

oder als indeterminiert und damit unkontrollierbar zu haben, von dem infiniten Regress ganz abgesehen, den der Rekurs auf den Willen hier offensichtlich eröffnen würde.

Der Sonderweg Kants, die Rücknahme des Freiheitsgrundes aus der Welt in die „intelligible“ Sphäre reiner Begriffe, scheidet nicht nur an seinen inneren Inkonsistenzen, sondern vor allem an der Unmöglichkeit, von einer solchen transzendentalen Freiheit aus eine kausale Brücke in die Welt der „Erscheinungen“ zu schlagen. Wer will mag einen Begriff „noumenaler Freiheit“ aus dem sittengesetzlichen „Factum der Vernunft“ deduzieren. Aber „Kausalität aus Freiheit“ *dieser* Provenienz ist eine Fata Morgana.

Das Resultat unserer Analysen zum libertären Inkompatibilismus lässt sich nicht bündiger fassen als in die Formel eines seiner bedeutendsten Vertreter: „Free will remains a mystery“.<sup>125</sup> Ein genauerer Blick zeigt den prinzipiellen Grund: eine Art Hase-und-Igel-Verhältnis zwischen dem Willen und seiner neuronalen Basis. Welche gedankliche Position mittels welcher begrifflichen Feinheiten man für den mentalen Hasen des Wollens oder Entscheidens auch reklamieren möchte, um dessen Freiheit zu behaupten – der Igel des Neurophysiologischen ist immer schon da. Denn ohne ihn, ohne es, ist überhaupt nichts Mentales vorstellbar, ein Wille sowenig wie irgend ein sonstiger Vorgang des Bewusstseins.<sup>126</sup> Daher scheint es unumgänglich, sich bei der Begründung von Freiheit genau damit zu arrangieren.

## V. Die kompatibilistischen Positionen: Grundlagen und Grenzen

Das tun die Kompatibilisten. Sie erkennen sowohl die Unlösbarkeit des Zusammenhangs von Gehirn und „Geist“ als auch einen dabei

125 *van Inwagen* (Anm. 116); gegen ihn wenig überzeugend *Griffith*, Does Free Will Remain a Mystery?, in: *Philosophical Studies* 124 (2005), 261 ff.

126 Genau besehen nicht einmal sein akausales Dasein als Noumenon in einer Welt reiner Vernunft: Auch diese ist nicht ohne Denken und dieses seinerseits nicht ohne seinen „unreinen“ Ursprung im Gehirn vorstellbar.

(irgendwie) maßgeblichen Vorrang des Gehirns an. Der erste gemeinsame Nenner ihrer Theorien ist deshalb die Zurückweisung der Annahme, der Wille sei in irgendeinem Sinn souverän gegenüber seinen neuronalen Grundlagen. Vielmehr deuteten alle verfügbaren empirischen Evidenzen auf einen determinierten Zusammenhang zwischen beiden.<sup>127</sup> Doch bedrohe dies, und das ist der zweite gemeinsame Nenner, weder eine richtig verstandene Freiheit noch die Möglichkeit der Zuschreibung von Verantwortung und Schuld. Will man das überprüfen, so muss man sich zunächst über diesen Zusammenhang von Geist und Gehirn, von Mentalem und Neuronalem Klarheit verschaffen.

### *1. Das Gehirn-Geist-Problem und seine Bedeutung für die Frage der Willensfreiheit*

Wenn „der Geist“, nämlich alle mentalen Phänomene, irgendwie vom Funktionieren des Gehirns abhängig ist, was könnte dann noch die Bezeichnung „freier Wille“ verdienen? Unsere bisherigen Erörterungen haben schon erkennen lassen, dass diese Frage zwei Seiten hat:

(1) *Das Problem der autonomen Urheberschaft:* Kann ein derart vom physikalischen System des Gehirns hervorgebrachter Wille noch als hinreichend autonome Äußerung einer freien Person und damit als Grundlage ihrer Verantwortlichkeit begriffen werden?

(2) *Das Problem der mentalen Verursachung:* Wie kann ein geistiges Phänomen wie der Wille (oder eine Entscheidung) Wirkungen auf die physische Welt des Körpers ausüben, insbesondere Handlungen auslösen und steuern?

