A das Ereignis B bewirkt, so muß A dem Ereignis B vorangegangen sein. Damit hat man eine allgemeine Definition des Determinismus:

Definition 1 *Die Beziehung zwischen den Ereignissen A und B heißt* deter-

ministisch*, wenn zu jedem Ereignis B ein vorausgehendes Ereignis A existiert derart, dass das Gesetz $G(A \to B)$ gilt.*

Gadenne (p. 125) verweist darauf, dass es hierbei wesentlich ist, dass mit dem Ausdruck 'Ereignis' wirklich *alles* gemeint ist, "auch die kleinste Änderung des kleinsten Teilchens, und weiterhin jedes mentale Ereignis einschließich aller menschlicher Gedanken, Entscheidungen und Handlungen. Sie können z.B. determiniert sein durch vorangehende mentale Ereignisse oder durch Gehirnvorgänge."

Über die Definition 1 lassen sich auch die Begriffe *Erklärbarkeit* und *Vorhersagbarkeit* bestimmen. Ist ein Gesetz $G(A \to B)$ gegeben, so kann man das Auftreten von B *erklären*, indem man zeigt, dass A vorangegangen ist. Wenn überdies A das Ereignis B logisch impliziert, so kann man aus der Beobachtung von B logisch folgern, dass A vorgelegen haben muß. Das ist aber im Allgemeinen nicht der Fall: im Allgemeinen gilt der Schluß $B \to A$ *nicht*, im Allgemeinen kann man nur von $\neg B$ (nicht-B: B wurde unter gegebenen Bedingungen nicht beobachtet) auf $\neg A$ folgern, denn A ist u.U. nur eine von mehren Bedingungen, die B implizieren. Jedenfalls bedeutet der Determinismus, dass aus der Existenz von Gesetzen der Form $G(A \to B)$ eine *prinzipielle Erklärbarkeit* (Gadenne, p.125) folgt, – aber eben nur eine prinzipielle, denn man muß das Ereignis A auch tatsächlich finden. Man kann das Auftreten von B *vorhersagen*, wenn man zeigt, dass jetzt A vorliegt. Kann man A nicht beobachten, so kann man B nicht vorhersagen und man hat auch keine Erklärung für die Beobachtung B. Es ist eben so, dass man im Allgemeinen kein erklärendes Ereignis A findet oder die relevanten Anfangsbedingungen A nicht vollständig bestimmen kann, – d.h. dass die Welt größtenteils unerklärlich erscheint. Zu Kants Zeiten schien die newtonsche Mechanik eine Möglichkeit für die vollständige Erklärbarkeit zumindest der physikalischen Wirklichkeit zu liefern. Mittlerweile hat die Entwicklung der Mechanik gelehrt, dass diese Vermutung sich nicht halten läßt. Insbesondere bei nichtlinearen Modellen (Gesetzen) physikalischer Prozesse – das sind Prozesse, bei denen sich die Effekte der verschiedenen, im Gesetz auftretenden Variablen nicht nur additiv überlagern – kann es zu Interaktionen kommen derart, dass sich kleinste Effekte zu immer größeren aufschaukeln und die Vorhersage des Zustands des Systems (der Zustand ($s(t)$ eines Systems zum Zeit ist durch die Werte $(w_1(t), \ldots, w_n(t))$ aller Variablen V_j, $1 \leq j \leq n$) zur Zeit t gegeben – trotz deterministischer Gleichungen nicht mehr möglich ist. Das schon seit Kepler und Kopernikus bekannte *Dreikörperproblem* ist ein Beispiel: die Aufgabe ist, den Bewegungsverlauf dreier Körper (Planeten) unter dem Einfluß des newtonschen Gravitationsgesetzes zu finden. Das gelingt bestenfalls numerisch, und jedenfalls nur approximativ. Probleme dieser Art werden u.a. in der Chaos-Theorie diskutiert, hier gibt es trotz deterministischer Gleichungen grundsätzliche Grenzen bei der Erklärbarkeit und Vorhersagbarkeit.

Für einen großen Bereich von Naturgesetzen gilt, dass sie gar nicht strikt sind, weil die Prozesse, auf die sie sich beziehen, nicht deterministisch, sondern *probabilistisch* sind. So läßt sich schon die Bildung von Aktionspotentialen ("Spikes")

eines Neurons nicht deterministisch, sondern "nur" probabilistisch darstellen. Die Aktivierung ist ein stochastischer Prozess: die Stimulierung durch einen auf das Neuron treffenden Impulses versetzt die Natrium-Kalium-Pumpe in eine andere Gangart. Generell werden Natrium-Ionen aus dem Inneren der Zelle herausgepumpt und Kalium-Ionen werden hineingepumpt, bis ein kritisches Potential von - 70 mV erreicht wird, das ein Aktionspotential auslöst, woraufhin der Prozess aufs neue beginnt, aber diese Aktivität kann intensiviert oder eher inhibiert werden. Der Verlauf des Potentials vom Zeitpunkt der Stimulation (Depolarisation: + 30 mV) bis - 70 mV ist "zufällig"[7], wie er bei thermodynamischen Prozessen zu erwarten ist[8]. Gesetzesaussagen haben dann die allgemeine Form

$$P(B|A) = p, \quad p \in [0, 1], \tag{1}$$

wobei $P(B|A)$ die bedingte Wahrscheinlichkeit des Ereignisses B, gegeben das Ereignis A, ist, und p ist ein spezieller Wert aus dem Intervall $[0, 1]$. Man muß stets bedenken, dass 'zufällig' nicht automatisch "akausal" bedeutet, 'zufällig' bedeutet nur, dass sich der tatsächliche Verlauf der Berechenbarkeit entzieht. Gadenne (p. 16) spricht denn auch von einem *aufgeklärten Determinismus*, wenn probabilistische Gesetze in die Menge möglicher Gesetzesaussagen einbezogen werden.

Oft wird der Begriff der Kausalität um den Begriff der *probabilistischen Kausalität* erweitert: Das Ereignis A ist (probabilistisch) Ursache von B, wenn

$$p(B|A) > p(B|\neg A), \tag{2}$$

wobei $\neg A$ = nicht-A, wenn also B eine größere Wahrscheinlichkeit hat, wenn A gegeben ist, als wenn A nicht gegeben ist. Die "einfache" Kausalität ist offenbar ein Spezialfall der probabilistischen Kausalität: A verursacht B genau dann, wenn $p(B|A) = 1$.

Dazu ist zu sagen, dass diese Definition von Kausalität, egal ob einfach oder probabilistisch, die Komplexität dieses Begriffs ein wenig verschleiert. Bereits David Hume argumentierte[9], dass wir tatsächlich nur eine zeitliche Abfolge zweier Ereignisse A und B beobachten und wir nur aus reiner Gewohnheit von A auf B schließen, der Schluß von A als Ursache von B aber logisch gar nicht gerechtfertigt sei, man könne nicht logisch schließen, dass die Zukunft gleich der Vergangenheit sein wird, – dies ist das berühmte Induktionsproblem. Hermann von Helmholtz (1821 – 1894) hat 1878[10] das Kausalgesetz als "notwendiges Gesetz unseres Denkens" bezeichnet, das er allerdings im Unterschied zu Kant, der es als a priori gegeben annahm, als erworben betrachtete. Einen sehr einflußreichen Aufsatz[11]

[7] d.h. ist nicht exakt berechenbar

[8] s. a. Alan Turings Reaktions-Diffusions-Gleichungen (1952), beschrieben u.a. in Murray (1989), Abschnitt 9.2

[9] in A Treatise of Human Nature, (1739/1740), und überarbeitet in *An Enquiry concerning Human Understanding (1748)*

[10] In: Die Tatsachen in der Wahrnehmung

[11] On athe notion of cause (1912)

hat Bertrand Russell[12] im Jahr 1912 verfasst. Nach Russell ist der Begriff (bzw. das Gesetz) der Kausalität ein Relikt aus vergangener Zeit[13]. Man muß aber beachten, dass sich Russell besonders auf die Physik bezog. Physikalische Gesetze werden im Allgemeinen in Form mathematischer Gleichungen geschrieben, in denen, wie er schreibt, die Kausalität gar nicht als Variable auftaucht, – es gehe um funktionale Zusammenhänge zwischen Variablen. 'Kausalität' sei nur eine ungenaue Beschreibung ohne klare Bedeutung. Diese Argumentation läßt sich leicht auf neurophysiologische Prozesse übertragen.

Natürlich ist Russells Kritik nicht unkritisiert geblieben. Russell habe den Begriff zu eng ausgelegt; Heidelberger (1992) hat Russells Arbeit in Bezug auf neuere Ansätze ausführlich diskutiert. Für die Diskussion um den freien Willen sind noch folgenden beiden Ansätze relevant:

(i) Lewis (1973, 1979) Kontrafaktische Kausalitätstheorie: ein Ereignis C verursacht E, wenn E ohne C nicht eingetreten wäre. Das heißt: Wenn $\neg C$ (kontrafaktische Aussage), dann $\neg E$.

(ii) Strukturelle Kausalmodelle (Strukturgleichungsmodelle): Hier werden durch Gleichungen komplexe Beziehungen zwischen Variablen modelliert.

Die Liste ist nicht vollständig, es soll nur ein erster Eindruck der Möglichkeiten vermittelt werden, kausale Beziehungen zu charakterisieren, auch wenn philosophischerseits behauptet wird, bei miteinander wechselwirkenden Kognitionen seien nur Gründe wirksam.

In folgenden Abschnitt wird auf die Rolle der Physik in der Diskussion um die Freiheit des Willens eingegangen. Die Frage ist insbesondere, ob die Physik notwendig durchgängig kausal strukturiert ist oder nicht: einige Philosphen wie etwa J. Habermas (s. Abschnitt 1.2.5.2 gehen davon aus, dass die Naturwissenschaften generell von der Gültigkeit des Kausalprinzips ausgehen, und da sie – die Philosophen – von der Freiheit des Willens überzeugt sind, lehnen sie eine Diskussion im Rahmem der Naturwissenschaft ab.

1.1.2 Physik, Kausalität, und freier Wille

Der Physiker Werner Heisenberg (1923), hat zuerst sich selbst und dann eine Mehrheit der Physiker davon überzeugt, mit seiner Unschärferelation die "Kausalität abgeschafft" zu haben (v. Weizsäcker (1999), p. 315). Es mag nützlich sein, das Argument W. Heisenbergs zu rekapitulieren:

> "Aber an der scharfen Formulierung des Kausalgesetzes: "Wenn wir die Gegenwart genau kennen, können wir die Zukunft berechnen", ist nicht der Nachsatz, sondern die Voraussetzung falsch, denn wir können die Gegenwart

[12] (1872 – 1970)

[13] "I believe, like much that passes muster among philosophers, is a relic of a bygone age, survivinig, like the monarchy, only because it is erroneously supposed to do no harm." p. 1

> in allen Bestimmungsstücken prinzipiell nicht kennenlernen. Deshalb ist alles Wahrnehmen eine Auswahl aus einer Fülle von Möglichkeiten und eine Beschränkung des zukünftig Möglichen. Da nun der statistische Charakter der Quantentheorie so eng an die Ungenauigkeit aller Wahrnehmung geknüpft ist, könnte man zu der Vermutung verleitet werden, dass sich hinter der wahrgenommenen statistischen Welt noch eine "wirkliche" Welt verberge, in der das Kausalgesetz gilt. Aber solche Spekulationen scheinen uns, das betonen wir ausdrücklich, unfruchtbar und sinnlos. Die Physik soll nur den Zusammenhang der Wahrnehmungen formal beschreiben. Vielmehr kann man den wahren Sachverhalt viel besser so charakterisieren: Weil alle Experimente den Gesetzen der Quantenmechanik und damit der Gleichung (1)[14] unterworfen sind, so wird durch die Quantenmechanik die Ungültigkeit des Kausalgesetzes definitiv festgestellt." Heisenberg (1927) , p. 175, Z 1

Heisenbergs Argumentation ist nicht nur deswegen bemerkenswert, weil sie eine radikale Aussage über das bis dahin sakrosankte Kausalitätsprinzip ist (von Aristoteles bis Leibniz: Jedes Ereignis hat seine Ursache), sondern auch, weil es eigentlich keine physikalische Begründung für seine Interpretation der experimentellen Daten gibt. Heisenberg will eine metaphysische Aussage über eine wirkliche Welt, in der das Kausalgesetz gilt, vermeiden, weil metaphysische Aussagen sinnlos seien, und postuliert: "Die Physik soll nur den Zusammenhang der Wahrnehmungen formal beschreiben". Diese Aussage korrespondiert zu den Auffassungen Ernst Machs und denen des (frühen) Wiener Kreises. Noch im Jahr 1958 führt er aus:

> Einer bestimmten Wirkung eine bestimmte Ursache zuzuordnen hat nur dann einen Sinn, wenn wir Wirkung und Ursache beobachten können, ohne gleichzeitig in den Vorgang störend einzugreifen. Das Kausalgesetz in seiner klassischen Form kann also seinem Wesen nach nur für abgeschlossene Systeme definiert werden. In der Atomphysik ist aber im Allgemeinen mit jeder Beobachtung eine endliche, bis zu einem gewissen Grade unkontrollierbare Störung verknüpft, wie dies in der Physik der prinzipiell kleinsten Einheiten auch von vorn herein zu erwarten war. Da andererseits jede raum-zeitliche Beschreibung eines physikalischen Vorganges die Beobachtung des Vorgangs voraussetzt, folgt, dass die raum-zeitliche Beschreibung von Vorgängen einerseits und das klassische Kausalgesetz andererseits komplementäre, einander ausschließende Charakteristika des physikalischen Geschehens darstellen. Dieser Sachlage entspricht der Sachverhalt, dass zwar ein mathematisches Schema der Quantentheorie existiert, dass dieses Schema aber nicht als einfache Verknüpung von Dingen in Raum und Zeit gedeutet werden kann. Durch diese Komplementarität der Raum-Zeit-Beschreibung einerseits und der kausalen Verknüpfung andererseits tritt ferner eine eigenartige Unbestimmtheit des Begriffs "Beobachtung" auf, indem es der Willkür anheimge-

[14] Es geht in Heisenberg (1927) (p. 174) um die Frage der Ortsbestimmung eines Elektrons, wozu Licht mit der Wellenlänge λ verwendet wird. Es sei p der Impuls (= Masse mal Geschwindigkeit) eines Teilchens (Korpuskel), h sei die Plancksche Konstante; dann gilt $\lambda = h/p$. Der Ort q eines Teilchens ist mit einer Genauigkeit $\Delta q > 0$ bekannt, p mit einer Genauigkeit Δp; dann gilt $\Delta p \Delta q \geq h$, oder, in Heisenbergs (1927) Notation $\Delta p \Delta q \sim h$. Dies ist die Beziehung (1) auf Seite 175 von Heisenberg (1927)

> stellt bleibt, welche Gegenstände man zum beobachtenden System rechnen oder als Beobachtungsmittel betrachten soll.
>
> (Heisenberg 1958/2001, p. 48), Z 2

Im Zitat 1 spricht er direkt von der Ungültigkeit des Kausalprinzips (... so wird durch die Quantenmechanik die Ungültigkeit des Kausalgesetzes definitiv festgestellt."), in Zitat 2 etwas differenzierter. Philosophen und Physiker (etwa von Weizsäcker (1985, 1999) haben darauf hingewiesen, dass Heisenberg hier wohl doch zu radikal argumentiert. Denn implizit scheint Heisenberg, wenn er vom Kausalgesetz spricht, den klassischen, Laplaceschen Determinismus zu meinen, demzufolge Kausalität äquivalent zu deterministischer Berechenbarkeit ist. Kennt man den Ort $x(t)$ eines Teilchens zur Zeit t und seine Geschwindigkeit $v(t)$ – bzw. seinen Impuls $p(t) = mv(t)$, m die Masse des Teilchen – so erlaubt die klassische (newtonsche) Physik die genaue Berechung der Bahn des Teilchens. Damit man die Berechung durchführen kann, muß man aber $x(t)$ und $v(t)$ kennen, und das heißt, man muß sie im Prinzip messen können, und zwar im Prinzip beliebig genau. Wie in Fussnote 14 schon angemerkt wurde, gilt aber die Heisenbergsche Unschärferelation[15]

$$\Delta p \Delta q \geq h,$$

wobei Δp und Δq,die "Unschärfen" sind, mit denen die Größen p und q gemessen werden (es läßt sich zeigen, dass die Unschärfen Standardabweichungen von Wahrscheinlichkeitsverteilungen sind), und h ist eine Konstante. Die Relation besagt, eine kleine Unschärfe für eine Größe eine große Unschärfe für die andere bedeutet, und umgekehrt: misst man etwa den Ort genau, wird die Bestimmung der Geschwindigkeit ungenau. Darüber hinaus darf keine der Größen Δp und Δq gleich Null werden, weil dann das Produkt $\Delta p \Delta q$ gleich Null wird, obwohl es doch nicht kleiner als h werden kann. Wenn x und v aber nicht genau gemessen werden können, so ist die Trajektorie des Teilchens nicht genau berechenbar, womit das klassische Kausalprinzip nicht mehr gilt. Es folgt aber nicht, dass keinerlei Kausalität angenommen werden darf: die Begriffe Kausalität und Determinismus haben keinen identischen Bedeutungsumfang. Die Unschärferrelation besagt "nur", dass die Größen p und q (es muß sich dabei nicht notwendig um Ort und Geschwindigkeit eines Teilchen handeln) *nicht gleichzeitig* scharf gemessen werden können, mit der Folge, dass nur probabilistische Voraussagen gemacht werden können. Es folgt, dass es für die Frage nach der Willensfreiheit nicht genügt, zur Begründung einer möglichen Akausalität in der Natur auf die Heisenbergsche Unschärferelation zu verweisen.

Einstein stand bekanntlich der probabilistischen Interpretation der Quantenphysik Heisenbergs skeptisch gegenüber, insbesondere der Interpretation Heisenbergs, dass die Unschärferelation die Akausalität der involvierten Prozesse impliziere. Es müsse versteckte Größen (hidden variables) geben, die dem Kausalgesetz entsprechend das physikalische Geschehen bestimmen, und 1964 publizierte J. S.

[15] Δ ist der griechische Buchstabe Delta, der in der Mathematik oft für "Unterschied" (**D**ifferenz) steht.

Bell die dann nach ihm benannte Ungleichung, die einen Test der Kopenhagener Deutung der Quantenmechanik ermöglicht; die in den frühen 70er Jahren durchgeführten Experimente (Freedman & Clauser (1972)) lieferten Daten, die als die Widerlegung der einsteinschen Bedenken gelten (s. a. Kapitel IV, Determinismus, wo ausführlicher auf diese Diskussion eingegangen wird). Allerdings sind diese Ergebnisse anscheinend nicht so eindeutig, wie man den Eindruck haben kann. Bei der Interpretation der Daten anhand der Bellschen Ungleichung hat man, so einige Physiker (z.B. Hossenfelder & Palmer (2020)), anscheinend übersehen, dass in die Herleitung der Ungleichung eine Annahme eingeht, die nicht gelten *muß* (Brans(1988), Larsson (2014)). Sollte eine besonders strenge Theorie des Determinismus gelten, derzufolge alles Geschehen vom Beginn des Universums an bestimmt ist (dies ist die Annahme des *Superdeterminismus*), so haben Rauch et al. (2018) geschätzt, dass seit mindestens 7.8 Milliarden Jahren jedes Ereignis vorherbestimmt ist (die Autoren sind keine Vertreter der des Superdeterminismus, sie zeigen nur die Implikationen dieser Auffassung auf). Hossenfelder & Palmer (2020) sehen allerdings in der Theorie des Superdeterminismus eine Möglichkeit, die Probleme der QM zu überwinden, die ihrer Ansicht nach durch Bells Annahmen zementiert werden: es ist aus Hossenfelder et al.s Sicht die Annahme der statistischen Unabhängigkeit der versteckten Variablen einerseits und der Messeinstellungen andererseits. Die Unabhängigkeit dieser Größen sei eine intuitive Grundannahme der Physik, weshalb der Superdeterminismus üblicherweise schnell verworfen werde. Die Autoren verweisen auf die Binsenwahrheit, dass Intuitionen nicht notwendig wahr sind. Läßt man die Annahme der Unabhängigkeit fallen, so kann man den Superdeterminimus akzeptieren, der aber, nach Hossenfelder et al., die Freiheit des Willens ausschließt. Insbesondere Hossenfelder ist eine durchaus radikal anmutende Vertreterin der These, dass der Wille nicht frei ist. In ihrem Buch *Existential Physics* (2022) argumentiert sie allgemeinverständlich, also ohne auf die Details der Bellschen Ungleichung einzugehen, für die Determiniertheit des menschlichen Willens: "the future is fixed except for occasional quantum events that we cannot influence" (p. 126) Ihr Argument für diese Auffassung ist die eher rhetorische Frage, wie man denn den freien Willen als mit den Naturgesetzen kompatibel ansehen könne, wenn diese doch einen Determinismus implizieren. Eine Antwort könnte sie bei den Philosophen Julian Nida-Rümelin, Jürgen Habermas und insbesondere bei Alva Noë finden, die jeweils verchiedene Formen des Superdualismus als Gegenpol zum Superdeterminismus konzipiert haben, vergl. Abschnitt 1.2.5.4.

Der niederländische Physiker Gerard t'Hooft (1946 –) (Nobelpreis 1999 für seine Arbeit auf dem Gebiet der Quantenfeldtheorie) vertritt eine ähnliche Ansicht: [16]

> "Whether Divine Intervention takes place or not, and whether our actions are controlled by "free will" or not, will never be decidable in practice. This author suggests that, where we succeeded in guessing the reasons for

[16] Die folgenden Zitate von t'Hooft und Zeilinger sind "Reddit PhilosophyofScience, Free Will required for Science or not" entnommen worden.

> many of Nature's laws, we may well assume that the remaining laws, to be discovered in the near or distant future, will also be found to agree with similar fundamental demands. Thus, the suspicion of the absence of free will can be used to guess how to make the next step in our science." -Gerard 't Hooft, 1999 Nobel Laureate in Physics, Z 3

t'Hooft hat sich wiederholt gegen die These des freien Willens geäußert (vergl. t'Hooft (2017, 2020), und seine Argumente beruhen – ebenso wie die von Hossenfelder et al. – nicht nur auf intuitiven Mutmaßungen.

Der österreichische Physiker Anton Zeilinger (1945 –) (Nobelpreis 2022 für seine Arbeit auf dem Gebiet der Quanteninformation und der Verletzung der Bellschen Ungeichung) argumentiert dagegen

> "The second important property of the world that we always implicitly assume is the freedom of the individual experimentalist. This is the assumption of free will. It is a free decision what measurement one wants to perform. In the experiment on the entangled pair of photons, Alice and Bob are free to choose the position of the switch that determines which measurement is performed on their respective particles. It was a basic assumption in our discussion that that choice is not determined from the outside. This fundamental assumption is essential to doing science. If this were not true, then, I suggest, it would make no sense at all to ask nature questions in an experiment, since then nature could determine what our questions are, and that could guide our questions such that we arrive at a false picture of nature." In: Zeilinger (2010). Z 4

Wenn zwei renommierte Physiker zu derart unterschiedlichen Aussagen kommen, kann man vermuten, dass sie von unterschiedlichen Konzeptionen des Begriffs 'freier Willen' haben. Es ist hier nicht möglich, in die Details der Theorien zu gehen, die möglicherweise hinter diesen Aussagen stehen, aber unterschiedliche Definitionen des freien Willens lassen sich schon diskutieren. Bevor auf diese Frage eingegangen wird, sei noch einmal auf grundsätzliche Aspekte des Begriffs des Determinismus eingegangen: Das Konzept scheint sich einfach genug aus dem Kausalitätsprinzip 'Jedes Ereignis hat seine Ursache' zu ergeben[17].

Aber so einfach ist die Sache nicht. Bereits David Hume (1789) hat sich skeptisch über die Begriffe Ursache und Wirkung geäußert, obwohl man in der empirischen Forschung kaum ohne sie auskomme. Aber es handele sich um Kon-

[17]Es sollte hier auf Conway & Kochens Arbeit *The Free Will Theorem* (2006) zumindest hingewiesen werden, in der die Autoren beweisen, dass, wenn Experimentatoren eine Entscheidung über eine vorzunehmende Messung fällen, die frei ist in dem Sinne, dass die Entscheidung keine Funktion irgendwelcher, den Experimentatoren verfügbaren Informationen über die Partikel ist, dann das Verhalten vermessenen Partikel ebenfalls keine Funktion irgendwelcher, vor der Messung verfügbarer Informationen sind. Der Beweis beruht auf Axiomen, deren Wahrheit evident sei. Der freie Wille der Experimentoren wird hier vorausgesetzt, nicht bewiesen. Damit machen sie u.a. eine Aussage über die Beziehung zwischen physischer und mentaler Welt: Sollte für die physische Welt ein strikter Determminismus gelten, so kann der Wille nicht frei sein. Das Theorem zielt, wie schon die Bellsche Ungleichung, auf die Nichtexistenz versteckter Variablen; Kritik wurde zum Beispiel von Goldstein et all (2009) formuliert.

struktionen des Verstandes, so dass ihnen eine gewisse Beliebigkeit eigen sei. Bertrand Russell hat in seinem Essay "On causation" den Begriff der Kausalität gewissermaßen aufgelöst, und John Earman liefert in seinem Buch *A Primer on Determinism* (1986) weitere Einblicke in die Komplexität dieses Begriffs; in Kapitel IV[18], insbesondere Abschnitt 2.4 ist bereits darauf eingegangen worden. Wie Earman zeigt ist schon die newtonsche Physik nicht das Paradies für den Determinismus, für das sie normalerweise gehalten werde: der Begriff des Naturgesetzes sei "crucial, but troublesome".

van Kampens Betrachtungen: Quantenfluktuationen werden von Philosophen oft angeführt, um sich nicht einem strengen Determinismus ausgesetzt zu sehen. In diesem Zusammenhang ergibt sich die Frage, ob diese Fluktuationen überhaupt auf dem Aktivtätsniveau von Neuronen relevant sind, weshalb noch einmal auf van Kampens Arbeit "Determinism and predictability" (1991), p.273, eingegangen werden soll. van Kampens Analysen beziehen sich auf den Laplaceschen Determinismus[19], aber sie gelten im Prinzip auch für die Prozesse im atomaren bzw subatomaren Bereich. Es wird ja angenommen, dass die Interaktionen zwischen Teilchen zunächst einmal deterministischen Gesetzen folgen, schließlich ist Schrödingers Gleichung eine deterministische Gleichung. Aber wie weit schlagen Fluktuationen in diesem Bereich durch? Betrachtet man molekularbiologische Vorgänge, so sind sie relativ zu den subatomaren gewissermaßen grobkörnig, aber sie – etwa Stoßprozesse – werden wieder durch deterministische Gleichungen beschrieben, auch wenn diese im Prinzip Approximationen darstellen. Auf makroskopischem Niveau wie in der statistischen Mechanik wiederholt sich das Prinzip. Die tatsächlich beschriebenen Variablen zeigen stochastische Fluktuationen. Die Gleichungen beschreiben die durchschnittliche Dynamik, und, wie im Gesetz der Großen Zahlen beschrieben, die Fluktuationen sind klein relativ zum Durchschnitt der Fluktuationen. Auf der mesoskopischen Ebene erhält man stochastische Prozesse, und die korrespondierenden Differentialgleichungen liefern Wahrscheinlichkeiten, wobei die Verteilungen für Zustände oft gegen Gauss-Verteilungen konvergieren. In dem Maße, in dem analoge Ansätze für (Teil-)Populationen von Neuronen gelten, kommt es zu "rauschinduzierten" Zustandsübergängen, d.h. zu Übergängen, die nicht deterministisch, sondern eben nur probabilistisch beschreibbar sind (Horsthemke & Levever (1984)). Geht man davon aus, dass Kognitionen, insbesondere Entscheidungsprozesse, das Resultat von Hirnprozessen sind, so kann man vermuten, dass derartige Zustandsänderungen auch Entscheidungen beeinflussen können. Aber man kann so zustande gekommene Entscheidungen kaum "frei" nennen, weil der Mensch sie eben nicht "frei" getroffen hat, denn sie sind ja durch Fluktuationen, über die er keine Kontrolle hat, beinflusst worden (Rubinov, Sporns, & Breakspear (2011)): die Rede ist von neuronalen Bursts durch selbst-organisierte Kritikalität (Bak, Tang & Wiesenfeld (1988)). Bei diesen Betrachtungen werden die von

[18] Wissenschaftstheorie IV – Kausalität und Zufall

[19] van Kampens Theorem: "Der ontologische Determinismus à la Laplace kann anhand von empirischen Beobachtungen weder bewiesen noch widerlegt werden." Beweis: van Kampen (1991), oder Wissenschaftstheorie IV – Kausalität und Zufall, p. 44

der Unschärferelation implizierten zufälligen Ereignisse nicht genannt, allerdings tauchen sie in der im Folgenden referierten Arbeit von Ellis auf.

George F.E. Ellis[20]: From Chaos to free will Ellis (2020) liefert eine Beschreibung der neuronalen Aktivität, die von Anfang an darauf abzielt, die Zufälligkeit in dieser Aktivität zu betonen. Die Bewegungsgleichungen der klassischen Dynamik sind deterministische Differentialgleichungen, aus denen folgt, dass die Dynamik des Systems zumindest im Prinzip durch die Startbedingungen bestimmt ist. Das heißt aber nicht, dass notwendigerweise exakte Vorhersagen möglich sind, wie schon im Zusammenhang mit dem Laplaceschen Dämon erwähnt wurde. Die Nichtvorhersagbarkeit auch deterministischer dynamischer Prozesse wird gelegentlich von Philosophen als Ausgangspunkt für eine Argumentation für den freien Willen genommen, aber Unvorhersagbarkeit bedeutet ja nicht Freiheit der Entscheidungen. Die Nichtvorhersagbarkeit tritt schon bei relativ einfachen dynamischen Abläufen auf, z.B. beim bereits genannten Dreikörperproblem: es gelingt nicht, die Bahnen von drei Himmelskörpern unter Berücksichtigung der wechselseitigen gravitationalen Effekte zu bestimmen analytisch darzustellen, d.h.diese Bahnen sind nicht mehr mit elementaren Funktionen beschreibbar, die Bahnen müssen mit numerischen Methoden approximiert werden. Das Gleiche gilt für das Doppelpendel. Das Verhalten von drei künstlichen Neuronen in einem künstlichen neuronalen Netz ist unter bestimmten Bedingungen nicht mehr analytisch berechenbar (an der Heiden (1996), zitiert nach Roth (2014), p. 165). Die Dynamik von Hirnprozessen ist ungleich komplexer als die dieser Systeme, so dass es überhaupt nicht verwundert, dass die Vorgänge im Gehirn nicht im Detail vorhersagbar sind. Andere Philosophen beziehen sich auf die Stochastizität quantenmechanischer Prozesse, – die aber nicht hilfreich ist, denn wenn der Zufall den Willen regiert, kann man ihn kaum als frei ansehen. Ganz abgesehen davon weist Roth (2014) noch einmal darauf hin, dass Prozesse, bei denen quantenmechanische Zufälligkeit eine Rolle spielen (die Ausschüttungen von Transmittervesikeln an Synapsen, das Öffnen und Schließen von Ionenkanälen in den Zellmembranen von Neuronen, etc) könnten sich "um Größenordnungen unterhalb der verhaltensrelevanten Ebene" abspielen, – die Komplexität der Hirnvorgänge stellt ein viel größeres Problem dar als der quantenmechanische Indeterminismus (Roth (2014), p. 164). Ellis (2020) jedenfalls ist gleichwohl von der Existenz des freien Willens überzeugt, und so macht er, trotz aller Betrachtungen zur Rolle der Physik, auf Seite 10 einen gedanklichen Sprung und spricht von "echtem mentalen Funktionieren":

> "Genuine mental functioning and the ability to make decisions in a rational way is a far more persuasive explanation of how books get written".

Mit den Büchern meint er die Dramen William Shakesspears oder die Encyclopedia Britannica: würde der Determinismus alles Geschehen lenken, so wäre die

[20] George R. R. Ellis, (1939 –), Mathematiker, Kosmologe

Produktion solcher Bücher schon von Beginn des Universums vorprogrammiert. Das erscheint als absurd, weshalb er ohne weitere Umschweife die Existenz der Fähigkeit zu *freien*, rationalen Entscheidungen postuliert; warum es dabei um *rationale* Entscheidungen geht, ist unklar, denn wenn Entscheidungen frei sind, können sie gleichwohl irratonial sein.

Dagegen wird vorgebracht, dass zumindest aus der Kopenhagener Interpretation der Quantenmechanik folge, dass es einen absoluten Zufall gibt, also einen Zufall, der nicht nur mangelndes Wissen oder mangelnde Berechenbarkeit reflektiert. Das Kausalprinzip ist aufgehoben, so zumindest die Annahme. Ellis argumentiert weiter, dass es zwar richtig sei, dass bei deterministischen Differentialgleichungen für gegebene Anfangsbedingungen die Menge der Zustände, die durchlaufen werden determiniert ist, dass aber Störungen von außerhalb des Systems diese Menge verändern. Als Beispiel hierfür betrachtet Ellis die Dynamik eines Pendels: einmal ausgelenkt liegt die Bewegung des Pendels fest, – bis es zu einer Störung kommt, die die Amplitude und die Phase des Pendelbewegung ändert. Ellis kommt dann auf die grundlegenden molekularen Prozesse zu sprechen, ´die die neuronale Aktivität ausmachen, z. B. die Aktivität in den Ionenkanälen der Membranen der Neuronen, die sich wegen der Wechselwirkungen zwischen den Neuronen ständig verändert, und die die "höheren" Prozesse, die die kognitive Aktivität ausmachen, beeinflussen. Diese höheren Prozesse wirken zurück auf die basalen Vorgänge, wo wiederum Heisenbergs Unschärferelation die, so Ellis, "unvermeidbare Unschärfe" (unavoidable fuzziness) in den Prozessen erzeuge. Nicht zuletzt die diversen Rückkopplungen in dem hochkomplexen System, das das Gehirn ist, stören einmal begonnene Teilprozesse, die dann eben nicht mehr den Anfangsbedingungen entsprechend weiterlaufen;

> "... because of what happens at the microscale. At molecular scales, the processes at work forget initial data due to billions on billions of collisions between molecules every second. Biology thrives on that disorder – a molecular storm ..." (p. 9), Z 5

Ellis verweist dann auf Arbeiten aus der Chaosforschung, in der genauere Vorstellungen darüber, wie sich aus den "molekularen Stürmen" geordnete Strukturen entwickeln können, hergeleitet werden. Aus dem Begriff des Determinismus könne diese Erklärung nicht entwickelt werden[21], – und so kommt er zu der Aussage, dass die "außergewöhnliche hierarchische Struktur" der Hirnprozesse diese geordnete kognitive Aktivität und damit auch freie Entscheidungen ermögliche:

> "If you seriously believe that fundamental forces leave no space for free will, then it's impossible for us to make choices as moral beings. ... That's a devastating conclusion. We can be grateful it's not true." (p. 10), Z 6

Diese Schlußfolgerung ist – wie schon bei Kant – ein gedanklicher Sprung, der ein

[21] Ellis geht nicht darauf ein, dass mit dem Begriff des Chaos in der Physik i. A. das *deterministische Chaos* gemeint ist, das in perfekt deterministischen Systemen auftreten kann.

Evidenzerlebnis bewirkt[22]; es ist nicht klar, warum der Sprung aus den vorangegangenen Äußerungen folgen soll. Ellis argumentiert, ein unfreier Wille impliziere die Unmöglichkeit moralischen Handelns, und die Tatsache, dass Menschen moralisch handeln können, wenn sie es denn wollen, soll dann dann die Existenz eines freien Willens implizieren. Diese Argumentation wirkt vielleicht auf den ersten Blick logisch, weil die Schlußregel *modus tollens*[23] korrekt angewendet wird. Nur muß die Aussage "Ist der Wille unfrei (p), so kann nicht moralisch gehandelt werden (q)" nicht notwendig wahr, so dass die Schußfolgerung, dass Möglichkeit, moralisch zu handeln (nicht-q) impliziert, dass der Wille nicht unfrei (nicht-p), also frei ist, nicht wahr sein muß. Denn man könnte ebensogut argumentieren, dass erst ein unfreier Wille moralisches Handeln ermöglicht, wenn nämlich moralisches Handeln in sozialen Normen festgeschrieben ist uns gruppendynamische Prozesse die Einhaltung der Normen erzwingen. Die Tatsache, dass es Menschen gibt, die sich nicht an soziale Normen und damit nicht an moralische Gesetze gebunden fühlen, ändert nichts an diesem Argument, das gleichermaßen für den Fall gilt, dass es einen freien Willen gibt: dass ein freier Wille moralisches Handeln ermöglicht, heißt das noch nicht, dass eine Person mit freiem Willen auch tatsächlich stets moralisch handelt. Der Wille wäre nicht frei, erlaubte er nicht die Möglichkeit, unmoralisch zu handeln.

Man kann sagen, dass Ellis überhaupt nicht in Rechnung stellt, dass die von ihm erwähnten "höheren" kognitiven Prozesse Sozialisierungsprozesse umfassen, die die allermeisten Menschen durchlaufen und die soziales, prosoziales und damit moralisches Verhalten generieren. Diese Entwicklung setzt nicht *notwendig* einen freien Willen voraus. Man kann argumentieren, dass viele Entwicklungen von Formen des Zusammenlebens ein Produkt der Evolution sind: Gruppen von Menschen, in denen sich soziale ("moralische") Normen entwickelt haben, haben eine höhere Überlebenswahrscheinlichkeit als Gruppen, in denen sich keine derartigen Normen entwickelt haben[24]. Wie oben bereits argumentiert wurde, kann man mit einem Hauch von Polemik sagen, dass moralisches Verhalten möglich wird, gerade weil der Wille nicht frei ist, denn ein von den Sinnen diktiertes (Kant) asoziales Verhalten kann zwar einen unmittelbaren Vorteil haben: ich schlage meinen Mit-Jäger tot, weil ich dann das erlegte Wild ganz für mich habe und damit kurzfristig meine Überlebenswahrscheinlichkeit erhöhe. Langfristig ist ein solches Verhalten aber nachteilig, weil ich nicht mehr versorgt werde, wenn ich mir auf der Jagd ein Fußgelenk verstaucht habe, also weder jagen noch vor dem Bären fliehen kann und deshalb sterbe. Als Mitglied einer Gruppe von Jägern,

[22]Der trügerische Effekt von Evidenzerlebnissen wird in Wissenschaftstheorie III (2), Abschnitt 2.2, diskutiert

[23]modus tollens, auch modus tollendo tollens (lat. für Weise, durch Verneinen zu verneinen), ist die Schlußform "Wenn die Aussage p die Aussage q impliziert, so impliziert nicht-q (die Negation von q) die Aussage nicht-p. Beispiel: Wenn es regnet (p), so ist die Strasse nass (q). Aber die Strasse ist nicht nass (nicht-q, also folgt, dass es nicht regnet) (nicht-p).

[24]Folgt man den Superdeterministen, so können wir aus der Tatsache, dass sich Normen entwickelt haben, folgern, dass diese Entwicklung vom Big Bang an "eingeplant" war, will sagen, dass die ablaufenden physikalische Prozesse diese Entwicklungen notwendig implizieren.

die erlegtes Wild miteinander teilen, lebt man länger, weil man wegen eines verstauchten Gelenks oder oder anderer Unbill für eine gewisse Zeit von den anderen Gruppenmitgliedern durchgefüttert wird. Die Einsicht in die Notwendigkeit der Kooperation *determiniert* das Verhalten, und das Postulat, dass die Akzeptanz dieser Norm ein Resultat der Existenz des freien Willens sei, ist gar nicht notwendig, um ein durch Einsicht geprägtes Verhalten zu erklären. Ein Beispiel für die verhaltensbestimmende Kraft von Normen ist das Taboo bei polynesischen Gruppen: hat ein Mitglied einer Gemeinschaft gegen ein Taboo verstoßen, so kann es erkranken, auch wenn noch gar nicht bekannt geworden ist, dass diese Person ein Taboo gebrochen hat,– das Verhalten der Gruppenmitglieder vermittelt der Person, die unwissentlich (!) einen Taboo-Bruch begangen hat, auf indirekte Weise, dass sie einen Taboo-Bruch begangen hat, ohne dass er explizit genannt wird[25]. Jedes Mitglied der Gemeinschaft wird sehr darauf achten, keines der verschiedenen Taboos zu brechen (Taboos werden zum Beispiel für eine beschränkte Dauer ausgesprochen, – z.B. das Ernten von Kokosnüssen – um das Überleben der Gemeinschaft auf einer Insel zu sichern). Man benötigt keine Willensfreiheit, um Taboos zu entsprechen, sie wirken nahezu automatisch.

Man kann natürlich argumentieren, dass die Annnahme eines freien Willens mit einsichtigem Verhalten kompatibel sei. Der Punkt ist, dass diese Annahme nicht notwendig aus einsichtigem Verhalten folgt und deshalb dem Ockhamschen Rasierer überlassen werden kann. Es folgt dann, dass weder moralische noch unmoralische Entscheidungen etwas über die Freiheit oder Unfreiheit des Willens aussagen. Wichtig ist allerdings der Befund, dass Menschen *sich als frei erleben*, gerade auch dann, wenn sie eine moralische Entscheidung treffen. Nur bedeutet das Erleben von Freiheit nicht notwendig auch tatsächliche Freiheit.

1.1.3 Libets Experimente, die Rolle des Bereitschaftspotentials

Bis hierher ist der Begriff des freien Willens nur als Theorie diskutiert worden. Man mag dann fragen, ob sich empiriebasierte Aussagen, d.h. Daten, über die Möglichkeit eines freien Willens finden lassen. Experimente, deren Ergebnisse relevant für die Hypothese der Existenz des freien Willens sein könnten, werden in diesem Abschnitt besprochen.

Geht man von der Annahme aus, dass das Bewußtsein aus der Aktivität des Gehirns entsteht, so muß diese Aktivität durch Integration vieler Teilaktivitäten bestehen. Eine erste Stufe besteht in der Integration von Information der unmittelbaren Umgebung, d.h. aus der Verarbeitung visueller, auditiver, taktiler Informationen, die mit Zentren gekoppelt werden, die diese Informationen interpretieren. Dabei handelt es sich um komplexe Prozesse, die u.a. motorische Aktivitäten implizieren können. Diese Dynamik entwickelt sich in der Zeit – keine Dynamik ist instantan im strengen Sinn des Wortes, und was im Bewußtsein erscheint, wird

[25]Lit: R.L Stevenson – Life in Western Samoa.
https://www.uniaktuell.unibe.ch/2022/eating-taboos-the-cultural-construction-of-disgust/index-eng.html)

relevanten Teilen dieser Dynamik nachgeschaltet sein. Man kann also vermuten, dass die Dynamik begonnen hat, bevor sie im Bewußtsein repräsentiert wird. Walter et al. (1964) sowie Kornhuber und Deecke (1965) fanden, dass willentlichen Hand- oder Fußbewegungen negative kortikale Potentiale in der Größenordnung von 10 – 15 μV (Mikrovolt, 1 μV = 0.000 001 Volt) vorausgingen; sie nannten sie *Bereitschaftspotentiale* (readiness potentials). Dieses Potenial wächst mit der "Aufmerksamkeit und intentionalen Beteiligung der Versuchsperson und nimmt bei Gleichgültigkeit ab".

Der Physiologe Benjamin Libet (1983, 1985) berichtete Experimente, die mögliche neuronale Aktivitäten nachweisen sollten, die einer bewußten Entscheidung bzw. Handlung vorangehen. Sollte es derlei Aktivitäten geben, so sollten sie im EEG aufscheinen. Die Aufgabe eines Experimentators besteht darin, die Versuchsperson (VP) zu einer von ihr selbst gewählten Zeit eine einfache Handlung durchführen zu lassen. Bei Libet mußte die Vp einfach die Hand heben. Insgesamt bestand die Aufgabe der Vp darin, (i) auf einen Monitor zu blicken, auf den ein Punkt zu sehen war, der sich auf einem Kreis im Uhrzeigersinn bewegte. (ii) zu einem von der Vp frei wählbaren Zeitpunkt sollte sie eine Hand heben. In dem Moment, in dem ihr der Wille oder der Drang (the urge) kam, die Hand zu heben, sollte sie einen Knopf drücken, der die Bewegung des Punktes stoppte. Nun sollte sie den Punkt auf die Position bewegen, der dem Zeitpunkt entsprach, an dem der Wille zum Handheben bewußt geworden war. Nach einer kleinen Pause begann der nächste Versuchsdurchgang, der wie der eben beschriebene verlief. Es wurden nun nicht nur die von der Vp angezeigten Zeitpunkte ausgewertet, sondern diese Daten wurden zu den evozierten Potentialen eines gleichzeitig erhobenen EEG in Beziehung gesetzt.

Das wesentliche Resultat der Untersuchung war, dass stets ein der bewußten Entscheidung vorausgehendes Bereitschaftspotential gemessen werden konnte. Hatte die Vp den Zeitpunkt geplant, begann das Bereitschaftspotential im Durchschnitt 1050 msec vor der bewußten Entscheidung, wählte sie die Handlung spontan, so setzte es 550 msec vorher ein[26]. Zwischen dem Zeitpunkt des Bewußtwerdens und dem Beginn der Muskelaktivität lagen im Durchschitt 200 msec.

Dieses Ergebnis legt die Vermutung nahe, dass die Entscheidungen vorbewußt gebahnt werden. In der Realität sind die jeweils gestellten Aufgaben natürlich sehr viel komplexer – soll man sich überhaupt ein Auto kaufen und wenn ja welches, soll man das Fach A studieren oder besser das Fach B, etc., und moralische Entscheidungen erfordern oft längere kognitive Aktivitäten. Man wird hier Hypothesen aufstellen können, die im Prinzip eine durch kognitive Aktivität – die selbst eine neuronale Basis hat – auf andere kognitive Aktivitäten einwirkt, die ihrerseits Entscheidungen bewirken, für die dann dasselbe Prinzip gilt: bevor die Entscheidung bewußt wird geht ein jeweiliges Bereitshaftspotential voran. Ein wichtiger Punkt dabei ist, dass eine bereits "getriggerte" Entscheidungsbildung

[26] 1 msec = eine Millisekunde = 1 Tausendstel Sekunde

noch vor der Ausführung einer korrespondierenden Handlung gestoppt werden kann, oder, wie Libet sich ausdrückte, es kann ein Veto eingelegt werden. Jedenfalls wird es bei komplexeren Entscheidungprozessen zu einer Wechselwirkung zwischen vorbewußten und bewußten Prozessen kommen.

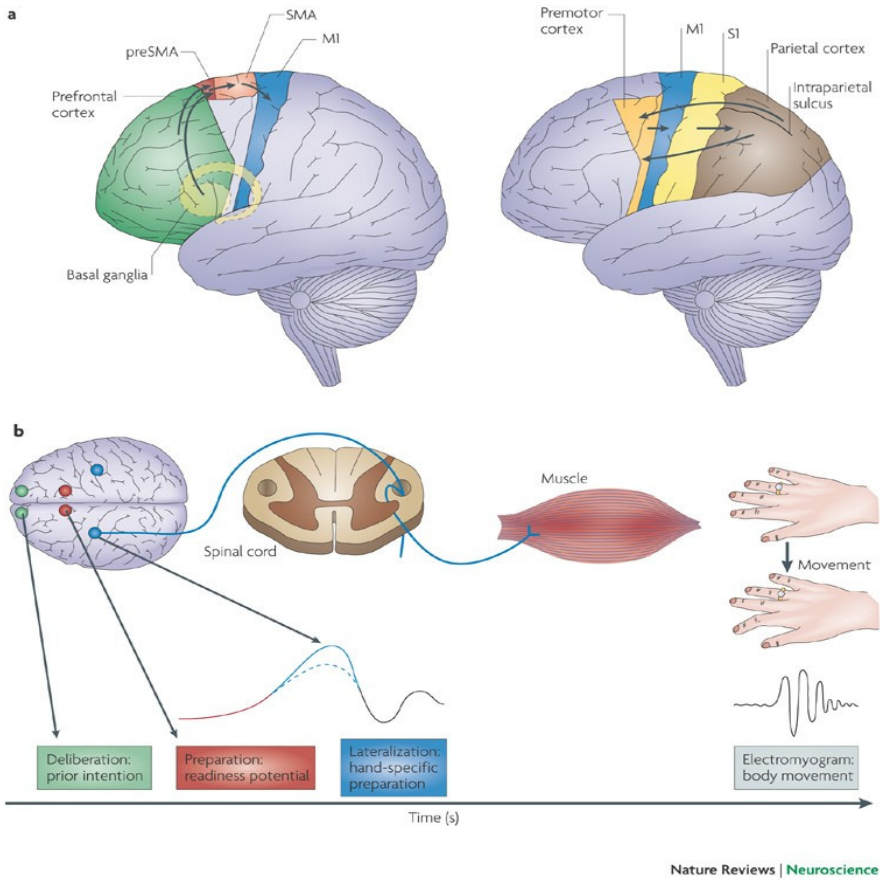
Natürlich sind die Libet-Experimente nicht unkritisiert geblieben. Philosophischerseits ist betont worden, dass die Ergebnisse noch nicht zwingend zeigen, das mentale Ereignisse wie das Fällen von Entscheidungen Ausdruck neuronaler Aktivität sind, und dass insbesondere die kausale Deutung der Bereitschaftspotentiale durch Libet nicht zutreffen müsse; Braun et al. (2021) haben eine Metaanalyse zu Untersuchungen von Libet durchgeführt und fanden, dass einige von Libet et al.s Befunden "fragiler erscheinen als man im Licht der substantiellen, darauf aufbauenden wissenschaftlichen Arbeiten erwarten würde"[27]. Überdies handele es sich nicht um echte Entscheidungen zwischen zwei (oder mehr) Alternativen, da es ja nur darum ginge, eine Hand zu heben oder nicht zu heben. Auf die Details der Kritik muß hier nicht eingegangen werden (Libet hat 2004 ein ganzes Buch dazu veröffentlicht). Wie zu erwarten wurden Folgeuntersuchungen mit verbesserten Versuchsanordnungen publiziert, etwa von Haggard & Eimer (1999), die zu einem differenzierteren Bild kommen: so scheinen die von Libet beoachteten Bereitschaftspotentiale in der Tat keine *kausale* Rolle zu spielen, dafür aber die *lateralisierten Bereitschaftspotentiale*:.

> We investigated the relation between neural events and the perceived time of voluntary actions or the perceived time of initiating those actions using the method of Libet. No differences were found in either movement related potentials or perceived time of motor events between a fixed movement condition, where subjects made voluntary movements of a single finger in each block, and a free movement condition, in which subjects chose whether to respond with the left or the right index finger on each trial. We next calculated both the readiness potential (RP) and lateralised readiness potential (LRP) for trials with early and late times of awareness. The RP tended to occur later on trials with early awareness of movement initiation than on trials with late awareness, ruling out the RP as a cause of our awareness of movement intiation. However, the LRP occurred significantly earlier on trials with early awareness than on trials with late awareness, suggesting that the processes underlying the LRP may cause our awareness of movement initiation. Z 7

Desmurget et al. (2009) berichten Daten aus neurologischen Untersuchungen, bei denen einzelne Areale des Gehirns durch Elektroden stimuliert wurden, weshalb ein genaueres Bild der neuronalen Aktivitäten entstand, s. a. Haggard (2009). Die Stimulation der rechten inferior parietalen Bereiche erzeugte eine ausgeprägte Intention, die kontralaterale Hand zu heben. Stimulation der linken parietalen Bereiche erzeugte die Absicht, den Mund zu bewegen und zu sprechen. War die Stimulierung hinreichend stark, so waren die Patienten überzeugt, tatsächlich

[27] Braun et al.: "We conclude that some of Libet et al.'s findings appear more fragile than anticipated in light of the substantial scientific work that built on them".

Abbildung 1: Lokalisationen für (Haggard (2008))



Hand oder Fuß erhoben oder tatsächlich gesprochen zu haben. Die Stimulierung der prämotorischen Bereiche dagegen löste zwar Mund- und Hand/Fußbewegungen aus, aber nun waren die Patienten überzeugt, sich nicht bewegt zu haben (Haggard (2008)). Haggard macht einen wichtigen Punkt: wenn man den freien Willen untersuchen will, kann man nicht einen Stimulus präsentieren und die folgende Response beobachten – Handlungen oder Entscheidungen, die nach freiem Willen geschehen, müssen unabhängig von irgendwelchen Stimuli sein. Die Versuchsanordnung soll also so konzipiert werden, dass nur teilweise vom Experimentator evozierte Entscheidungen vorgegeben werden. Es gibt mehr motorische neuronale Subpopulationen ("circuits"), die an freiwilligen Handlungen beteiligt sein können, die auf den primären Motorkortex M1 konvergieren, vergl. Abb. 1. M1 führt motorische Anweisungen aus, indem er "Anweisungen" an die Muskeln gibt. Soon et al. (2008) lieferten weitere Evidenz für die Hypothese, dass spezifische neuronale Aktivität dem Bewußtwerden einer als frei erlebten Entscheidung vorangeht. Statt EEG- wurden fMRI-Untersuchungen durchgeführt, die die Identifikation verschiedener Bereiche des Gehirns, die an den Entschei-

dungen beteiligt sind, ermöglichten. Den Vpn wurden Folgen von Buchstaben präsentiert und die Aufgabe war, sobald sie den Drang spürten einen von zwei möglichen Knöpfen mit entweder dem rechten oder dem linken Finger zu drücken. Auf dem Bildschirm erschien eine Anordnung von vier Alternativen, und die Vpn mußte denjenigen Buchstaben angeben, der ihr präsentiert worden war, als sie den Entschluß fassten, den rechten oder linken Finger zu drücken. Danach begann einer neuer Versuchsdurchgang. Die Zeit bis zum ersten Drücken eines Knopfes betrug im Durchschnitt 21.6 Sekunden, – Zeit genug für den Aufbau eines Bereitschaftspotentials. Es zeigte sich, dass beide Knöpfe mit annähernd gleicher Häufigkeit gedrückt wurden. Es konnten spezifische Bereiche des Gehirns identifiziert werden, in denen Aktivierungen beobachtet wurden, die schließlich eine Entscheidung im Bewußtsein manifestierten. Der erste Bereich liegt im frontopolaren Kortex (BA10), 7 ms bevor die Vp die motorische Antwort (Drücken eines Knopfes) produziert wurde. Da die BOLD-Reaktionen[28] des fMRI relativ langsam erfolgen vermuten die Autoren, dass die "prädiktive Reaktion" - also der Aufbau eines BP, bereits 10 msec vor der Entscheidung beginnt. Der zweite Bereich liegt im Parietalkortex, aber schwache Reaktionen wurden auch in anderen Bereichen gefunden. Auf eine detailliertere Beschreibung der Resultate kann hier nicht eingegangen werden, es genügt, zu berichten, dass die Autoren eine mögliche kausale Struktur der Entscheidungsfindung andeuten, derzufolge die Vorbereitung einer Entscheidung im präfrontalen Kortex beginnt und dann in Aktivitäten im supplementärmotorischen Kortex (SMA[29], ein Teilbereich des motorischen Kortex) übergeht. Der Bewußtwerdung der Entscheidung ging eine nahezu 10 Sekunden dauernde unbewußte Vorbereitung voran, wobei es sich nicht um eine diffuse, sondern um eine spezifische Aktivität handelt, die die spezielle Entscheidung charakterisiert. Soon et al.: "Thus, a network of high-level control areas can begin to shape an upcoming decision long before it enters awareness".

Fried et al. (2011) haben Libet-inspirierte Untersuchungen im Rahmen von neurochirurgischen Eingriffen durchgeführt, indem sie die Aktivität von insgesamt 1019 Neuronen im motorischen Kortex bei selbst-induzierten Fingerbewegungen bestimmt haben. Es wurde also die Aktivität einzelner Neurone direkt durch Elektroden gemessen, während die Vpn (in diesem Fall Patienten) Angaben über ihre bewußten Intentionen gemacht haben. Anhand der Daten konnen Fried et al. Vorhersagen über die Entscheidungen der Personen berechnen 1000 ms bevor der Drang ("urge") bzw. der Wunsch zu handeln bewußt wurde. Die Rekrutierung der Neurone insbesondere im SMA, dem Prä-SMA und dem anterior cingulären Kortex geschah dabei in entweder aktivierender oder inhibierender Weise. Die Daten bestätigen die EEG-Befunde, sind aber einerseits sehr viel genauer und lassen andererseits eine viel detailliertere Analyse als die EEG-Daten zu.

In der Tat haben sich Psychologen schon länger mit der Frage beschäftigt, wie

[28]"blood-oxygen-level-dependent (BOLD) signal, detected in fMRI, reflects changes in deoxyhemoglobin driven by localized changes in brain blood flow and blood oxygenation, which are coupled to underlying neuronal activity by a process termed neurovascular coupling".

[29]SMA: **S**upplementary **M**otor **A**rea

Entscheidungen getroffen werden, ohne dabei speziell "freie" Entscheidungen im Blick zu haben. Die allgemeine Annahme ist, dass stets nach Maßnahme irgendwelcher Kriterien entschieden wird. Für "einfache" Entscheidungen, wie sie im Rahmen psychophysischer Untersuchungen von den Vpn gefällt werden müssen, gibt es Modelle, die auf die stochastische Natur der Entscheidungsfindung fokussieren, etwa Marley & Colonius (1992), oder Vorberg & Ulrich (1987), bei denen Reaktionszeiten als Wartezeiten bis zu einer manifesten Entscheidung bestimmt wurden; die Analyse von Reaktionszeiten kann Aufschuß über Details des Entscheidungsprozesses liefern. Busemeyer & Townsend (1993)schlugen ein im Rahmen der Theorie dynamischer Systeme formuliertes Modell für Entscheidungsprozesse vor. In dieser Theorie werden die Wechselwirkungen zwischen verschiedenen Variablen in ihrem zeitlichen Ablauf modelliert, indem die Wechselwirkungen durch Differentialgleichungen repräsentiert werden. Dieser Ansatz ist eigentlich der natürliche Ansatz, Hypothesen über die jeweils untersuchten Prozesse zu formulieren, ist aber in der Psychologie relativ selten verfolgt worden, weil die Details der Prozesse oft nicht direkt beobachtet werden können, was den Test der Hypothesen erchwert. Tatächlich ist Busemeyer & Townsends Modell zu komplex, um an dieser Stelle auf seine Details eingehen zu können, aber es sollte auf jeden Fall für an mathematischen Modellen interessierte Leser erwähnt werden. Auch hier wird auf die Frage nach der Freiheit der Entscheidungen nicht explizit, sondern allenfalls implizit eingegangen, indem angenommen wird, dass Entscheidungen, wie in allen Modellen der hier angeführten Artn, nach Maßgabe eines persönlichen Präferenzsystems gefällt werden. Die so getroffenen Entscheidungen kann man als "frei" betrachten, weil das Präferenzsystem eben ein persönliches ist: die entscheidende Person handelt 'frei', weil sie nassch *ihren* Präferenzen handelt, oder man interpretiert sie als "unfrei", weil die Präferenzen ja die Entscheidungen bestimmen, – das entscheidende "Ich" stellt die Präferenzen der Person ja nur fest. Die Diskussion über die Freiheit oder Unfreiheit der Entscheidungn wird zu einem Streit um des Kaisers Bart. Die auf Kant zurückgehende Charakterisierung freier Entscheidungen, *nicht* Folge irgendwelcher Ursachen und dazu korrspsndierender Wirkungen zu sein, taucht in diesen Modellen gar nicht auf. Gemeinsam ist allen Modellen, dass stochastische Effekte explizit berücksichtigt werden, vergl. Braun (2021), denn gerade bei komplexeren Entscheidungen wird ja nicht durch Anwendung eines rigiden Algorithmus entschieden, sondern oft erst nach längeren Erwägungen ("Deliberationen") verschiedener Aspekte der jeweiligen Situation, z.B.: Soll ich das Angebot auf eine interessante, aber zeitlich begrenzte Stelle im Ausland annehmen, oder das alternative Angebot einer eher langweiligen, dafür aber zeitlich nicht begrenzten Position im Inland? Verschiedene Aspekte der Alternativen liegen i.A. zu Beginn des Entscheidungsprozesses nicht explizit vor, sondern ergeben sich als Resultate einzelner Deliberationsprozesse, die nicht als deterministische algorithmische Prozesse beschrieben werden können. Die Entscheidungsfindung gleicht jedenfalls eher einem allgemeinen *random walk*: es werden Aspekte der Alternativen deutlich, die erst im Delibertionsprozess erkannt werden, etc. Unfreie Entscheidungen sind dann solche, bei denen die entscheidende Person den Eindruck hat, eigentlich keine Wahl zu haben, sich also für eine

der Alternativen entscheiden zu müssen. Wesentlich ist hier die Annahme, dass alle Prozesse eine neuronale Basis haben, von deren Aktivität aber nur ein Teil überhaupt bewußt wird. Dabei muß angenommen werden, dass die Prozesse eine hierarchische Struktur haben, da die Alternativen oder zumindest Aspekte der Alternativen im Laufe eines Prozesses zum Gegenstand bewußter Betrachtungen werden. Das Gehirn muß Ergebnisse seiner eigenen Aktivität bewerten. Man kann der Frage nachgehen, ob diese Selbstreferentialität ein grundsätzliches Hindernis für die Selbsterkenntnis ist oder nicht. Auf Fragen dieser Art soll hier nicht eingegangen werden, weil sie bereits im Rahmen der Thesen von Lucas (1961) und Penrose (1989, 1995) in Teil 1 diskutiert wurden.

1.2 Einige philosophische Ansätze

1.2.1 David Hume

David Hume (1711 - 1776) geht von einer empirisch näherliegenden Konzeption des Begriffs des freien Willens aus. Handlungen – und damit Entscheidungen – sind nach Hume frei, wenn sie auf der Basis des Willens und der Wünsche des Akteurs herbeigeführt werden[30]. Das ist die im vorangegangenen Abschnitt betrachtete Möglichkeit, eine Entscheidung als 'frei' anzusehen, wenn die entscheidende Person nach *ihrem* Präferenzsystem entscheidet. Da man, wie oben angedeutet, derartige Entscheidungen auch als 'unfrei' betrachten kann, kann man die Humesche Definition auch als willkürliche Definition ansehen. Jedenfalls kann man sagen, dass der Humesche Akteur damit auch verantwortlich für seine Entscheidungen ist, denn sie sind eben nach *seinem* Willen und *seinen* Wünschen gefällt worden[31]. Der Punkt hierbei ist, dass dem *Willen* und nicht irgendwelchen Ursachen entsprechend entschieden wird. Man kann es vielleicht so formulieren: der Wille erscheint in Humes Interpretation als Teil des Ich-Bewußtseins, weshalb sich die entscheidende Person mit der Entscheidung als *ihre* Entscheidung identifizieren kann, – deshalb erlebt sie ihre Entscheidung als als 'frei'. Damit kann sich die Person auch als verantwortliche Person erleben.

So kommt es auf die Bildung des jeweiligen Willens an, wobei Hume allerdings nicht auf die Details den Prozesses der Willenbildung eingeht, – die Hirnforschung war zu seiner Zeit noch nicht hinreichend fortgeschritten, um neurophysiologische Prozesse in Betracht ziehen zu können. Der Focus seiner Bschreibungen liegt dementsprechend auf der Phänomenologie des Willens. Die Kausalität natürlicher Prozesse scheint für Hume kein Problem für die Anahme des freien Willens gewesen zu sein, da er die Kausalität eher als Resultat einer kognitiven Täuschung sah.

Eine Entscheidung ist nach Hume nicht frei, wenn sie nicht nach eigenem Willen und eigenen Wünschen gefällt wird (in *A Treatise of Human Nature* (2.3.1–2),

[30] free actions are those that are caused by the agent's willings and desires (Russell (2017))

[31] Hume, D.: An Enquiry Concerning Human Understanding (1748), Abschnitt VIII: of Liberty and Necessity.

bzw. in *Enquiry concerning Human Understanding*). Damit sind freie Entscheidungen mit dem auch von Hume angenommenen Determinismus der physischen Welt kompatibel. Der wesentliche Punkt der Humeschen Definition ist aber, dass die freie Entscheidung nicht akausal entsteht: die Ursache oder der Grund der Entscheidung ist der Wille (oder ein Wunsch) des entscheidenden Akteurs. Wie dieser Wille zustande kommt wird in Humes Betrachtungen zur Freiheit und Notwendigkeit (liberty and necessity) nicht weiter diskutiert.

1.2.2 Kant, Schopenhauer, Nietzsche

1.2.2.1 Immanuel Kant Während David Hume von freiem Willen spricht, wenn eine Entscheidung ohne äußeren Zwang nur nach Maßgabe des eigenen Willens getroffen werden kann, ist für Immanuel Kant (1724 – 1804) eine Entscheidung oder Handlung dann frei, wenn sie nicht dem Kausalitätsprinzip unterliegt. Offenbar gehen Hume und Kant von verschiedenen Definitionen des Freiheitsbegriffes aus. Da nach Kants Ansicht alle natürlichen Prozesse dem Kausalprinzip genügen – jedes Ereignis hat eine Ursache – muß Kant folgern, das menschliche Entscheidungen keine "natürlichen" Prozesse sind. Während Hume schon dem Begriff der Kausalität skeptisch gegenber stand, ist dieser Begriff für Kant ein reiner Verstandesbegriff: wir können Veränderungen in der Welt nur wahrnehmen, wenn wir sie nach einer Regel, der Ursache-Wirkung-Regel, miteinander verbinden. Die Regel resultiert nicht aus der Erfahrung, sondern geht der Erfahrung voraus, sie ist eine notwendige Bedingung für Erfahrungen. Aus der Regel ergibt sich auch die "Anschauungsform" der Zeit: die Zeit ist notwendig, um zwischen einem Vorher (der Ursache) und einem Nachher (der Wirkung) unterscheiden zu können. Kants Betrachtungen zum freien Willen basieren also auf grundlegenden Überlegungen, die man in seiner *Kritik der reinen Vernunft* findet.

Mit Verstand meint Kant die Fähigkeit zu denken, und denken heißt bei ihm, Begriffe zu bilden, Aussagen zu formulieren und von Aussagen auf andere Aussagen zu schließen. Die Vernunft ist dagegen die Fähigkeit, zu apriorischer Erkenntnis zu gelangen, – wobei der Verstand zwar notwendig, aber nicht hinreichend ist. Apriorische Aussagen sind nicht-empirische, analytische Aussagen, die tautologisch sind und deswegen notwendig wahr sind und ohne Ausnahme gelten. Weiter untersucht Kant die Möglichkeit *synthetischer Urteile*, wie zum Beispiel die Aussage "Der Wille ist frei", oder "Die Seele ist unsterblich". Synthetische Aussagen sind nicht analytisch, weil sie sich nicht tautologisch sind, sie sind *synthetisch* bzw. *synthetisch a priori*. Kant geht von der Frage nach der Möglichkeit synthetischer Urteile a priori aus. So nehmen wir Objekte in Raum und Zeit wahr, und Aussagen über Raum und Zeit sind nicht analytisch, weil sie nicht durch Prädikate definiert sind, aus denen die Struktur des Raumes gefolgert werden kann. Für Betrachtungen über die Möglichkeit solcher Aussagen führt Kant den Begriff *transzendental* ein:

> "ich nenne als Erkenntnis *transzendental*, die sich nicht sowohl mit Gegenständen, sondern mit unserer Erkenntnisart von Gegenständen, sofern

> sie a priori möglich sein soll, überhaupt beschäftigt." ((B 25)[32] s. Tetens, p. 34). ... Es geht um "Bedingungen der Möglichkeit von Erfahrung ... dass wir unsere Erkenntnisfähigkeiten ausüben" (Tetens, p. 35). Z 8

Tetens (p. 35, Fußnote) weist darauf hin, dass 'transzendental' u.a. für 'notwendig' steht. Transzendentale Aussagen spezifizieren die Bedingungen für empirische Erkenntnis. Im Allgemeinen sind transzendentale Aussagen metaphysische Aussagen der reinen Vernunft. Nach Kant sucht der Verstand nach Regeln in der Erfahrung, während die Vernunft die obersten Prinzipien hinter den Regeln und damit nach der Einheit des Wissens sucht[33]. Es ist hier wichtig, zu verstehen, was diese Unterscheidung bedeutet. Im allgemeinen Sprachgebrauch wird kaum zwischen diesen beiden Begriffen unterschieden, aber im Rahmen der kantschen Philosophie ist die Unterscheidung denknotwendig, weil er von einer grundsätzlichen Beschränktheit des menschlichen Wissens ausgeht: die Prinzipien der Vernunft haben, anders als die Befunde des Verstandes, keine direkte Beziehung zur Erfahrung. So erzeugt die Idee der Freiheit als notwendige Kategorie menschlichen Handelns, wobei diese Kategorie, wie alle Katgorien der Venunft, nicht beweisbar ist.

Kant postulierte die uneingeschänkte Gültigkeit des Kausalprinzips für die empirische Welt, d.h. ohne dieses Prinzip könnten wir die Welt nicht objektiv erfahren. Da ein uneingeschänktes Kausalprinzip nicht mit einem freien Willen kompatibel ist, muß Kant für die Existenz einer Autonomie argumentieren; er spricht von einer *vernünftigen Autonomie*. Tetens, p. 245, führte dazu die folgende Notation ein, um die kantsche Argumentation für heutige Leser besser zugänglich zu machen, wobei U_j für ein Ereignis steht, das Ursache für U_{j+1} ist:

$$\cdots \Rightarrow U_{n+2} \Rightarrow U_{n+1} \quad \Rightarrow \quad U_n \Rightarrow \cdots \Rightarrow U_1 \Rightarrow W \text{ (natürliches Ereignis)} \quad (3)$$

$$A \quad \Rightarrow \quad U_n \Rightarrow \cdots \Rightarrow U_1 \Rightarrow W \text{ (freie Entscheidung)}, \quad (4)$$

wobei $\Rightarrow$ hier für "bewirkt" steht. Die drei Punkte links in (3) stehen für eine beliebig lange Folge $U_{n+1}, U_{n+2}, \ldots$ von jeweils vorangegangenen Ereignissen in der natürlichen Welt, die sich irgendwo in den Quantenfluktuationen der Welt vor dem Urknall verliert, während A und (4) für ein Ereignis ohne vorangegehende Ursache steht und dem Ereignis W in den meisten Fällen zeitlich nahe ist. A ist eine "unverursachte Ursache". Die Ereignisse U_n bis U_1 sind Ereignisse, die die freie Entscheidung Realität werden lassen, wie etwa eine sprachliche Artikulierung, oder eine motorische Handlung. Dieses Vermögen, eine "unverursachte Ursache" zu generieren, ist, was Kant die *Vernünftige Autonomie* nennt, Kant spricht auch von *Kausalität aus Freiheit*. Da es in der Welt der Erscheinungen (der natürlichen, physikalischen Welt) nur Ursache-Wirkungs-Ketten der Art (3)

[32] Zur Zitierweise: Es gibt zwei, von Kant verfasste Ausgaben der *Kritik der Reinen Vernunft*: Die Ausgabe A von 1781 und die Ausgabe B aus dem Jahr 1787. B 25 heißt Ausgabe B, Seite 25,

[33] in: *Kritik der reinen Vernunft*, Abschnitt *Transzendentalen Dialektik*, insbesondere B 355 ff.

gibt folgt für Kant, dass es unmöglich ist, Freiheit in der empirischen Welt nachzuweisen. Kant argumentiert schon in der *Kritik der reinen Vernunft* (1788) in der Einleitung zum ersten Kapitel, dass der Begriff der Freiheit kein empirischer Begriff, sondern ein reiner Vernunftbegriff sei, der nicht aus der Erfahrung stamme, der aber denknotwendig sei, um moralisches Handeln erklären zu können. Es ist nämlich die Vernunft, so Kant, die die Idee einer Spontaneität (= Akausalität) schaffe (B 561, Tetens p. 246) und die ohne vorangehende Verursachung entsteht. Kant spricht auch von einer *Kausalität aus Freiheit*. Die Vernünftige Autonomie impliziert Kant zufolge

1. die Möglichkeit, unverursachte Ereignisse durch Nachdenken zu erzeugen. Die Möglichkeit des Nachdenkens ist eine Komponente der Vernunft, die somit ebenfalls nicht dem Kausalprinzip unterliegt.

2. Dies bedeutet, dass das Nachdenken nicht nach naturgesetzlichen Regeln abläuft und mithin nicht allgemeinen Kausalgesetzen folgt.

3. Da Empirie sich immer auf die Welt der Erscheinungen und damit auf kausale Strukturen bezieht, und da die Freiheit als Produkt autonomer Vernunft der empirischen Welt nicht angehört, kann Freiheit nicht empirisch bewiesen werden.

Der freie Wille: Eine Antinomie der reinen Vernunft Nun sind aber die Begrifflichkeiten, auf denen metaphysische Aussagen beruhen, häufig nicht hinreichend scharf definiert, so dass sich aus ihnen Thesen ebenso wie deren Negation herleiten lassen. Dies sind nach Kant die *Antinomien der reinen Vernunft*. Kant schreibt (Kursivsetzung von Kant):

> Wenn wir unsere Vernunft nicht bloß, zum Gebrauch der Verstandesgrundsätze, auf Gegenstände der Erfahrung anwenden, sondern jede über die Grenze der letzteren auszudehnen wagen, so entspringen *vernünftelnde* Lehrsätze, die in der Erfahrung weder Bestätigung hoffen, noch Widerlegung fürchten dürfen, und deren jeder nicht allein an sich selbst ohne Widerspruch ist, sondern sogar in der Natur der Vernunft Bedingungen seiner Notwendigkeit antrifft, nur dass unglücklicherweise der Gegensatz eben so gültige und notwendige Gründe der Behauptung auf seiner Seite hat. ((B 448/449), s. Tetens, p. 229). Z 9

Die dritte Antinomie zeigt Widersprüchlichkeiten der Struktur der Begrifflichkeit des freien Willen auf:

> **These:** Die Kausalität nach Gesetzen der Natur ist nicht die einzige, aus welcher die Erscheinungen der Welt insgesamt abgeleitet werden können. Es ist noch eine *Kausalität der Freiheit* zur Erklärung derselben anzunehmen notwendig.
>
> **Antithese:** Es ist keine Freiheit, sondern alles in der Natur geschieht lediglich nach Gesetzen der Natur. Z 10

Beweis der These:[34] Man nehme an, es gebe keine andere Kausalität, als nach Gesetzen der Natur; so setzt alles, was geschieht, einen vorigen Zustand voraus, auf den es unausbleiblich nach einer Regel folgt.

Nun muß aber der vorige Zustand selbst etwas sein, was geschehen ist (in der Zeit geworden, da es vorher nicht war), weil, wenn es jederzeit gewesen wäre, seine Folge auch nicht allererst entstanden, sondern immer gewesen sein würde.

Also ist die Kausalität der Ursache, durch welche etwas geschieht, selbst etwas Geschehenes, welches nach dem Gesetz der Natur wiederum einen vorigen Zustand und dessen Kausalität, dieser aber eben so einen noch älteren voraussetzt usw. Wenn also alles nach bloßen Gesetzen der Natur geschieht, so gibt es jederzeit nur einen subalternen, niemals aber einen ersten Anfang, und also überhaupt keine Vollständigkeit der Reihe auf der Seite der voneinander abstammenden Ursachen. (A 446/B 474)

Nun besteht aber eben darin das Gesetz der Natur: dass ohne hinreichend a priori bestimmte Ursache nichts geschehe. Also widerspricht der Satz, als wenn alle Kausalität nur nach Naturgesetzen möglich sei, sich selbst in seiner unbeschränkten Allgemeinheit, und diese kann also nicht als die einzige angenommen werden.

Diesem nach muß eine Kausalität angenommen werden, durch welche etwas geschieht, ohne daß die Ursache davon noch weiter, durch eine andere vorhergehende Ursache, nach notwendigen Gesetzen bestimmt sei, d. i. eine absolute Spontaneität der Ursachen, eine Reihe von Erscheinungen, die nach Naturgesetzen läuft, von selbst anzufangen, mithin transzendentale Freiheit, ohne welche selbst im Laufe der Natur die Reihenfolge der Erscheinungen auf der Seite der Ursachen niemals vollständig ist. Z 11

Beweis der Antithese: Setzet es gebe eine Freiheit im transzendentalen Verstande, als eine besondere Art von Kausalität, nach welcher die Begebenheiten der Welt erfolgen könnten, nämlich ein Vermögen, einen Zustand, mithin auch eine Reihe von Folgen desselben, schlechthin anzufangen; so wird nicht allein eine Reihe durch diese Spontaneität, sondern die Bestimmung dieser Spontaneität selbst zur Hervorbringung der Reihe, d. i. die Kausalität, wird schlechthin anfangen, so daß nichts vorhergeht, wodurch diese geschehende Handlung nach beständigen Gesetzen bestimmt sei.

Es setzt aber ein jeder Anfang zu handeln einen Zustand der noch nicht handelnden Ursache voraus, und ein dynamisch erster Anfang der Handlung einen Zustand, der mit dem vorhergehenden eben derselben Ursache gar keinen Zusammenhang der Kausalität hat, d. i. auf keine Weise daraus erfolgt.

Also ist die transzendentale Freiheit dem Kausalgesetze entgegen, und eine solche Verbindung der sukzessiven Zustände wirkender Ursachen, nach welcher keine Einheit der Erfahrung möglich ist, die also auch in keiner Erfahrung angetroffen wird, mithin ein leeres Gedankending.

Wir haben also nichts als Natur, in welcher wir den Zusammenhang und Ordnung der Weltbegebenheiten suchen müssen. Die Freiheit (Unabhängigkeit)

[34] http://12koerbe.de/phosphoros/antinom3.htm

von den Gesetzen der Natur ist zwar eine Befreiung vom Zwange, aber auch vom Leitfaden aller Regeln. Denn man kann nicht sagen, daß, anstatt der Gesetze der Natur, Gesetze der Freiheit in die Kausalität des Weltlaufs eintreten, weil, wenn diese nach Gesetzen bestimmt wäre, sie nicht Freiheit, sondern selbst nichts anderes als Natur wäre. (A 447/B 475) Z 12

Kants Kommentar: Natur also und transzendentale Freiheit unterscheiden sich wie Gesetzmäßigkeit und Gesetzlosigkeit, davon jene zwar den Verstand mit der Schwierigkeit belästigt, die Abstammung der Begebenheiten in der Reihe der Ursachen immer höher hinauf zu suchen, weil die Kausalität an ihnen jederzeit bedingt ist, aber zur Schadloshaltung durchgängige und gesetzmäßige Einheit der Erfahrung verspricht, dahingegen das Blendwerk von Freiheit zwar dem forschenden Verstande in der Kette der Ursachen Ruhe verheißt, indem sie ihn zu einer unbedingten Kausalität führt, die von selbst zu handeln anhebt, die aber, da sie selbst blind ist, den Leitfaden der Regeln abreißt, an welchem allein eine durchgängig zusammenhängende Erfahrung möglich ist. Z 13

Kants Argumentationen zu dieser Antinomie sind etwas mühsam zu lesen, weshalb hier Tetens Re-Formulierung vorgestellt wird (Tetens, p. 249):

1. These: Annahme um des Arguments willen – Es gibt keine andere Kausalität als die naturgesetzliche, d.h. für jedes Ereignis X gilt: es gibt ein Ereignis Y, so dass nach einem Naturgesetz aus der Beschreibung von Y logisch eine Beschreibung von X folgt.
2. Jede naturgesetzliche Ursache eines Ereignisses ist selber ein Ereignis.
3. Also ist die regressive Ursachenkette für jedes Ereignis E unabgeschlossen und hat die Form $\ldots U_n \Rightarrow \ldots \Rightarrow U_1 \Rightarrow E$.
4. Wenn die Ursachenkette $\ldots U_n \Rightarrow \ldots \Rightarrow U_1 \Rightarrow E$ unabgeschlossen ist, so ist sie unvollständig.
5. Die unvollständige Ursachenkette $\ldots U_n \Rightarrow \ldots \Rightarrow U_1 \Rightarrow E$ ist keine hinreichende Ursache für E.
6. Nach Voraussetzung hat aber jedes Ereignis eine hinreichende gesetzliche Ursache.
7. Also ist die Annahme falsch, dass es nur naturgesetzliche Kausalität gibt, d.h. es muß abgeschlossene Ursachenketten $U_n \Rightarrow \ldots U_1 \Rightarrow E$ mit einem Anfangsglied, also mit einer ihrerseits nicht verursachten Ursache, also Kausalität aus Freiheit geben. Z 14

Dazu die Gegenthese:

2. Gegenthese: Annahme um des Arguments willen: Es gibt Freiheit im Sinne einer Ursachenkette $U_n \Rightarrow \ldots \Rightarrow U_1 \Rightarrow E$ mit einem Anfangsglied U_n.
2. Alle Glieder einer Ursachenkette sind erfahrbare Zustände.
3. Also ist U_n ein erfahrbarer Zustand.

4. Ist ein Zustand A der absolute Anfang einer Kausalkette, muß per definitionem für alle A vorhergehenden, erfahrbaren Zustände gelten, dass in X nichts vorhanden ist, das nach einem Naturgesetz zum Zustandekommen von A beiträgt (sonst begännen die durch A ausgelösten Wirkungen nicht mit A sondern schon früher).
5. Also ist der Anfangszustand U_n der Ursachenkette $U_n \Rightarrow U_1 \Rightarrow E$ mit keinem der ihm vorhergehenden erfahrbaren Zustände auf irgendeine Weise naturgesetzlich verbunden.
6. Jeder erfahrbare Zustand muß aber aus transzendentalen Gründen als Teil einer Erfahrungswelt mindestens mit einigem ihm vorhergehenden Zuständen als auch mit ihm nachfolgenden Zuständen nach Kausalgesetzen verbunden sein; der Zustand muß sich in den zeitlichen Fluß der Weltzustände kausal einfügen lassen.
7. Also ist der Anfang U_n und damit eine solche Ursachenkette $U_n \Rightarrow \ldots U_1 \Rightarrow E$ gar nicht erfahrbar, mithin ein bloßes "Gedankending". In der empirischen Welt läßt sich "Kausalität aus Freiheit" nicht entdecken. Z 15

Man hat, Kants Argumentationen zufolge, die Befunde
(i) Kausalität aus Freiheit läßt sich metaphysisch nicht beweisen,
(ii) Kausalität aus Freiheit läßt sich aus transzendentalphilosophischen Gründen in der empirischen Welt niemals entdecken.

Also steht man eigentlich mit leeren Händen da, – aber Kant hat noch ein As im argumentativen Ärmel:

"Die Aufgabe, die wir aufzulösen hatten, verbindet uns hiezu gar nicht, denn sie war nur diese: ob Freiheit der Naturnothwendigkeit in einer und derselben Handlung widerstreite, und dieses haben wir hinreichend beantwortet, da wir zeigten, daß, da bei jener eine Beziehung auf eine ganz andere Art von Bedingungen möglich ist, als bei dieser, die Gesetze der letzteren jene nicht bestimmen. Da nun diese Antinomie auf einem bloßen Scheine beruhe, und, daß Natur der Causalität aus Freiheit wenigstens nicht widerstreite, das war das einzige, was wir leisten konnten, und woran es uns auch einzig und allein gelegen war." (B 586) Z 16

Das heißt, Kant will gezeigt haben, dass die Annahme einer Kausalität aus Freiheit zumindest möglich ist, ohne in einen Widerspruch zu geraten.

Die Aufgabe de3 Philosophie sei es nun, diese Antinomie aufzulösen. Dazu geht Kant einerseits davon aus, dass alles, was in der Natur geschieht, strikt dem Kausalitätprinzip folgt. Kant argumentiert, dass es keine Freiheit geben könne, sollte das Kausalprinzip uneingeschränkt gelten, denn dann sei jedes Ereignis durch Ursache-Wirkungs-Ketten bestimmt. Die Annahme, dass es Freiheit gibt, impliziert dann, dass das Kausalprinzip nicht uneingeschränkt gilt, d.h. es

folgt dann die Existenz von "nicht verursachten Ursachen", also die Ursache eines Entscheidungs- oder Willensprozesses, die ihrerseits keine Ursache hat. Kant schreibt dementsprechend:

> "dagegen verstehe ich unter Freiheit, im kosmologischen Verstande, das Vermögen, einen Zustand von selbst anzufangen, deren Kausalität also nicht dem Naturgesetze wiederum einer anderen Ursache steht, welche sie der Zeit nach bestimmte."[35] (B 560/61, s.a. Tetens, p. 244).
>
> Z 17

Kant muß erklären, wie es zu nicht verursachten Ursachen kommt. Dazu formuliert er

Die Zwei-Charakter-Theorie Dieser Theorie zufolge hat jeder Mensch sowohl einen *empirischen* als auch einen *intelligiblen* Charakter. Der Mensch ist Teil der empirischen Welt und unterliegt insofern dem Kausalprinzip. Unser empirischer Charakter unterliegt psychologischen Gesetzmäßigkeiten und kann deswegen von der empirischen Psychologie erforscht werden. Der empirische Charakter ist in die kausale Struktur der empirischen Welt eingebunden, weshalb er als nicht frei erscheint. Sind wir dagegen *vernünftig*, d.h. handeln wir nach Maßgabe vernünftiger Gründe, so agiert der intelligible Charakter, der Teil der *noumenalen Welt* ist, d.h. des "Noumenon"[36]. Tetens (p. 256)) fasst die Kantsche Argumentation so zusammen:

> "Handeln wir aus vernünftigen Gründen, so sind wir frei und der Naturkausalität nicht unterworfen. Vernunft ist das Vermögen zur Kausalität aus Freiheit."
>
> Z 18

Der Begriff der "noumenalen Welt" impliziert also keinen Dualismus à la Descartes, sondern gehört zur Definition des Begriffs 'Vernunft'. Natürlich hat Kant das Recht, Vernunft als Vermögen zur Kausalität aus Freiheit zu definieren[37]. Die Frage ist nur, ob Menschen tatsächlich über diese Art von Vernunft verfügen. Kant stellt diese Frage nicht, ebensowenig stellt er die Frage, welche Art von kognitiven Prozessen, die hinter dem Begriff 'Vernunft' stehen, eine Kausalität aus Freiheit ermöglichen.

Die Zwei-Welten-Lehre: Die Begriffe, mit denen Kant den Freiheitsbegriff charakterisiert, sind konstituierende Bestandteile den Kantschen Zwei-Welten-Lehre,

[35] der Text ist an dieser Stelle etwas verwirrend : 'im kosmologischen Verstande' meint eben die Fähigkeit, eine Kausalkette "von selbst" anzufangen.

[36] von griech. für 'denken'

[37] Für Kant scheint die Möglichkeit derartiger Definitionen im Rahmen seiner 'rationalen' Philosophie nicht in Frage zu stehen. Dann aber ergibt sich sofort die Frage, was 'rational' bedeuten soll, – und warum soll 'Vernunft' gerade in *diesem* Vermögen bestehen? Für den Rationalismus scheint es hinreichend zu sein, dass philosophische Theorien als konsistent erscheinen. Aber dann bleibt die Frage nach der Wahrheit solcher Theorien offen.

derzufolge es außer der empirischen Welt noch die Welt der *intelligiblen Dinge an sich* gibt. Diese Welt ist das Noumenon. Das Noumenon existiert außerhalb der Kausalität, und da Kausalität eine zeitliche Folge der Ereignisse bedeutet, folgt, so Kant, dass das Noumenon außerhalb der Zeit ist. Die Formulierung ist etwas unklar: wie gerade angemerkt wurde, ist nicht gemeint, dass das Noumenon in einer ontologisch anderen, zur empirischen Welt dualen Welt "lebt". Gemeint ist vielmehr, das der Mensch ein *noumenales Wesen* ist, es ist eine Kompetenz des Menschen, vernünftig handeln zu können. Dazu gehört auch der Prozess der *Apperzeption*, das ist die Interaktion von Verstand und Vernunft mit dem sinnlich Wahrgenommenen, die zu einer Gesamtvorstellung von Wahrnehmung und Gedanklichem führt. Die Apperzeption ermöglicht u.a. dem Menschen, sich bis zu einem gewissen Grade selbst erkennen zu können.

Freiheit und Moral: Moralische Regeln und Normen ergeben sich nach Kant aus der Vernunft. Moralisch verhalten wir uns, wenn wir uns nach den Kriterien der Vernunft und nicht etwa nach Kriterien der unmittelbarer Nützlichkeit oder, wie Kant sich ausdrückt, nach "sinnlichem Antrieb" handeln, je mehr "sinnliche Antriebe" in die Entscheidungen eingehen, desto mehr wird die moralische Qualität der Entscheidung reduziert. Um moralisch handeln zu können, müssen wir uns von den "natürlichen Trieben", d.h den Kausalitäten der Natur befreien können, und das gelingt nur, wenn wir nach den Kriterien der reinen Vernunft und der damit verbundenen Einsicht in das jeweils Notwendige geschieht[38]. Die Freiheit des Willens wird somit zur notwendigen Bedingung für moralisches Handeln. Handeln wir also vernunftgemäß und damit moralisch, so sind wir frei von den Zwängen des kausalen Naturgeschehens; nur wenn der Mensch über die transzendentale Freiheit verfügt kann er die dunklen Triebe aus der Welt der naturhaften Sinnlichkeit überwinden.

In der *Kritik der praktischen Vernunft* von 1788 fasst Kant seine Charakterisierung des Freiheitsbegriffs zusammen: er kann nicht aus der Erfahrung abgeleitet werden, sondern

> "Der Begriff der Freiheit ist ein reiner Vernunftbegriff und eben darum transzendent für die theoretische Philosophie, d. i. ein solcher, dem kein entsprechendes Beispiel in irgendeiner möglichen Erfahrung gegeben werden kann, und der also keinen Gegenstand einer möglichen theoretischen Erkenntnis ausmacht, sondern bloß eine notwendige Voraussetzung der praktischen Vernunft ist; und die objektive Realität desselben kann durch keine theoretische, sondern allein durch praktische Vernunft dargetan werden." ((KpV, AA V, S) Z 19

Die Vorrede zur *Kritik der praktischen Vernunft* (1788) liefert dann diese Einsicht:

> Der Begriff der Freiheit, sofern seine Realität durch ein apodikti-

[38] Kants Kategorischer Imperativ: "Handle nur nach derjenigen Maxime, durch die du zugleich wollen kannst, dass sie ein allgemeines Gesetz werde."

> sches Gesetz[39] der praktischen Vernunft bewiesen ist, macht nun den Schlussstein vom ganzen Gebäude eines Systems der reinen, selbst der spekulativen Vernunft aus, und alle anderen Begriffe (die von Gott und Unsterblichkeit), welche als bloße Ideen in der letzteren ohne Stütze bleiben, schließen sich nun an ihn an und bekommen mit ihm und durch ihn Bestand und objektive Realität, d. i. die Möglichkeit derselben wird dadurch bewiesen, dass Freiheit wirklich ist; denn diese Idee offenbart sich durchs moralische Gesetz. Z 20

Die Möglichkeit zu moralischem Handeln impliziert also, nach Kant, die Existenz eines freien Willens, die apodiktischen Gesetze der praktischen Vernunft beweisen diese Implikation, – und damit das Prinzip der Kausalität aus Freiheit, die absolute Spontaneität.

Kritik: In der 3ten Antinomie betrachtet Kant zuerst die These, dass es keinen freien Willen gibt und alles Geschehen dem Kausalitätsprinzip gehorcht: jedes Handeln und Entscheiden hat eine hinreichende Ursache. Die Ursache ist ebenfalls Wirkung einer vorangegangenen Ursache, und so fort. Da *jede* Wirkung eine Ursache hat, gibt es keine erste Ursache, mit der eine Kausalkette beginnt. Die Kausalkette geht beliebig weit zurück, sie hat keinen Anfang, d.h. sie ist unabgeschlossen und damit unvollständig. Da andererseits hat jedes Ereignis eine hinreichende Ursache hat, kann das Kausalprinzip nicht unbedingt gelten, - es muß eine 'Kausalität der Freiheit' existieren, so dass $U_n \Rightarrow \dots U_1 \Rightarrow E$, n endlich, gilt. E sind dann freie Entscheidungen und Handlungen, da für sie das naturgesetzliche Kausalprinzip nicht gilt.

Die Argumentation, die auf das Postulat der Kausalität aus Freiheit führt, ist im Kern rein formal: wenn eine Ursachenkette nicht "unvollständig" sein darf, um frei zu sein, dann muß sie eben "vollständig" sein, d.h. sie kann dann nur endlich viele Glieder $U_n, U_{n-1}, \dots, U_1 \Rightarrow E$ enthalten. U_n entsteht autonom, spontan, – aber was soll man sich darunter vorstellen? Wenn U_n keine Ursache hat, so könnte man auf die Idee kommen, es sei der "reine Zufall", der regiert, – was gelegentlich der Fall sein kann und sicher von Kant so nicht gemeint war.

Hinzu kommt, dass bei naturgesetzlichen Ereignissen die Unvollständigkeit der Kausalketten anscheinend kein Problem ist; Unvollständigkeit bedeutet bei ihnen nicht, dass es keine hinreichende Ursache gibt. Kant postuliert für natürliche Prozesse einen vollkommenen Determinismus, aber es gibt neuere Argumente, dass ein derartiger Determinimus nicht nachweisbar ist (vergl. van Kampens Theorem in Abschnitt 1.1.2[40]

Es gibt einen weiteren kritischen Punkt in Bezug auf Kants Argument, dass

[39] Apodiktische Gesetze sind Aussagen, die unbedingt, notwendig und allgemeingültig wahr sind. Sie sind Produkte der Vernunft oder des Verstandes.

[40] Die Problematik der Unvollständigkeit, die sich aus dem infiniten Regress der Ursache-Wirkungs-Folgen ergibt, ist tiefer als hier angedeutet. ES würde den Rahmen dieses Textes sprengen, hierauf detaillierter einzugehen.

naturgesetzliche Kausalketten "unvollständig" seien, freie Entscheidungen und Handlungen aber spontan und autonom generiert würden. Dieser Punkt resultiert aus der natürlich erscheinenden Annahme, dass Ursache-Wirkungsketten aus diskreten Ereignissen bestehen. Diese Annahme führt auf ein Problem, das bereits von Zenon von Elea (ca 490 - 430 v. Chr.) beschrieben wurde: Achilles kann die Schildkröte, der wegen ihrer relativen Langsamkeit ein Vorspung gewährt wurde, nicht einholen. Denn Achilles benötigt eine Zeit t_1 bis zum Ort, von dem aus die Schildkröte den Wettlauf beginnt. Aber in dieser Zeit hat sich die Kröte um eine Strecke s_1 bewegt. Achilles benötigt für die Strecke s_1 die Zeit t_2, in der die Schildkröte die s_2 zurücklegt. So geht das beliebig weiter. Die t_i, $i = 1, 2, \ldots$ streben zwar gegen Null und sind, so Zenon, am Ende gleich Null, aber Achilles kommt nicht an der Kröte vorbei, weil die Gesamtstrecke aus lauter aneinander gereihten Nullen besteht, und die Summe von noch so vielen Nullen ist doch Null!. Natürlich irrte sich Zenon denn die Gesamtstrecke ist in Wirklichkeit

$$1 + \frac{1}{2} + \frac{1}{4} + \cdots = \sum_{j=0}^{\infty} \left(\frac{1}{2}\right)^j = 2,$$

was Zenon noch nicht wußte, das fand erst Archimedes ca 200 Jahre nach ihm heraus. Geht man davon aus, dass Bewegungen, allgemein Veränderungen, in einem raum-zeitlichen Kontinuum stattfinden, so läßt sich zeigen, dass "unendlich" viele Schritte in endlicher Zeit ablaufen können. Betrachtet man Ursache-Wirkungs"ketten" auf einem Kontinuum, so folgt, dass unendlich viele Ursache-Wirkungs-Elemente durchaus in endlicher Zeit für eine endliche Strecke ablaufen können. Daraus folgt, dass es keinen Grund gibt, eine mysteriöse 'Kausalität der Freiheit' anzunehmen. Kant hat sich mit der Frage nach der Kontinuität natürlicher Prozesse durchaus beschäftigt. Sie gelten für jeden Prozess, – also auch für Entscheidungsprozesse oder freiwillige Akte[41] (in *Metaphysische Anfangsgründe der Naturwissenschaft (1786)*, s. a. McNulty (2019), hat seine Gedanken zu dieser Frage aber nicht mehr auf seine Argumentationen im Zusammenhang mit der dritten Antinomie angewandt. Die Problematik kann und soll hier nicht weiter verfolgt werden, – das Wichtige ist, dass Kants Überlegungen zur Begründung einer Kausalität der Freiheit nicht so stichhaltig sind, wie sie im ersten Moment erscheinen mögen.

Der Beweis der Gegenthese, nämlich dass es Freiheit gibt, so dass Handlungen und Entscheidungen E möglich sind mit $U_n \Rightarrow U_1 \Rightarrow E$ und U_n hat keine Ursache (vergl. Zitat Z 17) liefert einen Hinweis: naturgesetzliche Ereignisse sind *erfahrbar*, sie gehören zur empirischen Welt. Die Kausalität aus Freiheit dagegen ergebe sich aus Vernunftsgründen, sie lebe, wie schon das "Ding an sich", in der Zweiten Welt. Kant greift auf die in der 'Kritik der reinen Vernunft' von ihm entwickelte Begrifflichkeit zurück: handelt ein Mensch 'vernünftig', so handelt er frei,

[41] Kant hat bereits in seiner *Inaugural Dissertation* (1770) Betrachtungen zur Kontinuität allen Geschehens angestellt, wobei er auf Überlegungen Leibniz' zurückgreift ("Die Natur macht keine Sprünge"). Bei seinen Argumentationen bezüglich von Ursache-Wirkungs-Ketten hätte er auf diese Überlegungen zurückkommen können, – aber er ist es nicht. Jankowiak (2020) hat Kants Behandlung dieses Themas ausführlich diskutiert.

Vernunft ist das "Vermögen zur Kausalität aus Freiheit", s. Zitat 18. Er befindet, dass Freiheit ein "reiner Vernunftbegriff" und als solcher eine notwendige Voraussetzung für die praktische Vernunft ist (vergl. Zitat Z 19), und schließlich schreibt er, dass der Begriff der Freiheit durch apodiktische, also denknotwendige, Gesetze *bewiesen* wird (Zitat Z 20). Kant schafft ein sich selbst tragendes System von Begriffen, um seine Idee der Existenz des freien Willens zu begründen; er selbst schreibt, dass "Der Begriff der Freiheit, sofern seine Realität durch ein apodiktisches Gesetz der praktischen Vernunft bewiesen ist, macht nun den Schlussstein vom ganzen Gebäude eines Systems der reinen, selbst der spekulativen Vernunft aus" (Z 20).

Ein weiterer Kritikpunkt ergibt sich aus Kants Argumentation (in der 'Kritik der praktischen Vernunft'), dass die Fähigkeit zu moralischen Urteilen die Existenz eines freien Willens 'praktisch' beweise (vergl. Zitat 20). Kant folgert aus dem Bewußtsein moralischer Gesetze die Existenz der Freiheit. Dieser Schluss ist nicht logisch zwingend. Die Idee hinter Kants Argument ergibt sich aus einer Interpretation der Forderung eines moralischen Gesetzes, diesem Gesetz folgen zu sollen. Kant ist der Ansicht, dass dieses "sollen" bedeutet, dass man dem Gesetz auch folgen "kann", – *kann* man dem Gesetz nicht folgen, so mache es keinen Sinn, zu fordern, dass man ihm folge. Aus der Fähigkeit zu folgen folgt aber nicht notwendig die Freiheit des Willens, denn es kann ja Bedingungen geben, die einen Menschen *determinieren*, dem moralischen Gesetz zu folgen, weil das Nicht-Befolgen Konsequenzen haben kann, die für den Menschen einen Betrag von Schaden haben kann, der größer als der Betrag des Nutzens ist, der sich aus dem Nicht-Befolgen ergibt.

Kants Philosophie ist ohne Zweifel beeindruckend, es ist ein argumentativ geschlossen erscheinendes System von Begriffen, – argumentativ zwingend ist es deshalb noch nicht. Man muß bedenken, dass Kant Vertreter einer Philosophie des Idealismus ist, deren Wurzeln bis in die Philosophie Platons zurückreichen. Kants Begrifflichkeiten scheinen hier ihre Basis haben. Bei jeder Philosophie kann man fragen, worin ihre nicht hinterfragte Basis liegt. Dies ist das Problem der Letztbegründung[42], das im Rahmen der Kritik an Poppers Falsifikatinstheorie diskutiert wurde. Die Letztbegründung führt auf den Begriff der Evidenz, also in die Einsicht in die Wahrheit der Grundannahmen einer Philosopie. Stegmuller (1954) (p. 102) kam nach einer kritischen Betrachtung zu dem Schluß

> *Das Evidenzproblem ist absolut unlösbar, die Frage, ob es Einsicht gibt oder nicht, ist absolut unentscheidbar.* *Z* 21

(s. Wissenschaftstheorie III (2), p. 61). Deshalb haben sich alsbald kritische Philosophen gemeldet, die Kants Theorie in Frage stellten. Hier werden zwei von Kants Kritikern vorgestellt.

[42] Wissenschaftstheorie III (2), http://uwe-mortensen.de/PhilSciIII2Neu.pdf

1.2.2.2 Arthur Schopenhauer Das zentrale Werk Schopenhauers (1788 – 1860) ist das Buch "Die Welt als Wille und Vorstellung" (1818/1819), das in immer wieder veränderten Versionen in mehreren Auflagen erschien. Schopenhauer präsentiert hier eine schon fast mystisch wirkende allgemeine Metaphysik des Willens, der ziellos und blind wirkt, – ähnlich wie die Darwinsche Evolutionstheorie, nach der Veränderungen zufällig bewirkt werden (Richard Dawkins prägte dafür den Ausdruck "the blind watchmaker"). Der menschliche Wille ist denn auch keine bewußte Kraft, sondern eher Ausdruck einer Triebstruktur. Wir erleben also auch den wirklichen Grund unserer Entscheidungen nicht, sondern nur die jeweilige Entscheidung selbst, folglich ist die erlebte Freiheit unserer Entscheidung eine Illusion. Es ist klar, dass diese Positionen nicht mit der Kantschen Philosophie verträglich sind. Tatsächlich enthält "Die Welt als Wille und Vorstellung" (Ausgabe 1887) eine 144-seitige Kritik der Philosophie Kants. In diesem Anhang betont er immer wieder, dass er Kant für einen außerordentlich scharfsinnigen und verdienstvollen Philosophen hält; gleichwohl habe er, Schopenhauer, "allein seine [Kants] Fehler und Schwächen im Auge, stehe ihnen feindlich gegenüber und führe einen schonungslosen Vertilgungskrieg gegen sie ..." (p. 494). Auf diese Kritik und auf seine allgemeine Willenstheorie kann und muß hier nicht eingegangen werden, es genügt, auf den Kern seiner Kritik an Kants Theorie des freien Willens zu fokussieren. 1838 publizierte er "Die beiden Grundprobleme der Ethik", eine Preisschrift über die Freiheit des Willens, für die er den ersten Preis[43] bekam. Das zentrale Argument darin ist die Aussage

> "Der Mensch kann zwar tun was er will, aber er kann nicht wollen, was er will." Z 22

Die Aussage ist deswegen für die Diskussion des freien Willens so wesentlich, weil sie auf ein logisches Problem der Annahme eines freien Willen verweist. In einem Interview für die Zeitung "Die Zeit"[44] kommentierte der Hirnforscher Gerhard Roth Schopenhauers Arbeit zum freien Willen, – sie sei

> "wohl die beste [ist], die je über den freien Willen geschrieben wurde. Der Mensch hat zwar einen Willen, aber er kann diesen Willen nicht selbst willentlich beeinflussen. Das ist auch logisch unmöglich: Wenn wir unseren Willen beeinflussen könnten – wodurch würde der Wille, der unseren Willen treibt, beeinflusst? Wieder durch einen Willen, einen dritten, vierten, fünften? Schon seit dem Mittelalter haben kluge Menschen dieses Problem der willentlichen Willenssteuerung erkannt". Z 23

Natürlich ist Schopenhauers Argumentation kritisiert worden. Ein immer wieder auftauchendes Gegenargument richtet sich gegen den von Schopenhauer bereits implizit schon in "Die Welt als Wille und Vorstellung" postuliertem Determinismus. So argumentierte Robert Chisholm in seiner Arbeit *Human freedom and the*

[43] Königlich Norwegische Societät der Wissenschaften in Drontheim
[44] https://www.zeit.de/campus/2008/02/interview-freier-wille

self (1964), das Menschen "unbewegte Beweger"[45] seien, d.h. sie können Kausalketten initiieren; tatsächlich hat der Neurobiologe Martin Heisenberg einen empirischen Nachweis für die biologische Möglichkeit eines unbewegten Bewegers zu finden versucht (s. Abschnitt 1.2.3, Seite 41). Chisholms Argumentation ist weitgehend phänomenologisch: wir haben einen freien Willen, weil wir uns als frei Entscheidende und als für unsere Entscheidungen Verantwortliche erleben. Diese Argumentation ist sicherlich schon vor ihm vorgestellt worden, und Chisholm merkt gleich zu Anfang der Arbeit an, dass es unmöglich sei, zu dieser Frage [46] etwas Relevantes zu sagen, dass nicht schon von jemand anderem gesagt worden ist. Tatsächlich gelingt es Philosophen immer wieder, den Schluß von der erlebten Freiheit auf die tatsächliche Freiheit zu motivieren, wie die Arbeiten von Habermas und Nida-Rümelin zeigen (s. die Abschnitte 1.2.5.2, Seite 49, und 1.2.5.3, Seite 67).

Auch die hier in Abschnitt 1.3 vertretene These, das Problem des freien Willens sei ein Scheinproblem, ist bereits diskutiert worden (z.B. Cambell (1951)), aber philosophische Widerlegungen von Thesen sind ihrerseits nicht immun gegen Widerlegungen. Ein immer wieder vorgetragenes Argument gegen den Determinismus ist die beobachtbare – d.h. nicht nur postulierte – Fähigkeit des Menschen, sich moralisch zu verhalten. Das Argument konfligiert mit den Befunden der empirisch arbeitenden Psychologie, denen zufolge soziales und damit auch moralisches Verhalten gelernt wird, – woraus wiederum *nicht* folgt, dass die Gesellschaft nicht gegen Formen des unmoralischen Verhaltens vorgehen dürfe, aber das ist ein anderes Problem, das hier nicht diskutiert werden soll. Auch neuere Philosophen argumentieren für die These der Existenz eines freien Willens, z.B. J. Habermas und J. Nida-Rümelin, auf die in gesonderten Abschnitten weiter unten eingegangen wird, die aber auf speziellen Definitionen des Begriffs des freien Willens beruhen.

1.2.2.3 Friedrich Nietzsche Friedrich Nietzsche (1844 – 1900) war gelernter Altphilologe, was sich in seinem philosophischen Stil niederschlägt: seine Schriften wirken eher essayistisch als systematisch deduktiv. Zuerst war er der Schopenhauerschen Philosophie nahe, entfernte sich aber von ihr, weil Schopenhauer ihm zu pessimistisch war, er suchte lieber nach einer optimistischen Interpretation der Welt. Die folgende Beschreibung der Philosophie Nietzsche konzentriert sich auf seine Auffassung des Problems des freien Willens und ist an Röd (2000) orientiert.

Nietzsche begann als Kulturkritiker und Kulturphilosoph; in seinem Werk "Die Geburt der Tragödie aus dem Geiste der Musik" (1872) die Polarität 'dionysisch' versus 'apollinisch' ein, wobei das Dionysische "vitale Energie" bezeich-

[45] Dieser Begriff geht auf Aristoteles zurück: weil die Annahme, dass jedes Ereignis eine Ursache hat, sofort in einen infiniten Regress führt, denn die Ursache muß dann ja selbst eine Ursache haben, etc, argumentierte Aristoteles, es müsse ein unbewegter Beweger existieren, der Kausalketten initiiere.

[46] Perhaps it is needless to remark that, in all likelihood, it is impossible to say anything significant about this ancient problem that has not been said before.

net, die, wenn sie nicht gebremst wird, in einen Rauschzustand führen kann. Das Apollinische ist durch Prinzipien des Maßes, der Harmonie und der Harmonie definiert. Das Begriffspaar charakterisiert seiner Ansicht nach die griechische Kultur: damit Kunstwerke wie die griechische Tragödie entstehen können muß die dionysische Kraft durch apollinische Kontrolle gebändigt werden.

Wie Röd (p. 373) ausführt, steht hinter Nietzsches Polarität Schopenhauers Unterscheidung des von Kant eingeführten Begriffs des Dinges an sich und dem der Erscheinung, wobei bei Schopenhauer das An-sich ein blinder, zielloser Wille ist, die Erscheinung aber durch die Formen des Raumes, der Zeit und der Kausalität definiert wird. Die Erscheinung verbirgt gewissermaßen das An-sich. Nietzsche war mit Schopenhauer einer Meinung: das An-sich sei vollkommen sinnlos. Da aber der Mensch nicht im Bewußtsein dieser Sinnlosigkeit leben kann, braucht er den Schein und die Illusion einer geordneten Welt, Dies sei eine romantisch-mystische Deutung, die Nietzsche aber bald zugunsten einer psychologisch-physiologischen Auffassung der Kunst aufgab[47]. Das Leben sei nicht auf Wahrheit, sondern auf Täuschung angewiesen. Nietzsche nimmt eine anti-intellektualistische Position ein, indem der dem Inellektualisten untertellt, "dass das Denken, an dem Leitfaden der Kausalität, bis in die tiefsten Abgründe des Seins reiche, und das das Denken nicht nur das Sein erkennen, sondern sogar zu *korrigieren* imtande sei. Wie Röd ausführt, wollte Nietzsche das wissenschaftlich-rationale Denken kritisieren, allerdings habe er es gründlich missverstanden, da die Wissenschaftler seiner Zeit nicht, die "Tiefen des Seins am Leitfaden der Kausalität" ausloten wollten, sondern einfach "Regelmäßigkeiten in Bereich der Erscheinungen" ermitteln und in Form von Gesetzeshypothesen ausdrücken wollten[48]. Röd sieht hier einen der Anfänge von irrationalistischen, wissenschaftskritischen bzw. - feindlichen Stömungen in der Philosophie, die im 20. Jahrhundert auftraten (Lebensphilosophien, mystische Betrachtungen, etc).

Diese Auffassung habe Nietzsche allerdings später revidiert. Wissenschaftliche Erklärungen beziehen sich, so Nietzsche, nur auf Erscheinungen und seien nicht mehr "auf Transzendentes bezogen", wie Nietzsche in seinem Werk "Die Fröhliche Wissenschaft" elaboriert. Aber die Wissenschaft beruhe gleichwohl auf der metaphysischen Annahme, dass es eine objektive Wahrheit gebe. Andererseits bekannte er sich nun zur Wissenschaft und wechselte zu einer psychologischen Sichtweise, mit einer "Chemie der Begriffe" wollte er die der Entstehung metaphysischer Vorstellungen, Gefühlen und Wünschen herausfinden. So gelangt er zu einer kritischen Bewertung der These, der Wille sei frei. Seine Bemerkungen dazu sind über verschiedene Werke verteilt.

Er beginnt zu zweifeln, dass Denkprozesse nicht dem Kausalprinzip unterliegen (die Seitenangaben beziehen sich auf Nietzsche (1958)). In "Jenseits von Gut und Böse", § 17, stellt er fest,

> dass ein "Gedanke kommt, wenn *er* will, und nicht, wenn "ich" will; so

[47] Röd, p. 374
[48] Röd, p. 375

> dass es eine *Fälschung* des Tatbestandes ist, zu sagen: das Subjekt "ich" ist die Bedingung des Prädikats "denke". *Es* denkt: aber dass die "es" gerade jenes alte berühmte "Ich" sei, ist, milde geredet, nur eine Annahme, eine Behauptung, vor allem keine "unmittelbare Gewißheit." Z 24

Nietzsche stellt eine Beziehung zwischen Denken und sprachlichen Strukturen her:

> Zuletzt ist schon mit diesem "es denkt" zu viel getan: schon dieses "es" enthält eine *Auslegung* des Vorgangs und gehört nicht zum Vorgang selbst. Man schließt hier nach der grammatischen Gewohnheit "Denken ist eine Tätigkeit, und zu jeder Tätigkeit gehört einer, der tätig ist, folglich ...". Z 25

Hier deutet er an, dass er denkt, dass unser Denken durch die sprachlichen, insbesondere grammatischen Strukturen unserer Sprache zumindest mitbestimmt wird; später wird Habermas unser Denken und Urteilen mit dem 'objektiven Geist' in Verbindung bringen (s. Abschnitt 1.2.5.2). Nietzsche sieht hier ein Denkschema wirken: die ältere Atomistik habe nach der "Kraft" gesucht, "die wirkt, noch jenes Klümpchen Materie, worin sie sitzt, aus der heraus sie wirkt, aus der heraus sie wirkt, das Atom; strengere Köpfe lernten endlich ohne diesen 'Erdenrest' auszukommen" (p. 581).

Das heißt, Denkvorgänge – und damit Deliberationen – werden nicht akausal von einem autonomen "Ich" gesteuert, sondern von einer unterliegenden Dynamik (§ 19):

> "In jedem Willensakte gibt es einen kommandierenden Gedanken (p. 582) – und man soll ja nicht glauben diesen Gedanken von dem "Wollen" abscheiden zu können, wie als ob dann nach Wille übrig bliebe! ... Der Wille ist nicht nur ein Komplex von Fühlen und Denken, sondern vor allem noch ein *Affekt*: und zwar jener Affekt des *Kommandos*. Das, was "Freiheit des Willens" genannt wird, ist wesentlich der Überlegensheit-Affekt in Hinsicht auf den, der gehorchen muß: "ich bin frei, 'er' muß gehorchen – dies Bewußtsein steckt in jedem Willen ... Z 26

Nietzsche fährt fort, den Begriff des Wollens zu analysieren und kommt zu der Aussage

> "Ein Mensch der *will* –, befiehlt einem Etwas in sich, das gehorcht oder von dem er glaubt, dass er gehorcht. Nun aber beachte man, dass das Wunderlichste am Willen ist – an diesem so vielfachen Dinge, für welches das Volk nur *ein* Wort hat: insofern wir im gegebenn Falle zugleich die Befehlenden *und* Gehorchenden sind, und als Gehorchende die Gefühle des Zwingens, Drängens, Drückens, Widerstehens, und Bewegen kennen, welche sofort nach dem Akte xes -willens zu beginnen pflegen; insofern wir andrerseits die Gewohnheit haben, uns über diese Zweiheit vermöge des synthetischen Begriffs "ich" hinwegzusetzen, hinwegzutäuschen, hat sich an das Wollen noch eine ganze Kette von irrtümlichen Schlüssen und folglich von

> falschen Wertschätzungen des Willens selbst angehängt – dergestalt, dass der Wollende glaubt, Wollen *genüge* zur Aktion. Z 27

Etwas weiter unten schreibt Nietzsche dann

> ... der Wollende glaubt, mit einem ziemlichen Grad von Sicherheit, dass Wille und Aktion irgendwie eins seien –, er rechnet das Gelingen, die Ausführung des Wollens noch dem Willen selbst und genießt dabei einen Zuwachs an Machtgefühl, welches alles Gelingen mit sich bringt. "Freiheit des Willens" – das ist das Wort für jenen vielfachen Lust-Zustand des Wollenden, der befiehlt und sich zugleich mit dem Ausführenden als eins setzt, ... Z 28

Das Argument, ein freier Wille könne sich nicht in einem deterministischen System entfalten, führt üblicherweise zu dem Schluß, dass ein freier Wille ein akausales System benötige. Wird dann von der (Hypo-)These ausgegangen, der Wille sei frei, so wird gefolgert, dass er sich sozusagen selbst verursache. Seit man philosophische Aufsätze auf Lateinisch schrieb ist der terminus technicus dafür *causa suis*, – Ursache seiner selbst oder aus sich selbst.

In § 21 von "Jenseits von Gut und Böse" nimmt Nietzsche zu diesem Begriff Stellung. Er schreibt (p. 584):

> "Die *causa suis* ist der beste Selbst-Widerspruch, der bisher ausgedacht worden ist, eine Art logischer Notzucht und Unnatur: aber der ausschweifende Stolz des Menschen hat es dahin gebracht, sich tief und schrecklich gerade mit diesem Unsinn zu verstricken". 29

Der Rest des Paragraphen sind größten Teils Explikationen des Begriffs des Willens (und zum Teil polemische Ausführungen zum Begiff des Willens – er spricht z.B., von der "bäurischen Einfalt des berühmten Begriffs 'freier Wille' ") und insbesondere des Begriffs des unfreien Willens, die hier übergangen werden können. Neuroanatomisches und neurphysiologisches Wissen, wie wir es heute haben, stand zur Zeit Nietzsches noch nicht zur Verfügung, so dass er nicht darauf zurückgreifen konnte, so dass er sich auf introspektiv-phänomenologische Überlegungen und begriffsanalytische Betrachtungen beschränken musste. Deutlich wird jedenfalls seine Überzeugung, dass sich der Mensch i.a. nicht bewußt und rational entscheiden könne, – stets spielen bei den Entscheidungen unbewusste Triebe und Wünsche eifv ne Rolle. Nietzsche ist der Ansicht, dass unser Handeln und Entscheiden Ausdruck "innerer" Kräfte und Lebensbedingungen ist. So steht der vielkritisierte Begriff des "Willens zur Macht" nicht für den Willen, andere zu unterdrücken und Macht über sie ausüben zu wollen, sondern für den Willen eines jeden Individuums, das eigene Potential zu entwickeln, d.h. 'Macht' steht hier für die innere Kaft, die für dieses Ziel aufgebracht werden muß. Dabei ist der Mensch durch und durch ein Naturwesen, es gibt keine zweite Welt der Noumena. Vernunft bestimmt nicht, was man will, sondern Triebe bestimmen die Vernunft.

Es sei angemerkt, dass Nietzsche argumentiert, der freie Wille sei eine Erfindung der Religion, insbesondere des Christentums, da man nun den Begriff der

Schuld und damit auch der Strafe begründen könne; so spricht er[49]

> "Der freie Wille ist das unerquicklichste Theologen-Kunststück." Z 30

Das kann man so sehen.

1.2.3 Martin Heisenbergs Befunde

Der Genetiker und Neurobiologe Martin Heisenberg[50] (1983) argumentiert auf der Basis empirischer Befunde einerseits und der Betrachtungen Immanuel Kants zur Frage nach dem freien Willen andererseits, dass der Wille frei sei. Im Folgenden wird Martin Heisenberg als Heisenberg_M notiert, um ihn von seinem Vater Werner Heisenberg (Heisenberg_W) zu unterscheiden).

Was Heisenberg_M (1983) von Kant unterscheidet ist Kants Ansicht, dass die für den freien Willen vorausgesetzte Akausalität notwendig impliziert, dass eine nicht im Bereich des Physischen existierende Vernünftige Autonomie bzw. Kausalität aus Freiheit existiert. Denn er gibt an, eine biologisch generierte *initiale Aktivität* U_0 gefunden zu haben, so dass ohne die Annahme einer dualistischen Basis für die Hypothese eines freien Willen argumentiert werden könne. In Bezug auf die in (3) und (4), Seite 26, eingeführte Notation beschreibt die durch $\Rightarrow \cdots \Rightarrow$ repräsentierte Folge von Ereignissen nicht mehr die Entscheidungsfindung, sondern die Ausführung der Entscheidung, z.B. die Aktivierung der Muskeln, die den Arm heben, wenn die Entscheidung darin bestand, *jetzt* den Arm zu heben und nicht eine oder 10 Sekunden später.

Heisenberg_M weist darauf hin, dass Entscheidungen nicht notwendig bewußte Entscheidungen sind. Heisenberg_M ist der Ansicht, dass eine Art von initialer Aktivität bereits aus zufälliger Aktivítät resultiert, sofern sie nicht durch äußere Einflüsse generiert wird. Diese zufällige Aktivítät könne schon bei einzelligen Lebewesen beobachtet werden, etwa bei der Bewegung des Bakteriums *Escheria coli*. Das Bakterium verfügt über ein *Flagellum*, dass in zwei Richtungen um die Längsachse des Bakteriums rotieren kann. Bei der einen Richtung wird eine Vorwärtsbewegung erzeugt, bei der anderen eine andere Bewegungsrichtung generiert. Die Orientierungen werden zufällig gewählt, so dass ein zufälliger Pfad erzeugt wird, auf dem Futter und die richtige Temperatur gefunden werden kann. Der wichtige Punkt dabei ist, dass diese Aktivität nicht das Resultat eines externen Einflusses ist. Bei der Entwicklung eines Organismus enstünde das motorische System *vor* dem sensorischen System. Dies gelte auch für komplexere Organismen, bei denen Verhalten auf der Basis von Modulen erzeugt wird: z.B. der Greifreflex bei Neugeborenen, die rythmische Bewegung der Beine beim Gehen, das Zuschnappen eines Krokodils. Die Entwicklung derartiger Module sei das Resultat zufälliger und gesetzmäßiger Prozesse im Gehirn, deren physiologische Details bisher nur unzureichend untersucht worden seien. Aber es gebe hinreichend viele

[49] I: Götzen-Dämmerung, § 7: Die vier großen Irrthümer, (1889

[50] 1940 –

Hinweise, dass die Entwicklung von Verhaltensweisen nicht nur das Resultat externer Einflüsse sei. Heisenberg$_M$verweist auf Untersuchungen an Fruchtfliegen, die ihre Erwartungen über die Konsequenzen ihres Verhaltens auch in Situationen verändern können, die sie in ihrer Evolutionsgeschichte vorher noch nie erlebt haben; Fruchtfliegen können, so Heisenberg$_M$, Probleme lösen, vor denen sie noch nie gestanden haben[51].

Natürlich stellt sich hier die Frage, was mit einem 'zufälligen Ereignis' gemeint ist. Heisenberg$_M$ verweist auf die Ionenkanäle in Zellmembranen, von deren Zustand es abhängt, wann ein Aktionspotential von einem Neuron ausgelöst wird. Diese Kanäle öffnen oder schließen sich "zufällig", wobei Heisen berg$_M$ u.a. an die Prozesse denkt, die radioaktiven Zerfall ausmachen. Nach Ansicht seines Vaters Werner Heisenberg (Heisenberg$_W$) folgt aus der von ihm gefundenen Unschärferelation, dass diese Prozesse akausal und in diesem Sinne "zufällig" sind (Akausalität ist keine notwendige Bedingung für Zufälligkeit im Sinne der Wahrscheinscheinlichkeitstheorie), so dass letztlich das Universum nicht deterministisch beschrieben werden kann. Heisenberg$_M$s Argumentation ist allerdings ein wenig schnell, denn Heisenberg$_W$s Theorie ist zwar Meinungsumfragen unter Physikern zufolge die dominierende Theorie in der Quantenmechanik, aber es gibt Gegenstimmen. Abgesehen davon diskutiert Heisenberg$_M$nicht die naheliegende Frage, ob stochastische Fluktuationen in den quantenmechanischen Prozessen überhaupt eine Rolle bei eher makroskopischen Prozessen der Entscheidungsfindung spielen. Die empirisch gut belegte Tatsache, dass die Aktivität einzelner Neurone stochastische Komponenten enthält läßt sich auch durch die Zufälligkeit der Stoßprozesse (im Sinne der Nicht-Berechenbarkeit der Trajektorien individueller Moleküle) unter Molekülen erklären, die eine zentrale Rolle bei chemischen Interaktionen spielen und die haufig als Diffusionsprozesse charakterisiert werden können (Turings Reiz-Reaktions-Diffusionssprozesse[52]). Weiter muß das Skalen-Problem gelöst werden, d.h. es muß die Frage nach den Zeitskalen, auf denen die verschiedenen Prozesse laufen, beantwortet werden, vergl. He (2014), Kobaleva et al. (2021)). In Bezug auf den freien Willen muß diskutiert werden, ob er noch frei genannt werden kann, wenn eine Entscheidung von einem zufälligen Ereignis abhängt. Tatsächlich ist die Entscheidung dann ja nicht mehr das Resultat eines Willensaktes eines Individuums, sondern eines Prozesses, der das zufällige Ereignis *bewirkt*, – 'zufällig' heißt ja nicht, dass es keine Ursache für das Ereignis gab. Die Relevanz dieser Frage ergibt sich aus der Vermutung, dass das Adjektiv 'frei' nicht mehr sinnvoll verwendet werden kann, weil seine Bedeutung im gegebenen Zusammenhang unklar geworden ist. Einerseits werden eher makroskopische Prozesse betrachtet, zB die Evaluation einer Situation in Bezug auf moralische Aspekte, andererseits sollen quantenmechanische Prozesse eine Rolle spielen. Man

[51]im Originial: "For exanple, my lab has demonstrated that fruit flies in situations they have never encountered, can modify their expecations about the consequences of their actions. They can solve problems that no individual fly in the evolutionary history of the species has solved before. Our experiments show that they actively initiate behaviour."

[52]Eine sehr klare Darstellung findet man in Murray, J.D.: Mathematicial Biology, Springer-Verlag, Berlin, Heidelberg etc 1989

wird auf die grundlegende Frage zurückgeführt, ob das Naturgeschehen letztlich ein insgesamt deterministischer Prozess ist oder nicht.

Heisenberg$_M$s Ansatz ist typisch für Denker und Denkerinnen, deren Hintergrund eine Naturwissenschaft ist. Für die meisten von ihnen ist es zunächst natürlich, anzunehmen, dass sich Entscheidungen auf die Aktivitäten des Gehirns zurückführen lassen. Handelt es sich bei den Denkern aber um professionelle ("studierte") Philosophen, so haben viele von ihnen Bedenken, Entscheidungen und allgemein kognitive Aktivitäten als Ausdruck von ausschließlich neurobiologischen Prozessen zu sehen. Zu einem descartesschen Dualismus möchte allerdings nur eine radikale Minderheit zurückkehren (z.B. Alva Noë , Abschnitt 1.2.5.4), aber eine schon von Aritoteles eingeführte Unterscheidung von Ursachen und Gründen hilft ihnen, auf ausschließlich neurobiologische Erklärungen zu verzichten. Auf diese Unterscheidung soll kurz eingegangen werden.

Um den Bezug zu den Entscheidungen komplexerer Organismen wie etwa Menschen herzustellen fragt Heisenberg$_M$ nach einer Definition von 'frei' und verweist[53] dabei auf Kants Ausführungen zur dritten Antinomie: eine Person handelt frei, wenn sie aus eigenem Antrieb tut, was getan werden muß, nur verwendet Heisenberg$_M$diese spezielle Definition von 'frei' gar nicht weiter, es geht ihm nur um die Akausalität. Tatsächlich ist es denkbar, dass eine Person, die dem kantschen Kriterium entsprechend entscheidet, auch als determiniert betrachtet werden kann: sie hat moralische Vorstellungen entwickelt, die es nicht zulassen, dass sie sich gegen sie entscheidet, und die moralische Entwicklung muß ihrerseits nicht das Resultat einer freien Entscheidung sein, sondern kann ihr durch ein Bestrafungs- und Belohnungssystem oktroyiert oder zumindest antrainiert worden sein. Dass die introjizierten moralischen Normen es im Allgemeinen erschweren, sich gegen sie zu entscheiden, resultiert aus einem komplexen Wechselspiel emotionaler, kognitiver Prozesse. Letztlich bleibt unklar, warum eine initiale Aktivität Freiheit bedeuten soll, - es ist am Ende eine Frage der Definition des Begriffs 'freie Entscheidung'. Heisenberg$_M$ hat sich entschieden, Freiheit so zu definieren, wie er es eben getan hat, – und ob diese Entscheidung eine freie Entscheidung war oder nicht, bleibt eine offene Frage. Womöglich konnte er gar nicht anders, als sich so zu entscheiden, weil seine begrifflichen Strukturen bereits so angelegt waren, dass es ihm erschien, gar keine andere Wahl zu haben.

1.2.4 Ursachen, Gründe, Wittgenstein

Die Unterscheidung zwíschen Ursachen und Gründen hat eine lange Geschichte. Sie scheint zuerst von Aristoteles vorgenommen worden zu sein (Audi (1999), p. 44), so in Aristoteles: Physik II, 3, wo Ursachen als Antworten auf die Frage "Warum?" angegeben werden. Dabei wird 'kausal' als Gegenteil von 'nur zufällig' gesehen. Ursachen beziehen sich auf naturgesetzliche bzw. mechanische Zusam-

[53] let's establish what I mean by freedom. One acknowledged definition comes from Immanuel Kant, who resolved that a person acts freely if he does of his own accord what must be done. Thus, my actions are not free if they are determined by something or someone else.

menhänge (David Hume: Ursachen sind der *Zement des Universums*[54], (Audi (1999), p. 125). Kant unterschied zwischen Gründen der Vernunft und Naturkausalität; letztere entspricht der Wirkung von Ursachen, während die Vernunft Erklärungen und Argumente liefert.

Zu den *Gründen* gehört dementsprechend das, was handelnde Personen eben als Begründung für ihr Handeln angeben, sie haben den Charakter von Erklärungen. So hat Elizabeth Anscombe[55] in ihrem Buch *Intention* (1957) zwischen Handlungsgründen ("reasons") für intentionales, also durch Absichten geleitetes Handeln, und Ursachen ("causes") für bloßes Verhalten unterschieden. Das bloße, durch Ursachen erzeugte Verhalten ist in diesem Sinne nicht rational verstehbar.

Großen Einfluß auf die Debatte um Ursachen und Gründe hatte insbesondere Ludwig Wittgenstein. Er kritisierte in den *Philosophischen Untersuchungen* den insbesondere in der Psychologie vorgenommenen Versuch, Handeln, Denken und das Verstehen von Sprache durch zugrunde liegende, korrespondierende, innere Prozesse kausal zu verstehen. Nach Wittgenstein beruht dieser Ansatz auf einem Mißverständnis: bei diesen Aktivitäten handele es sich um *Rechtfertingungen innerhalb einer Praxis*. Damit sind Gründe ein Teil der *Sprachspiele*, und Regeln gelten als Gründe, aber auch empirische Gesetzmäßigkeiten (PU[56], § 217). Wittgenstein begründet dies mit der Beobachtung, dass die Frage nach dem Grund eines Handelns sich normalerweise nicht auf den inneren Mechanismus, der dem Handeln zugrunde liegt, bezieht. Der 'Grund' ist die Regel selbst. Wittgenstein führt aus (PU, §211): "Wenn ich eine Regel befolge, so tue ich das nicht, weil ich sie interpretiere, sondern ich tue es blind". Wittgensteins Anmerkungen zu Ursachen einerseits und Ursachen andererseits verteilen sich über die Philosophischen Untersuchungen. Nach Wittgenstein ist die Befolgung einer Regel kein Akt, der durch mentale Vorgänge kausal hervorgerufen wird, denn wenn nach dem Grund eines Handelns gefragt wird, so heißt das nicht, dass man Auskunft über einen inneren Mechanismus erhalten müsste: der Grund ist die Regel selbst. Die Frage "Warum?" kann i.A. nicht mit einer kausalen Erklärung beantwortet werden, sondern mit einem Grund: "Warum mußte das Fenster geschlossen werden?" – "Weil es sonst hereingeregnet hätte." Wie Witttgenstein zusammenfasst: Ursachen erklären Ereignisse, wohingegen Gründe Handlungen rechtfertigen (PU, §154 - 155).

In seinem Frühwerk *Tractus Logico Philosopicus* (TLP) hatte Wittgenstein eine Bildtheorie der Sprache vorgeschlagen. Die besagt, dass Sätze logische Bilder von der Welt repräsentieren. Ein Sachverhalt hat eine logische Form, und der den Sachverhalt repräsentierende Satz muß dieser Form Ausdruck verleihen. Aber in PU §308 schreibt er nun: "Wie erklärt man die Bedeutung eines Wortes? – Indem man den Gebrauch erklärt." Allgemein gilt für Wittgenstein, dass Bedeutungen, Absichten und Gründe *keine* inneren, kausal wirkenden Zustände sind, dass diese

[54] the cement of the universe

[55] Gertrude Elizabeth Margaret Anscombe (1919 – 2001, britische Philosophin, Schülerin von Wittgenstein

[56] PU = Philosophische Untersuchungen

vielmehr im öffentlichen Gebrauch erklärt werden.

In diesem Kapitel geht es um die Frage nach dem freien Willen, die Unterscheidung von Gründen und Ursachen wird gewissermaßen nur zur Vorbereitung der Darstellungen der Ansichten von Dretske, Habermas und Nida-Rümelin eingeschoben. Ludwig Wittgensteins Konzeption des Unterschieds zwischen Gründen und Ursachen nimmt darin allerdings eine prominente Rolle ein, weil sich insbesondere Letztere immer wieder zumindest indirekt auf Wittgenstein bezogen haben, indem sie in ihren Argumentationen vom Begriff des Sprachspiels Gebrauch gemacht haben, der vor allem in Wittgensteins Philosophischen Untersuchungen einge wichtige Rolle spielt. Deshalb mag man fragen, ob, und wenn ja wie, sich Wittgenstein explizit über den freien Willen geäußert hat. Wie es scheint, hat die Frage nach dem freien Willen nicht Wittgensteins besonderes Interesse erzeugt; im *Tractatus* findet man nur einen Satz, den Satz 5.1362, über den freien Willen:

> Satz 5.1362, "Die Willensfreiheit besteht darin, dass zukünftige Handlungen jetzt nicht gewußt werden können. Nur dann können wir sie wissen, wenn die Kausalität eine *innere* Notwendigkeit wäre, wie die des logischen Schlusses. – Der Zusammenhang von Wissen und Gewußtem, ist ja der der logischen Notwendigkeit.
>
> ("A weiß, daß p der Fall ist" ist sinnlos, wenn p eine Tautologie ist".
>
> L. Wittgenstein, Tractatus logicus philosophicus Z 31

Der Satz wirkt auf den ersten Blick völlig isoliert, und seine Aussage ist nicht unmittelbar klar. Auf den zweiten Blick bemerkt man, dass er in Aussagen über das logische Folgern eingebettet ist. Wenn man weiß, was geschieht, hat man keine Freiheit mehr, – denn egal, wie man sich entscheidet bzw. was man will, das, von dem gewußt wird, dass es geschehen wird, wird eben geschehen, mein Wollen kann daran nichts ändern, und von Willensfreiheit kann nicht die Rede sein. Auch der logischen Notwendigkeit kann man sich nicht entziehen, so dass aus diesem Blickwinkel heraus die Freiheit der Entscheidung nicht gegeben zu sein scheint. Damit scheint eine notwendige Bedingung für die Freiheit des Willens das Nichtvorhandensein logischer Randbedingungen zu sein.

Der Satz steht im *Tratcatus* ohne weiteren expliziten Bezug auf die vorangehenden oder nachfolgenden Sätze: nach Satz 5.1361 können wir die Ereignisse der Zukunft nicht aus den gegenwärtigen erschließen, und aus Satz 5.1363 *folgt* daraus, dass ein Satz uns einleuchtet, und nicht, dass er auch wahr ist, was erklären mag, warum in den Diskussionen um die Willensfreiheit auf ihn nicht weiter Bezug genommen wird.

Der Satz Z 31 ist aber gleichwohl hilfreich, weil er einen, wenn auch kargen, Aufschluß über die Art gibt, wie zumindest der frühe Wittgenstein über den Freien Willen dachte: der Freie Wille ist für ihn eher ein sprachprogrammatisches Phänomen, das in speziellen Sprachspielen auftauchen kann. In den

Philosopheíschen Untersuchungen (PU) wird der Freie Wille gar nicht systematisch diskutiert und dementsprechend auch nicht explizit genannt, Wittgensteins Ansichten scheinen aber bei der Diskussion anderer, damit verwandter Aspekte von Kognitionen durchzuschimmern, wenn er nämlich z.B. in Teil I, §§611 - 628 über Begriffe wie 'Absicht', 'Entscheidung', und 'wollen' spricht. Hierbei geht es Wittgenstein nicht um innere Ereignisse, die einere Handlung vorausgehen und sie verursachen, was bedeutet, dass ihm nicht – wie bei Kant – um eine metaphysisch begründete Freiheit geht. Für Wittgenstein ist die 'Freiheit' (u.U. des Willens) einfach ein Teil der Grammatik der Begriffe, womit das System von Regeln gemeint ist, das den Gebrauch von Wörten – und damit deren Bedeutung – in Sprachspielen bestimmt. Dies heißt, dass die Bedeutung von Wörtern gewissermaßen in assoziativen Feldern schwebt, ohne notwendig eine feste Bedeutung haben. Dies gilt dann auch für den Ausdruck "freier Wille". Obwohl Wittgenstein diese Auffassung erst in seiner Spätphilosophie entwickelt, zeigt doch seine Aussage Z 31 schon an, dass sie weder in Ursachen verankert noch metaphysisch begründet sieht.

Wittgenstein stellt insbesondere fest, dass *kausalpsychologische Ansätze* für das Verstehen von Handeln, Denken, und Sprache fehlgeleitet seien, – innere Vorgänge seien keine Ursachen. Nach Wittgenstein gilt: fragt man nach einem Grund, so fragt man nicht nach der Ursachen, sondern nach einer *Rechtfertigung einer Praxis*. Gründe beziehen sich auf Sprachspiele, Regeln, auf das Sinnvolle, Richtige, Gerechtfertigte. Ursachen gehören zur Beschreibung der Natur, zu empirischen Gesetzmäßigkeiten, zur Erklärung von Ereignissen

Es gibt demnach einen Unterschied zwischen Handeln und Verhalten. Andere Autoren (Donald Davidson, vergl. Roth (2005)) haben weitere grundsätzliche Betrachtungen zum Unterschied von Gründen und Ursachen angestellt, auf die hier aber nicht weiter eingegangen werden muß, weil sie im Folgenden nicht berücksichtigt werden; hier soll nur der Hintergrund aufgezeigt werden, vor dem Philosophen wie F. Dreske, J. Nida-Rümelin und und J. Habermas ihre jeweiligen Gedanken zum freien Willen entwickeln.

1.2.5 Externalismus, Enaktivismus, und freier Wille

Theorien zur Existenz oder Nichtexistenz des freien Willens werden im Allgemeinen vor dem Hintergrund allgemeiner philosophischer Annahmen formuliert. Sehr spezielle Annahmen werden im Rahmen der Philosophie des *Externalismus*, und da wieder in der Unterklasse der Theorien des *Enaktivismus* (*enacted mind* (Varela et al. (1991)), oder des *erweiterten Geistes* (*extended mind*) (Clark & Chalmers (1998)) gemacht. Kernthese dieser Philosophien ist das Postulat, dass sich Bewußtsein durch eine "transformierende" Interaktion mit der Umwelt ergibt, wobei das Gehirn nicht mehr notwendig der Sitz des Bewußtseins ist. Eine zusammenfassende Übersicht findet man bei Rowlands, Lau, & Deutsch (2020); diesen Autoren zufolge gilt, dass mit 'körperlich' alles, was innerhalb der Haut eines Individuums geschieht, gemeint ist, und in diesem Sinne ist das Gehirn

'körperlich', wobei aber körperliche Prozesse *nicht immer* die mentalen Prozesse bestimmen. So gibt es etwa den inhaltlichen Externalismus (content externalism, der sich auf den Inhalt (die Bedeutung) des mentalen Ereignisses bezieht, der (oder die) nicht immer innerhalb des Körpers (d.h. des Gehirns) bestimmt wird. Dann gibt es noch den "erweiterten Geist" (extended mind, oder enaktivistische Positionen). Einer der ersten, die eine solche Position vertraten, scheint Wittgenstein (1953) gewesen zu sein. Ihm zufolge ergibt sich die Bedeutung eines Zeichens oder Symbols nicht im Rahmen einem eines inneren Prozesses, sondern hängt wesentlich von Gebräuchen oder Praktiken ab, bei denen das Symbol in bestimmter Weise verwendet wird. Von einer Reihe von Philosophen wird diese Charakterisierung wörtlich genommen (ob Wittgenstein seine Aussagen wörtlich genommen hat, scheint Sache der Interpretation seiner Aussagen zu sein), – das ist merkwürdig, denn es wird nicht angegeben, *wo* denn die Bedeutung von Aussagen gebildet und wahrgenommen wird, wenn nicht im Gehirn.

Auf die Details des Externalismus soll hier nicht weiter eingegangen werden, im Folgenden werden nur beispielhaft noch die Gedanken von J. Habermas, J. Nida-Rümelin, und Alva Noé) vorgestellt. Allen gemein ist die Zurückweisung der Annahme, dass Entscheidungen im bzw. vom Gehirn getroffen werden. Begründet wird diese Zurückweisung u.a. mit der Aussage, neurobiologische Prozesse seien ihrer Natur nach deterministisch, wohingegen Entscheidungsprozesse in Wirklichkeit in einem nicht-deterministischen "Raum der Gründe" und nicht in einem kausal strukturierten "Raum der Ursachen" stattfänden und dementsprechend außerhalb des Gehirns ablaufen müssten. Die im Folgenden vorgestellten Denker beziehen sich auf Gründe im Unterschied zu Ursachen, ohne diese Differenzierung weiter zu erläutern, – sie wird als bekannt vorausgesetzt. Deswegen wird die Differenzierung kurz vorgestellt.

Ursachen und Gründe Die Unterscheidung von Ursachen und Gründen in der neueren Philosophie des Geistes geht auf Ludwig Wittgensteins *Philosophische Untersuchungen* zurück, hat aber – wie nicht anders zu erwarten – Wurzeln in der griechischen Philosophie, und da insbesondere bei Aristoteles. Bereits bei ihm sind Ursachen in erster Linie physische Ursachen, während Gründe eher für Prinzipien stehen[57] Spät hat David Hume zwischen theoretischen und praktischen Gründen – letztere korrespondieren zu den Ursachen – unterschieden. Theoretische Gründe korrespondieren zu Hypothesen, die u.U. einer empirischen Überprüfung unterzogen werden müssen, während praktische Gründe Handlungen rechtfertigen sollen. Im Verlauf der weieren Entwicklung bis hin zum (Neo-)Positivismus einerseits und der analytischen Philosophie andererseits wurden die praktischen Gründe mehr als Wirk-Ursachen aufgefasst, die im Sinne einer kausalen Beziehung interpretiert wurden, bis hin zur Auffassung, dass derlei Beziehungen nur noch als regelmäßige Aufeinanderfolge gedacht werden können, – wohl als Folge von Bertrand Russells Aufsatz *On the notion of cause* (1912), in dem er den Begriff der Kausalität gewissermaßen auflöste.

[57]Seine Unterscheidungen sind sehr viel differenzierter, aber hier ist nicht der Platz für eine ausführliche Darstellung.

Gründe sind eher kognitive Konstruktionen; im extremen Fall sind Beweise, wie sie in der Mathematik geführt werden, Begründungen. Bei Wittgenstein dienen Gründe in erster Linie zur Erklärung von Handlungen, wobei eine Beziehung zum Begriff der Vernunft entsteht; im Englischen wird für 'Grund' das Wort *reason* verwendet, das ebenso für 'Argument' und 'Vernunft' steht. Spezielle Auffassungen der Begriffe Ursache und Grund ergeben sich bei verschiedenen Philosophen. Weiter ergeben sich aus der Unterscheidung zwischen Ursache und Grund Betrachtungen zum Unterschied von Verstehen und Erklären.

1.2.5.1 Frederik I. Dretske Frederik Dretske[58] hat aus der Unterscheidung von Gründen und Ursachen (Dretske (1989)) weitgehende Folgerungen gezogen, wobei die Unterscheidung aus einer allmählichen Explikation der Begriffe, insbesondere des Begriffs der Bedeutung besteht. Dretske ist einerseits bekennender Naturalist[59], der gleichwohl feststellt:

> "Meaning certainly isn't an intrinsic property of meaningful things, something you could discover by looking in the head, by taking the measure of the tracks, or by studying the light under a magnifying glass", p. 4, Z 32

Damit meint er vermutlich, dass man durch Beobachtung der neuronalen Aktivität eines Gehirns nicht erfährt, welchen bewußten Zustand dieses Gehirn gerade erzeugt, und man deshalb aus der neuronalen Aktivität auch nicht ablesen könne, warum sich ein Mensch in einer bestimmten Weise entscheidet. Er stellt fest,

> "Meaning itself is causally inert, powerless to initiate, modify, or influence behavior. The best we can hope for is that it supervenes[60] von something, presumably electrical and chemical events in the nervous system, that packs a real behavioral punch. I find this result quite unacceptable." (Dretske (1989), p. 4), und "Even if the things that have meaning are in the head, the meanings themselves aren't in the head." (p. 5), Z 33

Die möglicherweise naive Frage ist, wo die Bedeutung denn sein soll, schließlich kann man davon ausgehen, dass das Gehirn bei der Entwicklung von Bedeutungen eine Rolle spielt. Was soll es heißen, wenn Dretske schreibt: "Die Bedeutung selbst ist kausal inaktiv und kann kein Verhalten initiieren, ändern oder beeinflussen." Wenn jemand im Haus ruft: "Es brennt", ist die Bedeutung unmittelbar klar und kann den jeweiligen Umständen entsprechende Verhaltensweisen auslösen, – aber derartige Trivialitäten hat Dretske vermutlich nicht gemeint, es fragt sich nur, was er denn gemeint hat, – eine Antwort auf diese Frage wird in Dretskes Text nicht geliefert. Es gibt kein Argument für die Aussage, dass Bedeutungen nicht

[58] Frederick Irwin Dretske (1932 – 2013), US-amerikanischer Philosoph

[59] "I am a materialist who thinks that we sometimes do things because of what we believe ...", p. 1

[60] supervenieren: X (physisch) und Y (mental) seien Eigenschaftsmengen. "X superveniert über Y", wenn jede Änderung von X eine Änderung von Y bedeutet, eine Änderung von Y bedeutet aber nicht notwendig eine Änderung von Y.

im Kopf sind, es gibt nur Behauptungen, dass es so sei. Es ist wahr, dass wir nicht wissen, wie aus neuronaler Aktivität Bedeutung entsteht, aber die Aussage, die Bedeutung sei außerhalb des Kopfes erklärt noch nicht, *wie* und *wo* es außerhalb des Kopfes entsteht. Es sei an Singers Betrachtung erinnert:

> "Die Aufklärung der neuronalen Grundlagen höherer kognitiver Leistungen ist mit epistemischen Problemen behaftet. Eines folgt aus der Zirkularität des Unterfangens, da Explanandum und Explanans eins sind. Das Erklärende, unser Gehirn, setzt seine eigenen kognitiven Werkzeuge ein, um sich selbst zu begreifen, und wir wissen nicht, ob dieser Versuch gelingen kann." (Erster Absatz in Singer (2004), Z 34

Singers Aussage charakterisiert den Kern der Frage, wie Bewußtsein entsteht, impliziert aber nicht die (Hypo-)These, dass es nicht im Gehirn entsteht. Medizinische Befunde und empirische Forschung legen die Gegenthese nahe, nämlich dass das Bewußtsein ein Produkt *innerhirnlicher* Aktivität ist. Die Behauptung, es sei ein außerhirnliches Ereignis, resultiert dagegen aus reiner Spekulation, deren Charme wohl in der anti-rationalen Negation des Occamschen Prinzips der Sparsamkeit für Theorien liegt: man weiß nicht, wie genau das Bewußtsein entsteht, aber die naheliegende Annahme, dass es im Gehirn generiert wird, erscheint als "trivial", so dass der Prozess der Bewußtwerdung nach außen verlagert wird. Wie und wo der Prozess "da draußen" erzeugt wird, muß anscheinend nicht weiter erläutert werden.

1.2.5.2 Jürgen Habermas J. Habermas (1929 – 2026) Habermas war einer der weltweit meistrezipierten Philosophen und Soziologen der Gegenwart[61]. Sein Aufsatz "Freiheit und Determinismus" aus dem Jahr (2004) ist die Ausarbeitung seiner Dankesrede für die Verleihung des Kyoto-Preises 2004, was dazu beigetragen haben mag, dass er, wie Roth (2005) feststellte, zu einer der prominentesten Arbeiten zu diesem Thema wurde. Auf Habermas' Arbeit wird hier relativ ausführlich eingegangen, weil seine Weise, für die Freiheit des Willens zu argumentieren, typisch nicht nur für ihn, sondern für eine Klasse von Begründungen für diese These ist. Habermas will kein metaphysisches Argument vorstellen, sondern einen "richtigen"[62] Naturalismus vertreten, wobei die These, dass das Gehirn die Basis für alles Denken und Handeln ist, für ihn eine Provokation darstellt (p. 872)[63]. Habermas kommt nicht auf die Idee, zu begründen, warum das "alltägliche Selbstverständnis handelnder Subjekte" ein Maßstab für

[61] wikipedia.org/wiki/J%C3%BCrgen-Habermas

[62] Habermas erklärt nicht, was ein 'richtiger' Naturalismus im Unterschied zu einem 'falschen'ist.

[63] "Dieser Determinismus ist allerdings mit dem alltäglichen Selbstverständnis handelnder Subjekte unvereinbar. Im Alltag kommen wir nicht umhin, uns gegenseitig und bis auf weiteres die verantwortliche Urheberschaft für unsere Handlungen zuzuschreiben. Die in Aussicht gestellte wissenschaftliche Aufklärung über die naturgesetzliche Determination unseres Handelns kann das intuitiv verankerte und pragmatisch bewährte Selbstverständnis von zurechnungsfähigen Aktoren nicht ernsthaft infrage stellen. Die objektivierende Sprache der Neurobiologie mutet dem "Gehirn" die grammatische Rolle zu, die bisher das "Ich" gespielt hat, aber sie findet damit

eine "richtige" Philosophie sein soll, obwohl doch hinlänglich bekannt ist, dass derlei Selbstverständnis fundamental falsch sein kann. Allerdings scheint eine Teilmenge der Philosophen diesen Sachverhalt nicht als problematisch zu sehen, wie etwa Julian Nida-Rümelin, auf den in Abschnitt 1.2.5.3 eingegangen wird. Habermas' Theorie ist gleichwohl von Interesse, weil sie durch seine Kunst der sprachlichen Kodierung von Theorien in höhere Sphären des Tiefsinns transformiert zu werden scheint. Wer schon an den freien Willen glaubt, wird es auch weiterhin tun, und wer nicht daran glaubt, wird keines Besseren belehrt. Sehr kurz gefasst sind seine Thesen:

- Wir alle schreiben uns "gegenseitig und bis auf weiteres die *verantwortliche Urheberschaft* für unsere Handlungen zu" (p. 871).
- Der freie Wille ergibt sich u.a. aus der "Evidenz eines in allen unseren Handlungen performativ mitlaufenden Freiheitsbewußtseins" (p. 872), das für Habermas neurobiologisch nicht erklärbar ist.
- Die Reduktion menschlicher Kognitionen auf neuronale Aktivität impliziert, so Habermas, einen Epiphänomenalismus[64], der uns handlungsunfähig machen würde. Aber wir sind nicht handlungsunfähig, also entsprechen mentale Operationen nicht neurobalen Aktivitäten, sondern operieren auf Gründen und nicht auf Ursachen
- Der freie Wille läßt sich durch einen *Perspektivendualismus* erklären bzw. charakterisieren. Die eine Perspektive sei der naturalistische Ansatz, d.h. der Ansatz, mentale auf neurobiologische Prozesse zurückzuführen. Die zweite sei die Interpretation der menschlichen Interaktion über Fragen der sozialen Normen, also Fragen zur Verantwortlichkeit und der Moral. Dieser Bereich menschlichen Handelns unterliege nicht dem Kausalprinzip und könne demnach nicht im Rahmen der Naturwissenschaft diskutiert werden; er erschließt sich nur durch philosophische Reflexion.

Dazu beginnt er mit dem kantschen Argument, es müsse einen freien Willen geben, damit verantwortliches Handeln möglich sei: wir alle schreiben uns "gegenseitig und bis auf weiteres die *verantwortliche Urheberschaft* für unsere Handlungen zu" (p. 871). Die Ergebnisse der Hirnforschung legen aber nahe, dass unsere Kognitionen ihre Basis in den Aktivitäten unseres Gehirns haben. Habermas will diesen Sachverhalt nicht bestreiten, ist aber der Ansicht, dass "höhere" Kognitionen, wie etwa das Treffen von Entscheidungen, nicht durch biologische Prozesse *bestimmt* seien.

---

an die Sprache der Alltagspsychologie keinen Anschluss. Die Provokation, die darin besteht, dass "das Gehirn" statt meiner "selbst" denken und handeln soll, ist gewiss nur eine grammatische Tatsache; aber so schirmt sich die Lebenswelt gegen kognitive Dissonanzen erfolgreich ab." Habermas, p. 872

[64] Der E. ist die Theorie, dass Gefühle, Gedanken, etc. zwar ihre Basis in der Hirntätigkeit haben, aber nicht kausal auf den Körper oder allgemein auf die physische Welt zurückwirken.

Die Frage ist, was 'bestimmt' in diesem Zusammenhang heißt. Es gibt keine knappe Definition, die Antwort erschließt sich gewissermaßen allmählich. Dazu stellt er fest, dass neurobiologiche Prozesse als natürliche Prozesse deterministisch seien, so dass die diese Prozesse beschreibenden naturwissenschaftlichen Gesetze notwendig deterministische Gesetze seien, was wiederum mit einem freien Willen nicht kompatibel sei. Es geht ihm um die seiner Ansicht nach neurobiologisch nicht erklärbare "Evidenz eines in allen unseren Handlungen performativ mitlaufenden Freiheitsbewußtseins" (p. 872), und er fragt, ob die "deterministische Auffassung überhaupt eine naturwissenschaftlich begründete These [ist], oder "ob sie nur Bestandteil eines naturalistischen Weltbildes [ist], das sich einer spekulativen Deutung naturwissenschaftlicher Erkenntnisse verdankt?" (p. 872). Dass es auch eine probabilistische Kausalität gibt, scheint ihm bekannt zu sein, aber er findet sie nicht relevant:

> "Die deterministische These ist unabhängig davon, ob wir Naturgesetze probabilistisch interpretieren." Dazu die Fußnote: "Denn Willkür lässt sich nicht auf Zufall zurückführen." (p. 871) Z 35

Diese beiden Sätze bilden einen Stolperstein auf dem Weg zum Vertändnis des Habermasschen Textes, denn ihr Sinn verbirgt sich um so mehr, je länger man über sie nachdenkt. Der zweite Satz (die Fußnote) soll wohl den ersten erklären, aber dazu muß man wissen, was Habermas unter 'Willkür' versteht. Nimmt man an, dass er die umgangssprachliche Beliebigkeit und damit Zufälligkeit meint, so macht der Satz kaum Sinn. Aber man kann mit 'Willkür' auch die Fähigkeit, bewußt zwischen Handlungsalternativen zu wählen, sehen – Willkür als Kür des Willens. Sollte er von dieser Interpretation von 'Willkür' ausgehen, könnten ihm Skeptiker der freien Willens allerdings entgegenhalten, dass "bewußt zwischen Handlungsalternativen zu wählen" nicht notwendig eine 'freie' Wahl bedeutet, denn diese bewußte Wahl kann ja durch die verschiedensten Faktoren determiniert sein. Habermas' Argumentationen bestehen zu einem großen Teil in Versuchen, diese Deutungsmöglichkeit auf sprachlich mesmerisierende Weise auszuschließen und zu begründen, dass die Freiheit des Willens sich aus dem kommunikativen Handeln ergibt und unabhängig von Zufall oder Determinismus im physikalischen Bereich ist.

Habermas will sich der Frage nach dem freien Willen "als eine Auseinandersetzung über die *richtige Weise* der Naturalisierung des Geistes" (p. 872, Habermas' Kursivsetzung) führen; er möchte "ohne die metaphysischen Hintergrundannahmen", wie sie von Descartes (Zwei-Substanzen-Theorie) und Kant (Zweite-Welt-Theorie) gemacht wurden, auskommen. Gleichwohl vertrete er einen nicht-szientistischen[65] Ansatz, den er auch "einen durch philosophische Reflexion

[65]Der Ausdruck 'Szientismus' bzw. 'szientistisch' wurde Ende des 19-ten Jahrhunderts vom französischen Biologen Félix le Dantec (1869–1917) eingeführt, der damit seiner Auffassung Ausdruck verleihen wollte, dass alle Fragen – sofern sie überhaupt sinnvoll sind – mithilfe (natur-)wissenschaftlicher Methoden beantwortet werden können. le Dantec meinte damit einen gewissermaßen verschärften Positivismus. Seit 'Erkenntnis und Interesse' (1973) ist bei Habermas

charakterisierten ··· weichen"-Naturalismus nennt (p. 872).

Habermas verfolgt also keinen Dualismus nach Art René Decartes oder Immanuel Kants, sondern einen Quasi-Dualismus, den er *Perspektivendualismus* (p. 880) oder auch *espistemischen Dualismus* nennt: das Gehirn bringt den individuellen, subjektiven Geist hervor, der über über die Vergesellschaftung des Individuums mit dem objektiven Geist verbunden wird. Der objektive Geist ist das "intersubjektiv geteilte kulturelle Hintergrundwissen", bzw. das "symbolisch gespeicherte objektive Wissen" (p. 885):

> "Ohne Intersubjektivität des Verstehens keine Objektivität des Wissens. Ohne die reorganisierende "Anbindung" des subjektiven Geistes und seines natürlichen Substrats, des Gehirns, an einen objektiven Geist, das heißt an symbolisch gespeichertes kollektives Wissen, fehlen propositionale Einstellungen zu einer auf Distanz gebrachten Welt." (p. 885) Z 36

Dass der subjektive Geist irgendwie mit dem objektiven Geist verbunden ist, ist plausibel, denn existierte keine solche Anbindung, wäre nicht klar, warum ein objektiver Geist überhaupt existieren soll. Diese Anbindung hat anscheinend mit "propositionalen Einstellungen zu einer auf Distanz gebrachten Welt" zu tun, diese Einstellungen sind anscheinend notwendig für die Anbindung. Spekulationen über das, was er unter propositionellen Einstellungen einerseits und einer auf Distanz gebrachten Welt versteht überläßt Habermas, gewissermaßen als hermeneutische Übung, dem Leser. Vermutlich glaubt er, dass naturwissenschaftliche Überlegungen den Menschen von sich selbst entfremden, – aber diese Deutung ist nur eine von vielen möglichen. Eine andere Deutung ist, dass die Aktivitäten des Gehirnes und des subjektiven Geistes Ursache-Wirkungsprozesse sind, die Interaktionen mit dem objektiven Geist aber einer Dynamik der Gründe entsprechen, von der behauptet wird, dass sie nicht der Naturgesetzlichkeit unterliegt.

Habermas ist der Ansicht, dass "reduktionistische Forschungsprogramme", also Forschungsprogramme, die davon ausgehen, dass Kognitionen ihre Basis in der neuronalen Aktivität haben, einen Epiphänomenalismus implizieren (p. 872), aber die Wechselwirkung zwischen dem individuellen, subjektiven Geist und dem objektiven Geist führe auf

> "Das komplexere Bild der Interaktion zwischen einem Gehirn, das den Geist determiniert, und einem Geist, der das Gehirn programmiert, ist Ergebnis philosophischer Reflexion und nicht selbst naturwissenschaftliche Erkenntnis." (p. 872). Z 37

Das ist elegant formuliert, aber wie viele Sätze aus der Habermasschen Sprachschmiede erfordert der Satz exegetische Bemühungen: Zunächst, so Habermas, determiniert das Gehirn den Geist. Das wäre der Teil des Bildes, an dem sich die Naturwissenschaft versucht. Die Frage ist nun, was er meint, wenn er schreibt,

'szientistisch' ein pejorativer Ausdruck für 'empirisch', 'empiristisch', oder 'positivistisch.

dass das Gehirn den Geist determiniert. Vielleicht will er sagen, dass das Gehirn, das seiner Ansicht nach als biologisches Organ deterministisch funktioniert, dem Geist eine bestimmte Struktur gibt, in deren Rahmen freie Entscheidungen möglich sind. Das würde dann implizieren, dass die Freiheit nicht unbedingt, sondern bedingt, also eingeschränkt ist. Nun muß geklärt werden, was es heißen kann, dass der Geist das Gehirn programmiert. Will er zumindest eine bedingte Freiheit des Geistes aufrecht erhalten, ist die Wahl einer Programmierung zu einem gewissen Grade frei, sie wird durch eine nichtdeterministische Interaktion mit dem objektiven Geist mitbestimmt. So ungefähr könnte das Zitat 37 gedeutet werden.

Es bleibt die Frage, was mit diesem Bild erklärt wird. Die Interaktion mit dem absoluten Geist enthält die Komponente 'Freiheit', – weil sie nichtdeterministisch ist. Aber woher weiß Habermas, dass sie nicht-deterministisch ist? Habermas leitet diese Eigenschaft nicht aus irgendwelchen überprüfbaren Grundsätzen ab. Wie es aussieht, behauptet er die Nicht-Determiniertheit einfach. Vielleicht reflektiert diese Behauptung nur eine intuitive Ahnung, die sich durch "philosophische Reflektion" ergeben hat. Jedenfalls hat die Aussage als Behauptung bestenfalls den Charakter einer Hypothese. Um sie zu substantiieren, müßte Habermas näher spezifizieren, wie es zu einer zumindest bedingt freien Interaktion zwischen subjektivem und objektiven Geist kommt, und wie diese Interaktion geartet ist. Seine Betrachtungen zur Interaktion sind derart abstrakt, dass sie kaum noch etwas aussagen, und darüber hinaus muß man bedenken, dass die Interaktion ja nicht in der freien Luft stattfinden kann. Der einzige Ort, der sinnvoller Weise in Frage kommt, ist das Gehirn. Womit die Interaktion ein neurobiologischer Prozess wäre.

Habermas selbst erläutert auf Seite 885 den Ausdrucks 'programmieren' als Kommentar zu Singers (2004) Aussagen zu kulturspezifischen Ausprägungen individueller Gehirne; Habermas schreibt

> "Diese Aussagen scheinen so etwas wie die "Programmierung" des Gehirns durch kulturelle Überlieferungen und gesellschaftliche Praktiken und damit auch eine Interaktion zwischen Geist und Natur nahe zu legen." p. 885, Z 38

Die Rede von der "Interaktion zwischen Geist und Natur" macht klar, dass "Geist" etwas ist, das nicht Natur ist. Damit macht er von der Mehrdeutigkeit des Geist-Begriffs in der deutschen Sprache Gebrauch, obwohl er doch ohne Metaphysik auskommen will. Mit dem Geist-Begriff wird auch das Prädikat 'nichtdeterministisch' assoziiert, weshalb sich Formulierungen wie Z 38 so glatt in seine Texte einfügen lassen. Die von ihm angestrebte "'richtige Naturalisierung" der Debatte um den freien Willen enthält eben doch eine unbestimmte Menge Metaphysik, weil der 'Geist' undefiniert in Habermas' Argumentationen eingeht. Es geht um "eine Einwirkung der grammatisch geregelten und kulturell gespeisten Prozesse bewussten Urteilens und Handelns auf die neuronalen Vorgänge", zu denen Habermas dann die Aussage Z (46), die hier auf Seite 61 wiedergegeben

wird, fügt. Eine Spezifikation des Interaktionsprozesses sind diese Amerkungen allerdings nicht, sie sagen u.a. nichts darüber aus, warum die Interaktionen nicht-deterministisch sind.

Habermas geht dann auf die "Nicht-Hintergehbarkeit der Verschränkung verschiedener Wissensperspektiven" ein, wobei die folgenden Ausführungen gelten sollen:

> "Tatsächlich "existiert" jedoch der Geist nur dank seiner Verkörperung in akustisch oder optisch wahrnehmbaren materiellen Zeichensubstraten, also in beobachtbaren Handlungen und kommunikativen Äußerungen, in symbolischen Gegenständen oder Artefakten. Neben der propositional ausdifferenzierten Sprache, dem Herzstück kultureller Lebensformen, gibt es viele andere symbolische Formen, Medien und Regelsysteme, deren Bedeutungsgehalte intersubjektiv geteilt und reproduziert werden. Diese Symbolsysteme können wir als emergente Eigenschaften verstehen, die sich mit jenem evolutionären Schub zur "Vergesellschaftung der Kognition" herausgebildet haben". (p. 886),
>
> Z 39

Hier räumt Habermas die zentrale Rolle des Gehirns ein, ohne deutlich zu machen, was sie bedeutet, indem er die Aufmerksamkeit auf die Sprache, auf Medien und Regelsysteme, und insbesondere auf emergente Eigenschaften lenkt. Vielleicht soll somit der Fokus auf die von ihm postulierten nicht-gehirnlichen Aspekte der Kognition gelenkt werden. Das Zitat 52 enthält eigentlich kein Argument zur Freiheit des Willens, es enthält nur allgemeine Aussagen, die auch mit einer Negation des freien Willens kompatibel sind, gerichtet ist es aber gegen neurobiologische Interpretationen kognitiver Aktivitäten, wie sie u.a. von Singer (2004) und Roth (2014) vorgebracht werden.

Habermas dagegen erklärt gerade die subjektiv erlebte Freiheit als Grund für die tatsächliche Freiheit des Willens; er geht davon aus (p. 871), dass:

1. Der Determinismus der biologischen Prozesse und die damit vermutete Unfreiheit des Willens ist mit dem "alltäglichen Selbstverständnis handelnder Subjekte unvereinbar", und

2. "Im Alltag kommen wir nicht umhin, uns gegenseitig und bis auf weiteres die verantwortliche Urheberschaft für unsere Handlungen zuzuschreiben." Diese Urheberschaft kann durch "objektivierende Sprache der Neurobiologie nicht ernsthaft infrage gestellt" werden,

3. Entscheidungsprozesse seien Abwägungen von Gründen, die die Interaktion zwischen subjektivem und objektivem Geist charakterisieren. Anders als Ursachen seien Gründe nicht Elemente eines kausalen Beziehungsgefüges, weshalb Entscheidungen auf der Basis von Gründen freie Entscheidungen seien.

Um bei seiner Begründung des freien Willens ohne 'metaphysische Hintergrundsannahmen' auszukommen verweist Habermas auf die analytische Handlungstheorie, die den "Weg zu einem nicht-deterministischen Begriff bedingter Freiheit" weise (p. 872). Es sei der "interne Zusammenhang mit Gründen, der Handlungen erst zu freien Handlungen macht" (p. 873/874). Wie es zu Handlungen und Willensbildung kommt beschreibt er wie folgt:

> "Ein Wille bildet sich, wie unmerklich auch immer, im Zuge von Überlegungen. Und weil ein Entschluss erst in der Folge von wie immer auch flüchtigen und unklaren Erwägungen heranreift, erfahren wir uns nur in den mehr oder weniger bewusst vollzogenen Handlungeng als frei." Habermas fügt weiter unten hinzu: "Frei ist nur der überlegte Wille" (p.874), Z 40

Mit dieser Beschreibung einer Willensbildung können auch Skeptiker des freien Willens ihre Skepsis begründen, – bis auf den letzten Satz, demzufolge nur der "überlegte Wille" frei sein kann. Habermas liefert keinen Grund, aus dem *folgt*, dass der so gebildete Wille ein freier Wille sein *muß*, er behauptet nur, dass der so gebildete Wille frei ist, vorausgesetzt, er ist überlegt. Nach Habermas orientieren sich freie Entscheidungen an Gründen. Entscheidet eine Person unüberlegt, wird ihr Handeln durch "Ursachen", also durch gesetzmäßig ablaufende neuronale Aktivitäten, gesteuert, und deshalb handelt sie unfrei. Also: sich etwas zu überlegen bedeutet die Möglichkeit der freien Entscheidung, – aber warum? Sich etwas zu überlegen kann ja ein determinierter Prozess sein, der die Entscheidung bestimmt, – woraus also schließt Habermas, dass dies keine Möglichkeit ist? Sätze wie "Frei ist nur der überlegte Wille" sind apodiktische Suggestivaussagen, keine Argumente.

Wenn die Freiheit einer Entscheidung sich aus ihrer Orientierung an Gründen ergibt, muß geklärt werden, was Gründe zu Gründen macht, und dass das Konzept eines Grundes die Freiheit von Entscheidungen und Handlungen impliziert, – und aus diesem Sachverhalt folgt, dass man wissen muß, was eine "freie Entscheidung" überhaupt ist. Wie schon in der Einleitung des Kapitels angemerkt wurde, ist diese Freiheit eigentlich nur negativ definiert: eine Entscheidung oder Handlung ist *nicht* frei, wenn die Auswahl aus einer Menge von Alternativen in irgendeine Weise eingeschränkt ist, vielleicht sogar nur auf nicht gewünschte Alternativen eingeengt ist. In Bezug auf das Zitat 40 kann man noch sagen, dass der resultierende Wille unfrei ist, wenn der Deliberationsprozess "unfrei" ist, d.h. ausgedrückt im Vokabular der Philosophie des Geistes zumindest Folgen von Ursachen enthält, wenn er nicht mindesten eine Teilfolge von Ursache-Wirkungs-Ketten enthält. Akzeptiert man diese Überlegung, so ist immer noch nicht klar, was es bedeutet, was "frei" – positiv definiert – bedeutet, außer, dass die Deliberation keine Teilfolge der determinierten Art enthält. Die Frage ist, welche Instanz zwischen verschiedenen Gründen im Rahmen der Deliberation entscheidet. Das "Ich", könnte man vermuten, aber diese Vermutung hat nur dann eine erklärende Funktion, wenn man über eine explizite Konzeption von dem hat, was das "Ich" eigentlich ist; Habermas parodierend könnte man sagen, das Ich

sei eine 'performativ mitlaufende Selbstreferenz', womit – im Sinne des habermasschen Philosophiestiles – alles gesagt und – im Sinne naturwissenschaftlicher Argumentation – nichts erklärt wird.

Unklar bleibt, was Gründe von Ursachen unterscheidet. Die Unterscheidung ist wichtig, weil nicht nur nach Habermas Gründe und mentale Operationen auf Gründen nicht neuronal konzipiert werden können, wohingegen neuronale Aktivitäten, weil biologisch, determiniert und ergo unfrei seien. Die Frage ist, woher Habermas die Überzeugung nimmt, dass Gründe grundsätzlich nicht mit neuronalen Aktivitätsstrukturen identifiziert werden können. Um seine Überzeugung zu begründen, müsste er zeigen, dass komplexe Einheiten wie Gründe durch keinerlei neuronales Netzwerk generiert werden können. Ein solcher Nachweis dürfte ungemein schwierig sein, der Hinweis, biologische Strukturen seien nun einmal deterministisch, genügt nicht, weil er einfach zirkulär ist: 'deterministisch' bedeutet unfrei, aber wir sind ja 'frei" – wegen des performativ mitlaufenden Freiheitserlebnisses und der verantwortlichen Urheberschaft, also sind 'Gründe' nichtbiologische, nicht-deterministische Wesenheiten[66].

Habermas argumentiert, dass Gründe eine "rationale Motivation" für eine Entscheidung liefern, und dass wir *Urheberschaft* erleben – d.h. wir müssen uns als diejenigen erleben, die die Entscheidung treffen. Dazu führt er aus:

> "können wir nur aus der Perspektive des Teilnehmers am öffentlichen Prozess des "Gebens und Nehmens von Gründen" (Robert Brandom[67]) erklären. Deshalb muss ein Beobachter das Diskursgeschehen in einer mentalistischen Sprache beschreiben, das heißt in einer Sprache, die Prädikate wie 'meinen' und 'überzeugen','bejahen' und 'verneinen' enthält. In einer empiristischen Sprache müsste er, aus grammatischen Gründen, alle Bezüge auf die propositionalen Einstellungen von Subjekten, die etwas für wahr oder falsch halten, eliminieren[68]. Aus dieser Sicht verwandelt sich das Diskursgeschehen in ein gleichsam hinter dem Rücken der Subjekte ablaufendes Naturgeschehen." (p. 875) Z 41

Also: Gründe werden mit Diskursen, die sich im Rahmen der Vergesellschaftung des Individuums abspielen, in Zusammenhang gebracht, d.h. der Hintergrund von

[66]Man ist versucht, auf KI-Programme wie ChatGpt oder Googles KI-Programm hinzuweisen: mit diesen Programmen kann man durchaus sinnvolle Diskussionen über die Unterschiede zwischen Humes und Kants Auffassungen zum freien Willen führen; auch wenn die Programme Gebrauch von Bayes' Theorem machen, können sie deterministisch operieren, z.B. gemäß der Regel, eine gedankliche Alternative (welche auch immer) mit Wahrscheinlichkeit 1 zu wählen, wenn sie die höchste a posteriori-Wahrscheinlichkeit hat. Dass die Programme auch Fehler machen, ist insofern kein Gegenargument, als Menschen ebenfalls Fehler machen.

[67]Robert Boyce Brandom (1950 –), US-amerikanischer Philosoph, der u.a. empiristisch/naturalistischen Bedeutungstheorien ablehnend gegenüber steht; zum Beispiel können Diskurse nur in mentalistischer Sprache beschrieben werden, die nicht in Aussagen der Neurobiologie übersetzbar sei. Habermas ist von Brandom beeinflußt worden.

[68]Diese Sätze werden klarer, wenn man weiß, dass Habermas sich hier auf Begriffe aus der analytischen Philosophie bezieht: eine 'propositionale Einstellung' drückt eine Haltung zu einer Aussage aus: glauben, dass (ich glaube, dass es morgen regnet), bezweifeln, dass (ich bezweifle, dass Trump den Friedens-Nobelpreis bekommt), etc

Habermas' Theorie des freien Willens liegt in Habermas' *Diskursethik*, die er in seiner *Theorie des kommunikativen Handelns* (1981) entwickelt. So räumt Habermas ein, dass eine handelnde, insbesondere entscheidende Person nicht unbedingt frei ist: "Einsichten entstehen nicht willkürlich, sie bilden sich nach Regeln. Hätte der zum Handeln Entschlossene anders geurteilt, hätte er anders gewollt." (p. 875) Einschränkungen der Freiheit entstehen also durch Regeln, an die sich die handelnde Person gebunden fühlt, aber

> "Andererseits lässt sich die handlungsmotivierende Rolle von Gründen nicht nach dem Modell der Verursachung eines beobachtbaren Ereignisses durch einen vorangehenden Zustand begreifen. Der Prozess des Urteilens *ermächtigt*[69] den Handelnden zum Autor einer Entscheidung. Durch einen kausal erklärbaren Naturprozess würde sich der Handelnde entmächtigt, nämlich seiner Initiative beraubt fühlen. Deswegen ist es nicht nur grammatisch falsch zu sagen: Wenn der Aktor anders geurteilt hätte, hätte er anders wollen "müssen". Der zwanglose Zwang des besseren Arguments, der uns zu , "Ja-" und "Nein"-Stellungnahmen motiviert, ist von dem kausalen Zwang einer auferlegten Einschränkung zu unterscheiden, die uns nötigt, anders zu handeln, als wir wollen: "Wenn wir die Urheberschaft vermissen, so liegt es daran, dass es uns misslingt, als Denkende und Urteilende Einfluss auf unser Wollen und Tun zu nehmen. Freiheit in diesem Sinne ist nicht nur mit Bedingtheit verträglich [...]; sie verlangt Bedingtheit und wäre ohne sie nicht denkbar." (p. 875) Z 42

So geht argumentieren: die Wahl eines Wortes ('ermächtigt') soll erklären und beweisen zugleich. Aber man kann doch fragen, warum ein Urteilsprozess überhaupt 'ermächtigt' und warum darüber hinaus dieses Wort die Nicht-Kausalität von Zustandsfolgen impliziert. Habermas' Aussage ist kein Argument, es ist ein Versuch der verbalen Überrumpelung. Es ist also der "Der zwanglose Zwang des besseren Arguments", der uns aus der physikalischen Determiniertheit befreit. Für Habermas kann das Entscheiden nach Gründen nicht als eine Folge von Verursachungen gesehen werden, – seine apodiktische Art, diese Ansicht als Einsicht zu präsentieren, ergibt sich aus seiner Definition des Freiheitsbegriffs: Willensfreiheit ist damit kein Merkmal, kein Aspekt kognitiver Aktivität, über das bzw. über die wir verfügen, oder auch nicht verfügen. Die Bedeutung von 'Willensfreiheit' wird bei Habermas durch eine spezielle Setzung festgelegt. Dabei insistiert er, der freie Wille könne nicht in unmittelbarer Relation zu Hirnprozessen gedacht werden, denn Gründe seien in der Hirnaktivität – wie man sie in EEG-, oder fMRI- oder MEG-Untersuchungen sichtbar machen kann – nicht zu finden. Habermas übersieht hier, dass derlei Untersuchungen mit begleitenden Verhaoltensbeobachtungen einhergehen können (worauf in Abschnitt 1.3 noch ausführlicher eingegangen wird, vergl. auch Roths Kritik weiter unten. Neuronale Prozesse seien, Habermas, *Ursachen* für Reaktionen, freie Entscheidungen resultieren dagegen aus der Abwägung von *Gründen*. Diese Aussage scheint für Habermas den Begriff der freien Entscheidung in der Tat zu definieren: eine Entscheidung ist dann

[69] Kursivsetzung durch Habermas

'frei', wenn sie das Resultat einer Abwägung von Gründen ist und nicht durch Ursachen bestimmt wird. Er unterstellt dabei, dass Gründe nicht determinierend sind – wären sie es, wäre die Entscheidung ja unfrei. Aber worin besteht dann die 'Abwägung von Gründen'? Das Abwägen ist der Prozess der Bildung eines Urteile, oder des Treffens einer Entscheidung. Habermas muß unausgesprochen annehmen, dass dieser Prozess aus freien Schritten (Teilprozessen) besteht, damit die Entscheidung 'frei' sein kann. Aber warum soll das so sein, woher "weiß" er das? Der bloße Ausdruck "Abwägung von Gründen" sagt noch gar nichts über die Freiheit des Prozesses aus, ist also kein Argument für die Freiheit des Willens.

Aber Habermas insistiert: Willensfreiheit sei keine Eigenschaft neuronaler Aktivität. Freiheit wird aber als solche erlebt, wozu das Erleben von Urheberschaft gehört (vergl. Zitat 42), – und deswegen existiert die Freiheit des Willens. Das wirkt widersprüchlich, und ist es vermutlich[70] auch. Man muß Habermas sorgfältig lesen, um zu merken, dass Habermas nicht wirklich begründet, dass 'Gründe' nicht durch neuronale Prozesse generiert werden, sondern dass der diese These nur suggeriert. Seine Aussage, dass Gründe in den oben genannten Untersuchungen nicht sichtbar werden, mag auf den ersten Blick als 'wahr' erscheinen, vernachlässigt aber wesentliche Komponenten der Hirnforschung, die darin bestehen, Hirnaktivität mit Verhalten zu korrelieren. So ist die Wachkraniometrie ein routinemäßig eingesetztes Verfahren, um mittels in das Hirngewebe eingeführten Elektroden Hirnaktivität und z.B. Sprachverhalten zu korrelieren[71], um die Schäden etwa bei der Entfernung eines Tumors zu minimieren. Das Argument, dass "Gründen" keine neuronale Aktivität zugrunde liegt, wird nicht nur von Habermas behauptet, allerdings ohne wirkliche Begründung. Diese würde etwa darin bestehen, dass Zustände realer neuronaler Netze, die einen 'Grund' repräsentieren, nicht existieren können. "Philosophische Reflexion" nach habermasscher Art wird einen solchen Beweis aller Wahrscheinlichkeit nach nicht liefern können.

Perspektivendualismus Nach Habermas gibt es zwei Sichtweisen, also Perspektiven, auf das menschliche Verhalten. Die erste, die Außenperspektive, ist die der Naturwissenschaft. Sie ist nach Habermas' Meinung deterministisch, weil von der allgemeinen Gültigkeit des Kausalitätsprinzips ausgegangen wird, weshalb es aus dieser Sicht keine Willensfreiheit geben kann.

Die zweite Perspektive sei die der teilnehmenden Beobachtung. Zur menschlichen Interaktion gehört der kommunikative Austausch unter anderem über moralische Fragen und damit über Fragen der Verantwortlichkeit für das Verhalten. Die Notwendigkeit, Verhalten nicht nur beobachten, sondern auch bewerten zu müssem, u.a. in der Strafjustiz, erfordert die Betrachtung von Gründen, die nicht dem Kausalprinzip unterliegen. Der Begriff der Verantwortlichkeit setzt den der Willensfreiheit voraus, die damit wesentlich für Demokratie und Rechtsstaatlich-

[70] Man kann sich bei Habermas nie sicher sein, ob man ihn richtig verstanden hat, vergl. die Bedeutung von 'Willkür' in Zitat 35, Seite 51.

[71] Die in der Statistik bekannte Binsenwahrheit, dass Korrelationen nicht notwendig Kausalität abbilden, gilt hier nicht!

keit wird. Diese Perspektive entziehe sich "den heute bekannten Naturwissenschaften" (p. 881).

Nach Habermas müssen beide Perspektiven berücksichtigt werden, soll Verhalten versteh– und beurteilbar sein, und keine der beiden Perspektiven könne auf die andere zurückgeführt werden. Zu jeder Perspektive gehöre ein spezifisches *Sprachspiel*[72]. Habermas formuliert

> Aus der anthropozentrischen Sicht einer Lebensform von problemlösenden Sprach- und Kooperationsgemeinschaften vergesellschafteter Individuen bleiben die beiden Vokabulare und Erklärungsperspektiven, die "wir" der Welt auferlegen, für uns "nicht-hintergehbar". Das erklärt die Stabilität unseres Freiheitsbewusstseins gegenüber dem naturwissenschaftlichen Determinismus. Andererseits können wir den organisch verwurzelten Geist nur so lange als eine Entität in der Welt verstehen, wie wir den beiden komplementären Wissensformen keine Geltung apriori zuschreiben. Der epistemische Dualismus darf nicht vom transzendentalen Himmel gefallen sein. (p. 881), Z 43

Der erste Satz soll wohl bedeuten, dass Menschen i.a. in sozialer, und das heißt insbesondere sprachlicher Interaktion mit anderen Menschen leben, die "beiden Vokabulare" beziehen sich auf (i) die objektive und (ii) die teilnehmende Perspektive, und 'nicht hintergehbar' heißt wohl einfach, dass wir uns nur in jeweils einer dieser beiden Perspektiven ausdrücken können, d.h. es gibt keine Metaperspektive hinter diesen beiden Perspektiven. Die beiden Perspektiven korrespondieren zu den Kategorien 'Erklären' (Naturwissenschaft) und 'Verstehen' (Sozialwissenschaft)[73]. "Die Stabilität unseres Freiheitsbewusstseins" ergibt sich eben aus der Existenz der beiden Komponenten des Dualismus; die teilnehmende Perspektive wird, so Habermas, durch die objektive nicht überflüssig.

Dass man nicht hinter den Perspektivendualismus zurück gehen kann begründet Habermas so: "Weil die korrespondierenden Sprachspiele und Erklärungsmuster nicht aufeinander reduziert werden können; Gedanken, die in einem mentalistischen Vokabular ausgedrückt werden, lassen sich nicht ohne semantischen Rest in ein empiristisches, auf Dinge und Ereignisse zugeschnittenes Vokabular übersetzen" (p. 882). Das liege daran, "weil in die Grammatik der beiden Sprachspiele unvereinbare Ontologien eingebaut sind. Seit Frege und Husserl wissen wir, dass sich propositionale Gehalte oder intentionale Gegenstände nicht im Bezugsrahmen kausal wirksamer, raumzeitlich datierbarer Ereignisse und Zustände individuieren lassen"[74]:

[72]Der Begriff des Sprachspiels wurde von Ludwig Wittgenstein in seinen *Philosophischen Untersuchungen* eingeführt. Ein Sprachspiel ist eine sprachliche Äußerung in einem bestimmten Kontext, der u.a. die jeweilige Bedeutung von Ausdrücken festlegt.

[73]Eine ausführliche Diskussion des Begriffspaares Erklärfen-Verstehen findet man in http://www.uwe-mortensen.de/RingvorlesungNeu.pdf

[74]Habermas bezieht sich hier auf Frege (1892), der argumentierte, dass propositionale Gehalte bzw. Gedanken nicht auf kausal-raumzeitliche Ereignisse zurückgeführt werden können, sondern unabhängig von individuellen psychischen Akten existierende objektive, abstrakte Inhalte dar-

> Das lässt sich auch an der Verschränkung des Kausalbegriffs mit dem Funktionskreis instrumenteilen Handelns erläutern. Indem wir die Aufeinanderfolge von zwei beobachteten Weltzuständen A und B als eine kausale Beziehung (in dem strengen Sinne, dass der Zustand A eine hinreichende Bedingung für das Auftreten von B ist) interpretieren, lassen wir uns implizit von der Vorstellung leiten, dass wir selbst den Zustand B hervorrufen könnten, indem wir instrumentell in die Welt eingreifen und den Zustand A herbeiführen. Dieser interventionistische Hintergrund des Kausalitätsbegriffs macht deutlich, warum sich mentale Zustände oder semantische Gehalte, die wir nicht auf instrumentellem Wege wie Dinge und Ereignisse manipulieren können, dieser Art von kausalen Erklärungen entziehen. (p. 882) Z 44

Jeweils auf Geistiges oder Physisches zugeschnittenen Sprachspiele lassen sich also, so Habermas, nicht aufeinander reduzieren. Also müsse sich die Beobachterperspektive, auf die uns das empiristische Sprachspiel beschränkt, "mit der eines Teilnehmers an kommunikativen und gesellschaftlichen Praktiken verschränken, um vergesellschafteten Subjekten wie uns den kognitiven Zugang zur Welt zu öffnen. Wir sind Beobachter und Kommnunikationsteilnehmer in einer Person" (p. 882). Habermas spricht auch von der *komplementären Verschränkung* von Teilnehmer- und Beobachterperspektive (p. 882). Das Wesen dieser Verschränkung hat er in einer nachgerade erhabenen Formulierung auf den Punkt gebracht, deren analytische Kommentierung das Wesen des Punktes als einem Ort ohne, in diesem Fall semantische, Ausdehnung decouvrieren würde:

> Die Perspektivenverschränkung wird im grammatisch geregelten Austausch der Kommunikationsrollen von Sprecher, Adressat und Beobachter fixiert, wenn das Kind im Zuge des Spracherwerbs das System der Personalpronomina beherrschen lernt. (p. 884-885) Z 45

Was Habermas hier sagt, ist insofern banal, weil wir alle wissen, dass der Spracherwerb auch mit dem Lernen von sozialen Normen einhergeht. Aber die pompöse Formulierung soll eine Allerweltsweisheit zur philosophisch reflexierten Tiefenaussage machen. Für den, der so etwas mag, ist das etwas Schönes. Nur, – was soll das?

Kritik: Die Kritik an Habermas' Theorie setzt zunächst an den Grundannahmen, von denen Habermas ausgeht an. Wie viele Philosophen, aber auch einige Neurowissenschaftler, geht er davon aus, dass die Hirnforschung auf deterministische Gesetzmäßigkeiten zielt, weil (neuro-)biologische Prozesse wegen des – angeblich – allgemein in der Natur gültigen Kausalprinzips eben deterministische Prozesse seien. Anscheinend setzt er dabei voraus, dass Gesetze einfache Wenn-dann-Aussagen über das Verhalten sind. Aber was ist, wenn die Gesetzmäßigkeit in der allgemeinen Charakterisierung der Dynamik eines Prozesses besteht und nicht in

stellen; man kann seinen Ansatz *platonistisch* nennen, und Husserl (1900/01) argumentierte, dass der Bedeutungsgehalt einer Aussage ist nicht identisch ist mit dem psychischen Akt, in dem jemand aussagt. Eine ausführliche Diskussion dieser Aussagen ist im hier gegebenen Rahmen nicht möglich.

Aussagen über den Zustand, in dem sich der Prozess zu einem bestimmten Zeitpunkt befindet? Selbst wenn der Prozess deterministisch ist, ist er deswegen noch lange nicht berechenbar, d.h. es können bestenfalls Wahrscheinlichkeitsverteilungen über der Zustandsmenge spezifiziert werden. Habermas unterstellt Vereinfachungen, die überhaupt nicht zutreffen müssen, womit der Bau seiner Philosophie auf unsicherem Fundament steht. In einer Fußnote (vergl. Fußnote 35) geht er zwar auf die Möglichkeit probabilistischer Gesetze ein, argumentiert aber eher kryptisch, dass derartige Gesetze für die Frage nach der Willensfreiheit irrelevant seien. Das kann er allenfalls folgern, wenn seine Basisannahmen tatsächlich gelten, – aber darüber verliert er kein Wort. Habermas hat recht, sollte er meinen, dass man den freien Willen nicht aus möglichen Zufälligkeiten in der Hirntätigkeit ableiten kann, weil dann eben der Zufall entscheidet und nicht die Person. Es ist aber nicht klar, ob er seine Aussage über die Irrelevanz des Zufalls so verstanden haben will. Habermas geht auch nicht auf den Sachverhalt ein, dass viele Gesetze in der Psychologie, allgemein in den Kognitionswissenschaften Ceteris-paribus-Gesetze (c.p.-Gesetze) sind, d.h. sie gelten nur unter bestimmten Randbedingungen. Man kann sagen, dass der Versuch, c.p.-Gesetze zu finden, darauf zielt, die Wirkung einzelner Variablen zu bestimmen. Kontrolliert man die Randbedingungen nicht, so können Ceteris-paribus-Effekte so wie zufällige Effekte wirken, weshalb man z. B. in der empirischen Psychologie selten ohne eine statistische Datenanalyse auskommt (Gadenne (2004), Kap. 6). Wenn die Randbedingungen nicht kontrolliert werden können die jeweils untersuchten Prozesse zwar nach wie vor determiniert ablaufen, sie wirken aber, als ob zufällige Effekte existieren, die sich in der Varianz etwa von Messwerten äußern, weshalb statistische Verfahren zur Evaluation der Daten erforderlich werden. Bei Entscheidungsprozessen äußern sich die variierenden Randbedingungen als mangelnde Prognostizierbarkeit, die jetzt eben nicht mangelnde Kausalität der Prozesse bedeutet. Tatsächlich ist das Postulat, dass Naturgesetze deterministische Aussagen sind, ein rein metaphysisches Postulat: nach van Kampens Theorem kann der Determinismus empirisch weder bewiesen noch widerlegt werden (van Kampen (1991), vergl. Mortensen (2025), p. 44) – d.h. man kann einen allgemeinen Determinismus annehmen, oder auch nicht. Damit verliert Habermas' These, der freie Wille könne gar nicht in einem biologischen System realisiert werden, weil solche Systeme eben deterministisch ablaufen, ihren argumentativen Wert.

Auf die Komplexitäten der Begriffe Kausalität und Determinismus geht Habermas nicht ein. Wichtig ist, dass Habermas, wie viele Philosophen, davon ausgeht, dass biologische Gehirne keinen freien Willen ermöglichen. Er übersieht anscheinend, dass gar nicht klar ist, was überhaupt ein "freier" Wille ist, wenn nicht nur die Abwesenheit von äußerem Zwang gemeint ist ("negative" Definition): gemeint ist die Frage nach der Bedeutung von 'frei'("positive" Definition). Habermas umgeht eine solche Definition bzw. ersetzt sie durch die Aussage, der freie Wille äußere sich in der Interaktion mit dem objektiven Geist; dies ist bestenfalls eine implizite Definition von 'frei':

"Der objektive Geist ist die Dimension der Handlungsfreiheit. Im perfor-

> mativ mitlaufenden Freiheitsbewusstsein spiegelt sich die bewusste Teilhabe an jenem symbolisch strukturierten "Raum der Gründe", worin sich sprachlich sozialisierte Geister gemeinsam bewegen. In dieser Dimension vollzieht sich die rationale Motivation von Überzeugungen und Handlungen nach logischen, sprachlichen und pragmatischen Regeln, die sich nicht auf Naturgesetze reduzieren lassen.", p. 886, Z 46

Der letzte Satz ergibt sich für Habermas aus seinem Hinweis auf Frege (und Husserl) (2004, p. 882):"Seit Frege und Husserl wissen wir, dass sich propositionale Gehalte oder intentionale Gegenstände nicht im Bezugsrahmen kausal wirksamer, raumzeitlich datierbarer Ereignisse und Zustände individuieren lassen". In der Fußnote 74, Seite 59, ist bereits darauf hingewiesen worden, dass Freges Ansichten zwar nicht identisch mit Platons Ansichten sind, aber die Rede ist von einem *logischen Platonismus* Freges: Gedanken existieren zeitlos, sind objektiv und immateriell. Offenbar ist Habermas der Ansicht, dass Freges Ansatz über jeden Zweifel erhoben wahr ist ("Seit Frege und Husserl wissen wir ..."). Aber das stimmt nicht. Eine moderne Gegentheorie ist zum Beispiel der Konnektionismus, demzufolge 'Bedeutung' sich als dynamische Entwicklung insbesondere neuronaler Systeme ergibt. Habermas liefert keinerlei Argument, warum etwa konnektionistische Ansätze *nicht* gelten können, er behauptet nur, dass derlei Ansätze falsch sein müssen, weil die Fregesche Position wahr sei[75].

Willensfreiheit ergibt sich nach Habermas aus der Fähigkeit des Menschen, im Raum der Gründe zu navigieren und moralische Verantwortung zu übernehmen, und nach pragmatischen Regeln zu entscheiden, die sich nicht auf Naturgesetze reduzieren lassen. Das impliziert für Habermas, dass Entscheidungsprozesse sich nicht notwendig auf Hirnprozesse zurückführen lassen. Was Habermas bietet ist eine sprachgewaltige Beschreibung menschlichen Verhaltens, das an biologischen Bedürfnissen und an der Interaktion mit gesellschaftlichen Bedingungen und Normen orientiert ist. Übersetzt in umgangssprachliches Deutsch ergäbe sich die Frage, warum Habermas' Text ein philosophischer Text ist, – in "normaler" Sprache sagt der Text nichts aus, was nicht Alltagswissen bzw. übliche Alltags*annahme* ist, und die Frage nach dem freien Willen bliebe offen. Diese Frage wird nicht beantwortet, wenn man die Umgangsprache in Philosophiesprech übersetzt. Was genau soll es bedeuten, wenn Habermas schreibt "Der objektive Geist ist die Dimension der Handlungsfreiheit"? Es ist, als ob semantische Dunkelheit philosophischen Tiefsinn erzeugen soll. Es ist ja nicht so, dass man keine Interpretation dieses Satzes finden kann, das Problem ist eher, dass man nicht weiß, ob eine Interpretation dem entspricht, was Habermas gemeint hat. Ist eine Entscheidung "frei", wenn sie sich an den Elementen des objektiven Geistes orientiert, etwa an dem, was in Büchern gefunden wurde, diese Elemente akzep-

[75](vergl.Habermas (2004, p. 882): "auf der grundbegrifflichen Ebene scheitern diese Versuche [d.h. etwa der Konnektionismus] einer Naturalisierung des Geistes an der erforderlichen Übersetzung. Die von diesen Theorien vorgenommenen Übersetzungen zehren entweder selber unausdrücklich vom Sinn der mentalistischen Ausdrücke, die sie ersetzen sollen; oder sie verfehlen wesentliche Aspekte des Ausgangsphänomens und gelangen so zu unbrauchbaren Umdefinitionen."

tierend oder ablehnend? Es ist keineswegs klar, warum der objektive Geist die Freiheit einer Entscheidung konstituiert. Es drängt sich der Verdacht auf, dass der Satz zum Habermasschen Hexeneinmaleins gehört, als das Karl R. Popper die Sprache Habermas' bezeichnet hat. Wenn sich eine Person im Restaurant für ein Gericht entscheiden muß, so wird ihr im Allgemeinen die freie Wahl überlassen, es hängt nur von ihr ab, wie die Entscheidung ausfällt, – und wie kommt der objektive Geist *als Dimension* der Willensfreiheit in diesen Entscheidungsprozess? Vielleicht spielen ethische Betrachtungen eine Rolle, die Person ist vielleicht Vegetarierin, so dass Fleischgerichte von vorn herein nicht in Betracht kommen, – ist diese ethische Einstellung ein Teil der Dimension des objektiven Geistes, der sie freiwillig auf Fleisch verzichten lässt? Habermas würde vielleicht auf die individuelle 'verantwortliche Urheberschaft' verweisen, die die Vegetarierin auf ein Fleischgericht verzichten läßt. Man kann sagen, dass ihre Entscheidung zumindest partiell determiniert ist (partiell, weil sie immer noch die Wahl zwischen veschiedenen vegetarischen Gerichten haben könnte), weil sie sich ihren Normen entsprechend nicht für ein Fleischgericht entscheiden kann. Tut sie es doch, ist sie ein determiniertes Opfer ihrer Lust auf Fleisch. Man kann argumentieren, dass die Norm einen Teil ihres Selbst definiert, mithin sei es ihre Entscheidung, die 'frei' ist, weil es *ihre* Entscheidung ist. Für welche dieser beiden Interpretationen man sich entscheidet, hängt anscheinend davon ab, wie man "frei" definiert, und es ist offen, in welcher Weise die Ausführungen des Zitats 46, Seite 61, hier weiterhelfen können. Habermas' Sprachkunst vernebelt eher als dass sie klärt.

So ist auch der unscharfe Begriff des "weichen" Naturalismus, ebenso wie der des Perspektivendualismus, nicht hilfreich. Sein nicht-szientisticher Ansatz könnte auch ein Ausdruck von Szientophobie sein. Mit 'weichem Naturalismus' scheint gemeint zu sein, dass die neurobiologischen Prozesse des Gehirns einen freien Willen nicht zulassen, weil sie kausal ablaufen und damit determiniert seien. Also findet ein Teil der Entscheidung irgendwo außerhalb des Gehirns statt: die Dimension der Handlungsfreiheit ist der objektive Geist. Was den aktuellen Ort der Aktivität, die das freie Entscheiden ist, unbestimmt läßt. Für Habermas scheint sprachliche Interaktion einen irgendwie außerhirnlichen Charakter zu haben, – aber warum das so sein soll, bleibt eine völlig offene Frage. Andererseits können wir sagen, dass alle sozialen Interaktionen nicht so ablaufen würden, wie sie ablaufen, hätte das individuelle Gehirn irgendwelche Defekte. Die abstrakten und intentionalen Begriffe wie Freiheit, Gleichheit und Brüderlichkeit würden für mich keinen Sinn machen, wäre etwas mit meinem Wernicke-Sprachzentrum nicht in Ordnung, auch wenn diese drei Begriffe noch so fester Bestandteil des mich umgebenden absoluten Geistes wären. Habermas erklärt an keiner Stelle, warum der kognitive Umgang mit Gründen nicht an eine dazu korrespondierende Hirntätigkeit gekoppelt ist, außer dass "sich Gründe nicht im Gehirn finden lassen" (zitiert nach Roth (2005), p. 692). Der philosophisch durchaus gebildete Neurowissenschaftler und Hirnforscher Gehrhard Roth hat die Habermassche Philosophie einer kritischen Betrachtung unterzogen (Roth (2005)), worauf kurz eingegangen werden soll. Seine Argumente sind auch relevant für die in den fol-

genden Abschnitten vorgestellten Philosophen des Geistes.

Roths Kritik: Roth beginnt seine Argumentation (in Abschnitt II) mit der Bemerkung, dass wir "beim Selbsterleben mentaler Prozesse keine neuronalen Prozesse und beim Studium neuronaler Prozesse keine mentalen Prozesse erfahren" (p. 693). Anhänger des ontologischen Dualismus gehen davon aus, dass alles Seiende in eine von zwei Kategorien gehört, entweder in eine physische oder in eine geistige. Roth verweist aber auf die Befunde, dass zwischen mentalen Zuständen und neuronalen Prozessen eine enge Beziehung besteht. Neurowissenschaftler gingen, so Roth, von verschiedenen Begriffen für "Hirnprozess" aus: einmal meinen sie damit messbare Erregungsausbreitung im Gehirn, oder Prozesse an einem einzelnen Neuron bzw. an einer einzelnen Synapse. Prozesse der ersten Art lassen lassen sich mit der Elektroenzephalografie (EEG), der Magnetenzephalograhie (MEG, hohe zeitliche Auflösung), der Positronen-Emissions-Tomographie (PET) sowie der funktionellen Kernspintomographie (fMRI, hohe räumliche Auflösung) untersuchen. Damit läßt sich die Bedeutung der Aktivitäten noch nicht bestimmen, – der neuronale Code ist neutral (Roth (2000)). Andererseits erzeugen die Aktivitäten Verhlten. Um die Bedeutung der neuronalen Aktivitätsmuster zu erfassen muß man das korrespondierende Verhalten mit erfassen. Präsentiert man Versuchspersonen (Vpn) bedrohliche Gesichter, so führt das zu bestimmten Aktivierungen der Amygdala. Fallen am Hippocampus Neuronen aus, so resultiert ein anterograde oder retrograde Amnesie. Neurobiologen oder Psychologen meinen dann mit "Hirnprozessen" die Kombination von neurophysiologischen und neurochemischen Prozessen einerseits und Verhaltensweisen oder mentalen Aktivitäten. Einige Philosophen verweisen dann auf die Möglichkeit, dass der Zusammenhang z.B. von fMRI-Bildern und Verhalten nur Korrelationen seien, und Korrelationen bedeuten bekanntlich noch keine Kausalität. Roth weist diese Kritik zurück: man wisse mittlerweile, warum z.B. die Amygdala etwas mit angeborener oder erlernter Furcht zu tun habe. Es gibt also Hintergrundinformationen, die auf kausale Beziehungen schließen lassen. Es gibt sogar Untersuchungen, die den Zusammenhang zwischen Hirnprozessen und bewusstem Erleben charakterisieren. Noesselt et al (2002) kombinierten die MEG- und fMRI-Verfahren, um die bewußte visuelle Gestaltwahrnehmung zu untersuchen: von der Retina über den lateralen Kniehöcker zum primären visuellen Kortex weiter zu den sekundären assoziativen visuellen Arealen zurück zu den primären visuellen Arealen. Man kann auch die zeitliche Reihenfolge der Aktivitäten verschiedener Hirnzentren bestimmen, so dass man herausfinden kann, was geschehen muß, damit bestimmte Verhaltensweisen bzw. Bewußtseinszustände eintreten (Roth (2005), (p. 696). Dabei findet man, dass u.a. auch soziale Erfahrungen zu Veränderungen im Gehirn führen; auf diese Weise läßt sich untersuchen, wie Bedeutungen im Gehirn entstehen, was der Behauptung von Philosophen widerspricht, neuronale Netzwerke hätten nichts mit Bedeutung zu tun.

Auch intentionale Zustände lassen sich neuronal erklären. Roth (p. 695) beschreibt einen Versuchsaufbau. Eine Vp liegt in einem Kernspintomographen. Eine Fingerkuppe wird ein leichter Schmerz zugefügt. In ihrem sensorischen Kor-

tex und im insulären, anterioren cingulären Kortex entstehen Aktivitäten genau dann, wenn die Vp eine Schmerzerfahrung berichtet bzw auf eine Knopf drückt, so dass diese Erfahrung registriert wird. Es zeigt sich, dass man von der Aktivierung im sensorischen Kortex auf eine Schmerzerfahrung schließen kann, und umgekehrt von der Schmerzerfahrung auf die Aktivierung im sensorischen Kortex. Manchmal berichten Vpn, dass zwar ein Schmerzreiz verabreicht wurde, sie aber *keine* Schmerzerfahrung haben. Es zeigt sich, dass zwar im sensorischen Kortex ein sensorischer Reiz wahrgenommen wird, und zu ihrer Aussage "Es tut nicht weh" korrespondiert der Sachverhalt, dass im anterioren cingulären Kortex *keine* Aktivität beobachtet wird. Das heißt, dass die Vp zwischen dem Stimulus und dem Schmerzerlebnis unterscheiden kann. Daraus folgt, "dass die limbischen Kortexanteile das "Schmerzhafte" am Schmerz erzeugen" (Roth, p. 697), – diese Kortexanteile erzeugen das Erleben von Qualia, die in der Philosophie des Geistes als Paradebeispiele für angeblich neuronal nicht erklärbare mentale Phänomene betrachtet werden. Roth schreibt

> "Wir sind nur deshalb zu intentionalen Zuständen und zur Gesellschaftlichkeit befähigt, weil unser Gehirn Netzwerke besitzt, die sich unter dem Einfluss sozialer Erfahrung, einschließlich sprachlicher Kommunikation, verändern und ihrerseits direkt oder indirekt auf das Handeln einwirken". (Roth (2005), p. 698), Z 47

Reduktion und Epiphänomene: Roth (2005) geht auch auf die Behauptung Habermas (2004, p. 872) ein, der Reduktionismus impliziere, Bewußtsein und bewußtes Handeln (freie Entscheidungen sind bei Habermas stets bewußte Entscheidungen) seien Epiphänomene. Das sind Phänomene, die kausal erzeugt werden, selbst aber nicht kausal wirken können. Die diesem Argument unterliegende Vorstellung ist, dass neurobiologische und mentale Prozesse grundsätzlich separate Prozesse sind. Es ist aber überhaupt nicht klar, warum das so sein soll: geht man von einer systemtheoretischen Vorstellung aus, was immer möglich ist – das Gehirn ist ein Paradebeispiel für ein dynamisches System[76] – , so wird man zwar sagen können, dass uns bestimmte Prozesse im Gehirn, nämlich diejenigen, die tatsächlich Bewußtsein erzeugen, noch nicht bekannt sind, dass man aber davon ausgehen kann, dass sie von bestimmten neuronalen Aktivitäten "angetrieben" werden, was durch neurophysiologische Untersuchungen, die mit simultan ablaufenden Verhaltensbeobachtungen gekoppelt sind, nahegelegt wird. Man hat es dann insgesamt mit *einem* System zu tun, so dass bewußtes Erleben und Handeln durch bestimmte, dazu korrespondierenden Systemzustände definiert sind. Wie Roth (2005), p. 702, anmerkt, läßt dieses Modell auch die Eigengesetzlichkeit des Mentalen zu, ohne dass damit ein Epiphänomenalismus impliziert würde, da Zustände des Systems auf andere Zustände rückwirken können. Diese Formulierung ist gewählt worden, weil der Epiphänomenalismus die Möglichkeit einer

[76] Gemeint sind hier Systeme, die sich durch Syteme von Differentialgleichungen beschreiben lassen, wobei gegenwärtig kaum jemand auf die Idee kommen wird, ein derariges System von Gleichungen für das gesamte Gehirn aufstellen zu wollen.

Rückwirkung vom Mentalen auf das Neuronale negiert, sie ist allerdings auch irreführend.

Habermas ist, wie andere Philosophen auch, davon überzeugt, dass Handeln nicht ein Resultat von Ursachen, sondern von Gründen sei. Wirft man einen Stein gegen eine Fensterscheibe, so zerspringt sie, wenn bestimmte physikalische Bedingungen erfüllt sind (Aufprallwinkel, Masse und Geschwindigkeit des Steins, Dicke des Glases etc müssen Werte in bestimmten Bereichen haben); diese Bedingunen sind die Ursachen des Zerspringens. Sie *erklären* das Zerspringen. Allerdings kann man auch fragen, warum der Stein geworfen wurde, womit man bei den Gründen für den Wurf angelangt ist. Vielleicht hat ein pubertierender 14-Jähriger einem diffusen Unmut über den Zustand der Welt Ausdruck verleihen wollem, eine solche Hypothese würde die Handlung des Jugendlichen *verstehbar* machen. Die Polarität von Ursachen und Gründen korrespondiert dann zur Polarität von Erklären und Verstehen. Roth betrachtet dazu ein Drei-Ebenen-Modell des Gehirns. Die *unterste* Ebene ist die der elementaren Reaktionen zur Lebenserhaltung und Bedürfnisbefriedigung, zu denen Roth auch die affektiven und emotionalen Zustände zählt. Roth nennt sie den *unbewussten Kern* der Persönlichkeit, sie werden im Gehirn durch die subkortikalen Anteile des limbischen Systems und Teile der Basalganglien repräsentiert. Die *mittlere* Ebene liefert die bewussten, ego-zentrierten Emotionen, Vorlieben, Überzeugungen, private Handlungsantriebe (Roth). Sie entwickelt sich aufgrund von emotional-kognitiven Erfahrungen in der Kindheit im jugendlichen Alter und bestimmt unsere Persönlichkeit. Im Gehirn wird diese Ebene durch kortikale limbische Anteile und den insulären, entorhinalen und anterioren cingulären Kortex etc, und damit mit Teilen der Großhirnrinde und der okzipitael, parietalen und temporalen Kortex repräsentiert. Als *oberste* Ebene schließlich nennt Roth die Ebene der sprachlich-kommunikativen sozialen Existenz, die durch den präfrontalen Kortex wobei die Sprachzentren (Broca-Wernicke-Areal) repräsentiert wird. Dies ist die Ebene, auf die Habermas sich bezieht. Wie Roth ausführt ist allen drei Ebenen gemeinsam, dass sie in neuro-motorische Programme umgesetzt werden müssen, um handlungsrelevant werden zu können. Bestimmte Aktivitäten des Gehirns werden zu Ursachen von Handlungsgründen. Das bedeutet, dass die Unterscheidung von Ursachen und Gründen entgegen der Habermasschen Vorstellung das Gehirn *nicht* von seiner sozialen Umwelt trennt: "Gründe – welcher Art auch immer – müssen, um handlungswirksam zu werden, neuromotorische Ursachen bewirken." (Roth, p. 710).

Roth geht noch auf weitere Aspekte der Organisation des Gehirns ein ("die alle Vorstellungen übersteigende Binnenverdrahtung des Kortex", "Der Kortex ist tatsächlich im Wesentlichen mit sich selbst beschäftigt" (p. 703), auf die hier nicht im Detail eingegangen werden kann; es geht hier nur darum, auf die Eigenschaften des Gehirns zu verweisen, die die Annahme von Dualismen gleich welcher Art unnötig machen.

1.2.5.3 Julian Nida-Rümelin Der Philosoph Julian Nida-Rümelin (2007)[77] geht anscheinend davon aus, dass der Wille notwendig frei ist. Alle folgenden Nida-Rümelin -Zitate beziehen sich auf diese Schrift. Nida-Rümelin stellt wie Dretske die Frage nach dem Verhältnis zwischen Freiheit und Kausalität (p. 2), d.h. wie sich die seiner Ansicht nach notwendig freie Entscheidung mit der "umfassenden kausalen Erklärbarkeit aller Ereignisse" in Einklang bringen läßt. Dazu erklärt er, dass "... nach meinem Verständnis Gründe eine zentrale Rolle für unseren Freiheitsbegriff spielen." und Entscheidungen finden nicht im Kopf, also nicht im Gehirn statt[78]. Konstitutiv für Nida Rümelins Ansatz ist sein Begriff des *Commonsense-Dualismus*, der durch den "Gegensatz zwischen dem von der Natur Gegebenen und menschlicher Setzung" chrakterisiert wird; diese Setzung – gemeint sind wohl soziale Normen und Regeln – bezieht sich auf die Kultur, soll aber nicht mit Kultur gleichgesetzt werden. Der Commonsense-Dualismus führe zu einer *humanistischen Perspektive*, die eine spezielle Definition von Freiheit erfordere: Nida-Rümelin will, wie er schreibt, erklären, "was Freiheit eigentlich ist" (p. 3): die "eigentliche Definition" der Freiheit sei, so Nida-Rümelin , die Annahme eines Spektrums,

> "dessen eines Ende durch die "Willkürfreiheit" definiert ist, und dessen anderes Ende "jene Form von Freiheit [ist], die zurückgeht auf grundlegende, langfristig wirksame und möglicherweise schon lange vor der Ausführung einer konkreten Handlung getroffene Festlegungen einer Person." (p. 15), Z 48

Diese Definition ist nicht mit der eingangs gegebenen Standarddefinition des Begriffs 'freier Wille' als Ausdruck für Entscheidungen ohne innere und äußere Zwänge kompatibel; hier geht Freiheit zurück auf "langfristig wirksame, möglicherweise schon lange vor der Ausführung einer konkreten Handlung getroffene Festlegungen einer Person"; bei Nida-Rümelin wird eine Theorie – die der humanistischen Perspektive – nicht in Abhängigkeit von der Definition von Willensfreiheit entwickelt, sondern umgekehrt die Definition nach Maßgabe der Theorie formuliert. Willkürfreiheit bestehe jedenfalls darin, sich im Falle völliger Indifferenz in Bezug auf die Entscheidungsalternativen eben rein willkürlich für die eine oder andere Alternative zu entscheiden. Nida-Rümelin erklärt nicht, was er mit 'willkürlich' meint: - 'zufällig', oder deterministisch, aber beliebig, etc. Das andere Ende der Skala repräsentiert die Form von Freiheit, die Entscheidungen auf lange Deliberationen, Grundeinstellungen etc zurückführt. Diese Entscheidungen seien "lebensstrukturierende oder existentielle Entscheidungen" (p. 15). Dass einige unserer Entscheidungen lebensstrukturierend oder existentiell sein

[77] Ausarbeitung einer Akademievorlesung am 12. April 2007 an der Berlin- Brandenburgischen Akademie der Wissenschaften

[78] In einem im Frankfurter Rundschaumagazin publizierten Streitgespräch am 03. 04. 2004 fragt der Philosoph Nida-Rümelin den Hirnforscher Wolf Singer, ob Singers Frau nicht mit ihm, Singer, schimpfe, wenn er ihr seine "berüchtigten Forschungsergebnisse präsentiere", denen zufolge der freie Wille eine Illusion sei, und ob es ihm, Singer, denn lieber sei, "wenn Ihr Gehirn Ihnen sagt, was Sie zu tun haben?".

können, wird niemand bestreiten, bestreiten kann man allerdings, dass Entscheidungen durch Deliberationen zu freien Entscheidungen werden: durch Deliberationen werden möglicherweise Entscheidungen für bestimmte Alternativen ausgeschlossen, andere kommen in eine engere Wahl, in die sie nicht gekommen wären, wären frühere Entscheidungen und Erfahrungen andere gewesen. Man könnte sagen, dass Nida-Rümelins Freiheit eigentlich gar keine ist, sondern dass hier eher von Unfreiheit die Rede ist, weil (i) die Deliberationen die Entscheidungen festlegen, und (ii) über den Prozess der Deliberation selbst nichts weiter ausgesagt wird. Es wird anscheinend angenommen, dass dieser Prozess "frei" abläuft, also wiederum nach dem Willen der deliberierenden Person, aber dafür liefert Nida-Rümelin keinerlei Argumente. Wie bei Habermas wird der Deliberierungsprozess selbst gar nicht analysiert. Kritiker der These, der Wille sei frei, verweisen nicht zu Unrecht auf den Sachverhalt, dass die Details der Deliberationen der entscheidenden Person oft gar nicht bewußt sind und auf diese Weise Zwänge entstehen, die die Entscheidung am Ende unfrei sein läßt (z. B. Harris (2012), Strawson (1998)). Nida-Rümelin liefert keine Analyse der Möglichkeit eines freien Willens, sondern einfach eine Umdeutung des Begriffs des freien Willens. Im ersten Ansatz des einleitenden Abschnitts 1.1 wurde davon ausgegangen, dass das Prädikat 'frei' in Bezug auf Entscheidungen die Abwesenheit von Zwängen meint, und bisher ist Martin Heisenberg der einzige, der diesen Aspekt des Freiheitsbegriffs ernst genommen hat, vergl. Abschnitt 1.1.2. Ob sein Ansatz überzeugend ist, ist eine andere Frage. Der nächste offene Punkt ist, woher sich Nida-Rümelin die Freiheit nimmt, die Definition Z 48 zur "eigentlichen Definition" der Freiheit zu erheben, denn sie ist eigentlich völlig willkürlich. Nida-Rümelin liefert kein Argument, aus dem hervorgeht, dass man sein Konzept der 'eigentlichen Freiheit' anderen Konzeptionen vorzuziehen sei. Wie Dretske und Habermas ist Nida-Rümelin der Ansicht, dass kognitive Aktivitäten wie das Deliberieren und Konzipieren von Gründen gar nicht im Gehirn stattfinden (vergl. Fußnote 78, Seite 67). Dann aber ergibt sich die unbeantwortet bleibende Frage, wo denn dann die zu Entscheidungen führenden kognitiven Prozesse ablaufen. Insbesondere schreibt Nida-Rümelin , dass ein neurobiologischer Ansatz wegen der "naturalistischen Unterbestimmtheit von Gründen" inadäquat sei, weil:

> "Gründe – beziehungsweise das Abwägen und Akzeptieren der letztlich besten Gründe – tatsächlich relevant sind, für das was ich tue. Wenn diese These gilt, so kann man Handlungen ohne Bezugnahme auf Gründe nicht erklären und erst recht nicht prognostizieren" (p. 17), Z 49

Diese These sei, so Nida-Rümelin , "mit einer bestimmmten Form von Determinismus, nämlich dem in seiner naturalistischen Form, unvereinbar"; es wäre für die Diskussion nützlich, wenn Nida-Rümelin spezifizieren würde, was genau er mit "Determinismus in seiner naturalistischen Form" meint. Jedenfalls müsse man anerkennen, "dass Gründe für das menschliche Leben, die menschliche Verständigungspraxis, die menschliche Interaktion etc. eine irreduzible Rolle spielen."

Liest man die Aussagen Nida-Rümelins als umgangssprachliche Aussagen, so wird man wenig einzuwenden haben: wollen wir eine Erklärung für ein gegebenes

Verhalten haben, so suchen wir nach Gründen dafür, und wenn wir keine Gründe für dieses Verhalten finden können (Nida-Rümelin: "so kann man Handlungen ohne Bezugnahme auf Gründe nicht erklären"), so werden wir – trivialerweise – nicht erwarten, dass wir dieses Verhalten hätten vorhersagen können.

Nida-Rümelin will aber mehr als nur umgangsspachliche Betrachtungen anstellen, er möchte begründen, dass wir über eine Art von freiem Willen verfügen. Dazu wird der Begriff des Grundes nicht mehr nur umgangssprachlich verwendet, wobei die Begriffe 'Grund' und 'Ursache' praktisch denselben Umfang haben, sondern 'Grund' und 'Ursache' werden unterschiedlich definiert. 'Gründe' sind mentale Einheiten, 'Ursachen' sind neuronale Prozesse ("naturalistische Einheiten"). Die Steuerung des Verhaltens durch Gründe wird im Rahmen der Theorie Nida-Rümelins zu einer notwendigen, wenn auch nicht hinreichenden Voraussetzung für Prognosen. Eine Möglichkeit, die Aussagen Nida-Rümelins auf den Punkt zu bringen, ist diese:
Nach Zitat 49 sind Gründe für Handlungen relevant, und es soll gelten:
1. Eine Entscheidung geht entweder auf Gründe oder auf Ursachen zurück,
2. Wird eine Handlung (z.B. eine Entscheidung) ohne Bezug auf Gründe vorgenommen, so ist sie nicht erklärbar und nicht prognostizierbar.
3. Folgerung: gilt "Eine Handlung wird ohne Bezug auf Gründe vorgenommen", so folgt einerseits (s. 1.), dass sie das Resultat einer Ursache ist, andererseits nach dem vorangehenden Satz 2., dass sie nicht prognostizierbar ist. Aber Ursachen gehören nach Nida-Rümelin zur naturalistischen Perspektive[79], und damit zum Regime der angeblich deterministischen Naturgesetzlichkeit, weshalb die Handlung nun prognostizierbar wird[80]. Dass 'deterministisch' nicht 'prognotizierbar' impliziert scheint Nida-Rümelin entgangen zu sein. Die Struktur der von Nida-Rümelin spezifizierten Begriffe erweist sich demnach als opaque[81].

Nida-Rümelin wird derartigen Betrachtungen vermutlich nicht zustimmen wollen und sie als unphilosophische Pedanterien zurückweisen. Der Punkt ist, dass die Vagheit der Begriffe 'Grund' und 'Ursache', die doch die Basis für Nida-Rümelins philosophischen Ansatz bilden, in der Tat verstörenden Inkonsistenzen impliziert, und das wiederum nicht ohne Grund: die Unterscheidung von "naturalistischen" und mentalen, "nicht-naturalistischen" Systemen ist angesichts der in Abschnitt 1.1.2 noch einmal angedeuteten problematischen Natur des Kausalitäts- und Determinismusbegriffs kontraproduktiv, – Determinismus bedeutet nicht notwendig auch Prognostizierbarkeit und ist darüber hinaus selbst nicht begründbar (s. van Kampens Theorem). Auch ist unklar, warum der Begriff der Erklärung an den des Grundes gekoppelt wird, denn Erklärungen können doch auch durch Verweis auf Ursachen zustandekommen. Darüber hinaus sagt Nida-Rümelin nichts über

[79] Perspektiven: "... die humanistische, die auf Gründe, sowie die naturalistische, die auf Realursachen fokussiert ist". Nida-Rümelin (2007), p. 3 unten.

[80] im Rahmen seiner Kritik des kantschen Ansatzes schreibt er: "Zum einen besteht die Möglichkeit einer naturwissenschaftlichen, kausalen Analyse, die den von außen objektiv beobachtbaren Verhaltensprozess einer Person ohne Rest prognostizieren kann", p. 6, d.h. 'kausal' impliziert 'prognostizierbar'.

[81] 'opaque' ist hier eine eher behutsame Umschreibung für 'widersprüchlich'.

den Deliberationsprozess: ist der Wille in welchem Sinne auch immer noch frei, wenn der dieser Prozess bis auf stochastische Effekte deterministisch verläuft? Zufällige Entscheidungen sind eigentlich keine freien Entscheidungen. Wenn sich der Prozess als eine Folge von Abwägungen von Kosten- und Nutzenbewertungen darstellen läßt, so ist er bis auf stochastische Effekte bei den Bewertungen deterministisch[82]. Was ist, wenn eine Person sich auf der Basis eine Deliberationsprozesses nicht für eine der zur Verfügung stehenden Alternativen entscheiden kann und sich deshalb für eine Entscheidung anhand eines Münz- oder Würfelwurfs bzw. allgemein eines Zufallsgenerators entscheidet? Entscheidet man sich frei für diese Art der Entscheidungsfindung? Eigentlich nicht, denn der Deliberationsprozess kann anscheinend nicht zur Wahl einer bestimmten Alternative führen. Wenn es die Möglichkeit gibt, sich nicht zu entscheiden, kann man sich dazu entscheiden, sich nicht zu entscheiden, *oder* zum Zufallsgenerator zu greifen, wenn man sich aber entscheiden muß, so muß man dies über den Zufallsgenerator tun. Frei ist die Entscheidung offenbar nur dann, wenn die Alternative, sie nicht zu entscheiden existiert. Die Entscheidung für diese Alternative kann wieder durch einen Zufallsgenerator gefällt werden, oder es gibt Gründe, die diese Entscheidung bewirken, und dann sind es diese Gründe, – was bedeutet dann noch, dass die Entscheidung "frei" ist? Ob der Begriff der humanistischen Perspektive hier noch eine klärende Funktion hat ist nicht offensichtlich.

Zur Illustration sei an die Mitte des 19-ten Jahrhunderts durch den Physiologen E.H. Weber (1795–1878) und den Physiker und Philsophen G. T. Fechner (1801 – 1887) entwickelte Psychophysik erinnert, bei der es zunächst um die Beziehung zwischen Stimulusintensität und dazu korrespondierender Empfindungsstärke ging. Weber bestimmte für eine jeweils gegebene Stimulusintensität I einen Zuwachs ΔI[83] der Intensität derart, dass die Intensitäten I und $I + \Delta I$ "gerade eben" unterschieden werden könnten. Dabei fand er, dass es bei keiner Sinnesmodalität und gegebener Intensität I ein scharf definierts Inkrement ΔI_0 gibt derart, dass für $\Delta I < \Delta_0$ die Intensitäten I und $I + \Delta I$ nicht unterschieden werden können, für $\Delta I \geq \Delta I_0$ aber I und $I + \Delta I$ stets voneinander unterschieden werden können. Vielmehr ist es so, dass für "hinreichend kleine" ΔI der Unterschied zwischen I und $I + \Delta I$ mit einer von I und ΔI abhängenden Wahrscheinlichkeit $p(I, \Delta I)$ entdeckt wird, wobei $p(I, \Delta I)$ monoton mit ΔI wächst, d.h. gegen 1 strebt. Die Entdeckung eines Unterschieds zwischen I und ΔI ist also kein deterministischer, sondern ein probabilistischer Prozess. Für einen gegebenen Versuchsdurchgang (der Versuchsperson (Vp) wird ein Stimulus s_1 mit der Intensität I und ein zweiter Stimulus s_2 mit der Intensität $I + \Delta I$ gezeigt,

[82] Die Statistik hält für Entscheidungen unter Unsicherheit, bei denen die "Outcomes" nur mit Wahrscheinlichkeiten kleiner als 1 eintreten eine Reihe von Größen wie relatives Risiko, Odds Ratio etc bereit, anhand derer Entscheidungen in deterministischer Manier getroffen werden können. Diese müssen einer konkreten Person gar nicht bekannt sein, sie hat möglicherweise ganz ideosynkratische Enscheidungregeln, die auf einen derministischen Entscheidungsprozess hinaus laufen.

[83] Δ ist der griechische Buchstabe Delta, er wird oft zur Bezeichnung eines Unterschiedes (Differenz) verwendet.

wobei entweder $\Delta I = 0$ oder $\Delta I > 0$, aber hinreichend klein ist), und die Vp muß anzeigen, ob $s_1 = s_2$ ($\Delta I = 0$) oder $s_1 \neq s_2$ ($\Delta I > 0$) ist. Dabei kann nicht mit Gewissheit vorausgesagt werden, welche Antwort die Vp geben wird, obwohl wir es doch mit "Realursachen", d.h mit neurobiologischen Zuständen zu tun haben. Der Grund dafür ist die grundsätzliche Stochastizität der neuronalen Prozesse, mit der Stimuli mit den Intensitäten I und $I + \Delta I$ verarbeitet werden. Diese Stochastizität gilt allgemein für das Gehirn und führt auch bei Kognitionen "höherer Ordnung" zu Zufälligkeiten, die u.a. zu Aufmerksamkeitsschwankungen, zu Denkfehlern etc. führen können. In der empirischen Kognitionsforschung werden derartige Prozesse durchaus untersucht, bei den philosophischen Betrachtungen spielen sie so gut wie keine Rolle.

Die beschriebene Unsicherheit gilt dementsprechend auch für komplexere Entscheidungen, wenn die Folgen einer Entscheidung nicht mit Sicherheit überblickt werden können: Prognosen gelten immer nur mit bestimmten Wahrscheinlichkeiten, man denke an die Entscheidungen von Experten über das zu erwartende Wirtschaftswachstum des kommenden Jahres, oder die Entscheidung, eine bestimmte Investition zu tätigen oder nicht: ob die *richtige* Entscheidung getroffen wird ist i. A. nicht mit Sicherheit prognostizierbar. Dann würde nach Nida-Rümelin folgen, dass derlei Entscheidungen nicht in Bezug auf Gründe zustandekommen und demnach unfrei sind. Dann fragt sich aber, ob man überhaupt von freien Entscheidungen sprechen kann, sofern sie von komplexerer Natur sind, und die ganze Argumentation von Nida-Rümelin kollabiert. Wenn Nida-Rümelin schreibt, man müsse anerkennen, "dass Gründe für das menschliche Leben, die menschliche Verständigungspraxis, die menschliche Interaktion etc. eine irreduzible Rolle spielen," so kann man ihm zustimmen, so lange man nicht nach der genauen Bedeutung von 'irreduzibel' fragt, und sofern man sich an die umgangssprachliche Interpretation des Ausdrucks 'Grund' hält, womit man aber den Boden der Philosophie Nida-Rümelins verläßt. Es stellt sich die Frage, welchen Sinn nun seine Aussage über die "naturalistische Unterbestimmtheit von Gründen" hat. Die bloße Behauptung dieser Unterbestimmtheit ist ja kein Beweis für die Ungültigkeit der Gegenthese, dass Gründe ebenfalls naturalistisch bestimmt sind, zumal Nida-Rümelins Begriff der Unterbestimmtheit im hier gegebenen Zusammenhang selbst unterbestimmt ist. Die Gegenthese hat wiederum den Vorteil, ohne das Postulat, Entscheidungen fänden nicht im Gehirn statt, auszukommen. Warum ist es humanistisch, die biologische Basis unserer Kognitionen zu bestreiten? Nida-Rümelin suggeriert, weil wir uns mit der Akzeptanz einer biologischen Basis den "Determinismus in seiner naturalistischen Form" einhandeln, der freie Entscheidungen unmöglich mache. Aber wie die Diskussionen über den Determinismus gezeigt haben, ist dieser keine notwendige Implikation des naturalistischen Ansatzes, der nicht weniger humanistisch ist als Nida-Rümelins . Wer sich über die inhumanen Implikationen des humanistischen Bekenntnisses zum freien Willen informieren möchte, werfe einen Blick in das Kapitel 14 (The joy of punishment) von Sapolsky (2023).

1.2.5.4 Alva Noë Der amerikanische Philosoph Alva Noë (1964 –) publizierte 2009/2010 ein Buch mit dem Titel *Du bist nicht Dein Gehirn – Eine radikale Philosophie des Bewußtseins*. Die Seitenzahlen für die folgenden Zitate beziehen sich auf die deutsche Ausgabe. Noë geht nicht explizit auf die Frage nach dem freien Willen ein, sein zentrales Thema ist die Frage, ob die Aktiviäten des Gehirns die Basis des Bewußtseins bilden. Sollten sie es sein, würde sich die Frage nach dem freien Willen stellen, weil dann ja materielle Abläufe das Bewußtsein bestimmen und der Wille dann möglicherweise nicht frei ist. Ist das Bewußtsein aber unabhängig vom Gehirn, so kann man annehmen (Noë), dass das Bewußtsein und damit bewußtes Handeln unabhängig von den kausalen Abhängigkeiten sind, die materielle Prozesse bestimmen, und damit kann dann auch der Wille frei sein. Betrachten wir also Noës Argumentationen.

Noë hat herausgefunden, wie wir die Antwort auf die Frage nach dem Bewußtsein finden können.

> Um das Bewusstsein besser zu verstehen, müssen wir Abstand nehmen von dem nach innen gerichteten, neuronalen Mikrofokus ... Das Bewusstsein ist in den dynamischen Lebensprozess des gesamten, in eine Umwelt eingebetteten Lebewesens angesiedelt. Und nur wenn wir diese ganzheitliche Sicht auf die Lebensaktivitäten eines Menschen oder Tieres einnehmen, können wir allmählich verstehen, was das Gehirn zu unserem bewussten Erleben beiträgt.
>
> Dies ist ein positives Buch. die menschliche Erfahrung ist ein Tanz, der sich in der Welt entfaltet. Wir sind nicht unser Gehirn. Wir sind nicht im Gefängnis unserer Vorstellungen und Empfindungen gefangen. Das Phänomen des Bewusstseins ist ebenso wie das des Lebens ein mit der Welt verwobener, dynamischer Prozess. Wir sind zuhause in der Welt, die uns umgibt. Wir stecken nicht in unserem Kopf." (p .13), Z 50

Auf Seite 19 schreibt er, dass er den Leser

> "... überzeugen[will], dass die Grundthese der Bewußtseinsforschung grundsätzlich verfehlt ist. Das Bewusstsein entsteht nicht im Gehirn. Und daran liegt es auch, dass bis jetzt niemand eine überzeugende Erklärung aus den neuronalen Grundlagen liefern konnte." (p. 19), Z 51

und für den Geist gilt nach Noë:

> "... der Geist scheint ein Paradoxon in sich zu tragen: Er ist Teil unserer Natur, und doch eignet er sich nicht zum Gegenstand der Naturwissenschaft." (p. 56), Z 52

Die Meinung. teilt er mit vielen Geisteswissenschaftlern. Wir brauchen seiner Ansicht nach eine "Perspektive, aus der sich der bedeutungstragende, nichtmechanische Charakter bewussten Lebens überhaupt erst zeigt" (p. 56). So bewegen sich Bakterien mit ihren Flagellen gezielt auf Zuckerkonzentrationen zu, wobei sie "maschinenartig mit ihrer Umwelt verzahnt erscheinen. Indem wir die Bakterie

derart beschreiben, sehen wir sie aber eigentlich nicht mehr als rein mechanisches und rein physikalisches Wesen, sondern als Einheit, deren Handlungen tatsächlich als *Handlungen* gewertet werden können" (Kursivsetzung durch Noë). Sie wird von der Zuckerkonzentration angezogen, weil sie Zucker zum Leben benötigt, und "das heißt, dass es in seiner Umgebung in einer Beziehung steht." (p. 57).

Noë fragt nun, "wie aus dem unverkennbaren Knacken, Knistern und Knallen der Neurone eine Erfahrung [wird], die eine bestimmte Beschaffenheit und nicht eine andere hat?" (p. 71-72), – wir hören etwas, oder wir sehen etwas. Noë versichert noch einmal, dass die Wissenschaft diese Frage nicht beantworten könne, bis zum heutigen Tag seien wir nicht in der Lage, die sogenannte Erklärungslücke (explanatory gap) zu schließen. Das sei auch nicht verwunderlich:

> Wir können die qualitative Beschaffeneit der Erfahrung nicht mit der intrinsischen Natur der Hirnvorgänge erklären, weil es an Hirnvorgängen nichts gibt, was beispielsweise besonders "visuell" wäre." (p. 72), Z 53

Würde man Noë sagen, dass jeder Augenarzt ihm sagen kann, welche Schäden im visuellen System zu welchen Wahrnehmungsdefiziten führen, so würde ihn das vermutlich wenig beeindrucken, denn für ihn impliziert Die Aussage des Zitats 53, dass dem Organismus gleichtzeitig "Geist" zugeordnet werden kann. Beim Bakterium ist das ein eher rudimentärer Geist, aber es ist Geist. Denn "die Frage nach dem Geist ist die Frage nach dem Leben" (p. 58). Die Bewegungen des Bakteriums bedeuten die Existenz von Handlungskompetenz, und damit habe es "ein zumindest rudimentäres Bewußtsein". Die Frage nach dem Bewußtsein sei die Frage nach dem Leben, und es gehe darum, wie Leben in der Natur entsteht (p. 59). Das Leben einer Bakterie liege nicht in ihrem Inneren verborgen, es sei vielmehr ein Kräftespiel, "an dem die in eine Umgebung eingebettete Bakterie selbst teilnimmt. Und das Gleiche gilt auch für das Bewußtsein in einem weiteren Sinne." (p.60).

Man wird protestieren wollen, denn nach allem, was wir über Bakterien wissen, haben Bakterien keine Neurone und damit auch kein Gehirn. Dieses Argument zieht bei Noë nicht, weil es seiner Ansicht nach, wenn es um das Bewusstsein geht, ja gar nicht auf die Neurone ankommt. Schließlich regeln Neurone auch die Darmtätigkeit, die uns im Allgemeinen nicht bewusst ist. Andererseits enthält das Gehirn höherer Lebenwesen sowohl Seh- wie Hörzentren, – wenn es auf die Neurone dieser Zentren gar nicht ankommmt – wozu gibt es diese Zentren dann? Man könnte annehmen, dass die Sinneszentren so etwas wie Sende- und Empfangszentren für die extracraniale Geisteswelt ist. Wissenschaftstheoretisch ist ist ein solcher Ansatz wenig attraktiv, weil Noë einfach behauptet, das "geistige" Aktivität deswegen keine neuronale Basis hat, weil man bis jetzt nicht weiß, wie neuronale Aktivität Bewußtsein erzeugt. Aber aus diesem Nichtwissen folgt ja nicht, dass es da nichts zu wissen gibt. Mit der Verlagerung psychischer Aktivität in eine rein psychische Welt außerhalb des Kopfes kehrt Noë einfach zur kantschen Zweiten Welt der Noumena zurück, von der man auch nichts weiter weiß als dass

Kant sie postuliert hat.

Man fragt sich auch, wozu der Mensch neuronale Sprachzentren hat. Primaten ohne derartige Zentren können keine Sprache auf humanem Niveau erlernen. Noë scheint hier kein Problem zu sehen: Sprache sei u.a. eine soziale Praxis innerhalb einer Sprachgemeinschaft, deshalb seien unsere kognitiven Fähigkeiten auf eine *soziolinguistische Umwelt* angewiesen: "Unsere Geister überqueren die Grenze des Schädels und werden von einem gemeinsamen soziolinguistischen Gerüst getragen" (p. 109). So kommt er zu der Feststellung

"Die Bedeutung ist nicht im Kopf" (p. 110), Z 54

Noë hat recht mit der Aussage, dass die Neurowissenschaft in den letzten Jahrzehnten in Bezug auf die Frage nach der Entstehung des Bewußtseins aus der neuronalen Aktivität nicht so recht vorangekommen ist. Wie oben schon gesagt wurde folgt daraus noch lange nicht, dass es keine neurowissenschaftliche Erklärung geben kann: die Behauptung, dass alles Psychische gar nicht im Kopf ablaufe, muß erst noch als 'wahr' bewiesen werden, damit man Aussagen über die Möglichkeit neurowissenschaftlicher Erklärungen aufstellen kann: wenn die Naturwisenschaft *noch* keine Erklärung gefunden hat, folgt ja noch lange nicht, dass sie *nie* eine findet. Man hat weit über 2000 Jahre gerätselt, was Gravitation "eigentlich" ist, – hat man daraus gefolgert, dass es keine physikalische Erklärung für die Gravitation gibt? Nein, Einstein hat im November 1915 eine akzeptable Theorie, also eine Erklärung, geliefert. Der Große Fermatsche Satz[84] konnte mehr als 350 Jahre nicht bewiesen werden, obwohl die jeweils besten Mathematiker versucht hatten, einen Beweis zu finden, – man hätte glauben können, dass kein Beweis existiert. Bis der britische Mathematiker Andrew Wiles 1993 einen Beweis vorlegte; dieser Beweis enthielt noch kleine Fehler, einen fehlerfreie Version des Beweises liegt seit 1995 vor. Die Philosophie des Enaktivismus liefert ebenfalls keine Anwort auf die Frage, wie das Bewußtsein entsteht. Dass wir mit der Umwelt interagieren ist eine Einsicht ohne jeden Neuigkeitswert, die Aussage des Zitats Z = 54 ist völlig unbewiesen, Noë kann auf keine Daten verweisen, die die Hypothese der Repräsentation der Umwelt im Gehirn in Frage stellen, und ebenso wenig liefert er theoretische Argumente, sondern nur suggestive Wiederholungen von bestenfalls Plausibilitätsbetrachtungen unter Vernachlässigung aller Evidenz für die Rolle neuronaler Systeme, die allein im medizinischen Bereich akkumuliert wurde. Noë behauptet, die klassische Vorstellung, dass das Nachdenken und Argumentieren etc im Kopf stattfindet, sei von der Philosophie der letzten sechzig Jahre "vollkommen widerlegt" worden (p. 111). Derartige Aussagen verblüffen eher als dass sie zu einem Aha-Erlebnis führen, insbesondere, wenn er zur Illustration auf die in der Neuro- und Kognitionswissenschaft berühmt gewordenen Arbeiten von David H. Hubel (1926 – 2013) und Thorsten N. Wiesel (1924 –) zur Struktur des visuellen Kortex von Katzen und Primaten eingeht, für die die beiden Forscher 1981 mit dem Nobelpreis für Medizin geehrt wurden. Noë erkennt

[84] Die Aussage $a^n + b^n = c^n$ mit $a, b, c, n \in \mathbb{N}$ gilt nur für $n \leq 2$

die Leistung der beiden Forscher insofern an, als sie "fleissig gemessen" haben, aber er sieht in diesen Messungen keine Relevanz für das Verständnis der Prozesse der visuellen Wahrnehmung: die neuronalen Aktivitäten im visuellen Kortex sagen seiner Ansicht nichts über deren Rolle im Wahrnehmungsprozess aus. Dabei wäre es doch nach des Philosophen Ritters Meinung die vornehmste Aufgabe von Philosophen, uns zu zeigen, welche Bedeutung die Ergebnisse von Hubel & Wiesel nun wirklich haben (vergl. Zitat 2, Seite 4). Aber Noë hat nichts dazu zu sagen.

Noës Theorie wurde in den Feuilletons sehr positiv aufgenommen[85], sie wird aber auch in Philosophenkreisen durchaus geschätzt, wenn auch störende Vagheiten in seinen Ausführungen kritisiert werden[86]. Nüchterner ist das Fazit einer im *Scientific American* erschienenen Besprechung von Noës Buch:

> "One comes away from the book without a definitive example of a conscious state that would require more than a brain."[87], Z 55

Wir wissen zwar nicht, wie das Gehirn Bewußtsein erzeugt, aber dass es ohne Gehirn *nicht* erzeugt wird kann kaum geleugnet werden, dafür sprechen allein schon die medizinischen Befunde zu kognitiven Ausfällen nach Unfällen oder Erkrankungen.

1.2.6 Max Plancks Betrachtungen

Die Beziehung zwischen einer deterministischen Auffassung des Naturgeschehens und der These, der Wille sei frei, ist 1932 von Max Planck (Planck (1932)) sehr genau analysiert worden, – mit einer überraschenden Schlußfolgerung.

Zunächst stellt Planck fest, dass das allgemeine Kausalgesetz wegen seines transzendentalen Charakters wahr sei (p. 157); transzendental bedeutet hier so viel wie notwendig, also eine Bedingung der Möglichkeit, eine kausale Folge von Ereignissen als solche zu verstehen. Planck argumentiert nun, dass das Kausalgesetz nur in makroskopischen, nicht aber in mikroskopischen Betrachtungen anwendbar sei, wobei mit 'mikroskopisch' die Prozesse auf atomarem und molekularem Niveau gemeint sind, die nur der berühmten, von Laplace in einem Gedankenexperiment eingeführten *intelligence* zugänglich sind, wenn sie den Kausalitäten der physikalischen Welt bis ins wirklich letzte Detail nachgehen will. Nun muß geklärt werden, ob es der später in 'Laplacescher Dämon' umbenannten *intelligence* möglich ist, Raum für einen freien Willen zu finden, denn physikalisch ist ja, so die Annahme, alles determiniert (p. 159).

Planck überlegt, ob "ein Punkt" existiert, in dem jede kausale Gedankenfolge aufhört; er liefert keine mathematische Definition des Punktes, er erklärt nur,

[85] So schreibt z.B. Helmut Mayer in der FAZ vom 21. 12. 2010: "Alva Noë führt exzellent vor Augen, warum sich in den Neuronen nicht finden lässt, was Bewusstsein ausmacht".

[86] Eine ausführliche Besprechung liefert Rachlin, H. (2012)

[87] J. Hoffman: MIND Reviews: Out of Our Heads, by Alva Noë, April 2009 – Recommendations from Scientific American MIND

dass ein solcher Punkt wissenschaftlich nicht erreicht werden könne, es gebe ihn nur in der "unmessbaren Welt von Geist und Materie" ("in the immeasurable world of mind and matter") (p. 161), in der keine kausale Forschungsmethode mehr anwendbar sei (Planck). Dann kommt die Erläuterung: dieser Punkt sei das individuelle Ego:

> "It is a small point in the universal realm of being; but in itself it is a whole world, embracing our emotional life, our will and our thought." (p. 161),
>
> Z 56

Diese Aussage ist bemerkenswert wenig analytisch für einen auf klare Begrifflichkeit trainierten Physiker vom Kaliber Max Plancks, sie ist eher eine Bildbeschreibung. Sobald man versucht, sie *genau* zu verstehen, zeigt sich ihre Mehrdeutigkeit. Eine Möglichkeit, sich ein klareres Bild zu machen von dem, was Planck als "small point" meint, ist den mentalen Zustand (einschließlich Emotionen, Willen und Gedanken) als Zustand eines hochdimensionalen, komplexen dynamischen Systems zu betrachten; der Begriff des dynamischen Systems ist hinreichend allgemein, um diesen Ansatz zu rechtfertigen: es kommt darauf an, wie dieses System zu konzipieren ist. Der Zustand eines solchen Systems zu einem Zeitpunkt t wird durch die Werte von n Funktionen $x_1(t), \ldots, x_n(t)$ eben zum Zeitpunkt t bestimmt[88]: man könnte sie durch den Vektor $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))$ repräsentieren. Dann entspricht einem Zustand ein Punkt im *Zustands-* oder *Phasenraum* des Systems, eben mit den Koordinaten $x_1(t), \ldots, x_n(t)$; der Zustandsraum ist demnach ein n-dimensionaler Raum. Der Vektor $\mathbf{x}$ durchläuft, als Funktion von t, eine Menge von Punkten; diese Menge ist eine i.A. kontinuierliche *Trajektorie* des Systems; es gibt beliebig viele, mögliche Trajektorien, sie hängen von den Werten von Parameter des Systems ab, die selbst Funktionen der Zeit sein können. Die Frage ist nun, ob man den Punkt mit den Koordinaten $x_1(t), \ldots, x_n(t)$ begrifflich mit dem Planckschen Punkt zusammen bringen kann. Man kann zeigen, dass für einen gegebenen Startwert die Trajektorie völlig deterministisch verläuft, wenn man kein stochastisches System annimmt, was aber für die folgende Argumentation nicht wichtig ist und deshalb hier zur Vereinfachung nicht angenommmen werden soll. Planck postuliert aber, dass in seinem Punkt jede *kausale* Gedankenfolge aufhöre. So kann er argumentieren, denn in einem einzelnen Punkt kann Kausalität im üblichen Sinn nicht existieren, weil es in *einem* Punkt kein Vorher und kein Nachher gibt, so dass von Ursache und Wirkung nicht die Rede sein kann. Tatsächlich ist, wenn von der Trajektorie eines Systems gesprochen wird, von Kausalität nicht die Rede – einer der Sachverhalt, der Bertrand Russell (1913) in seinem berühmten Aufsatz "On the notion of cause" folgern ließ, dass der Kausalitätsbegriff umgangssprachlich zwar nützlich, aber philosophisch übeflüssig sei. Das Kausalitätsprinzip steckt gewissermaßen implizit in der Definition des Trajektorienbegriffs, was nicht bedeuten soll, dass Determinismus und

[88] Es mag plausibler sein, ein Feldmodell, d.h. ein System mit verteilten Parametern zu betrachten, die durch partielle Diff'gleichungen beschrieben werden. Es kommt hier aber nur auf eine Skizze der Grjuundgefaneken an; der Einfachheit halber werden hier nur durch gewöhnliche Differentialgleichungen beschreibbare Systeme betrachtet.

Kausalität identische Begriffe sind. Jedenfalls kann man sagen, dass die Kausalität nicht in einem Punkt – oder in einer diskreten Folge von Punkten – steckt. Die mathematische Konzeption des Zustandsraums für dynamische Systeme kollidiert mit Plancks Idee eines philosophischen Punkts. Vielleicht sollte man auch nicht von einem philosophischen Punkt sprechen: Sprachbilder wie die "unmessbare(n) Welt von Geist und Materie" sind unverträglich mit dem so nüchternen Ansatz, das kognitive System des Menschen als dynamisches System modellieren zu wollen. Der Descartesche Dualismus scheint in Plancks Denken noch eine nicht hinterfragte Rolle zu spielen, die im systemtheoretischen Ansatz nicht mehr vorhanden ist. Sätze wie das Zitat 56, Seite 76, können trotz einer gewissen Anschaulichkeit nur eine höchst diffuse, aber ähnlich wie gute Lyrik eine anregende Bedeutung haben. Denn schon das Suchen nach einer Bedeutung kann zu neuen Ideen führen.

Planck befindet weiter, dass es keinen Grund gebe, sich nicht zum Beobachter dessen, was in ihm geschieht, zu machen. Es sei logisch, wenn auch nicht praktisch, möglich, im Rahmen eines Gedankenexperiments zum Laplaceschen Dämon zu werden, also alle kausalen Abhängigkeiten zu erkennen, – allerdings nur bis auf die eigenen. Die eigenen kausalen Abhängigkeiten könne man nicht erkennen, – denn kein Auge könne sich selbst sehen, so Plancks Illustration seines Arguments. Skeptiker des freien Willens können dieses Argument zur Basis eines Arguments gegen die These der Existenz eines freien Willens machen: die angebliche Freiheit des Ichs sei nur das Resultat eines Mangels an vollständigem Verstehen aller kausalen Relationen. Aber Planck weist diese Argumentation sofort zurück: sie sei falsch, sie gleiche dem Argument, ein Läufer könne seinen eigenen Schatten nicht überholen, weil es ihm an Geschwindigkeit mangele. Die Wahrheit der Freiheit des Individuums resultiere aus der Logik:

> "The fact that the individual here and now, in regard to his own living present act, cannot be subject to the law of causation is a truth that is based on a perfectly sound logical foundation of an a priori kind, such as the axiom that the part is never greater than the whole." (p. 163-164).
>
> Z 57

Die Frage ist, ob dieses Argument tatsächlich so zwingend ist, wie es Planck anscheinend erschienen ist, als er es niederschrieb. Man kann vermuten, dass das *Bild* einer strikt kausalen Einbettung die Vorstellung erzeugt, dass der Mensch einem Automaten gleich durchs Leben geht. Aber wir erleben uns nicht als Automaten, sondern als Wesen, das sein Handeln – einschließlich des Nachdenkens über die Frage nach der Möglichkeit der Existenz eines freien Willens – durch seine eigenen Willensakte bestimmt wird. Natürlich in Interaktion mit der Umwelt. Ich stelle fest, dass der Behälter für Darjeeling-Tee leer ist, *also* schreibe ich 'Darjeeling kaufen' auf die Tafel für einzukaufende Dinge, etc, d.h. ich finde immer Gründe, deretwegen ich etwas getan oder gedacht habe, – wobei ich nicht immer weiß, *warum* mir vorhin der Gedanke kam, ich sollte doch mal wieder eine betimmten alten Freund anrufen. Wenn ich dann anrufe, scheint es mir, dass

diese Handlung "frei" ist, weil mich niemand dazu gedrängt hat und es keinen bestimmten Grund hat, dass ich anrufe, außer dem allgemeinen Gefühl, doch mal wieder anrufen zu sollen. Wir haben keinen Zugang zu den Prozessen, die dazu führen, dass wir dieses oder jenes denken oder fühlen und erleben uns deshalb als "frei". Man fragt sich auch, was Planck genau meint, wenn er das Axiom "that the part is never greater than the whole" anführt. Für sich genommen ist klar, was es aussagt, aber was bedeutet es in Bezug auf die zur Diskussion stehende Freiheit des Willens? Vielleicht meint er, dass man den eigenen Kognitionen nicht zusehen kann, denn das Zusehen selbst ist eine kognitive Aktivität, von der man sich nicht gut vorstellen kann, dass sie sich in ihrem momentanen Vollzug selbst wahrnimmt. Somit entzieht sich auch die momentane kausale Bedingtheit einer Kognition der Eigenwahrnehmung. Es fragt sich aber, ob daraus folgt, dass es deswegen *keine* kausale Bedingtheit gibt: in der Philosophie ist es erlaubt, eher suggestive Aussagen wie das Zitat 57 als analytische Aussagen zu machen. Glauben muß man sie deshalb nicht.

Planck bezieht die Superintelligenz des Laplaceschen Dämons in seine Argumentation ein: auch der Dämon müsste die Aktivitäten seines eigenen Ichs in dem Moment, in dem er sein Ich beobachtet, erkennen, was eben gerade nicht möglich sei (man erinnere sich an das Auge, das sich nicht selbst sehen kann);

> "We cannot possibly study ourselves at the moment or within the environment of an given activity. Here is the place where the freedom of the will comes in and establishes itself, without usurping the right of any rival." (p. 164) Z 58

Einige Zeilen weiter fügt Planck an, die Freiheit des Ichs und seine Unabhängigkeit von der Kausalkette resultiere aus dem Diktat des Bewußtseins (p. 165). Das "eiserne Gesetz der logischen Konsistenz" impliziere, dass wir die kausalen Grundlagen unserer eigenen Zukunft und deren Vorhersagbarkeit nicht bestimmen können (p. 167). Die Kommentare zu 57 lassen sich auch auf die Aussage 58 anwenden.

Planck hat sich viele Jahre mit der Frage nach dem freien Willen beschäftigt. In einem Vortrag (1946) schreibt er:

> "Von außen betrachtet ist der Wille kausal determiniert, von innen betrachtet ist der Wille frei. Mit der Feststellung dieses Sachverhaltes erledigt sich das Problem der Willensfreiheit. Es ist nur dadurch entstanden, dass man nicht darauf geachtet hat, den Standpunkt der Betrachtung ausdrücklich festzulegen und einzuhalten. Wir haben hier ein Musterbeispiel für ein Scheinproblem. Wenn diese Wahrheit gegenwärtig auch noch mehrfach bestritten wird, so besteht doch für mich kein Zweifel darüber, dass es nur eine Frage der Zeit ist, wann sie sich zur allgemeinen Anerkennung durchringen wird." (p. 9), Z 59

Das Problem des freien Willens ein Scheinproblem! Zu der Vermutung, dass es sich um ein Scheinproblem handeln könnte, hatten schon die intuitiven Betrachtungen zum Begriff des freien Willens in Abschnitt 1.1 geführt. Gleichwohl ist

nicht ganz klar, was Planck wirklich meint: bevor er zu der Aussage im Zitat 59 kommt, hat er ja festgestellt, dass die Entscheidungen im Ich verankert sind, das sich der Kausalität entzieht (Zitate 57 und 58), dass aber das Universum mit allem, was darin ist, gleichwohl determiniert sei. Planck wiederholt viele seiner Argumente aus den *Philosophical Essays* aus dem Jahr 1932. Das wirft die Frage auf, wann ein Problem ein Scheinproblem ist (zum Begriff des Scheinproblems vergl. Wissenschaftstheorie III (1), Abschnitt 6.3). Plancks Schlußfolgerungen sind oft *ad hoc* und von der Überzeugung geleitet, dass moralische Verantwortung nicht ohne die Freiheit des Willens möglich sei. Dass die Aussage, dass ein Individuum im Hier und Jetzt nicht dem Kausalgesetz unterliegt (Zitat 57), aus rein logischen Gründen wahr sein muß, erscheint als Behauptung und wird nicht weiter begründet. Eine detaillierte Kritik seiner Argumentation überschreitet den Rahmen dieser kurzen Darstellung, eine ausführliche Diskussion findet man bei Griffel (1998). Griffel findet in Plancks Vorträgen über die Willensfreiheit eine Reihe argumentativer Sprünge, die erst vor dem Hintergrund der Gesamtphilosophie Plancks verständlicher werden (vergl. Scheibe (2006)[89]). Plancks Aussage in Zitat 57 erinnert an ein Problem, das Singer (Zitat 34, Seite 49) auf den Punkt gebracht hat. Singers Aussage in Zitat 34 kann bedeuten, dass Planck mit seiner Deutung des "Ichs" recht hat[90]. Darüber ist aber noch nicht endgültig entschieden worden, und man sollte nicht vorschnell aufgeben, – mehr Detailkenntnisse über die neuronale Basis kognitiver Prozesse sind dazu dringend erforderlich.

Planck war erklärter Determinist, und bei dieser Grundposition ist es viel schwieriger, für die Willensfreiheit als für die Unfreiheit des Willens zu argumentieren. Die Frage, "ob es einen Punkt gibt, an dem jede kausale Gedankenfolge aufhört" und die dazu korrespondierende Aussage in Zitat 56 auf Seite 76 illustrieren seine Bemühungen, im Reich eines allgemeinen Determinismus einen Ort für die Freiheit der Entscheidung zu finden. Planck schreibt, dass wir uns selbst und unsere Aktivität in einem gegebenen Moment nicht erkunden (study) können und hieraus der freie Wille entstehe, diese Aussage sei logisch zwingend (vergl. Zitat 57 oben). Mit der Aussage, eine Selbstbeobachtung für einen gegebenen Moment einer kognitiven Aktivität sei nicht möglich, mag er recht haben (vergl. Singers Aussage in Zitat 34, Seite 49) – aber dass daraus die Freiheit des Willens folgt ist nicht offenkundig. Planck macht keinen Versuch, seine Argumentation zu formalisieren, um ihr Stringenz zu verleihen. Vermutlich hat Planck Alan Turings Arbeit aus dem Jahr 1936 nicht gekannt; sie liefert, wie im folgenden Abschnitt beschreiben wird, eine Basis für eine strengere Diskussion der Planckschen Betrachtungen.

[89] Einen guten Überblick findet man in
https://www.informationphilosopher.com/solutions/scientists/planck/.

[90] Die Anwendung der Theoreme von Gödel und Turing auf kontinuierliche Systeme wird gelegentlich kritisiert, weil die Theoreme sich auf diskrete Systeme beziehen. Neuere Arbeiten legen nahe, dass die Theoreme auf kontinuierliche Systeme übertragbar sind, z.B. Granicky, M.S. (1995), Vicente (2023)

1.2.7 Seth Lloyds rekursive Argumentation

Tatsächlich ist ein solcher Beweis von dem Physiker und Informatiker Seth Lloyd (2012) vorgelegt worden; eine Skizze des Beweises findet man im Anhang. Hier soll nur die Aussage des Lloydschen Befundes diskutiert werden[91]. Der Beweis basiert auf Turings (1936) Halte-Theorem, demzufolge es keinen Algorithmus T_2 gibt, mit dem festgestellt werden kann, ob ein gegebener Algorithmus T_1 für einen gegebenen Input stoppt oder in eine unendliche Schleife gerät; das Halte-Theorem wird im Anhang 2 vorgestellt.

Lloyd illustriert seine These zunächst informell. Er beschreibt sein Verhalten in einem Restaurant: er studiert die Speisekarte, ohne zu wissen, was er wählen wird. Seine Frau weiß es allerdings schon, weil er immer dasselbe Gericht wählt. Aber Lloyd ist nicht mit der Absicht ins Restaurant gegangen, genau dieses Gericht zu bestellen. Der Punkt ist, dass ihm sein Entscheidungsprozess nicht bewußt ist, – er weiß erst, was er möchte, nach einer gewissen Zeit des Studiums der Speisekarte. Er kann es nicht vor Ablauf des Entscheidungsprozesses wissen, denn könnte er es wissen, gäbe es einen einen zum Entscheidungsprozess parallel laufenden Prozess, der das Ergebnis des Entscheidungsprozesses vorhersagen könnte, – und dieser Prozess wäre dann der eigentliche Entscheidungsprozess, für den dann dasselbe Argument gelten würde. Jeder Prozess, der den Entscheidungsprozess "beobachtet", kann dem Entscheidungsprozess nur nachlaufen, andernfalls wäre der Beobachtungsprozess der Entscheidungsprozess. Deshalb kann der Beobachtungsprozess, sollte es ihn denn geben, den Entscheidungsprozess auch nicht beeinflussen. Moralische Aspekte, sollten sie für eine bestimmte Entscheidung von Bedeutung sein, werden beim Entscheidungsprozess bereits berücksichtigt.

Bei der Diskussion der Möglichkeit eines freien Willens geht es u.a. um die (Selbst-)Wahrnehmung und eventuelle Beeinflussung oder Steuerung des Prozesses, der zu einer Entscheidung führt. Aber die Beeinflussung bzw. Steuerung sind dann bereits Teil des Entscheidungsprozesses. Nimmt man, wie viele Kompatibilisten, an, dass diese Prozesse irgendwo außerhalb des Hirns ablaufen, so hat man die Diskussion von Entscheidungsprozessen kurz- und abgeschlossen, denn über den Ort und die Struktur derartiger extracranialer Abläufe ist nichts bekannt. Es wird nur *postuliert* und keineswegs bewiesen, dass diese Abläufe nicht durch die Fesseln des Kausalitätsprinzips eingeengt werden. Die Behauptung, Entscheidungsprozesse seien keine Hirnprozesse, liefert keinerlei Erklärung, weil es völlig unklar ist, was man sich unter außerhalb des Gehirns ablaufenden Prozessen überhaupt vorstellen soll.

Geht man aber davon aus, dass Entscheidungsprozesse Hirnprozesse sind, so muß man akzeptieren, dass es sich letzlich um physikalische Prozesse handelt. Lloyd führt aus (p. 11), dass derartige Prozesse durch Turingmaschinen simuliert werden können. Auch Folgen von Argumentationen können, wenn sie durch Regeln gesteuert werden, in ihrem Kern durch Algorithmen und deswegen durch

[91] Interessierte Leser können die Arbeit direkt aus dem Netz herunterladen: Lloyd, S.: A Turing Test for Free Will. https://arxiv.org/abs/1310.3225

Turingmaschinen abgebildet und deshalb auch simuliert werden. Lloyd nennt derartige Argumentationen *rekursive Argumentationen*. Er argumentiert, dass die genannten physikalischen Prozesse die rekursiven Argumentationen umfassen bzw. durch derartige Argumentationen repräsentiert werden können.

Die zentrale These ist, dass Entscheidungen für das entscheidende Subjekt nicht vorhersagbar sind: es gibt stets eine Deliberation, die ablaufen muß, damit entschieden werden kann. Dieser Prozess kann sehr kurz oder auch länglich sein, etwa wenn moralische Aspekte eine Rolle spielen. Für den Vegetarier benötigt die Entscheidung zwischen Wurst und Käse kaum Zeit, weil sie vorprogrammmiert ist, während die Entscheidung, ob das neue Eigenheim mit Gas oder mittels Wärmepumpe beheizt werden soll, länger dauern kann, weil unmittelbare ökonomische Aspekte einerseits und eher ethische Aspekte andererseits gegeneinander abgewogen werden müssen. Am Ende erscheint das Resultat des Prozesses im Bewußtsein als vom Ich erzeugt und *deshalb* als freie Entscheidung. Dabei ist es unerheblich, ob der Entscheidungsprozess deterministisch oder probabilistisch abgelaufen ist, jede während des Deliberationsprozesses getroffene Sub-Entscheidung ist "meine"', von meinem "Ich" getroffene und deshalb "freie" Entscheidung. Dass die bei der Deliberation ablaufenden, durch Aufmerksamkeitsfluktuationen, ästhetische und/oder moralische Präferenzen etc gesteuerte, aber nicht notwendig bewußte Prozesse gar nicht "frei" sind geht in das Erlebnis von Freiheit gar nicht ein. Hier ist der von Planck erwähnte Punkt (Zitat 56, Seite 76 oben), in dem Planck die Freiheit der Entscheidung angenommen hat. Es sieht aber so aus, als habe der Begriff der Freiheit in diesem Zusammenhang keinen Sinn mehr... es ist "meine" Entscheidung, denn ich bin so, wie ich nun mal bin.

Lloyds Ergebnisse lassen sich wie folgt zusammenfassen (den formalen Beweis findet man in Lloyd (2012), p. 10; er basiert wie der des Halte-Problems auf dem Cantorschen Diagonalverfahren (vergl. Anhang, Abschnitt 2.2):

1. Es existiert keine allgemeine Methode, mit der sich die Frage beantworten ließe, ob eine entscheidende Person überhaupt in endlicher Zeit zu einer Entscheidung kommt; diese Aussage ist Turings Halte-Problem.

2. Wenn die entscheidende Person sich innerhalb eines endlichen Zeitintervalls entscheiden muß, so kann jede allgemeine Methode, mit der die Entscheidung der Person bestimmt ("vorhergesagt") werden kann, für diese Bestimmung mehr Zeit benötigen, als der Entscheidungsprozess der Person selbst.

3. Jeder durch eine universelle Turingmaschine repräsentierbare Entscheider kann nicht alle Fragen über sein zukünftiges Verhalten beantworten.

4. Jeder zeitlich begrenzte Entscheider benötigt mehr Zeit, den Entscheidungsprozess zu simulieren als ihn direkt durchzuführen.

Diese Ergebnisse erklären, warum die meisten Menschen sich als *frei* entscheidend erleben: die Entscheidung erscheint im Bewußtsein, die entscheidende Per-

son kann sagen: "Ich habe die Entscheidung getroffen!", das "Ich" erscheint als die entscheidende Instanz. Dies gilt unabhängig davon, ob der Entscheidungsprozess deterministisch oder probabilistisch verläuft. Wir können den Entscheidungsprozess nicht mit einem parallel zu diesem Prozess laufenden Prozess beobachten und dementsprechend auch nicht auf "freie Weise" beeinflussen. Dieser Befund ist vermutlich das, was Planck meinte, dass ein Auge sich nicht selbst sehen und ein Läufer seinen Schatten nicht einholen kann, was ihn aber noch nicht dazu brachte, die (Hypo-)These der Existenz des freien Willens aufzugeben, – eine Metapher ist noch kein Beweis. Entscheidet man sich dazu, die These der Existenz des freien Willens beizubehalten, so nimmt man eine Art impliziter Umdefinition des Begriffs des freien Willens vor: der Wille ist frei, weil "Ich" die Entscheidungen treffe (sofern sie nicht von außen erzwungen werden) und der Entscheidungsprozess von "mir" durchgeführt wird, – der Entscheidungsprozess ist gewissermaßen in den Prozess, der das Bewußtsein generiert, integriert. Dass diese Prozesse von meinem Gesamtzustand abhängen und insofern auch determiniert werden, geht aber in diese Definition nicht ein. Die Erfahrung der "Freiheit" des Willens ist sehr überzeugend – man kann von einem Evidenzerlebnis sprechen –, so dass es schwer, wenn nicht unmöglich wird, sozusagen in Eigenbeobachtung die Rolle der nicht bewußt ablaufenden Prozesse zu reflektieren.

1.2.8 Freier Wille und moralische Verantwortung

Die zentrale Motivation von Philosophen, die für die Freiheit des Willens argumentieren, scheint die Frage nach der moralischen Verantwortung zu sein. Schon Kant konstatierte, dass es ohne einen freien Willen kein moralisches Handeln geben könne. Skeptiker des freien Willens sind der Ansicht, dass der freie Wille keine Voraussetzung für die Möglichkeit moralischen Handelns ist. Einigkeit herrscht aber über den Sachverhalt, dass wir, in Habermasscher Ausdrucksweise, vergesellschaftete Individuen sind, d.h. wir haben Regeln – wenn auch nicht notwendig alle – des sozialen Handelns gelernt (Bandura & Walters, (1963)) haben wohl als erste eine empirisch untermauerte Theorie des sozialen Lernens präsentiert, der aber von Philosophen einerseits und Wissenschaftlern andererseits verschieden ausgelegt werden kann. So schreibt der Neurowissenschaftler W. Singer (2004):

> "Genetische Faktoren, frühe Prägungen, soziale Lernvorgänge und aktuelle Auslöser, zu denen auch Befehle, Wünsche und Argumente anderer zählen, wirken stets untrennbar zusammen und legen das Ergebnis fest, gleich, ob sich Entscheidungen mehr unbewussten oder bewussten Motiven verdanken. Sie bestimmen gemeinsam die dynamischen Zustände der "entscheidenden" Nervennetze." (p. 254) Z 60

Natürlich sind Philosophen wie Habermas, Nida-Rümelin , Noë etc mit dieser Interpretation empirischer Daten nicht einverstanden, aber Singer weist auch darauf hin, dass eine Straftat von den Gerichten nach Maßgabe der psychischen Befindlichkeit des Täters beurteilt wird: wird nämlich festgestellt, dass ein Tumor die

im Frontalhirn angesiedelten Kontrollfunktionen reduziert hat, so wird der Täter mit Nachsicht beurteilt:

> "Der gleiche "Defekt" kann aber auch unsichtbare neuronale Ursachen haben. Genetische Dispositionen können Verschaltungen hervorgebracht haben, die das Speichern oder Abrufen sozialer Regeln erschweren, oder die sozialen Regeln wurden nicht rechtzeitig und tief genug eingeprägt, oder es wurden von der Norm abweichende Regeln erlernt, oder die Fähigkeit zur rationalen Abwägung wurde wegen fehlgeleiteter Prägung ungenügend ausdifferenziert. Diese Liste ließe sich nahezu beliebig verlängern. Keiner kann anders, als er ist." (Singer (2004), p. 254) Z 61

Zumindest bei den hier vorgestellten Philosophen werden derartige Sachverhalte noch nicht einmal erwähnt, die Idee, dass Entscheidungen nach Maßgabe von Gründen und außerhalb des Kopfes getroffen werden läßt alles, was mit dem Gehirn zusammenhängt, als irrelevant erscheinen; wie Noë im Zusammenhang mit den Untersuchungen von Hubel & Wiesel argumentiert hat: die Philosophie und nicht die Naturwissenschaft beantwortet alle relevanten Fragen.

Vielleicht auch nicht. Der Hinweis auf die Notwendigkeit von psychologischen und medizinischen Therapien mag genügen, um derart abgehobene philosophische Theorien mit Zurückhaltung zu betrachten. Denn psychische Störungen verweisen ja auf die Tatsache, dass Menschen sich eben nicht frei gegen ihre Störung entscheiden können, auch wenn die Störung die Lebensqualität herabsetzt, – sie benötigen eben therapeutische Hilfe, und die ist nicht notwendig nur psychologischer, sondern gelegentlich auch pharmakologischer Art, d.h. sie wirkt auf physiologische Weise[92]. Man fragt sich, welchen Erkenntniswert philosophische Theorien der hier vorgestellten Art überhaupt erzeugen. Der von Roth (2004) zitierte Philosoph J. Ritter, der der Ansicht war, dass Naturwissenschaftler die natürliche Welt zwar erforschen, aber nicht interpretieren können, weshalb sie die Interpretation besser den Philosophen überlaassen sollten, verfügte offenbar über ein gesundes Selbstbewußtsein. Ein Zitat des Physiker/Philosophen C.F. von Weizsäcker zur Ansicht des Philosophen Martin Heideggers über den epistemischen Wert der Naturwissenschaft illustriert dieses Selbstbewußtsein:

> "Die Philosophie billigt der Naturwissenschaft nur eine seinsvergessene Projektion der Wahrheit zu" (.v Weizsäcker (1988), p. 185). v. Weizsäcker fragte weiter, wie man in der Philosophie verständlich macht, dass Naturwissenschaft überhaupt gelingt. Heideggers Anwort: "Das weiß ich auch nicht." Z 62

Es soll hier nicht grundsätzlich gegen philosophische Reflexionen über empirische Befunde polemisiert werden, sondern nur der Ausschließlichkeitsanspruch der

[92] Habermas (2004) diskutiert die Rolle von Gründen und bedingter Freiheit im Unterschied zur Rolle von Ursachen ausführlich, allerdings am Ende nicht überzeugend, weil er nicht auf das, was bereits über die physische Basis bewußter Prozesse bekannt ist, eingeht.

Betrachtungen von Philosophen in Frage gestellt werden. Philosophische Theorien sind keine formalen Theorien, wie man sie aus der Mathematik kennt und wo man aus Axiomen mit wohldefinierten Schlußregeln und genau formulierten Definitionen "wahre" Aussagen deduziert. Tatsächlich würde ein Versuch, die Philosophie zu axiomatisieren, sofort in einem Selbstreferenzproblem enden, weil dann ja auch die Wahl der Axiome durch andere Axiome gerechtfertigt werden müsste, d.h. man wird in einen infiniten Regress geführt. Der polnische Mathematiker und Logiker Alfred Tarski (1901 – 1983) hat als erster die Beziehung zwischen dem allgemeinen Wahrheitsbegriff und dem Wahrheitsbegriff in formalen Sprachen ausführlich diskutiert (Tarski, A. (1933/1956)), und man kann sagen, dass Tarskis Analyse nicht trivial ist, worauf hier aber nicht eingegangen werden kann, aber es kann angemerkt werden, dass es scheint, dass viele Philosophen des Geistes sich über ihren Wahrheitsbegriff kaum Gedanken machen.

Eine Teilmenge von Aufsätzen aus der Philosophie des Geistes enthält in erster Linie begriffliche Explikationen, gepaart mit apodiktischen Aussagen (z.B. Dretske: "Meaning itself is causally inert, powerless to initiate, ...", Zitat 32, Seite 48), und "Even if the things that have meaning are in the head, the meanings themselves aren't in the head." Zitat 33, Seite 48). Selbst wenn die Elaborationen der Annahmen strenger und weniger apodiktisch vorgenommen würden bleibt immer noch das Letztbegründungsproblem, nach dem man stets gezwungen ist, für Theorien jeweils bestimmte Annahmen machen zu müssen, die zuguterletzt nicht weiter begründbar und deshalb angreifbar sind[93]. Da verschiedene Autoren nicht nur aus der Philosophie, sondern auch aus den Naturwissenschaften (man denke an die Physiker Anton Zeilinger (Zitat 4, Seite 13, und Gerard t'Hooft (Zitat3, Seite 12 in Abschnitt 1.1.2 und die Neurowissenschafter Peter Tse (Zitat 69, Seite 98 und Robert Sapolsky (Zitat 70, Seite 98) zu verschiedenen Ansichten bezüglich der Freiheit des Willens haben, scheint es unwahrscheinlich, dass man zu einer allgemein akzeptierten Axiomatik von Theorien über den Willen gelangt. Spätestens seit David Hume ist bekannt, dass der Begriff der Kausalität mehrdeutig ist. Der von vielen Kompatibilisten verwendete Begriff des Grundes rekurriert implizit auf den der Kausalität, so dass schon allein deswegen unklar ist, warum Entscheidungen auf der Basis von Gründen freie Entscheidungen sein sollen. Habermas' Ausführungen zum vergesellschafteten Individuum oder Nida-Rümelins Begriff der humanistischen Perspektive sind überzeugend nur für den, der sowieso schon in diesem begrifflichen Rahmen denkt oder die grundsätzliche Skepsis gegenüber dem naturwissenschaftlich orientiertem Denken teilt. Wenn ich mich auf dem Weg zur Konditorei entscheide, 20 Euro an Oxfam zu spenden statt, wie ursprünglich beabsichtigt, Kuchen zu kaufen, so kann ich sagen, ich hätte das aus freien Stücken getan, niemand hätte mich gezwungen, mich so zu entscheiden. Das Gegenargument ist aber, dass ich gar nicht anders konnte, als mich so zu entscheiden, – wenn nämlich *mein* Wertesystem kei-

[93] vergl. Wissenschaftstheorie III (2) Abschnitt 2.3

ne andere Entscheidung zuläßt[94]. Das Wertesystem ist die Ursache, der Grund[95] für meine Entscheidung; man kann sagen, dass meine Handlung kausal bedingt sei. Aber deswegen ist sie noch lange nicht vorhersagbar, wie die Diskussionen über den Determinismus gezeigt haben. Es ist bemerkenswert, dass diese Diskussionen von Philosophen mit "humanistischer Perspektive" nicht berücksichtigt werden. Die Gründe, die zur meiner moralischen Entscheidung führen, müssen mir zum Zeitpunkt meiner Entscheidung wegen stochastischer Fluktuationen der Hirnaktivität nicht alle präsent sein, oder die Randbedingungen können so sein, dass die "normalen" moralischen Werte durch andere Werte bzw. Wünsche zeitlich begrenzt dominiert werden, z.B. von dem Wunsch, die Geburtstagsfeier für meinen Enkel nicht zu verderben.

1.2.9 Galen Strawsons Folgerungen

Der britische Philosoph Galen Strawson[96] hat diese Argumentation straffer ausformuliert (Strawson, G. (2022)[97]):

> 1. You do what you do – in the circumstances in which you find yourself – because of the way you then are.
> 2. So if you're going to be ultimately responsible for what you do, you're going to have to be ultimately responsible for the way you are — at least in certain mental respects.
> 3. But you can't be ultimately responsible for the way you are in any respect at all.
> 4. So you can't be ultimately responsible for what you do.

Strawson elaboriert insbesondere die Aussage 3., derzufolge wir letztlich nicht für die Art und Weise, in der wir sind, verantwortlich sein können. Denn wir sind nicht für unsere genetische Struktur verantwortlich, da wir diese Struktur nicht vorgeburtlich selbst bestimmt haben, und darüber hinaus sind wir nicht für unsere frühen Erfahrungen verantwortlich. Wegen unseres genetischen Erbes und der frühen Erfahrungen können wir nicht in einem späteren Lebensabschnitt eine wahre moralische Verantwortung erlangen. Denn ob man bei dem Versuch, sich selbst so zu verändern, dass man mehr moralische Verantwortung übernehmen kann Erfolg hat, hängt bereits von der Art und Weise ab, in der man wegen seines genetischen Erbes und der frühen Erfahrungen schon existiert. Weiter gibt es eine Menge von Einflußfaktoren, u.a. zufällige Einflüsse, für die man nicht verantwortlich ist:

[94] Man kann hehre Werte haben, die aber bei der Erfüllung von Wünschen in einem Morast von Relativierungen verschwinden. Hier ist gemeint, dass ein Wertesystem im psychischen System so verankert ist, dass es auch handlungsrelevant ist.

[95] Ursache und Grund werden hier absichtlich als Synonyme angeführt!

[96] Der Sohn des renommierten britischen Philosophen Peter Frederick Strawson (1919 – 2006)

[97] (Strawson, G. (2022 Moral responsibility without free will. iai news, 27th April 2022, https://iai.tv/articles/moral-responsibility-without-free-will-auid-2115)

> Ultimate responsibility for how one is, is impossible, because it requires the actual completion of an infinite series of choices of principles of choice.

Strawson zitiert Nietzsche:[98]

> Die causa sui[99] ist der beste Selbst-Widerspruch, der bisher ausgedacht worden ist, eine Art logischer Notzucht und Unnatur: aber der ausschweifende Stolz des Menschen hat es dahin gebracht, sich tief und schrecklich gerade mit diesem Unsinn zu verstricken.[100]
>
> Nietzsche, F. (1886): Jenseits von Gut und Böse, Erstes Hauptstück: von den Vorurteilen der Philosophen, 11 - 20.

1.2.10 Albert Einsteins Bekenntnis

Albert Einstein (im Anhang von Plancks "Where is Science going?") hat Strawsons Gedanken offenbar schon vorher gehabt:

> "Honestly I cannot understand what people mean when they talk about the freedom of the human will. I have a feeling, for instance, that I will something or other; but what relation this has with freedom I cannot understand at all. I feel that I will to light my pipe and I do it; but how can I connect this up with the idea of freedom? What is behind the act of willing to light the pipe? Another act of willing? Schopenhauer once said: Der Mensch kann was er will; er kann aber nicht wollen was er will (Man can do what he wills but he cannot will what he wills)." Epilogue: A socratic dialogue Planck–Einstein–Murphy, in: Planck (1932), p. 201, Z 63

So kann man die Debatte auf den relevanten Punkt bringen, ganz ohne Perspektiven und Dualismen zu bemühen. Wahrscheinlich ist Einsteins *statement* für Philosophen wie Habermas und Nida-Rümelin eine Provokation.

1.2.11 Moritz Schlicks Argumentationen

Der Physiker und Philosoph Moritz Schlick (1882 – 1936) war der Begründer und die zentrale Figur des Wiener Kreises, eines philosophischen Zirkels, dessen Mitglieder sich das Ziel gesetzt hatten, eine 'wissenschaftliche Philosophie' zu entwickeln, in der die Rolle der Metaphysik minimal und die der Empirie maximal sein sollte. Im Jahr 1930 erschien Schlicks Buch 'Fragen der Ethik' in Wien

[98] http://www.zeno.org/Philosophie/M/Nietzsche,+Friedrich/Jenseits+von+Gut+und+B

[99] causa = Grund,Ursache, sui = selbst, causa sui = Ursache seiner selbst. In der Theologie Charakterisierung Gottes, im hier gegebenen Kontext die Freiheit, etwas aus sich selbst geraus tun zu können.

[100] Das Verlangen nach ≫Freiheit des Willens≪, in jenem metaphysischen Superlativ-Verstande, wie er leider noch immer in den Köpfen der Halb-Unterrichteten herrscht, das Verlangen, die ganze und letzte Verantwortlichkeit für seine Handlungen selbst zu tragen und Gott, Welt, Vorfahren, Zufall, Gesellschaft davon zu entlasten, ist nämlich nichts Geringeres, als eben jene causa sui zu sein und, mit einer mehr als Münchhausenschen Verwegenheit, sich selbst aus dem Sumpf des Nichts an den Haaren ins Dasein zu ziehn; etc

als vierter Band der *Schriften zur wissenschaftlichen Weltauffassung*, die vom Wiener Kreis herausgegeben wurde. Der Abschnitt 1 des VII-ten Kapitels hat die Überschrift "Das Scheinproblem der Willensfreiheit". Er mache sich nur "mit Widerstreben und Zögern" seiner Besprechung ethischer Fragen einzufügen. Die Frage nach der Willensfreiheit werde zwar noch oft als eine Grundfrage der Ethik angesehen, aber die Willensfreiheit sei nur durch ein Mißverständnis zu einem "vielerörterten Problem" geworden; "gescheite Köpfe" und insbesondere David Hume hätten diese Sachlage bereits aufgedeckt, es sei ein Skandal der Philosophie, dass "immer noch soviel Druckerschwärze und Papier auf diese Sache verschwendet" wird. In der Ethik komme es allein auf die Verantwortlichkeit an. In dem Zusammenhang müsse er auch etwas über die Freiheit sagen, – nicht zuletzt, um dem Skandal in der Philosophie ein Ende zu bereiten.

Schlick beginnt mit Anmerkungen zum Kausalprinzip. Egal um welche Wissenschaft es sich handele, es gelte generell, dass Erklären so viel bedeute wie die zur Diskussion stehenden Prozesse auf Gesetze zurückzuführen. Dieses Prinzip gelte auch für die Psychologie. Das Kausalprinzip besage, dass "alles Geschehen allgemeinen Gesetzmäßgkeiten untersteht", – Wissenschaft setze das Kausalprinzip voraus. Akzeptiert man diese Annahme, so postuliere man einen allgemeinen Determinismus, von dem man allerdings nicht sagen könne, ob er nun generell gelte oder nicht. Den Versuch, die Frage nach der allgemeinen Geltung des Determinismus durch reines Nachdenken entscheiden zu wollen, hält Schlick für "lächerlich": man müsse bedenken, "mit welch ungeheurem Aufwand experimenteller und logische Kunst die gegenwärtige Physik der Frage vorsichtig näher rückt, ob auch für die allerfeinsten Vorgänge im Innern der Atome der Kausalsatz aufrecht zu erhalten sei". Schlick bezieht sich hier wohl auf die Unschärferelation, die von Heisenberg im Jahr 1927 entdeckt worden war und von dem Heisenberg sagte, mit dieser Relation sei das Kausalprinzip "abgeschafft" worden (v. Weizsäcker (1999), p. 315). Anhänger der Willensfreiheit würden versuchen, das Kausalprinzip "mit verbrauchten Argumenten" zurückzuweisen, die Gegner würden versuchen, es mit ebensolchen Argumenten zu beweisen, und sowohl Befürworter wie Gegner würden vor der Größe ihrer Aufgabe (Beweis oder Widerlegung) nicht erschrecken, – weil ihnen diese Größe nicht bewußt sei. Andere Befürworter der Willensfreiheit würden zwischen zwei Reichen unterscheiden: in einem würde der Determinismus regieren, im anderen nicht; auch Kant habe diesen "allerbilligsten" Ausweg "leider" gewählt. Den Grund sieht Schlick in der Vorstellung, dass der Determinismus impliziere, dass er die Übernahme von Verantwortung unmöglich mache. Dieses Argument gelte aber nicht, es beruhe auf einer irrtümlichen Deutung des Wortes 'Gesetz'. Der Staat erläßt Gesetze, die den "natürlichen Wünschen" der Bürger oft widersprechen und deshalb als Zwang empfunden werden. In den Naturwissenschafften bedeutet 'Gesetz' aber etwas anderes: ein Gesetz ist keine Vorschrift, die bestimmt, wie sich etwas verhalten *soll*, sondern "eine Formel, die beschreibt, wie sich etwas tatsächlich verhält". Schlick veweist auf die Himmelsmechanik, die den Planeten nicht vorschreibt, auf welchen Bahnen sich die Planeten bewegen sollen, obwohl sie sich ganz anders

bewegen möchten. Die Gesetze der Himmelsmechanik besagen nur, auf welchen Bahnen sich die Planeten tatsächlich bewegen.

Dieser Sachverhalt müsse auf den Willen angewendet werden, – dabei würde einem "sofort ein Licht aufgehen": wenn man sage, der Wille gehorche psychologischen Gesetzen, so ist die Rede nicht von Staats- , sondern von Naturgesetzen, die "nur formulieren, welche Wünsche der Mensch unter bestimmten Umständen *tatsächlich* hat", sie "beschreiben die Natur des Willens, nicht anders als die astronomischen Gesetze die Natur der Planeten beschreiben". Zwang liege vor, wenn der Mensch an der Erfüllung seiner natürlichen Wünsche gehindert werde, und: wie kann man die Regel, nach der die natürlichene Wünsche aufsteigen, wieder als "Zwang" angesehen werden?

Dazu merkt Schlick an, dass die Frage, wann ein Mensch sich selbst für verantwortlich hält, wichtiger sei als die Frage, wann ein Mensch für verantwortlich erklärt wird. Dazu muß der Mensch das Bewußtsein haben, "selbstständig" und "aus eigenem Antrieb" Urheber seiner Handlung zu sein. Dies sei das Bewußtsein der *Freiheit* (Schlicks Kursivsetzung), und "dies ist einfach das Wissen, aus *eigenen* Wünschen heraus gehandelt zu haben." Diese Wünsche ergeben sich aus der Gesetzmäßigkeit des eigenen Charakters und nicht aus äußerer Gewalt; diese Abwesenheit drücke sich in dem Gefühl aus, die Möglichkeit gehabt zu haben, dass man auch hätte anders handeln können. Schlick schreibt, es sei ihm "unbegreiflich", diesen Sachverhalt als Grund für den Indeterminismus, der dem freien Willen unterliegen müsse, gehalten zu haben; es sei doch natürlich, dass man anders gehandelt hätte, hätte man denn *gewollt*, anders zu handeln. Hätte man anders handeln wollen, hätte man andere Motive gehabt. Das Bewußtsein des eigenen Wollens sage nichts über das Kausalprinzip aus, es sei kein Bewußtsein der Ursachelosigkeit, sondern eines der Freiheit, die darin bestehe, so handeln zu können, wie man es wünsche.

Für die Verantwortlichkeit sei die Kausalität sogar eine Voraussetzung, denn der Begriff der Verantwortung beruhe auf dem der Verursachung, – d.h. "der Gesetzlichkeit der Willensentschlüsse". Denn man stelle sich vor, ein Willensentscheid sei ursachelos, es gäbe keine Ursache für die Bildung eines Willens. Dann, so betont Schlick, sei die Willensbildung *zufällig* und deswegen eben nicht 'frei'. Einen Straftäter würde man um so strenger bestrafen, je mehr Gründe man für sein Handeln fände, und dem entsprechend wäre man um so nachsichtiger, je weniger Ursachen man fände. Dieses Prinzip findet man auch im heutigen Strafrecht. Man findet, so Schlick, die Ursachen für Handlungen im Charakter eines Menschen, kein Mensch würde deshalb das Kausalprinzip in Frage stellen. Die Metaphysiker verwechseln Freiheit von Zwang mit dem Fehlen von Ursachen, ein ursacheloses Geschehen sei schlicht zufälliges Geschehen, ein nicht determiniertes Wollen hebe jede Verantwortung auf. Verantwortliches Handeln setzt, so Schlick, das Kausalprinzip für Willensprozesse voraus.

Anregung: Man vergleiche den von Schlick vertretenen Begriff der Verantwortung mit dem von Habermas vetretenen Begriff der verantwortlichen Urheber-

schaft.

1.3 Philosophie und Empirie

1.3.1 Allgemeine Betrachtungen

Auf die Frage nach dem freien Willen kann man wie G. Strawson oder A. Einstein reagieren. Anhänger der These des freien Willens werden von Argumentationen vom Strawson-Typ nicht überzeugt. Ein möglicher Grund ist, dass die subjektive Erfahrung einer "freien" Entscheidung einem Evidenzerlebnis gleicht, weil "Ich" eine Entscheidung gefällt habe, ohne dabei irgendeine Art von Lenkung erlebt zu haben. Evidenzerlebnisse sind trügerisch (vergl. Kapitel III (2), Abschnitt 2.2)) und können deshalb kein Nachweis für die Freiheit der Entscheidung sein. Einige Philosophen argumentieren, dass Menschen zu rationalen Überlegungen imstande seien, weshalb es doch möglich sei, den Effekt genetischen Erbes und früher Erfahrungen hinter sich zu lassen. Aber die Fähigkeit zu moralischem Handeln kann auch als Resultat sozialen Lernens gedeutet werden, das ein Individuum dazu zwingt, Verantwortung "im eigentlichen Sinne" zu übernehmen, das heißt, die Möglichkeit moralischen Handelns impliziert nicht logisch die Existenz eines freien Willens, so dass die Existenz eines freien Willens allenfalls eine Hypothese ist. Aber Kant hat auch erklärt, der Begriff des freien Willens kein empirischer Begriff sei, sondern ein Vernunftbegriff, der sich nicht als logische Konsequenz aus der Möglichkeit des moralischen Handelns herleiten läßt (vergl. Abschnitt 1.2.2). Kant benötigt Vernunftbegriffe in seiner Philosophie[101], um den gedanklichen Widrigkeiten seiner Antinomie zu entkommen (verg. Die Punkte 1, 2 und 3, Seite 27).

Empirisch arbeitende Psychologen würden beginnen, über einen Versuchplan nachzudenken, der es erlaubt, derlei Hypothesen empirisch zu testen, und tatsächlich liefert Google Scholar auf die Anfrage "development of moral judgment" eine Fülle von Arbeiten zu diesem Thema. Philosophische Mutmaßungen über die Möglichkeiten moralischer Entwicklungen bieten allerdings wegen ihrer Abstraktheit kaum mehr als Allgemeinplätze. Darüber hinaus enthalten Arbeiten nach Art Habermas', Nida-Rümelins etc so gut wie keine Analysen, die auf die These, der Wille sei frei, führen, sie gehen im Gegenteil davon aus, dass der Wille frei ist. Begriffliche Explikationen u.a. des Konzepts 'freier Wille' sind oft hermeneutische Explorationen, die am Ende auf apodiktische Aussagen hinauslaufen, man denke an Habermas' Aussage[102], in der er die Alltagsspsychologie zur

[101] Kritik der reinen Vernunft

[102] "Die objektivierende Sprache der Neurobiologie mutet dem "Gehirn" die grammatische Rolle zu, die bisher das "Ich" gespielt hat, aber sie findet damit an die Sprache der Alltagspsychologie keinen Anschluss. Die Provokation, die darin besteht, dass "das Gehirn" statt meiner "selbst" denken und handeln soll, ist gewiss nur eine grammatische Tatsache; aber so schirmt sich die Lebenswelt gegen kognitive Dissonanzen erfolgreich ab. – Natürlich wäre das nicht die erste naturwissenschaftliche Theorie, die auf diese Weise am Commonsense abprallt." Habermas (2004), p. 872

Norm erhebt, an deren Sprache "die Neurobiologie keinen Anschluß" finde und er feststellt, dass "naturwissenschaftliche Theorien oft am Commonsense abprallen". Das tun sie gelegentlich in der Tat, aber sind Commonsense-Theorien deshalb auch wahr? Die historisch gut belegte Möglichkeit, dass der von Habermas beschworene Commonsense ebensogut an wissenschaftlichen Theorien abprallen könnte, hat er in seinem Aufsatz nicht in Betracht gezogen. Dass die Sonne im Osten auf- und im Westen untergeht hat zu der Commonsense-Theorie geführt, dass die Sonne sich um die Erde dreht, – sollte man deswegen Kopernikus, Galilei und andere vergessen? In der Geschichte der Philosophie des freien Willen ist "jeder Stein schon einmal gewendet" worden, wie Spohn (2008) treffend anmerkt; Habermas' Commonsense-Theorie des freien Willens ist wohl einer dieser Steine. Spohn erklärt knapp: "Jedermann bei Verstand glaubt an den freien Willen. Was immer wir genau darunter verstehen, er ist jedenfalls etwas, was wir haben." Hat Spohn recht, so war Einstein – und mit ihm alle anderen, die von der Hypothese der Existenz eines freien Willens nicht überzeugt sind – nicht bei Verstand, als er sich über den freien Willen äußerte (Zitat 63, Seite 86). Spohn vetritt eine Theorie des Willens, die nahe der von Hume und Schlick ist, wobei Schlick (wie Einstein und andere) zu dem Schluß kommt, dass das Problem des freien Willens eher ein Scheinproblem ist. So weit scheint Spohn nicht gehen zu wollen.

Die Definitionen von Anhängern der These des *unfreien* Willens gehen davon aus, dass Deliberationen unter anderem von nicht bewußten Aktivitäten des Gehirns beeinflusst werden können. Dieser Ansatz hat den Vorteil, mit empirischen Daten kompatibel zu sein. Auf Introspektionen beruhende Überlegungen zum freien Willen mögen als subjektiv wahr erscheinen, sind aber nicht notwendig valide (Nisbet & Wilson(1977)). Freie Entscheidungen gelten als bewußte Entscheidungen, aber wie das Gehirn es schafft, ein Ich-Bewußtsein zu erzeugen ist nicht bekannt. Noë, Dretske (vergl. Zitat 33, Seite 48), Habermas, Nida-Rümelin und andere haben ihre Aussage über die Außerhirnlichkeit des entscheidenden Ichs weder analytisch noch empirisch hergeleitet, sondern als Offenbarung präsentiert, die Habermassche Trennung von "Gehirn" und "Ich" soll wohl bedeuten, dass das Ich nicht vom Gehirn erzeugt wird. Behauptungen wie diese implizieren keinen Gewinn an Kenntnis oder Einsicht, sondern provozieren entweder irrationale Begeisterung über diese Absage an den Materialismus oder die verdutzte Frage, wo sich das Ich denn sonst konstituieren soll. Habermas bezieht sich in seiner Arbeit wiederholt auf den von Ludwig Wittgenstein eingeführten Begriff des Sprachspiels, um die Trennung von "Ich" und Gehirn zu motivieren, worauf im folgenden Unterabschnitt noch ausführlicher eingegangen wird; Habermas ist entweder eine Fehlinterpretation der wittgensteinschen Überlegungen unterlaufen, oder er beachtet sie einfach nicht weiter. Das Gehirn erzeugt Gedanken, Erinnerungen, Wahrnehmungen, und Gedanken über Gedanken. Gedankliche Operationen sind oft logisch, und sehr oft nicht logisch – pauschale Aussagen wie

> "In dieser Dimension vollzieht sich die rationale Motivation von Überzeugungen und Handlungen nach logischen, sprachlichen und pragmatischen Regeln, die sich nicht auf Naturgesetze reduzieren lassen." (s.

sind glatt und eingängig formulierte, aber irreführende Behauptungen: wenn sich die genannten pragmatischen Regeln "nicht auf Naturgesetze reduzieren lassen", so kann das einfach daran liegen, dass sie sich aus einer fehlerhaften Alltagshermeneutik ergeben, – man denke an Peter C. Wasons (1966) "Vier-Karten-Problem": die meisten Menschen können die elementare Schlußregel: "Wenn die Aussage $p \to q$ gilt, wobei p und q Aussagen sind, so folgt $\neg q \to \neg p$: wenn q nicht gilt, so folgt nicht-p (Beispiel: wenn es regnet (p), so ist die Strasse naß (q). Die Strasse ist aber trocken ($\neg q$), also regnet es nicht ($\neg p$)"). Überraschend viele Menschen versagen bei der korrekten Anwendung dieser Regel. Daniel Kahnemans (2011) hat mit seiner Unterscheidung zwischen schnellem und langsamem Denken[103] eine Erklärung für dieses Phänomen angeboten, d.h. er hat eine Systematik, d.h.eine Naturgesetzlichkeit, des fehlerhaften Denkens beschrieben. Habermas' Begriff des Naturgesetzes entspricht dem Standard des 19-ten Jahrhunderts, und er verliert kein Wort zur Frage, warum sich die genannten Regeln *nicht* auf Naturgesetze reduzieren lassen. Aus der Thermodynamik ist bekannt, dass sich die Trajektorien einzelner Gasmoleküle nicht berechnen lassen, was bekanntlich nicht bedeutet, dass die Thermodynamik nicht naturwissenschaftlich betrieben werden kann, wie ein Blick in ein Lehrbuch der Thermodynamik lehrt. Earman (1986, p. 243) verweist darauf, dass der Determinismus sehr nützlich als leitendes Prinzip gewesen ist, dass es aber "nicht den geringsten Grund gibt, anzunehmen, dass die Wissenschaft der Physik ohne die Annahme des Determinismus unmöglich sei"[104].

1.3.2 Der mereologische Fehlschluss

Der Homunkulus-Fehlschluß: Hermeneutik ist ein weites Feld, sie wird auch bei der Interpretation empirischer Befunde angewendet und findet sich dementsprechend in Äußerungen über neuronale Prozesse. Es ist vielleicht nützlich, an dieser Stelle einige Betrachtungen zur Rolle der Sprache in der Neuro- und Kognitionswissenschaft einzuflechten. Gerade bei Wissenschaftlern und Philosophen, die keine dualistische Position beziehen, finden sich Aussagen wie "das Gehirn sieht", oder "das Gehirn denkt", etc. Bereits A. Kenny (1971) hat darauf hingewiesen, dass derlei Aussagen einen Dualismus implizieren, er spricht von einem *Homunculus-Fehlschluss*, vor dem schon Descartes in seiner Schrift *Dioptrics* gewarnt habe: Descartes habe bereits darauf hingewiesen, dass die Linse des Auges ein Abbild der gerade existierenden visuellen Szene auf der Retina projiziere, von wo aus sie in das Innere des Kopfes übertragen wird. Aber es wäre ein Fehler, anzunehmen, dass dieses 'Abbild' ein Abbild im dem Sinne ist, wie ein Foto der Szene ein Abbild ist, das Abbild ist ein komplexes Muster neuronaler Aktivität,

[103] Kahneman, D.: *Thinking Fast and Slow* (2011)

[104] ... that the science of physics would be impossible without determinism, and we know from the many examples studied that denying determinism does not push us over the edge of the law and into the abyss of the utterly chaotic and non-lawful", p. 243

das auf wundersame Weise ein Abbild im Sinne eines Fotos im Bewußtsein erscheinen läßt[105]. Diese Sprechweise ist äquivalent der impliziten Annahme eines Homunculus in unserem Kopf, der die Szene "sieht". Eine solche Annahme ist absurd, da sie auf die Notwendigkeit führt, erklären zu müssen, wie der Homunkulus die Szene "sieht", – offenbar wird man in einen infiniten Regress geführt, denn der Homunkulus hat wieder einen Homunkulus, der sieht, etc. Kenny verweist dann auch gleich zu Beginn seines Aufsatzes auf eine Beobachtung Ludwig Wittgensteins in seinen *Philosophischen Untersuchungen:*

> Only of a human being and what behaves like a human being can one say: it has sensations; it sees; is blind; hears; is deaf; is conscious or unconscious. *Philosophical Investigations,* § 285. Z 65

Kenny führt aus, dass Wissenschaftler und Philosophen, die auf Wittgensteins Aussage hingewiesen werden, üblicherweise ablehnend reagieren: es sei kein Problem, man wisse ja, was gemeint sei. Aber Kenny besteht darauf, dass die Mißachtung der Wittgensteinschen These durchaus zu Denkfehlern führen könne. Das Homunkulus-Argument gleicht dem von Daniell Dennett in Dennett (1991), Kap. 5, eingeführten Begriff des *Cartesianischen Theaters*: bekanntlich hat Descartes mit seiner *res cogitans* eine immaterielle Basis für alle mentale Aktivität postuliert, die über die Zirbeldrüse (lat. glandula pinealis) mit dem Gehirn verbunden wird. Dennett (p. 107) spricht dementsprechend von einem *Cartesianischen Materialismus*, mit dem ein Punkt angenommen wird, in dem "alles zusammenkommt" (es sei an Planks Punkt erinnert, vergl. Zitat 56), Seite 76) – die Zirbeldrüse ist nach Descartes ein solcher Punkt, in dem das Cartesianische Theater stattfinde. Hierauf soll hier schon aus Platzgründen nicht weiter eingegangen werden, statt dessen soll kurz auf die Arbeit des Neurowissenschaftlers M.R. Bennett und des Philosophen P.M.S. Hacker (2003) (B & H im Folgenden), einem monumentalen Buch *Philosophical Foundations of Neuroscience* fokussiert werden.

B & Hs zentrales Anliegen sind Formulierungsfehler der besprochenen Art: die Aussage "Mein Gehirn fällt die Entscheidung *E*" ist ein Homunculus-Fehlschluss, – auch, wenn sie nicht als solche intendiert wird. B& H führen dafür einen speziellen Namen ein: es handele sich um einen *mereologischen Fehlschluss*. Allgemein wird unter einem mereologischen Fehlschluss ein Kategorienfehler[106] verstanden, der dadurch entsteht, dass einem Teil ein Prädikat zugeschrieben wird, dass nur dem Ganzen zugeordnet werden kann, im vorliegenden Fall wird dem Gehirn ein Prädikat zugesprochen, das nur einer Person zugesprochen wird, vergl. das Wittgensteinzitat 65, Seite 92. B & H wollen keinesfalls einem Dualismus das Wort reden, – im Gegenteil, sie wollen nachweisen, dass ein zu salopper Sprachgebrauch einen Dualismus suggeriert, wo keiner gemeint ist.

B & H elaborieren das Konzept des mereologischen Fehlschlusses ausgiebig

[105] "We must not think that it is by means of this resemblance that the picture makes us aware of the objects – as though we had another pair of eyes to see it, inside the brain", in Kennys Übersetzung des Descartesschen Textes.

[106] Der Begriff des Kategorienfehlers wurde von Gilbert Ryle (1949) eingeführt.

in Kapitel 3 ihres Buches, wobei sie renommierte Neurowissenschaftler zitieren, die keinesfalls irgendeine Art von Dualismus vertreten wollen. Zwei Beispiele aus diesem Kapitel genügen für den hier verfolgten Zweck:
(i) **Sir Colin Blakemore**[107]:

> "We seem driven to say that such neurons [as respond in a highly specific manner to , e.g., line orientation] have knowledge. They have intelligence, for they are able to estimate the probability of outside events – events that are important to the animal in question. And the brain gains its knowledge by a process analogous to the inductive reasoning of the classical scientific method. Neurons present arguments to the brains based on the specific features that they detect, arguments on which the brain constructs its hypothesis of perception."
> Aus: C. Blakemore, The Mechanics of the Mind. Cambridge 1977.
> Zitiert nach H & B, p. 69, Z 66

Es ist in der Tat schwer zu glauben, dass ein solcher Text von einem so kompetenten und renommierten Wissenschaftler wie C. Blakemore geschrieben wurde, wenn man annimmt, dass er seine Aussagen wörtlich gemeint hat. Welche anatomischen und physiologischen Kenntnisse implizieren, dass ein Neuron eine Wahrscheinlichkeit schätzen kann, und was soll es bedeuten, dass ein Neuron über Intelligenz verfügt? Aufschluß gibt vielleicht der Satz, demzufolge das Gehirn Wissen gewinnt, die *analog* dem induktiven Vorgehen der klassischen wissenschaftlichen Methode sei; auch der folgende Satz über Neurone, die Argumente repräsentieren muß als Aussage über eine Analogie verstanden werden, denn es ist ja überhaupt nicht klar, warum einzelne Neurone Wissen *im kognitiven Sinn* speichern können; nichts, was wir über Aufbau und Funktionsweise von Neuronen wissen, kann *direkt* mit einer kognitiven Funktion in Verbindung gebracht werden.

(ii) **David Marr**(1945 – 1980)[108]: David Marr gehörte zu den Stars unter den Wissenschaftlern, die sich in der Erforschung der neuronalen Grundlagen des Sehens engagierten. Wie bei Colin Blakemore kann man fragen, was diese Forschungen mit der Frage nach dem freien Willen zu tun haben, und die Antwort ist, dass das Sehen in sehr enger Beziehung zur Frage nach dem Bewußtsein steht, – nach Vorstellung mancher Philosophen und Wissenschaftler setzt der freie Willen Bewußtsein voraus. Bei der Untersuchung der visuellen Wahrnehmung drängen die Erfolge bei der Decodierung der neuronalen Prozesse im visuellen Cortex ergibt sich gewissermaßen automatisch die Frage, wo denn im Gehirn das eigentliche Sehen beginnt, also das bewußte Wahrnehmen der jeweiligen visuellen Szene. Über den Prozess der Umwandlung neuronaler Aktivierung in die eigentliche Wahrnehmung ist so gut wie nichts bekannt:

[107] 1944 – 2022, herausragender britischer Neuro- und Kognitionswissenschaftler, wurde 2014 für seine Verdienste von der englischen Königin geadelt ("knighted"

[108] britischer Psychologe und Informatiker, gilt als Begründer der Neuroinformatik

> From the pattern of stimulation on the retinas we perceive the world of objects; and this is nothing short of a miracle. (Gregory (1966), p. 7, Z 67

Kürzer kann man es kaum auf den Punkt bringen.

David Marr's großes Werk *Vision, a Computational Investigation into the Human Representation and Processing of Visual Information* heigt bereits im Titel, was B & H als Ausdruck fehlerhaften Denkens kritisieren. Marrs Darstellung beginnt mit der Darstellung des *primal sketch*, also einer ersten Skizze einer visuellen Szene auf der Basis elementarer neuronaler "Detektoren" für Merkmale wie Konturen, etc. die eigentiche Wahrnehmung ist das Resultat von "information processing", also Informationsverarbeitung, die zu einer symbolischen Beschreibung der Szene führen. B & H zitieren Marr: "vision is the *process* of discovering from images what is present in the world and where it is." Es handele sich, so B & H, um einen Prozess der Transformation der in der Szene vorhandenen INformation in eine explzite Beschreibung von dem, was gesehen wird (B&H, p. 143).

B & H kritisieren diesen Text (und andere, ähnlich abgefasste Texte) nicht zu Unrecht: etwas zu sehen bedeutet nicht, irgendetwas in einer Szene (Bild = image) zu entdecken, denn man könne nicht etwas entdecken, dass nicht wahrgenommen wird, was auch immer die Lichtverteilung der Szene darstellt. Diese Lichtverteilung sei kein Datum (im wörtlichen Sinne etwas "Gegebenes") oder ein "Stück Information", aus dem etwas hergeleitet oder erschlossen werden kann, denn das Gehirn habe weder Information noch könne es Information verstehen. Nicht alles Sehen sei Entdecken: wenn man etwas geschrieben habe und lese es dann, so *entdecke* man vielleicht Fehler, aber man 'entdecke' nicht, was man geschrieben habe. Beim Sehen sind zwar neuronale Prozesse involviert, so sei doch das Sehen selbst kein Prozess.

Tatsächlich ist es so: wie sich experimentell demonstrieren läßt, wirken sich schon sehr kleine Veränderungen eines Stimulusmusters auf die Wahrnehmung aus, und diese Veränderungen können mittels mathematischer Modelle des visuellen Cortex ziemlich genau vorhergesagt werden. Diese Modelle sind aber Beschreibungen der Neuronen samt ihrer rezeptiven Felder des visuellen Cortex und der Vernetzungen der Neuronen. Es ist klar, dass die Wahrnehmung von der Aktivierung dieser Neuronenstrukturen abhängt. Immer, wenn von der Wahrnehmung und anderen kognitiven Aktiväten gesagt wird, sie seien Hirnprozesse oder das Gehirn erzeuge sie, wird, sofern man neurowissenschaftlich orientiert ist, implizit auf derartige Aktivierungen von Neuronen Bezug genommen. Es ist aber sehr wahrscheinlich, dass es nicht die Aktivierungsmuster von Neuronenpopulationen selbst sind, die Bewußtsein erzeugen, sondern andere Prozesse, die wir noch nicht kennen, die aber von den Neuronenpopulationen angetrieben werden. So gesehen macht es keinen Sinn, zu sagen, dass Gehirn nimmt wahr, oder denkt, entscheidet, etc.

Oder doch? Schließlich wird man davon ausgehen müssen, dass das Gehirn auch diese Prozesse erzeugt. Die Alternative wäre, eben einen Dualismus zu pos-

tulieren.

B & Hs Argumente beziehen sich letztlich auf die Rolle der Sprache, und hier insbesondere auf Wittgensteins Begriff des Sprachspiels, wobei sie keineswegs irgendeinem Dualismus das Wort reden. Wittgenstein hat sich nicht direkt über die Möglichkeit eines freien Willens geäußert, sondern allenfalls indirekt, nämlich wenn er von Wllen etc sprach. Auf Wittgensteins Ansichten soll hier nicht weiter eingegangen werden – sie würden ein eigenes Kapitel rechtfertigen, das aber den Rahmen dieses Textes sprengen würde. Burgos & Donahoe (2006) haben einen ausführliche Kritik von B & Hs Kritik präsentiert.

1.3.3 Entscheidungsmodelle

In experimentalpsychologischen, insbesondere mathematisch formulierten Modellen über das Entscheidungsverhalten (etwa Busemeyer & Townsends (1993) Ansatz) erlauben die empirische Überprüfung von speziellen Hypothesen[109]. Die experimentalpsychologische Forschung zu Denk- und Entscheidungsprozessen zeigt unter anderem, dass kognitive Prozesse störanfällig sind (eine allgemeine Darstellung findet man in Kahneman (2011) und Kahneman, Sibony, Sunstein (2021), Tversky & Kahneman (1974) liefern eine knappe Zusammenfassung experimenteller Befunde zu kognitiven Täuschungen insbesondere auch im Entscheidungs-, d.h. im Deliberationsverhalten, denn Entscheidungen sind ja das Resultat von Deliberationen. Diese Befunde verweisen auf die letztlich neuronale Bedingtheit der Entscheidungsfindung: was subjektiv rational (im Sinne von Habermas) erscheint, kann in der Realität äußerst irrational sein.

Das Konzept des freien Willens wird im Allgemeinen mit der Möglichkeit moralischen Handelns gekoppelt: es wird argumentiert, dass die Möglichkeit moralisch verantwortlichen Handelns die Existenz des freien Willens voraussetzt bzw. diesen impliziert. Da die (Neuro-)Biologie eine Naturwissenschaft ist und diese wiederum auf der Annahme des Kausalitätsprinzips basiere, müssen Anhänger dieser Argumentation erklären, warum der Wille frei sein kann, was inbesondere viele Kompatibilisten neue Varianten des Dualismus annehmen lässt, z.B. Perspektiven- oder Commonsensedualismen (in den Abschnitten über Nida-Rümelin und Habermas in Abschnitt 1.2.5). Im Rahmen einer *experimentellen Philosophie* sind Umfragen durchgeführt worden, in denen die die intuitiven Vorstellungen der Bevölkerung über den Freien Willen und die moralische Verantwortung untersucht werden, und tatsächlich scheint kulturübergreifend eine dualistische

[109] "Decision field theory provides for a mathematical foundation leading to a dynamic, stochastic theory of decision behavior in an uncertain environment. This theory is used to explain (a) violations of stochastic dominance, (b) violations of strong stochastic transitivity, (c) violations of independence between alternatives, (d) serial position effects on preference, (e) speed-accuracy trade-off effects in decision making, (f) the inverse relation between choice probability and decision time, (g) changes in the direction of preference under time pressure, (h) slower decision times for avoidance as compared with approach conflicts, and (i) preference reversals between choice and selling price measures of preference. The proposed theory is compared with 4 other theories of decision making under uncertainty." Abstract zu Busemeyer & Townsend (1993)

Grundeinstellung den Glauben an einen freien Willen besser vorherzusagen als die Überzeugung, die Welt sei durchgängig deterministisch organisiert (Wisniewski et al. (2019)). Philosophen argumentieren oft, ihre philosophisch begründeten Ansichten entsprächen den Mehrheitsansichten in der allgemeinen Bevölkerung (Nahmias et al. (2005)). Dementsprechend wurden experimentalpsychologische Studien begrüßt, die nahelegen, dass bestätigende Aussagen über die Existenz des freien Willen und die damit einhergehende moralische Verantwortung einen positiven Effekt auf das Verhalten haben (konkret: die Ausssge, es gäbe einen freien Willen, haben die Schummelrate bei kognitiven Tests reduziert, während Aussagen der Art, der Wille sei determiniert, den gegenteiligen Effekt erzeugten (Vohs et al. (2009), Baumeister et al. (2009), Baumeister et al. (2011)). Wie so oft lassen kritische Analysen nicht lange auf sich warten, etwa Nahmias (2011), oder Iso-Ahola (2017). Nun geht es aber bei der philosophischen Diskussion der Frage nach dem freien Willen nicht darum, ob Befürworter der These des freien Willens die moralisch höherwertigen Aufsätze schreiben, sondern um die Frage, ob sie mit ihren Argumenten recht haben. Schon bei der Betrachtung der Argumente Nida-Rümelins wurde nach dem epistemischen Nutzen philosophischer Aussagen gefragt, die sowohl zur der Stützung der These der Existenz eines freien Willens, wie auch zur Stützung der Gegenthese verwendet werden können.

1.3.4 Philosophie versus Empirie

So kommt man zu der Frage, ob die Hypothese der Existenz eines freien Willens überhaupt allein mit philosophischen Methoden entschieden werden kann. Roth (2004) hat den philosophischen, die Empirie nicht in Rechnung stellenden Ansatz scharf kritisiert, u.a. weil die Philosophen empirische Ergebnisse zu diesem Thema entweder gar nicht erst zur Kenntnis nehmen, oder sie als irrelevant desavouieren. Insbesondere J. Habermas scheint sich von seiner schon in *Zur Logik der Sozialwissenschaften* (1970) und *Erkenntnis und Interesse* (1973) elaborierten Kritik an der Empirie und am "szientistischen" Denken nicht distanziert zu haben, und Alva Noës Glauben an den Primat der Philosophie hat die Grenze zur Bizarrerie deutlich überschritten.

Es ist nicht klar, was mit dem Begriff der philosophischen Methode überhaupt gemeint ist. Man könnte an den Philosophen Hans Georg Gadamer (1900 - 2002) erinnern, der in seinem Werk "Wahrheit und Methode" die Hermeneutik als eine allgemeine Methode der Erkenntnisgewinnung vorstellte. Gadamers Ansatz ist vielleicht zu allgemein. Die Behauptung etwa Habermas', Nida-Rümelins und Noës, das Gehirn habe nichts mit der Entscheidungsfindung zu tun, folgt nicht aus einer überprüfbaren Annahme, sie schwebt frei im Raum der assoziativen Möglichkeiten. Dies gilt auch für Max Plancks (Zitat 56, Seite 76) "Punkt in der unmessbaren Welt von Geist und Materie ... der in sich eine ganze Welt enthält und unser emotionales Leben, unseren Willen und unsere Gedanken enhält" und der nach Plancks damaliger Auffassung der Kausalität entzogen ist.

So hat man zwei Welten, die eine von Denkern wie G. Strawson, S. Harris,

R. Sapolsky, G. Roth , W. Singer etc, und die andere von Denkern wie Kant, Nida-Rümelin , Habermas, Noë , wobei man allerdings Kant und die letzten drei nicht in einer Reihe nennen sollte: Kant hat argumentiert, er hat versucht, ein konsistentes System von Aussagen zu generieren, während die übrigen Drei enfahc Behauptungen aneinandergereiht haben. Keiner der Denker der einen Welt kann einen Denker aus der anderen Welt überzeugen, weil sie, wie es scheint, von verschiedenen Interpretationen des Begriffs 'freier Wille' ausgehen. Skeptiker des freien Willens, d.h. Denker, die der These der Existenz eines freien Willens skeptisch gegenüber stehen, gehen anscheinend von der eingangs gegebenen, Standarddefinition von 'frei' aus, derzufolge der Wille 'unfrei' ist, wenn es Randbedingungen gibt, die letzlich die Entscheidungen bestimmen. Zu den Randbedingungen gehören auch zufällige Effekte, die in die jeweiligen Deliberationen eingehen, so dass unklar wird, was 'frei' überhaupt bedeuten soll. Die Argumente der Anhänger der These erscheinen eher als Rechtfertigung einer bereits existierenden Überzeugung.

Die Situation erinnert an das Postulat der klassischen Methodologie des rationalen Denkens[110], demzufolge nach hinreichenden Begründungen für eine Aussage gesucht werden soll, denn es gebe nur eine Wahrheit (Postulat des theoretischen Monismus). Angenommen, man möchte zeigen, dass die Aussage q: "Der Wille ist frei" wahr ist. Dazu benötigt man bereits als wahr erkannte Aussagen p, aus denen q folgt, $p \rightarrow q$. Die Aussagen p folgen wiederum aus anderen, bereits als wahr erkannten Aussagen, etc. Dann ergibt sich das von Albert so genannte Münchhausen-Trilemma: 1. Man gerät in einen infiniten Regress, oder 2. in einen logischen Zirkel in der Deduktion, d.h man greift auf Aussagen zurück, die schon begründungsbedürftig waren, oder 3. man muß den Begründungsversuch einfach abbrechen, dh man kann über einen bestimmten Punkt hinaus das Verfahren nicht weiterführen, obwohl noch keine letztliche Begründung vorliegt. Das Trilemma ist nicht ganz unproblematisch, weil es ja ebenfalls begründet werden muß und das Trilemmaprinzip damit auf sich selbst anwendbar ist. Nun können alle hier behandelten Philosophen, einschließlich Max Plancks, sagen, ihre Argumente seien "evident". Insbesondere Noë scheint so zu denken: er präsentiert sein Extracranialargument, als sei es für jeden, der seinen Ausführungen gefolgt ist, evident wahr, die eben genannte Folge von Argumenten, die auf die Aussage 54, Seite 74, führt, würde gar nicht benötigt. Aber auch der Begriff der Evidenz ist sorgfältig der philosophischen Analyse unterzogen worden. Der Evidenzbegriffs wurde von Karl R. Popper und dann insbesondere von Wolfgang Stegmüller (1954) kritisiert, der zu dem Schluß kam[111]

> Das Evidenzproblem ist absolut unlösbar, die Frage, ob es Einsicht gibt oder nicht, ist absolut unentscheidbar. Z 68

In diesem Zusammenhang sei auch daran erinnert, dass Philosophen – insbesondere Wissenschaftstheoretiker – die Unmöglichkeit des *experimentum crucis*

[110] Wissenschaftstheorie III(2), Abschnitt 2.3
[111] Wissenschaftstheorie III (2), Abschnitt 2.2

nachgewiesen haben (vergl. Wiss'theorie III(2), Abschn. 2.4). Dieses Argument ist sehr allgemein und man könnte deswegen vermuten, dass die Empirie sowieso keinen Sinn macht, – eine Sichtweise, die vor allem immer wieder bei Habermas angedeutet wird, für den die "philosophische Reflexion" die zentrale Methode der Annäherung an die Frage nach dem freien Willen ist. Natürlich kann man der Philosophie ein grundsätzliches Primat zuschreiben, indem man die Entscheidung, sich bestimmten Fragen nicht nur theoretisch, sondern auch empirisch zuzuwenden, als Resultat einer philosophischen Betrachtung ansieht. Leider ist dieses Argument nicht zwingend, es sind ja oft Experimente, deren Ergebnisse zu der Überzeugung geführt haben, dass eine systematische Empirie mehr Informationen über die Welt liefert als theoretische Überlegungen. Aber diese Diskussion muß an dieser Stelle nicht wiederholt werden, sie ist in Wissenschaftstheorie II und III einigermaßen ausführlich dargestellt worden. Schon in der Einleitung Abschnitt 1.1 wurde angemerkt, dass es müßig sein könnte, einen Streit um die Existenz oder Nicht-Existenz des freien Willens zu führen, einfach, weil zu jeder Aussage, die zur Begründung der These der Existenz eines freien Willens gemacht wird, eine Gegenthese formuliert werden kann, der zufolge es keinen freien Willen gibt.

Entgegengesetzte Ansichten in den Naturwissenschaften: Die, was den freien Willen angeht, gegensätzlichen Auffassungen der Physiker G. t'Hooft und A. Zeilinger sind bereits in Abschnitt 1.1.2 zitiert worden. Gegensätzliche Aussagen gleicher Art findet man aber auch bei Neurowissenchaftlern:

> Peter Tse (1962 –): "Just as Watson and Crick showed that there is a physical basis, a physical mechanism—namely DNA—that could account for evolution," Tse says, "I am arguing that rapid synaptic reweighting is the physical mechanism that gives humans the power to exercise free will." Tse (2013), Z 69

Der ebenfalls renommierte Neurowissenschaftler Robert Sapolsky (1957 –) kommt in Sapolsky (2023) zu dem entgegengesetzten Schluß:

> "The world is really screwed up and made much, much more unfair by the fact that we reward people and punish people for things they have no control over," Sapolsky said. "We've got no free will. Stop attributing stuff to us that isn't there." (2023)[112], Z 70

So stellt sich die Frage, ob es eine empirische Möglichkeit gibt, die Hypothese der Existenz eines freien Willens zu überprüfen. Dazu ergeben sich die folgenden Bemerkungen.

Verifikation, Falsifikation, und der freie Wille In Wissenschafstheorie III (2) ist die Möglichkeit, Hypothesen entweder zu verifizieren oder zu falsifizieren bereits diskutiert worden. Bei den folgenden Betrachtungen geht es darum, zu zeigen, dass die Hypothese des freien Willens grundsätzlich weder verifiziert noch falsifiziert werden kann; so ist ja schon von I. Kant (1788) argumentiert worden,

[112] https://phys.org/news/2023-10-scientist-decades-dont-free.html

vergl. Abschnitt 1.2.2; nach Kant kann der Begriff der Freiheit kein empirischer Begriff sein. Im Folgenden geht es um den freien Willen an sich, also nicht um den freien Willen als Konsequenz noch als Bedingung für die Fähigkeit zu moralischem Handeln. Offenbar ist die Möglichkeit einer Verifikation bzw. Falsifikation an gewisse Vorbedingungen geknüpft, nämlich dass irgendwelche Indikatoren I für eine Entscheidung E existieren, – oder nicht existieren. Man hat also das Schema

- *Verifikation:* Zeige, dass dem Indikator I von E *keine* kausalen Effekte K_I vorausgehen,
- *Falsifikation:* Zeige, dass I kausal erzeugt wird, d.h. dass kausalen Effekte K_I existieren.

Der Schluß, dass die Hypothese H_{fw} des freien Willens nicht testbar ist, ergibt sich, wenn gezeigt werden kann, dass weder die Verifikation – der Nachweis der Existenz kausaler Effekte – noch die Falsifikation – der Nachweis der Nichtexistenz kausaler Effekte – möglich ist. Grundsätzliches zur Problematik der Verifikation findet man schon bei David Hume und sehr ausführlich bei Karl R. Popper, und ebenso Grundsätzliches zur Problematik der Falsifikation ist bereits in der Kritik an Poppers Philosophie der Falsifikation formuliert worden. Hier soll noch auf weitere Aspekte der Verifikation-Falsifikation-Problematik hingewiesen werden.

Northcotts (2018) plausibler Argumentation zufolge ist der Nachweis der Nichtexistenz von Prozessen K_I kaum möglich, denn ein Nachweis der Nichtexistenz von Prozessen K_I bedeutet ja, dass das Zustandekommen von E nicht erklärbar ist, und für sich genommen beweise die Nichterklärbarkeit noch nichts, insbesondere beweise sie nicht, dass irgendwelche K_I-Prozesse nicht in der Zukunft gefunden werden können, – es sei denn, dass man *beweisen* kann, dass K_I-Prozesse nicht existieren können. Northcott zitiert dazu den britischen Physiker J.S. Bell (1964), der eine Ungleichung – die nach ihm benannte *Bellsche Ungleichung* – aufgestellt hat, die erfüllt sein muß, wenn Variablen existieren, die bestimmte Phänomene kausal erklären. Diese Variablen müssen nicht explizit bestimmbar sein, weshalb von *versteckten* Variablen (hidden variables) die Rede ist. In der klassischen Physik wird die Ungleichung stets erfüllt, nicht aber in der Quantenphysik. Die Ungleichung wird beim Phänomen der Quantenverschränkung verletzt, was als Nachweis der Nichtexistenz von versteckten, Kausalität reflektierenden Variablen gedeutet wird. Für die Nichtexistenz von K_I-Prozessen liegt aber kein derartiger Beweis vor, was als Hinweis auf die Nichtverifizierbarkeit von H_{fw} gedeutet werden kann. Es liegen auch keine Daten vor, die einen empirischen Hinweis auf Nichtkausalität liefern, wie man sie für die Widerlegung der Behauptung, die MMR-Impfungen würden Autismus erzeugen, hat (MMR: Masern, Mumps und Röteln)[113]. Northcott verweist auf die Vernetzheit

[113]Die methodischen Fehler, die zu dieser Behauptung führten, werden in https://www.chop.edu/centers-programs/vaccine-education-center/vaccines-and-other-conditions/vaccines-autism erläutert.

der Hirnprozesse, die den Nachweis der *Nichtexistenz* von Prozessen, die kausal zu Entscheidungen führen, schwierig, wenn nicht unmöglich machen.

Um H_{fw} zu falsifizieren muß man nachweisen, dass man stets Prozesse finden kann, die eine Indikatoraktivität I bewirken. Als ein solcher Nachweis wird oft das Bereitschaftspotential B_p genannt, das der Entscheidung E, die Hand zu heben, vorangeht. Nun wird aber B_p nicht in jedem Versuchsdurchgang beobachtet, d.h. B_p ist ein probabilistisches Phänomen. Schurger et al. (2016) haben dementsprechend argumentiert, dass ein B_p eher eine Art von Hintergrundrauschen (background noise) repräsentiere. Streng genommen müsste eine inkompatibilistische Position eingenommen werden: H_{fw} gilt als widerlegt, wenn für alles Geschehen der Determinismus gilt, d.h wenn nichts ohne Ursache geschieht. Aber hier greift die Argumentation van Kampens (Kap. IV, Abschnitt 2.4, p. 44), derzufolge ein solcher Determinimus weder bewiesen noch widerlegt werden kann. Deshalb kann man vermuten, dass H_{fw} prinzipiell nicht falsifiziert werden kann.

Die hier vorgetragene Argumentation zur H_{fw} ist sicherlich eher eine Skizze, Northcott hat sie weiter elaboriert, worauf hier nicht weiter eingegangen werden muß; es sei aber noch einmal auf Wissenschaftstheorie III (2), Abschnitt 2.4, verwiesen, wo die generelle Unterbestimmtheit empirischer Befunde (Duhem-Quine-These) diskutiert wird.

So ist es interessanter, zu untersuchen, *wie* Menschen jeweils zu ihren Entscheidungen gelangen. Die Behauptung einiger Philosophen, dass die (Natur-)Wissenschaften Fragen wie die nach dem freien Willen gar nicht beantworten können, weshalb ein nicht-"szientistischer" bzw. ein "humanistischer" Ansatz die geeignete Perspektive sei, um die Möglichkeiten eines freien Willens zu erforschen, unterliegt ebenfalls der oben erwähnten Trilemma-Problematik, die den Absolutheitsanspruch philosophischer Ansätze relativiert. Man könnte dem philosophischen Argument der Nichtwiderlegbarkeit auch entgegenhalten, es sei eine *petitio principii*[114], also eine "Vorwegnahme des Beweisgrundes"[115]: der Begriff des freien Willens wird so definiert, dass er als evident wahr, aber nicht weiter beweis- oder widerlegbar erscheint.

Weitere biologische Befunde Ein beliebtes philosophisches Argument zur Stützung von H_{fw} beschwört die Emergenz kognitiver Aktivitäten auf der Basis von neuronaler Aktivität, die Bewußtsein und die Freiheit des Willens ermögliche. Allerdings führt der Hinweis auf Emergenz kaum zu neuen Einsichten, so lange man sich keine expliziten Vorstellungen davon macht, welcher Art die biologischen Prozesse sind, die zur Emergenz etwa von Bewußtsein führen. In diesem Zusammenhang sei an die Arbeit von Wegner (2003) erinnert. Die von ihm zitierten Arbeiten legen nahe, dass das bewußte Erlebnis, eine Handlung herbei geführt zu haben, eine Täuschung ist (s. die Bemerkungen zum Reduktionismus bei Habermas). Es gibt weitere Ansätze, Prozessen der Emergenz auf die Spur zu

[114] gelegentlich mit "Erschleichung des Beweise" überesetzt

[115] so die Übersetzung ihm Rahmen der Logik: Kodakow, N.I.: Wörterbuch der Logik, Leipzig 1983, p. 373

kommen, etwa die *Dynamic Core* und *Global Workspace* Hypothesen: Edelman, Gally & Baars (2011) liefern einen Überblick. Die genannten Theorien basieren auf dem Ansatz des *Neural Darwinism* (Edelman (1978), 1987), Edelman & Tononi (2007)). Diesen Autoren zufolge besteht das Gehirn aus sehr vielen Gruppen von interagierenden Subpopulationen von Neuronen, deren Struktur – gemeint ist die Vernetzung der Neuronen untereinander – durch Wechselwirkungen mit der Umwelt zum Teil verstärkt, zum Teil geschwächt werden. Wesentlich für diese Entwicklung ist das Re-entry-Prinzip, demzufolge es einen stetigen, reziproken Signalaustausch zwischen den Gruppen gibt. Allgemein wird angenommen, dass das Bewußtsein u.a. das Resultat derartiger *reentranter neuronaler Aktivität* ist; dies ist insbesondere die zentrale These der *Dynamic Core-Hypothese* (Baars (1988)). Für spezifische bewußte Zustände bzw. Ereignisse sind anscheinend Aktivitäten des Thalamus und des cerebralen Kortex zuständig; Edelman, Gally, Baars (p. 4) führen den Begriff der *behavioral trinity* ein: (1) biologische Mechanismen und phänomenale Erfahrung, (2) eine Basis für das "subjektive Selbst", (3) die Beziehung zwischen den Signalen (Inputs) dem Gehirn und dem Körper und dem eigenen Inneren (Gedächtnis), – die verschiedenen Systeme des Gehirns müssen im *dynamischen Kern* (Dynamic Core) koordiniert werden. Im dynamischen Kern werden Inputs aus der sensorischen Wahrnehmung und der aus der eigenen Bewegung mit Gedächtnisinhalten kombiniert und durch Aktivitäten aus subkortikalen Strukturen (Cerebellum, Basalganglien) moduliert.

Alle diese Befunde und Interpretationen von Befunden beruhen im Wesentlichen auf den Ergebnissen von fMRI-Untersuchungen, wobei nicht verschwiegen werden soll, dass es Kritik insbesondere an früheren fMRI-Untersuchungen gibt, bei denen Korrelationen zwischen der Aktivität in bestimmten Voxeln[116] und kognitiver Aktivität berechnet wurden; diese Korrelationen können Methodenartefakte sein (Vul et al. (2009)). Alternativ dazu wurde versucht, Voxel zu bestimmen, bei denen die neuronale Aktivität für gegebene kognitive Aktivitäten besonders deutlich sind, wobei über räumliche Bereiche gemittelt wird. Bei diesem Vorgehen wird aber die *differentielle Aktivierung* in verschiedenen Voxeln nicht berücksichtigt. Es hat sich jedoch gezeigt, dass Kognitionen auf der Basis von Aktivitäten in Mustern von Voxeln klassifiziert werden können. Einen Überblick liefern Norman et al (2006), oder Mahmoudi et al. (2012), die anders als Norman et al. auch eine Einführung in die neueren Methoden der Klassifikation von Mustern wie etwa Support-vector-Machines präsentieren (eine detailliertere Einführung findet man in Bishop (2006)). Douglas et al. (2013) liefern fMRI-Daten zur Entstehung und Wechselwirkung von Emotionen und zur Bedeutung von Ausdrücken wie 'glauben' etc, die nach Philosophenmeinung keine Basis in neuronaler Aktivität haben.

Eine wichtige Frage bezieht sich auf das Erleben von Qualia und dem Selbst. Edelman et al. argumentieren, dass sie sich aus Unterscheidungen ergeben: Hitze ist nicht "grün", "grün" ist nicht "Berührung", etc. Das Erleben einer komplexen

[116]Ein Voxel ist ein 3-dimensionales *Volumenelement* (volume element) im Gehirn, im Unterschied zum *Pixel*, das ein 2-dimensinales Bildelemente (picture element) ist.

Szene ergibt sich demnach aus Unterscheidungen im Dynamischen Kern. Dabei ergibt sich das von Chalmers (1996) so genannte "schwierige Problem" (the hard problem), wissenschaftlich zu erklären, worin bewußtes Erleben besteht bzw. wie es sich entwickelt. Philosophen und Forscher, die sich die chalmersche Sichtweise zu eigen gemacht haben, sind der Auffassung, dass es sich grundsätzlich der wissenschaftlichen Forschung entzieht.

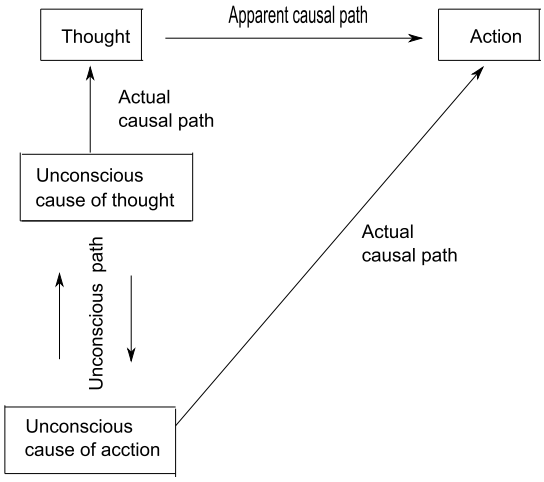
Grundsätzliche Aussagen dieser Art sind allerdings problematisch, da nicht klar ist, aus welchen als "wahr" erkannten Erfahrungen bzw. Annahmen sie folgen. Lange galt es als wahr, dass sich Leben nicht aus der bis dahin bekannten Physik und Chemie erklären lasse. Zur Erklärung des Lebens wurde eine spezielle *vis vitae*, eine Lebenskraft, postuliert. Wie so oft läßt sich diese Lehre bis in die Antike zurückverfolgen; sie hat sich bis ans Ende des 19-ten und den Beginn des 20-ten Jahrhunderts erhalten. Erst dann hatten sich Physik und Chemie so weit entwickelt, dass klar wurde, dass eine solche spezielle Kraft nicht postuliert werden muß. Chalmers hat nicht *bewiesen*, dass das Bewußtsein nicht biologisch erklärt werden kann, er hat diese These nur suggeriert, seine Argumente sind eher intuitiver Art. Man kann sich fragen, warum zum Beispiel Chalmers' Zombie-Argument zwingend sein soll[117], und dieselbe Frage stellt sich für Habermas' und Nida-Rümelins Thesen zur Extracranialität des Geistes.

Nach der Theorie des Neuronalen Darwinismus und des Dynamischen Kerns ergeben sich die Qualia aus Interaktionen höherer Ordnung; die Unterscheidung von 'warm' und 'blau' ist nur möglich, wenn die notwendigen neuronalen Strukturen vorhanden sind. Für Philosophen wie Habermas, Nida-Rümelin und vor allem Alva Noë muß allerdings gar nicht auf die Frage nach deren Existenz und Struktur eingegangen werden, weil ihrer Meinung nach das Gehirn für alles Psychische keine Rolle spielt. Die Hirnforschung findet aber Strukturen, die für diese Wahrnehmungen und Unterscheidungen in Frage kommen und im Rahmen des *Third-Person-View*-Ansatzes untersucht werden können, d.h. es ist möglich, die Korrelation zwischen subjektiver Erfahrung und kausalen neuronalen Prozessen zu entschlüsseln; Edelman et al. sehen darin die Auflösung der Barriere des chalmersschen *hard problems*. Um die Beziehung zwischen Hirnaktivität und Bewußtsein entschlüsseln zu können muß man aber nicht bei den Komplexitäten des menschlichen Bewußtseins beginnen.

Wegners Modell Wegner (2003) ist der Hypothese nachgegangen, dass willentliche Entscheidungen mit der Illusion einhergehen, die Entscheidungen seien frei, weil sie bewußt getroffen werden. Es ist aber denkbar, dass der Gedanke, eine Entscheidung zu treffen bzw. eine Handlung zu tun und die Entscheidung oder Handlung tatsächlich zu treffen ode zu tun eine gemeinsame neuronale Basis ha-

[117]Ein Zombie ist nach Chalmers eine molekulargenaue Kopie eines Menschen, die allerdings ohne das Bewußtsein, ohne phänomenale Erfahrung auskommt. Mit der begrifflichen Konstruktion eines solchen Wesens will Chalmers zeigen, dass das Bewußtsein keine physikalische Basis haben kann (!). Er begründet diese Argumentation mir der Aussage, Zombies seien konzipierbar (conceivable), – ergo metaphysisch möglich (Chalmers (1997)). Die philosophische Debatte darüber hält an.

Abbildung 2: Wegners Modell (nach Wegner, p, 66)



ben. Diese Basis – neuronale Prozesse – sind uns ja nicht bewußt, wir nehmen sie nicht wahr. Es ist auch nicht sinnvoll, sie wahrzunehmen: ihre Wahrnehmung würde die Wahrnehmung dessen, was wahrgenommen werden soll, nur stören, und abgesehen davon würde ein Problem der Selbstreferenz entstehen: kann ein Prozess, der zu Bewußtsein führt, sich gleichzeitig selbst wahrnehmen? Wie dem auch sei, wir nehmen diese Prozesse nicht wahr, und das bedeutet, dass wir nicht wahrnehmen, ob ein Gedanke die Handlung bewirkt hat oder nicht. Es ist dann denkbar, dass die neuronalen Prozesse einerseits Bewußtsein und andererseits, parallel dazu, die Ausführung der Handlung/Entscheidung bewirken: "Perhaps we experience conscious will when we infer´that our thought causes our action, although we can't really know that this is the causal path" (Wegner (2003), p. 65), s Abbildung 2. Das Gehirn erzeugt sowohl den Gedanken wie die Handlung, so dass die Person folgert, dass der Gedanke die Handlung bewirkt (Wegner, p. 66). Wegner liefert eine Reihe von Beispielen von Untersuchungen, die diese Hypothese stützen, oder doch zumindest kompatibel mit ihr sind.

So berichten Patienten mit Hirnschäden das "Fremde Hand Syndrom" (alien hand syndrome): eine Hand agiert "with a mind of its own". Die Hand führt Handlungen durch, ohne dass der Patient sie will oder oder sie willentlich kontrolliert. Ein Patient habe berichtet, dass die Hand von "jemandem auf dem Mond" kontrolliert werde. Schizophrene berichten, dass sie Stimmen hören, d.h. die eigenen Gedanken werden anderen Personen zugeordnet.

Ein anderes Beispiel sind spirituelle Sitzungen, bei denen sich eine Gruppe von Personen um einen leichten Tisch versammeln und die Hände auf die Tischplatte legen. Nach einer Weile macht sich ein überatürlicher Geist bemerkbar, indem er den Tisch in irgendwelche Bewegungen versetzt. Es war der bedeutende Empiriker Michael Faraday, der durch Anwendung von Messvorrichtungen nachwies, dass

es die Teilnehmer der Runde waren, die den Tisch in Vewegung setzten (Faraday (1853)). Die Erfahrungen des Willens war in Bezug auf die kausale Basis völlig irreführend.

Die vermutliche Ursache für die Erfahrung eigener, freier Entscheidungen und Handlúngen wird durch die 'Theorie der anscheinenden mentalen Verursachung[118] gedeutet. Wenn ein Gedanke im Bewußtsein erscheint kurz bevor eine Handlung durchgeführt wird, und die Handlung konsistent mit dem Gedanken ist, so erfahren wir einen bewußten Willen und wir schreiben die Autorschaft für die Handlung uns selbst zu; wir erfahren uns als Akteure, die bewußt ihre Handlungen verursachen, wenn unser 'mind' mit einer Vorschau von Handlungen versieht, die als genau gelten, die akkurat sind, wenn die Handlungen beobachten. Es gibt experimentelle Befunde, die diese Intrerpretation stützen. Langer & Roth (1975) berichten, dass Versuchspersonen (Vpn) in bei einer zuälligen Folge berichteten, sie hätten Kontrolle über die Folge, wenn eine große Zahl von Erfolgen bei der Vorhersage der Ereignisse hatten; analoge Befunde lieferten Jenkins & Ward (1965). Allgemein scheinen Menschen mehr ihren eigenen Willen zu erleben, wenn sie bei Vorhersagen häufiger recht haben. Dementsprechend haben depressive Patienten, die sich selbst als weniger erfolgreich erleben, weniger häufig das Erleben, Kontrolle über das Eintreten von Ereignissen zu haben (Alloy & Abramson (1979)). Wegner berichtet eine Reihe weiterer empirischer Befinde, die konsistent mit seinem Modell der Erfahrung eines anscheinend freien Willen sind. Dabei legt er Wert darauf, dass das Modell *nicht* besagt, dass Handlungen nicht durch bewußte Gedanken generiert werden. Die Bestimmung der kausalen Relationen zwischen bewußter Repräsentation und Handlungen sei die Aufgabe wissenschaftlicher Untersuchungen (s.a. Jack & Shallice (2001)). Wegner fasst zusammen:

> "The experience of conscious will is a marvelous trick of the mind, one that yields useful intuitions about our authorship – but it is not the foundation for an explanatory system that stands outside the paths of deterministic causation." p. 68, Z 71

Es kann sinnvoll sein, zunächst die Beziehungen zwischen sensorischem Input und korrespondierender Wahrnehmung zu untersuchen, wobei 'Wahrnehmung' das Resultat eines Prozesses der Bewußtwerdung ist. Wird eine visuelle Szene wahrgenomen, so beginnt der Wahrnehmungsprozess mit einer neuronalen Kodierung auf der Retina und einer "Weiterverarbeitung" im visuellen Kortex. Irgendwo im Hirn entsteht dann ein Bild der Szene, d.h. aus der neuronalen Kodierung entsteht ein uns bewußtes Bild der Szene. Wie dieser Prozess funktioniert ist (noch) nicht bekannt, aber es scheint, dass es sich um einen neuronalen Prozess handelt, der bestimmte neuronale Strukturen voraussetzt.

So gehen Cabanac et al. (2009) davon aus, dass die Dimension *angenehm – unangenehm eine* Dimension des Bewußtseins von mehreren sein könnte. Die visu-

[118] theory of apparent mental causation

elle Szene wir demnach in Kombination mit einem emotinoalen Zustand bewußt. Das Verhalten des Organismus wird so angelegt, dass sein Zustand möglichst 'angenehm' ist. Insgesamt gehen Cabanac et al. nehmen insgesamt vier Dimensionen an. Die erste Dimension wird als qualitativ angenommen: die Farbe rot, oder blau, oder was der Farbkreis sonst noch hergibt. Ein Geschmack ist süß, oder sauer, etc. Die zweite Dimension ist quantitativ, sie beschreibt die Intensität eines Stimulus: die Farbe ist hell oder dunkel, der Ton ist laut oder leise, und so weiter. Die dritte Dimension ist affektiv ("hedonisch": angenehm – unangenehm). Die vierte Dimension repräsentiert die die Zeit, die ein Stimulus einwirkt. Die Generierung von Bewußtsein erfordert neuronales Gewebe, also sollte die Größe des Gehirns eine Rolle spielen, – sie wird durch den *Enzephalisationsquotienten* (EQ) angegeben. Hier gibt es eine deutliche Differenz zwischen ektothermen und endothermen Organismen. Die Endothermen sind warmblütige, tachymetabolische[119] Tieren mit einem zehnmal größeren Gehirn als dem der kaltblütigen bradymetabolischen[120]. Ectothermische Tiere wie Lissanphibia (= rezente Amphibien wie Blindwühlen, Schwanz- und Froschlurche) haben 5 Vesikeln am cephalischen Ende eines neuronalen Tubus; bei den Lissanphibia behält das Telencephalon die Struktur eines embryonischen Vesikels bei, während man bei den Lepidosauria – eine Teilklasse von Reptilien, die zuerst im frühen Perm (Beginn ca vor 298 Millionen Jahren bis ca. vor 252 Millionen Jahren) erscheinen) – einen größeren Wechsel zu einer neuen Struktur findet, nämlich einen Kortex, d.h. zwischen Fröschen und Eidechsen gibt es einen qualitativen Sprung, der mit einem dazu korrespondierenden Sprung in der Produktion des Neurotransmitters Dopamin einhergeht. Dieser ist wiederum mit einer hedonistichen Erfahrung gekoppelt, und hedonistische Erfahrungen sind bewußte Prozesse. Wie Cabanac ausführt geht dieser Sprung mit Veränderungen im Verhalten einher. Weitere Details können hier nicht besprochen werden, aber zusammenfassend kann man sagen, dass es bei diesen Veränderungen nicht um das Bewußtsein eines Selbst geht, sondern nur um die Existenz eines mentalen Raumes (Cabanac et al. (2009), p. 271, Abbildung 6), wobei es sich um einen "quantitativen" (Cabanac et al.), nicht "qualitativen" Prozess handelt. Dieses erste Bewußtsein entwickelt sich in einem inkrementellen Prozess.

Mashour et al. (2013) haben ebenfalls eine Überblick über Arbeiten zur Frage, wann Bewußtsein entstanden ist, geliefert. Sie untersuchen u.a. die Möglichkeit, die Beziehung zwischen Bewußtsein und Anästhesie zu untersuchen, um Einblicke in die Phylogenie des Bewußtseins zu erlangen. Sie bestätigen bisherige Befunde, denen zufolge die Entstehung zumindest eines repräsentierenden Bewußtseins phylogenetisch weit zurückgreift. Die grundsätzlichen Mechanismen des menschlichen Bewußtseins sind phylogenetisch sehr früh in der Geschichte der Wirbeltiere entstanden. Die Unterschiede zwischen den Spezies bezüglich des Bewußtseins sind gradueller und nicht qualitativer Natur. Der Beginn der Entwicklung liegt wahrschenlich in der Kopplung eines internen Systems, das die

[119] tachymetabolisch = mit erhöhtem Grundumsatz

[120] bradymetabolisch = hoher aktiver Stoffwechsel und ein erheblich langsamerer Ruhestoffwechsel

Befriedigung von Notwendigkeiten wie Nahrungsaufnahme etc regelt, und einem situationsbezogenen Bewußtseinszustand (gemeint ist das englische 'awareness'). Es zeigt sich, dass der Ort der Entstehung von Bewußtsein in den phylogenetisch alten Hirnteilen wie dem Hirnstamm, dem Thalamus und dem Hypothalamus etc liegt, während der Neokortex relativ wenig am "primitiven" Teil des Bewußtseins teilhat. Bei "höherem" Bewußtsein zeigen PET-Untersuchungen zusätzlich eine verteilte Aktivität im gesamten Hirn auf; Mashour et al. liefern eine Menge von Details. Ob man aus immer genauer werdenden Untersuchungen Einsicht in die Mechanismen bekommen kann, die zumindest 'einfaches' Bewußtsein erzeugen, ist gegenwärtig noch nicht klar. Vielleicht läßt sich der Entstehungsprozess durch molekuarbiologische Prozesse erklären, oder – man kann es nicht voraussagen – durch eine Neufassung der quantenmechanischen Erklärung von Roger Penrose. Singers Aussage "Unser Gehirn setzt seine eigenen kognitiven Werkzeuge ein, um sich selbst zu begreifen, und wir wissen nicht, ob dieser Versuch gelingen kann" (Zitat 34, Seite 49) gilt nach wie vor. Aber dieses Unwissen sollte uns nicht davon abhalten, weiter zu versuchen, die Entstehung von Bewußtsein zu begreifen, und sollten wir es begreifen, werden wir vermutlich klarere Antworten auf die Frage nach dem freien Willen zu bekommen.

Eine Entscheidung ist frei, wenn zwischen einer Menge möglicher Alternativen ohne Zwang gewählt werden kann. Man kann diese Definition mit der Kantschen Definition verbinden, wenn man sagt, eine Entscheidung sei frei, wenn sie nicht kausal bestimmt wird, d.h. wenn sie nicht durch einen Grund oder eine Ursache bewirkt wird. Es werde einmal angenommen, dass die zur Verfügung stehenden Alternativen unterschiedlich präferiert werden, und dass es keine Zwangsbedingung gibt, die eine Entscheidung für die maximal präferierte Alternative verhindert. Die maximale Präferenz bestimmt die Entscheidung, weil sie wie eine Ursache oder ein Grund wirkt. Demnach wäre die Entscheidung unfrei. Allerdings kann man sagen, dass die eigentliche Entscheidung die Entscheidung über die Entscheidungsstrategie – hier die Entscheidung nach der maximalen Präferenz – ist, und *diese* Entscheidung könne doch frei sein. Aber auch hier kann argumentiert werdén, dass diese Entscheidung das Resultat eines Prozesses sein wird, der insgesamt nicht bewußt abläuft, – wir haben keinen Einblick in die neuronalen Prozesse, die den Entscheidungsprozess ausmachen.

Der Begriff der freien Entscheidung verliert deshalb seine Bedeutung, – es sei denn, man postuliert eine Form des Dualismus. Der Preis dafür ist, dass man nun nicht mehr sagen kann, wo und wie die Entscheidung getroffen wird. Max Planck ist nach langem Bemühen um eine Lösung zu dem Schluß gekommen, das Problem des freien Willens sei ein Scheinproblem. Schon Immanuel Kant hat gesehen, dass der Begriff des freien Willens kein empirischer Begriff ist, der Begriff sei ein weder beweisbares noch widerlegbares theoretisches Konstrukt. Da Kant andererseits überzeugt war, dass es wegen der Möglichkeit moralischen Handelns der Wille frei sein muß, schien er keine andere Wahl zu haben, als zu postulieren, dass der Ort der Gedanken und damit des freien Willens die kausalitätsfreie, raum- und zeitlose Welt der Noumena sein müsse. Die Annahme eines freien Willens

führt ihn unmittelbar zu einer unfreien Entscheidung[121], – wobei man bedenken muß, dass Kant nicht bedacht zu haben scheint, dass moralisches Handeln auch ohne freien Willen möglich ist. Es ist schwer vorstellbar, dass er diese Möglichkeit nicht bedacht hat. Aber er hatte ja seine Metaphysik des Sollens, so dass er diese Möglichkeit nicht als Gegenargument sehen mußte, etwa wenn aufgrund von Erziehung und sozialen Normen zumindest bestimmte amoralische Entscheidungen gar nicht in Betracht gezogen werden und nur moralische Entscheidungen als zulässig gelten. Unversehens entsteht ein Bild von Kant als einem gelenkten Automaten.

Dieses Bild ist sicher nur eine ironische Zuspitzung ohne philosohische Bedeutung. Aber die Kopplung von Moral an die Freiheit des Willens ist keine Denknotwendigkeit, sondern allenfalls eine Hypothese. Dass Kant hat diese Möglichkeit gesehen hat, legen seine Antinomien 9 und 10, Seite 27 nahe. Kant löst die Antonomie auf, indem er den Begriff der "Freiheit im kosmologischen Verstande, [als] Vermögen, einen Zustand von selbst anzufangen etc (s. Zitat 17), 31, einführt. Das Problem ist, dass er nicht zeigen kann, dass es dieses Vermögen auch wirklich gibt, so dass die Definition 17 nur eine weitere Spezifikation des begrifflichen Konstrukts "freier Wille" ist. Ähnliche Argumentationen lassen sich bei den anderen Befürwortern der Existenz eines freien Willens finden. Das Merkwürdige an der Diskussion der These ist, wie scharf sich Befürworter und Gegner der These voneinander unterscheiden (z.B. t'Hooft und Zeilinger). Anscheinend gehen sie von unterschiedlichen 'Selbstverständlichkeiten' aus, also von Annahmen, die stillschweigend gemacht und nicht weiter hinterfragt werden. Man kommt auf die Idee, das "Problem" des freien Willens als Scheinproblem zu sehen.

Macht man sich an die Elaboration dieser Vermutung, so findet man recht bald, dass sie schon elaboriert worden ist: der britische Philosoph Gilbert Ryle (1900 – 1976) entwickelte in seinem Buch "The Concept of Mind" den Begriff des Kategorienfehlers. Den begeht man, wenn man Dinge, die zuverschiedenen logischen Kategorien gehören, in ein und dieselbe Kategorie steckt. Ryles berühmtes eigenes Beispiel ist das des Besuchers von Oxford: er läßt sich alle Gebäude zeigen, die zur Universität Oxford gehören, um dann zu fragen, wo denn nun die Universität Oxford sei. Die Frage beruht auf einem Kategorienfehler, der auf der Annahme beruht, eine Universität müsse in *einem* Gebäude untergebracht sein. So erklärt er auch Willensakte (Kapitel III: Volition in 'The Comcept of Mind'). Sie würden dualistisch als sozusagen geheime, geistige Akte gesehen, die Handlungen vorangehen. Will man seine Hand heben, so die Annahme, gehe der Handlung ein willentlichier ("geistiger") Akt voran: man *will*, dass man die Hand hebt. Ryle spricht in diesem Zusammenhang vom *Ghost in the Machine*. Der "Wille" sei aber kein separater Antrieb, sondern nur eine Beschreibung einer Handlung, die als "frei" *beschrieben* wird, wenn sie nicht unter einem erlebten Zwang ausgeführt wird. Insofern sei der freie Wille ein Mythos. In Wirklichkeit ist der freie Wille ein Prädikat, das einem Verhalten zugeschrieben wird. Demnach gibt es keinen freien Willen "an sich" (Ryle benutzt diesen kantschen Begriff nicht), d.h. es gibt

[121] Die begriffliche Struktur seiner Philosophie läßt ihm – anscheinend – keine andere Wahl!

keine mentale Funktion, die den freien Willen verkörpert. In diesem Sinne kann man das Problem des freien Willens als Scheinproblem sehen.

Bleibt man bei Ryle, so kann man das Problem des freien Willens als sprachphilosophisch gelöst betrachten. Aber diese Betrachtung wird nicht jeden, der über die Wirkungsweise des Gehirns nachdenkt, befriedigen. Warum die philosophischen Argumente für die These, das Gehirn sei nicht die Basis für Entscheidungen, überzeugender sein sollen als die philosophischen Argumente, die auf die möglicherweise erhellende Rolle empirischer Untersuchungen verweisen, ist unklar. Rein philosophisch wird man aber kaum weiterkommen: wie komplexe neuronale Strukturen Kognitaionen erzeugen, ist hauptsächlich ein empirisches Problem. Dementsprechend sei das Schlußwort Patricia Smith Churchland überlassen:

> The philosophical lesson is this: when not much is known about the topic, don't take terribly seriously someone else's heartfelt conviction about what problems are scientifically tractable. Learn the science, do the science, and see what happens.
>
> Patricia Smith Churchland, *Brainwise* (2002), p. 180

2 Anhang

2.1 Turingmaschinen

Die meisten Menschen beantworten die Frage, wieviel 13 mal 17 ist, nicht aus dem Gedächtnis, sondern sie rechnen, meistens im Kopf, etwa $10 \cdot 13 = 130$, plus $7 \cdot 13 = 70 + 21 = 91$, also $130 + 91 = 221$. Oder $17 \cdot 10 = 170$, plus $3\dot{1}7 = 30 + 21 = 51$ und $170 + 51 = 221$. Das Wichtige dabei ist, dass jedes Mal eine Folge von Operationen durchgeführt wird, Zwischenergebnisse abespeichrt werden, etc. Diese Folgen können sich von Person zu Person ein wenig unterscheiden, eine gegebene Person wird vermutlich stets ein und dieselbe Folge wählen, aber sie kann auch zwischen Folgen wählen. bei größeren Zahlen wird sie mit Papier und Bleistift arbeiten und dabei die Operationen in stets der gleichen Reihenfolge vornehmen. Folgen dieser Art sind Algorithmen. Turing (1936) hat diese Art, Aufgaben zu lösen, als Vorbild für eine allgmeine Art, Algorithmen zu definieren, gewählt. Der Begriff des Algorithmus geht auf den iranisch-stämmigen Mathematiker und Universalgelehrten al-Chwarizmi zurück, der um das Jahr 800 herum im Haus der Weisheit in Bagdad wirkte. Die Definition eines Algorithmus genügte aber nicht den Anforderungen an die Genauigkeit in der modernen Mathematik, so dass der englische Mathematiker Alan Turing ((1912 – 1954) eine präzise, formale Definition vorlegte. Dazu konzipierte er eine, heute nach ihm bekannte Maschine – eben die Turing-Maschine. Sie besteht zunächst aus einem in Felder eingeteilten, unendlichen (Papier-)Band. In edes Feld kann ein Zeichen aus einem Arbeitsalphabet geshrieben werden. Über dem Band befindet sich ein Schreib- und Lesekopf, der jeweils ein Zeichen liest, oder lscht, oder schreibt, wobei die Aktivität des Kopfes durh ein Program gesteuert wird. Der Kopf kann seine Position um ein Feld nach links oder nach rechts veändern. Die Definitin der Turing-Maschine ist dann

Definition 2 *Eine Turing-Maschine (TM) ist durch das 7-Tupel*

$$M = (Z, \Sigma, \Gamma, \delta, z_0, \Box, E)$$

gegeben, wobei Z eine endliche Zustandsmenge ist,
2. Σ das Eingabealphabet,
3. $\Gamma \supset \Sigma$ das Arbeitsalphabet,
4. $\delta : Z \times \Gamma \longrightarrow Z \times \Gamma \times L, R, NP')$ die Überführungsfunktion,
5. $z_0 \in Z$ der Startzustand,
6. $\Box \in \Gamma - \Sigma$ ein Leerzeichen, 7. $E \supseteq Z$ ist die Menge der Endzustände,

Dann bedeutet

$$\delta(z, a) = (z', b, x)$$

δ beschreibt die Funktionsweise der Turing-Maschine. Befindet sich M im Zustand z und unter dem Lesekopf das Zeichen a steht, so geht M im folgenden Schritt in den Zustand z' über, schreibt das Zeichen b auf die Position, auf der vorher

a stand und führt dann eine Bewegung des Lesekopfes $x \in \{L, R, N\}$ aus $L =$ Bewegung nach links, $R =$ Bewegung nach rechts, $N =$ neutral, d.h. der Kopf wird nicht bewegt. Die TM ist eine allgemeine Beschreibung dessen, was ein Mensch macht, wenn er eine Aufgabe einer Folge von Regeln löst, und sie ist deswegen ein generisches Modell für einen Computer. Und sie ist, wie am Anfang dieses Abschnitts schon erwähnt, eine allgemeine Beschreibung von Prozessen.

s. Schöning, p. 72.

Algorithmen und der Begriff der Turing-Maschine, eher intuitiv, der Begriff der Berechnung.

2.2 Turings Halte-Problem

Man stelle sich vor, dass man ein Programm P für irgendetwas geschrieben hat. Man habe dabei irgendwo einen Fehler gemacht, der bewirkt, dass das Programm in eine Schleife ("loop") läuft und deswegen ewig weiterläuft, sofern man es nicht abbricht. Man weiß aber nicht, ob es nur so lange läuft, weil es in einer Schleife läuft, oder ob ob die Berechnung, die das Programm ausführen soll, so langwierig ist. So entsteht der Wunsch nach einem Programmm P_0, das prüft, ob P in eine Schleife laufen wird, oder nicht. In diesem Zusammenhang ist es wichtig, einen weiteren Begriff einzuführen, den der *universellen Turing-Maschine*. Dies ist eine Turing-Maschine, die eine beliebige Turing-Maschine für einen beliebigen Input simuliert.

Nun ist P_0 selbst ein Programm, das einen Fehler enthalten kann, der impliziert, dass P_0 in eine Schleife laufen kann. Man könnte auf die Idee kommen, dass ein weiteres Programm P_1 geschrieben werdem muß, das prüft, ob P_0 in eine Schleife läuft. Dieser Ansatz ist aber keine Lösung, weil sofort klar ist dass dann ein Programm P_2 benötigt wird, das prüft, ob P_1 in eine Schleife läuft, etc, d.h. man gerät in einen infiniten Regress. Dann aber bleibt nur die Möglichkeit, P_0 mit sich selbst zu prüfen, d.h. P_0 zum Input von P_0 zu machen. P_0 muß sich selbst prüfen, – womit man sich aber die Falle der Selbstreferenz einhandelt.

Ein Programm ist die Implementierung eines Algorithmus, d.h. eines Verfahrens, das ein Problem in einer endlichen Folge von Schritten löst. Das Verfahren kann deterministisch oder probabilistisch sein. Bei einem deterministischen Algorithmus liegt nach jedem Schritt eindeutig fest, welche Operation als nächste durchgeführt wird, bei einem probabilistischen Algorithmus wird die jeweils folgende Operation mit einer bestimmten Wahrscheinlichkeit aus einer Menge von möglichen Operationen gewählt.

Turing (1936), Abschnitt 8., hat gezeigt, dass ein Programm P_0, das bestimmt, ob es in ene Schleife gerät oder nicht nicht existiert. Diese Aussage ist das Halte-Theorem.

Turings Beweis ist indirekt, dh es wird angenommen, dass ein solches Programm P_0 (Turing (1939) hat dafür den Ausdruk 'Orakel' eingeführt) existiert,

um dann zu zeigen, dass diese Annahme auf einen Widerspruch führt, d.h. das Orakel existiert nicht. Eine grundlegende Annahme für die folgenden Betrachtungen ist, dass die Arbeit eines Programms stets aus einer diskreten Folge von Operationen oder Schritten besteht; es kann gezeigt werden, dass diese Annahme keine Einschränkung darstellt , d.h. das Halte-Theorem gilt auch für kontinuierliche Prozesse (s. unten).

Definitionsgemäß liefert das Orakel P_0 auf die Frage, ob ein gegebenes Programm P_s für einen gegebenen Input P (der selbst ein Programm sein kann) hält (H für 'hält') oder in eine Schleife gerät (L für 'Schleife' (loop)) stets die korrekte Antwort. Die hier interessierenden, möglichen Antworten von eines Programms (einschießlich P_0 und P_s) sind also

$$P = \begin{cases} L, & P \text{ gerät in eine Schleife,} \\ \text{H}, & P \text{ hält nach endlich vielen Schritten} \end{cases}$$

Die Frage von Ps an das Orakel P_0 ist, ob P, auf sich selbst angewandt, hält oder nicht. Es soll dann die folgende Regel gelten:

Regel R: Wenn das Orakel sagt, dass P loopt, stoppt P_s. Wenn das Orakel sagt, P wird stoppen, loopt P_s.

Nun kommt die Frage: was gschieht, wenn P_s sich selbst als Input bekommt? Die möglichen Antworten sind

1. Das Orakel sagt, P_s wird für immer loopen. Nach der Regel R muß P_s dann stoppen.
2. Das Orakel sagt, P_s wird stoppen. Nach der regel R muß P_s dann loopen.

Resultat: Das Orakel gibt in jedem Fall die falsche Anwort, entgegen der Annahme, dass es stets die Wahrheit spricht. Damit ist die Annahme, dass das Orakel stets die Wahrheit spricht widerlegt. In anderen Worten: das Orakel existiert nicht, - es gibt kein Programm, das voraussagen kann, ob ein Programm einen Fehler enthält oder nicht, d.h. es existiert kein Programm, dass sich selbst "erkennt"-

Man kann diesen Sachverhalt anhand von Cantors Diagonalverfahren illustieren. Dazu werde die folgende Notation eingeführt: $P_1(P)$ heißt, dass P_1 das Programm P prüft. Für $P = P_1$ erhält man $P_1(P_1)$: P_1 prüft sich selbst.

Man hat nun zwei Möglichkeiten: $P_1(P_1)$ stoppt (Stop), d.h. $P_1(P_1) = H$, oder läuft in eine Schleife (Loop), $P_1(P_1) = L$. Die Menge der möglichen Programme ist abzählbar unendlich (da es aus abzählbar vielen Elementen besteht). Jedes Programm kann als Input für jedes Programm gewählt werden und stoppt dann entweder (H für hält), oder läuft in eine Schleife (L für loop). Man kann dann eine Tabelle anschreiben: Es sei noch einmal betont, dass die Tabelle *alle* möglichen Programme mit allen möglichen Inputs enthalten soll. Die i-te Zeile, etwa die für P_3 anthält die "Reaktionen" des dritten Programms auf alle möglichen Inputs: für P_1 stoppt es (H), für P_2 loopt es (L) etc. Die Tabelle scheint also alle möglichen Reaktionen, – wie es scheint. Nun betrachte man die Reaktionenn in

	P_1	P_2	P_3	P_4	P_5	$\cdots$
P_1	L	H	H	L	L	$\cdots$
P_2	H	L	H	H	L	$\cdots$
P_3	H	L	H	L	H	$\cdots$
P_4	H	L	H	L	L	$\cdots$
P_5	H	L	H	H	H	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$
Diagonale	L	L	H	L	H	
inv. Diag	H	H	L	H	L	

den Diagonalzellen der Tabelle: L, L, H, L, H, $\cdots$ Dies sind die Reaktionen der Programme auf sich selbst. Nun ersetze man jedes Symbol durch das alternative Symbol, also L durcg H und H durch L. Man erhält die Folge H, H, L, H, L; $\cdots$ Diese Folge kann nicht in der Tabelle vorkommen, denn sie unterscheidet sich von den Folgen in der Tabelle in Bezug auf mindestens ein Element, nämlich das Element in der Diagonalzelle. Dies bedeutet, dass es unmöglich ist, ein Programm zu schreiben, für das sich diese neue Folge ergibt. Dieser Sachverhalt spricht gegen die Existenz des Orakel-Programms, – Weder P_s noch das Orakel können existieren! Damit ist das Halte-Theorem bewiesen.

2.3 LLoyds Beweis

First: check notion of Universal Turing machine! Es sei T eine monoton wachsende Abbildung der natürlichen Zahlen in die natürlichen Zahlen, also $T : \mathbb{N} \to N$, $\mathbb{N}$ die Menge der natürlichen Zahlen. Weiter seien $|d|$ und $|k|$ die Längen (in bits) der Zahlen d und k. $d_T(k)$ bezeichne einen zeitbegrenzten Entscheider (d von decider), mit

$$d_T(k) = \begin{cases} d(k), & d \text{ liefert Output auf Input } k \\ & \text{in } T(|d| + |k|) \text{ oder weniger Schritten} \\ 0, & \text{sonst} \end{cases}$$

Keine Anwort in $T(|d|+|k|)$ bedeutet 'Nein'. Wendet man das Diagonalargument auf die Frage an, ob zeitbegrenzte Entscheider mit 'ja' oder 'nein' in weniger als T Zeiteinheiten entscheiden, so zeigt sich, dass sie im Allgemeinen *mehr* als T Zeiteinheiten benötigen. Zeit wird dabei in Termen von komputationalen Schritten repräsentiert.

Es werden nun zeitbegrenzte Turingmaschinen $\mathcal{T}_d^T$ mit dem Input k und dem Output $d_T(k)$ betrachtet. Es sei $f(d_T, k) = d_T(k)$; f ist die Antwort auf die Frage, welche Entscheidung durch einen typischen Entscheider innerhalb der Zeit T gegeben wird. Es läßt sich nun eine Turingmaschine konzipieren, die (d, k) als Input akzeptiert und die $f(d_T, k)$ als Output liefert; f kann als Antwort auf die Frage, wie ein typischer Entscheider in der Zeit T handelt, gesehen werden. Die Funktion f ist, wenn nicht explizit auf die Zeit T Bezug genommen wird, *berechenbar*

in dem Sinne, dass sie durch eine Turingmaschine mit Input (d, k) und Output $f(d_T, k)$ berechnet werden kann. Sie ist aber *nicht* in der Zeit T berechenbar, d.h. jede Berechnung für die Entscheidung, die ein Entscheider treffen wird, nimmt mehr Zeit in Anspruch als die Entscheidung des betrachteten Entscheiders.

Um zu zeigen, dass die Berechnung von f mehr als T Einheiten erfordert, werde die Matrix A_T betrachtet, mit dem (d, k)-ten Element $f(d_T, k)$; dies ist die Matrix ller Entscheidungsfunktionen $d_T(k)$, die in der Zeit T berechnet werden können. Es sei $g(k) = 0$ wenn $f(k, k) = 1$, und umgekehrt. Dann folgt $g(k) = \neg f(k, k)$ ($\neg$ = nicht). Ist f in der Zeit T berechenbar, so ist es auch g.Wenn aber g in der Zeit T berechenbar ist, dann ist notwendig $g(g) = f(g, g)$. Damit hat man aber einen Widerspruch, da ja $g(g) = \neg f(g, g)$. Folglich sind weder f noch g in der Zeit T berechenbar.

Betrachte einen Entscheider d (d von *decider*). Versehe d mit einem Input: die Beschreibung eines anderen Entscheiders d'. d' entspricht einer Turingmaschine T_d mit zwei Input-Bändern, einem für d' und das andere für die Spezifikation des Entscheidungsproblems k. $d(d', k)$ sei der Output von T_d, wenn T_d stoppt, und $d(d', k) = d(k)$, wenn $d' = 0$, und $d(d', k) = d'(k)$ sonst. Wenn T_d nicht hält, ist der Output undefiniert. Das heißt, der universelle Entscheider d kann entweder die "straight decision" $d(k)$ machen, oder die Entscheidung eines anderen Entscheiders d' (einschließlich von sich selbst, $d = d'$, machen.

Anwendung auf die Selbstbeobachtung eines Entscheiders;

Unentscheidbarkeit ergibt sich, wenn der universelle Entscheider Aussagen über seine eigene Zukunft machen soll. Es werde eine 3-dimensionale Matrix mit den Elementen $d(d', k)$ betrachtet, wobei

$$d(d', k) = \begin{cases} d(d', k), & d(d', k) \text{ ist definiert,} \\ F, & d(d', k) \text{ ist nicht definiert} \end{cases}$$

Für festes k werden nun die Diagonalelemente $d' = d$ betrachtet. Dies sind die Reaktionen des Entscheiders d auf sich selbst. Hier gilt das Diagonalisierungsargument für das Halte-Problem (s .Anhang). Es impliziert, dass die Funktion $f_k(d) = d(d', k)$ gilt, und $f_k(d) = F$ sonst, und $f_k(d)$ ist nicht berechenbar. Das heißt, der Entscheider d kann keine Aussage machen, wenn es um seine eigene, zukünftige Entscheidung geht.

Mit T_d^T werden nun zeitabhängige Turingmaschinen bezeichnet. Diese liefern Outputs $d_T(d'k)$, wobei

$$d_T(d', k) = \begin{cases} d(d', k), & T_d^T \text{ hält nach Zeit } T(|d| + |d'| + |k|) \\ 0, & \text{sonst} \end{cases}$$

Nun betrachte man die 3-dimensionale Matrix mit den Elementen $d_T(d'k)$. Für festes k betrachte man weiter die Diagonalelemente mit $d = d'$. Aus dem Hartmanis-Stearns Diagonalverfahrern folgt, dass für diesen Fall die Outputs nicht in der Zeit T berechnet werden können. Dies bedeutet, dass ein universaler Entscheider, der Ausagen über seine eigenen Entscheidungen machen soll,

intrinsisch wenige Effizient ist als ein Entscheider, der Entscheidungen ohne die Introspektion treffen soll.

Literatur

[1] Albert, H. (2009) Die dualistische Metaphysik von Jürgen Habermas – Eine kritische Untersuchung seines nachmetaphysischen Denkens'. RMM Vol. 0, Perspectives in Moral Science, ed. by M. Baurmann & B. Lahno, 2009, 109-120

[2] An der Heiden U., Chaos und Ordnung, Zufall und Notwendigkeit, in: Küppers G. (Hrsg.), Chaos und Ordnung - Formen der Selbstorganisation in Natur und Gesellschaft, Reclam, Stuttgart 1996

[3] Alloy, L.B. and Abramson, L.Y. (1979) Judgment of contingency in depressed and nondepressed students: sadder but wiser? *J. Exp. Psychol. Gen.*, 108, 441–485

[4] Audi, R.: The Cambridge Dictionary of Philosophy, Sec. Edition (1999), Cambridge University Press

[5] Baars, B.J. A cognitive theory of consciousness. Cambridge University Press 1988

[6] Bak P, Tang C, Wiesenfeld K (1988) Self-organized criticality. *Phys Rev A* 38, 364–374

[7] Bandura, A., & Walters, R.H.: Social learning and personality development. Holt Rinehart and Winston: New York 1963

[8] Baumeister, R.F., E.J. Masicampo, and C.N. DeWall, (2009) Prosocial benefits of feeling free: Disbelief in free will increases aggression and reduces helpfulness. *Personality and Social Psychology Bulletin,* 35, 260–268

[9] Baumeister, R. F., Crescioni, A. W., & Alquist, J. L. (2011). Free will as advanced action control for human social life and culture. *Neuroethics,* 4, 1-11.

[10] Becker, O.: Grundlagen der Mathematik in geschichtlicher Entwicklung. Freiburge/München 1954

[11] Bell, J. S. (1964) On the Einstein Podolsky Rosen Paradox. In: *Physics,* Band 1, Nr. 3, 1964, S. 195–200

[12] BennetM.R, hacker, P.M.S.: Philosophical Foundatins of Neuroscience. Oxford 2003

[13] Bishop, C.M.: Pattern Recognition and Machine Learning. Springer Verlag 2006

[14] Blackmore, S., & Troscianko, E. T. (2024). Consciousness: an introduction. Routledge.

[15] Block, Ned (2005), Action in Perception by Alva Noë *The Journal of Philosophy,* Vol. 102, No. 5, S. 259–272.

[16] Brans, C.H. (1988) Bell's Theorem Does Not Eliminate Fully Causal Hidden Variables. *International Journal of Theoretical Physics,*27 (2), 219–226

[17] Braun, M. N., Wessler, J., & Friese, M. (2021). A meta-analysis of Libet-style experiments. *Neuroscience & Biobehavioral Reviews,* 128, 182-198.

[18] Braun, H. A. (2021). Stochasticity versus determinacy in neurobiology: From ion channels to the question of the "free will". *Frontiers in Systems Neuroscience,* 15, 629436.

[19] Burgos, J.E., Donahoe, J.W. (2006) of what value is philosophy to science? A review of Max R. Bennett und P.M.S. Hacker's *Foundations of Neuroscience*, *Behavior and Philosophy,* 34, 71-87

[20] Busemeyer, J. R., & Townsend, J. T. (1993) Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review,* 100(3), 432-59

[21] Cabanac, M., Cabanac A.J., Parent, A. (2011) The emergence of consciousness in phylogeny. *Behavioral Brain Research*, 198, 267 – 272x

[22] Campbell, C. A. (1951) Is 'free Will' a Pseudo-Problem? *Mind*, N.S. Vol. LX, 441-465

[23] Carnap, R.: Scheinprobleme in der Philosophie und andere metaphysikkritische Schriften. Philosophische Bibliothek 560 (1928)

[24] Caston, V. (2021). Aristotle and the Cartesian theatre. In Encounters with Aristotelian Philosophy of Mind (pp. 169-220). Routledge.

[25] Chalmers, D. J. (1997). The conscious mind: In search of a fundamental theory. Oxford Paperbacks.

[26] Chisholm, R. M. (1964/2018): Human freedom and the self. In *Agency And Responsiblity*, (pp. 126-137). Routledge.

[27] Clark, A., Chalmers. D.J. (1998), The extended mind. *Analysis,* 58 (1): 7–19.

[28] Conway, J., & Kochen, S. (2006). The free will theorem. *Foundations of Physics,* 36(10), 1441-1473.

[29] Davidson, D. (1963). Actions, reasons, and causes. *The journal of philosophy,* 60(23), 685-700

[30] Desmurget, M., Reilly, K. T., Richard, N., Szathmari, A., Mottolese, C., & Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science,* 324(5928), 811-813.

[31] Douglas, P. K., Lau, E., Anderson, A., Head, A., Kerr, W., Wollner, M., ... & Cohen, M. S. (2013). Single trial decoding of belief decision making from EEG and fMRI data using independent components features. *Frontiers in human neuroscience,* 7, 392.

[32] Dretske, F. (1989). Reasons and causes. *Philosophical Perspectives,* 3, 1-15

[33] Earman, J. (1986). A primer on determinism (Vol. 37). Springer Science & Business Media.

[34] Edelman, G. M., Gally, J. A., & Baars, B. J. (2011). Biology of consciousness. *Frontiers in psychology,* 2, 8906

[35] Epilogue: A socratic dialogue Planck–Einstein–Murphy, in: Planck (1932)

[36] Ellis, G. (2020). From chaos to free will. *Aeon Magazine,* 12. https://aeon.co/essays/heres-why-so-many-physicists-are-wrong-about-free-will

[37] Faraday, M. (1853) Faraday, M. (1853) Experimental investigation of table turning. *Athenaeum* (July), 801–803

[38] Frankfurt H.G. (1971). Freedom of the Will and the Concept of a Person. *The Journal of Philosophy,* Vol. 68, No. 1 (Jan. 14, 1971), pp. 5-20

[39] Frege, G.: Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik,* 100(1), 25-50. (1892)

[40] Freedman, S.J., Clauser, J. F. (1972) Experimental Test of Local Hidden-Variable Theories. *Physical Review Letters*, 28 (14) 938–941

[41] Fried, I., Mukamel, R., & Kreiman, G. (2011). Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron,* 69(3), 548-562.

[42] Gadenne, V.: Philosophie der Psychologie. Bern, 2004

[43] Goldstein, S., Tausk, D. V., Tumulka, R., & Zanghì, N. (2010). What does the free will theorem actually prove. *Notices of the AMS,* 57(11), 1451-1453.

[44] Granicky, M.S. (1995) Universal computation and other capabilities and continuous dynamical systems. *Theoretical Computer Science,*

[45] Gregory, R.L.: The Eye and the Brain, Princeton 1966

[46] Griffel,F. A.: Gesellschaft und Verantwortung: Zu Vorträgen von Max Planck über Willensfreiheit. *Archiv für Rechts- und Sozialphilosophie / Archives for Philosophy of Law and Social Philosophy,* 1998, Vol. 84, No. 4 (1998), 517-528

[47] Habermas, J. (2004) Freiheit und Determinismus, *Deutsche Zeitschrift für Philosophie,* Berlin 52, 6, 871-890

[48] Habermas, J. (2006). Das Sprachspiel verantwortlicher Urheberschaft und das Problem der Willensfreiheit: Wie lässt sich der epistemische Dualismus mit einem ontologischen Monismus versöhnen?. *Deutsche Zeitschrift für Philosophie,* 54(5), 669-708.

[49] Hacker, P.M.S. (2016) Seeing, representing and describing: an examination of David Marr's computational theory of vision. In: *Investigating psychology* (pp. 119-154). Routledge.

[50] Haggard, P., & Eimer, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. *Experimental brain research,* 126, 128-133

[51] Haggard, P. (2008). Human volition: towards a neuroscience of will. *Nature Reviews Neuroscience,* 9(12), 934-946.

[52] Haggard, P. (2009). The sources of human volition. *Science,* 324(5928), 731-733.

[53] Harris, S. Free Will, Free Press, New York 2012

[54] Hartmanis, J., & Stearns, R. E. (1965). On the computational complexity of algorithms. *Transactions of the American Mathematical Society,* 117, 285-306.

[55] He, B. J. (2014). Scale-free brain activity: past, present, and future. *Trends in cognitive sciences,* 18(9), 480-487.

[56] Heidelberger, M. (1992). Kausalität. Eine Problemübersicht. *Neue Hefte für Philosophie,* 32(33), 130-153.

[57] Heisenberg, W. (1927). Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik,* 43(3-4), 172-198.

[58] Heisenberg, W.: Physikalische Prinzipien der Quantentheorie (1958), 2. unveränd.Nachdr., Berlin Spektrum, Akad. Verlag 2001

[59] Heisenberg, M. (1983) Initiale Aktivität und Willkürverhalten bei Tieren, *Naturwissenschaften,* 70, 70 – 78

[60] Heisenberg, M. (2009) Is free will an illusion? *Nature,* 459(7244), 164–165.

[61] Jenkins, H.M. and Ward, W.C. (1965) Judgments of contingency between responses and outcomes. *Psychol. Monogr.*, 79 No. 1

[62] Hofstadter, D.R., Dennett, D.: The Mind's I. Fantasies and Reflections on Self and Soul. Penguin Books, Harmondsword, England, 1981

[63] Horsthemke, W., Levever, R. (1984). Noise induced transitions. Theory and applications in phhsics, Chemistry, and biology. Springer-Verlag, Berlin, Heidelberg, New York, Tokyo 1984

[64] Hossenfelder, S., Palmer, T. (2020). Rethinking superdeterminism. *Frontiers in Physics,* 8, 1 – 24

[65] Hossenfelder, S.: Existential Physics – A Scientist's Guide to Life's Biggest Questions. Viking, New York 2022

[66] Hume, D.:"Of liberty and necessity", in Book 2 von *A Treatise of Human Nature* (1739) sowie in *An Enquiry concerning Human Understanding* (1748)

[67] Hume, D. (1789). Of probability; and of the idea of cause and effect (L. A. Selby-Bigge, Ed.). In D. Hume & L. A. Selby-Bigge (Ed.), A treatise of human nature (Reprinted from the original ed., pp. 73–78). The Clarendon Press

[68] Husserl, E.: Logische Untersuchungen: 1. T. Untersuchungen zur Phänomenologie und Theorie der Erkenntnis (Vol. 2). M. Niemeyer. (1900/01)

[69] Iso-Ahola, S. E. (2017). Reproducibility in psychological science: When do psychological phenomena exist?,*Frontiers in Psychology,* 8, 237147.

[70] Jack, A.I. and Shallice, T. (2001) Introspective physicalism as an approach to the science of consciousness. In: The Neuroscience of Consciousness (Dehaene, S., ed.), pp. 161–196, MIT Press

[71] James, W. (1884). The dilemma of determinism (pp. 114-140). Kessinger Publishing.

[72] Jankowiak, J. (2020) Kant on the continuity of alterations. *Canadian Journal of Philosophy,*, 50 I, 49 — 66

[73] Johnson-Laird, P. N. (1997). Rules and illusions: A critical study of Rips's The Psychology of Proof. *Minds and Machines,* 7(3), 387-407.

[74] Kahneman, D. Thinking fast and slow, New York 2021

[75] Kahneman, D., Sibony, O., Sunstein, C.R.: Noise – a flaw in human judgment. Dublin 2021

[76] Kampen van, N. G. (1991) Determinism and Predictability, *Synthese,* 89 (2) p.273

[77] Keil, G.: Willensfreiheit und Determinismus. Reclam, Stuttgart 2018

[78] Keil, G. (2005) Eine fulminante Lehrstuhlkritik der Neurowissenschaften. *DTZPhil*, Berlin 53, 6, 951 – 972

[79] Kenny, A. (1971) The Homunculus Fallacy. In: *Interpretations of Life and Mind: Essays around the Problem of Reduction,* edited by Marjorie Grene. London: Routledge and Kegan Paul Books, 65–74.

[80] Kenny, A. (2016). The homunculus fallacy. In: *Investigating psychology* (pp. 155-165). Routledge.

[81] Kobeleva, X., López-González, A., Kringelbach, M. L., & Deco, G. (2021). Revealing the relevant spatiotemporal scale underlying Whole-Brain dynamics. *Frontiers in neuroscience,* 15, 715 – 861

[82] Kornhuber, H. H., & Deecke, L. (1965) Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere,* 284, 1-17

[83] Langer, E.J. and Roth, J. (1975) Heads I win, tails it's chance: the illusion of control as a function of the sequence of outcomes in a purely chance task. *J. Pers. Soc. Psychol.*, 32, 951–955

illusion of control as a function of the sequence of outcomes in a purely chance task. J. Pers. Soc. Psychol. 32, 951–955

[84] Larsson, Jan-Åke (2014), Loopholes in Bell inequality tests of local realism. *Journal of Physics A: Mathematical and Theoretical*, 47(42): 16. arXiv:1407.0363

[85] Lewis, David (1973). Causation. *Journal of Philosophy*, 70(17), 556–567.

[86] Lewis, David (1979). Counterfactuals. Oxford: Blackwell.

[87] Libet, B. Mind Time. The temporal factor in consciousness. Harvard University Press 2004

[88] Libet B, Gleason CA, Wright EW, Pearl DK (1983) Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): the unconscious initiation of a freely voluntary act. *Brain*, 106, 623–642, 1985

[89] Libet, B.: Unconscious cerebral initiative and the role of conscious will in voluntary action. *The Behavioral and Brain Sciences,* 9 8, 1985, S. 529–566.

[90] Libet, B.: Mind Time. The temporal factor in consciousness. Harvard University Press 2004

[91] Lloyd, S., A Turing test for free will. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* 2012, 370(1971), 3597-3610.

[92] Lucas, T. (1961) Minds, Machines and Gödel. *Philosophy*, 36 (137) 112-127

[93] Mahmoudi, A., Takerkart, S., Regragui, F., Boussaoud, D., & Brovelli, A. (2012). Multivoxel pattern analysis for FMRI data: a review. *Computational and mathematical methods in medicine,* 2012(1), 961257.

[94] Marley, A. A. J., & Colonius, H. (1992) The "horse race" random utility model for choice probabilities and reaction times, and its comparing risks interpretation. *Journal of Mathematical Psychology,* 36(1), 1-20ß

[95] Marr, D., Vision, a Computational Investigation into the Human Representation and processing of Visual Information. San Francisco 1980

[96] McNulty, M.B. (2019) Continuity of Change in Kant's Dynamics, *Synthese*, 196 (4):1595-1622

[97] Mashour, G. A., & Alkire, M. T. (2013). Evolution of consciousness: phylogeny, ontogeny, and emergence from general anesthesia. *Proceedings of the National Academy of Sciences,* 110 (supplement$_2$), 10357 – 10364

[98] Moral responsibility without free will. *iai news,* 27th April 2022, https://iai.tv/articles/moral-responsibility-without-free-will-auid-2115

[99] Mortensen, U., Wissenschaftstheorie IV – Kausalität und Zufall, 2025, http://www.uwe-mortensen.de/WissTheorieIVDeterminismus.pdf

[100] Murray, J.D.: Mathematicial Biology, Springer-Verlag, Berlin, Heidelberg etc 1989

[101] Nahmias, E., Morris, S., Nadelhoffer, T., Turner J. (2005) Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility. *Philosophical Psychology,*, (18(5), 561 — 584

[102] Nahmias, E. (2011) Why 'Willusionism' Leads to 'Bad Results': Comments on Baumeister, Crescioni, and Alquist. *Neuroethics,* 4, 17–24

[103] Nida-Rümelin, J., Freiheit und Kausalität. Redigierte Tonband-Mitschrift der frei gehaltenen Akademievorlesung am 12. April 2007 an der Berlin-Brandenburgischen Akademie der Wissenschaften. Redaktion: Christine Bratu, M.A., Universität München.

[104] Nietzsche, F.: www.zeno.org/Philosophie/M/Nietzsche,+Friedrich/Jenseits+von+Gut+und+B

[105] Nietzsche, F. (1886): Jenseits von Gut und Böse, Erstes Hauptstück: von den Vorurteilen der Philosophen. München 1954

[106] Nietzsche, F.: Friedrich Nietzsche – Werke in Drei Bänden, Zweiter Band. Europäischer Buchclub, Stuttgart, Zürich, Salzburg 1958

[107] Nisbet, R.E., Wilson, T.D. (1977) Telling more than we can now: verbal reports on mental processes. *Psychological Review,* 84, 231 — 259

[108] Noë, A.: Du bist nicht dein Gehirn – Eine radikale Philosophie des Bewußtseins. München 2010, Originaltitel *Out of our heads*, New York 2009

[109] Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in cognitive sciences,* 10(9), 424-430.

[110] Noesselt, T., Hillyard, S. A., Woldorff, M. G., Schoenfeld, A., Hagner, T., Jäncke, L., Tempelmann, C., Hinrichs, H., & Heinze, H.-J., (2002), Delayed striate cortical activation during spatial attention, *Neuron,* 35, 575-587.

[111] Northcott, R. (2019). Free will is not a testable hypothesis. *Erkenntnis,* 84(3), 617-631

[112] O'Connor, Timothy and Christopher Franklin, "Free Will", *The Stanford Encyclopedia of Philosophy* (Winter 2022 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = ¡https://plato.stanford.edu/archives/win2022/entries/freewill/¿.

[113] Penrose, R.: The Emperor's New Mind. Oxford University Press, New York 1989

[114] Penrose, R.: Shadows of the Mind. London 1995

[115] Planck, M.: Where is Science going?, New York 1932

[116] Planck, M.: Scheinprobleme der Wissenschaft. Von Max Planck – Vortrag gehalten in Göttingen am 17. Juni 1946

[117] Rachlin, H. (2012). Is the mind in the brain? A review of: Out of our heads: Why you are not your brain, and other lessons from the biology of consciousness by Alva Noë (2009). *Journal of the Experimental Analysis of Behavior,* 98(1), 131.

[118] Rauch, D., Handsteiner, J., Hochrainer, A., Gallicchio, J., Friedman, A. S., Leung, C., ... & Zeilinger, A. (2018). Cosmic Bell test using random measurement settings from high-redshift quasars. *Physical review letters, 121(8),* 080403

[119] Röd, W.: Der Weg der Philosophie, Zweiter Band, 17. bis 20. Jahrhundert. Becksche Reihe; München 1996/2000

[120] Roth, G. (2004). Wir sind determiniert. Die Hirnforschung befreit von Illusionen.
https://homepage.univie.ac.at/henning.schluss/seminare/028freiheit/texte/4.pdf

[121] Roth, G. (2005), Gehirn, Gründe, Ursachen. *DZPhil* 53, 5, 691-705

[122] Roth, G. (2014), Willensfreiheit, Physik und Hirnforschung. *Studium generale: Auf dem Weg zu einem allgemeinen Teil der Wissenschaften,* 163-174.

[123] Rowlands, M., Lau, J., Deutsch, M.: "Externalism About the Mind", The Stanford Encyclopedia of Philosophy (Winter 2020 Edition), Edward N. Zalta (ed.),
URL = ¡https://plato.stanford.edu/archives/win2020/entries/content-externalism/¿

[124] Rubinov, M., Sporns, O., Breakspear, J.P.M. (2011) Neurobiologically Realistic Determinants of Self-Organized Criticality in Networks of Spiking Neurons *PLOS, Computational Biology*, 7(6), 1–14

[125] Ryle, G.: The Concept of Mind. The University of Chicago Press 1949/2000

[126] Russell, B. (1912, January). On the notion of cause. In *Proceedings of the Aristotelian society* (Vol. 13, pp. 1-26). Aristotelian Society, Wiley.

[127] Russell, B.: On the notion of Cause, with Application to the Free-Will-Problem. In: Russell, B.: Our knowledge of the external world. 1914/ 2005

[128] Russell, P., "Hume on Free Will", The Stanford Encyclopedia of Philosophy (Fall 2017 Edition), Edward N. Zalta (ed.),
URL = ¡https://plato.stanford.edu/archives/fall2017/entries/hume-freewill

[129] Ryle, G.: The Concept of Mind. Routledge 1949

[130] Sapolsky, R. Determined. The Science of Life without Free Will. Vintage Penguin-Random House 2023

[131] Sapolsky, R. https://phys.org/news/2023-10-scientist-decades-dont-free.html

[132] Scheibe, E.: Die Philosophie der Physiker. München 2006

[133] Schopenhauer, A.: Schopenhauer, A. (1838/1854). Die beiden Grundprobleme der Ethik: behandelt in zwei akademischen Preisschriften. Joh. Christ. Hermannsche Buchhandlung.F

[134] Schöning, U. : Theoretische Informatik – kurz gefasst. Spektrum Akademischer Verlag, Heidelberg 2009

[135] Schroeder, S.: Are Reasons Causes? A Wittgensteinian Response to Davidson. In: S. Schroeder (ed.), Wittgenstein and Contemporary Philosophy of Mind, Basingstoke: Palgrave 2001; 150-70.

[136] Schurger, A., Mylopoulos, M., & Rosenthal, D. (2016). Neural antecedents of spontaneous voluntary movement: A new perspective. *Trends in Cognitive Sciences,* 20(2), 77–79

[137] Schaal, D. W. (2005). Naming our concerns about neuroscience: A review of Bennett and Hacker's Philosophical Foundations of Neuroscience. *Journal of the Experimental Analysis of Behavior,* 84(3), 683-692.

[138] Schlick. M.: Fragen der Ethik. Frankfurt, 1930/2002

[139] Singer, W. (2004) Selbsterfahrung und neurobiologische Fremdbeschreibung. *Deutsche Zeitschrift für Philosophie*, 52, 2, 235-256.

[140] Soon, C. S., Brass, M., Heinze, H. J. & Haynes, J. D. (2008). Unconscious determinants of free decisions in the human brain. *Nature neuroscience,* 11(5), 543-545.

[141] Spohn, W. (2008) Der Kern der Willensfreiheit. In: Sturm, D. (Hg.) Julian Nida-Rümelin über Vernunft und Freiheit. de Gryter 2008

[142] Srinivasan, R., Russell, D. P., Edelman, G. M., & Tononi, G. Increased synchronization of neuromagnetic responses during conscious perception. *Journal of Neuroscience,* 1999, 19(13), 5435-5448.

[143] Srinivasan, R., Winter, W. R., Ding, J., & Nunez, P. L. EEG and MEG coherence: measures of functional connectivity at distinct spatial scales of neocortical dynamics. *Journal of neuroscience methods,* (2007), 166(1), 41-52.

[144] Strawson, G. (2022 Moral responsibility without free will. iai news, 27th April 2022,
https://iai.tv/articles/moral-responsibility-without-free-will-auid-2115)

[145] R.L Stevenson – Life in Western Samoa bzw In the South Seas

[146] Tarski, A. (1933/1956). The concept of truth in formalized languages.

[147] Tetens. H.: Kants Kritik der Reinen Vernunft - Ein systematischer Kommentar. Reclams Universal Bibliothek, Stuttgart 2006

[148] t'Hooft, G. T. (2017), Free will in the theory of everything. arXiv preprint arXiv:1709.02874.

[149] 't Hooft, G. (2020). Deterministic quantum mechanics: the mathematical equations. *Frontiers in Physics,* 8, 253.

[150] Tse, P.U.: The Neural Basis of Free Will: Criterial Causation. MIT-Press 2013. Auch:
https://pbs.dartmouth.edu/news/2013/03/neuroscientist-says-humans-are-wired-free-will

[151] Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *J. of Math,* 58(345-363), 5.

[152] Turing, A. M. (1939). Systems of logic based on ordinals. *Proceedings of the London Mathematical Society, Series 2*, 45, 161-228.

[153] Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science,* 185(4157), 1124-1131

[154] Van Kampen, N. G. (1991). Determinism and predictability. *Synthese,* 89, 273-281.

[155] Vicente, J. G. (2023). Emergence of universal computations through neural manifold dynamics. bioRxiv, 2023-02.

[156] von Weizsäcker, C.F.: Errinnerungen an Martin Heidegger. In: Fetzer, G. (Hrsg.) Das Lesebuch der Deutschen. München 1988, 182 – 189

[157] von Weizsäcker, C.F.: Grosse Physiker – von Aristoteles bis Werner Heisenberg. München 1999

[158] von Weizsäcker, C.F.: Aufbai der Physik. München 1985

[159] von Weizsäcker, C.F.: Große Physiker – von Aristoteles bis Werner Heisenberg. München 1999

[160] Kampen van, N. G. (1991) Determinism and Predictability, *Synthese,* 89 (2) p.273

[161] Varela, F. J., Thompson, Rosch, E.: The Embodied Mind: Cognitive Science and Human Experience. MIT Press 1991

[162] Vohs, K.D. and J.W. Schooler. 2008. The value of believing in free will: Encouraging a belief in determinism increases cheating. *Psychological Science*, 19: 49–54

[163] Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Voodoo correlations in social neuroscience. *Perspectives on psychological Science,* 4(3), 274-290.

[164] von Weizsäcker, C.F., Boone, C.C. (1977) Werner Heisenberg. *CrossCurrents*, Winter 1977-8, 27 (4), 417-425

[165] von Weizsäcker, C.F.: Große Physiker – Von Aristoteles bis Werner Heisenberg. München 1999

[166] Vorberg, D., & Ulrich, R. (1987). Random search with unequal search rates: Serial and parallel generalizations of McGill's model. *Journal of Mathematical Psychology,* 31(1), 1-23

[167] Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. *Nature,* 203(4943), 380-384.

[168] Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), New Horizons in Psychology (pp. 135–151). Harmondsworth, UK: Penguin

[169] Wegner D.M. (2003) The mind's best trick: how we experience conscious will. *Trends in cognitive Sciences*, 7(2w), 65 – 69

[170] Wisniewski D, Deutschländer R, Haynes, J-D (2019) Free will beliefs are better predicted by dualism than determinism beliefs across different cultures. PLoS ONE 14(9): e0221617. https://doi.org/10.1371/journal.pone.0221617 (open source)

[171] Wittgentein, L., Tractatus-logico-philosophicus /Logisch-philosophische Anbhandlung. edition suhrkamp 12, 1963

[172] Wittgenstein, L., Philosophische Untersuchungen, Edition Suhrkamp 2003

[173] Zeilinger, A. The Dance of the photons. 2010

Index

Noë, A., 72

Algorithmus, 109
Antinomien
    der reinen Vernunft, 27
Apperzeption, 32
Argumentationen, rekursive, 81
Aussagen
    apriorische, 25
Automonie
    vernünftige, 26
Autonomie
    vernünftige, 26

behavioral trinity, 101
Bellsche Ungleichung, 99
Bereitschaftspotential, 19
    lateralisiertes, 20

Cartesianisches Theater, 92
causa suis, 5, 40
Commonsense-Dualismus, 67
content externalsim, 47

Determinismus, 91
    Definition, 7
dionysisch-appollinisch, 37
Diskursethik, 57
Dreikörperproblem, 7, 15
Dretske, F., 48
Dualismus
    Common sense, 67
    epistemischer, 52
    ontologischer, 52
Duhem-Quine-These, 100
Dynamic Core-Hypothese, 101
dynamischer Kern, 101
Dämon
    Laplacescher, 75

Einstein, Albert, 11
Ellis, George F.E., 15
Emergenz, 100
Enzephalisationsquotient, 105
EQ, 105
Erklärbarkeit, 7
Erklärungslücke, 73
Evidenzerlebnis, 89
experimentum crucis, 97
extended mind, 47
Externalismus, 46

Fehlschluss
    mereologischer, 92
Freiheit
    Blendwerk von Freiheit, 29

Gerard t'Hooft, 12
Gesetz
    Kausal-, 6
    Zustands-, 6
Gesetze
    probabilistische, 7
    Sukzessions-, 6

Habermas, J., 49
Heidegger, M., 83
Heisenberg, Werner, 9
Homunculus-Fehlschluss, 91

Induktionsproblem, 8
ingtellligible
    Dinge an sich, 32
initiale Aktivität, 41

Kategorienfehler, 107
Kausalität
    aus Freiheit, 26, 27
    probabilistische, 8
Kausalität aus Freiheit, 31
Kausalität der Freiheit, 27
Kern
    dynamischer, 101
komplementäre Verschänkung, 60
Konnektionismus, 62

Lehrsätze, vernünftelnde, 27
Letztbegründungsproblem, 84

Libets Experimente, 18
logischer platonismus, 62

Materialismus
    Cartesianischer, 92
modus tollens, 17
Monismus
    theoretischer, 97

Naturgesetz
    crucial, but troublesome, 14
    Determinismus, 12
    pragmatische Regeln, 90
Naturgesetze
    Ceteris-paribus, 6
    striktes, 6
Neural Darwinism, 101
neuronale Aktivität
    reentrante, 101
Nida-Rümelin, J., 67
Nietzsche, 86
noumenale Welt, 31
Noumenon, 31, 32

Perspektive
    humanistische, 67
Perspektivendualismus, 52, 58
petitio principii, 100
Phasenraum, 76
Philosophie
    experimentelle, 95
Plancks Punkt, 76
platonistisch, 60
primal sketch, 94
Psychophysik, 70

Quantenfluktuationen, 14
Quantenverschränkung, 99

Raum der Gründe, 47
Raum der Ursachen, 47

Scheinproblem
    nach M. Planck, 78
Spontaneität
    absolute, 28
    Idee einer, 27
Sprachspiel, 59, 95
Sprachspiele, 44
Stevenson, 18
subjektives Selbst, 101
Superdeterminismus, 12
synthetisch a priori, 25
Szientophobie, 63

Taboo-Bruch, 18
Third Person View, 102
Turing-Maschine, universelle, 110

Umwelt
    soziolinguistische, 74
Urheberschaft, 56
Ursache
    unverursachte, 26
Urteil
    synthetisches, 25

v. Weizsäcker, 9, 83
van Kampens Theorem, 14
Verantwortung
    Kausalprinzip (Schlick), 88
Vernunft, 27
    Kant, 25
vernünftelnde Lehrsätze, 27
Verstand
    Kant, 25
vis vitae, 102

Wesen
    noumenales, 32
Willkürfreiheit, 67

Zustandsraum, 76
Zustandsübergänge
    rauschinduzierte, 14
Zwei-Charakter-Theorie, 31
Zwei-Welten-Lehre, 31