#### *1.1 Identitätstheorien*

Viele Philosophen glauben, beide Fragen ließen sich auflösen, wenn man eine bestimmte Form der Identität von Geist und Gehirn anneh-

127 Zur Erinnerung: „determiniert“ meint hier keinen strikt gesetzlichen Determinismus, sondern nur einen Zusammenhang stabiler, naturgegebener Regularitäten, die für Menschen nicht veränderbar sind; vgl. oben II.

me.<sup>128</sup> „Identität“ heißt hier nicht „Ununterscheidbarkeit“ im Sinne der Logik; gemeint ist vielmehr die *Reduzierbarkeit* des Mentalen auf das Neurophysische. Einigen Autoren geht es dabei lediglich um eine epistemische These, nämlich um die logische Reduktion mentalistischer Beschreibungen auf neurophysiologische. Andere vertreten einen weiterreichenden, ontologischen Reduktionismus: Die mentalen Eigenschaften selber seien ohne Rest auf Eigenschaften des Gehirns zurückzuführen, nämlich aus diesen empirisch vollständig erklärbar.<sup>129</sup> Ein einfaches Modell mag das Gemeinte verdeutlichen: Man stelle sich ein Raster mit einer Million winziger Punkte (sog. Pixels) vor, deren jeder entweder hell oder dunkel ist. In ihrer Gesamtheit ergeben sie das naturgetreue Panorama einer Stadt mit einer Fülle spezifischer Details. Dieses Stadtbild und alle seine besonderen Eigenschaften (Formen, Konturen, Linien) existieren selbstverständlich wirklich. Will man aber wissen, wie und warum sie existieren, so wird eine zutreffende Erklärung ihre Existenz vollständig auf das spezifische Arrangement der hellen und dunklen Pixels reduzieren.<sup>130</sup> Reduktionistische Theorien verfolgen also nicht, wie manchmal befürchtet wird, das absurde Anliegen, irgendwelche Dinge der Welt zu „eliminieren“. Was sie eliminieren, ist lediglich die Annahme einer ontologischen Unabhängigkeit dieser Dinge in einer wahren Erklärung der Welt. Genau dieses Anliegen verfolgen Reduktionisten auch für die Erklärung des „Geistes“.

128 Das ist ersichtlich eine starke physikalistische Position (s. oben unter IV.1.2, S. 32 f.). Man muss aber, wie wir noch sehen werden, als Physikalist kein Identitätstheoretiker sein; es gibt auch einen nicht-reduktiven Physikalismus.

129 Aus der unüberschaubaren Flut der Literatur zum Reduktionismus nur *E. Nagel*, *The Structure of Science*, 1961, S. 336 ff. (zur allgemeinen Wissenschaftstheorie); zum psychoneuralen Reduktionismus nur *Crooks*, *Intertheoretic Identification and Mind-Brain Reductionism*, in: *The Journal of Mind and Behavior* 23 (2002), 193 ff. (*contra* Redukt.; mit vier kritischen Kommentaren), sowie *D. Lewis*, *Reduction of Mind*, in: *Guttenplan* (ed.), *A Companion to the Philosophy of Mind*, 1994, S. 412 ff. (*pro* Redukt.); umfassend *Bickle*, *Psychoneural Reduction: The New Wave*, 1998. – Ursprung der Debatte zur psychoneuralen Identität war *Smart*, *Sensations and Brain Processes*, in: *Philosophical Review* 69 (1959), 141 ff. (*pro* Identität).

130 Beispiel von *D. Lewis* (Anm. 129), S. 413.

Es ist leicht zu sehen, dass für Identitätstheoretiker die Frage (1) nach einer autonomen Urheberschaft an der Willensentstehung weitgehend obsolet ist: Unabhängig vom Gehirn sei so etwas nicht vorstellbar, weil der Wille eben mit bestimmten neuronalen Prozessen identisch sei. Doch berühre das eine vernünftig verstandene Autonomie der Person nicht. Schließlich sei es *ihr* Gehirn, dessen Funktionieren ihren Willen hervorbringe. Dass es dabei naturgegebenen Regularien folge, die keiner Willenssteuerung, vielmehr nur der Selbststeuerung des Gehirns unterlägen, mache den Willen nicht unfrei, wenn nur das Gehirn selber ohne manifesten neuro-psychologischen Defekt funktioniere. Im Gegenteil: diese Identität des Willens mit seinen neuronalen „Realisierern“ verbürge gerade das Authentische, Höchstpersönliche seiner Beschaffenheit. Denn alles, was den Willenden selbst ausmache, seine Charakterzüge, Präferenzen, kognitiven, emotionalen und motivationalen Eigenschaften, residiere in der neuronalen Architektur seines Gehirns. In Entscheidungssituationen werde es durch sensorische „Inputs“ umfassend aktiviert, integriert, synchronisiert, und führe genau deshalb zur Bildung eines höchstpersönlichen, zurechenbaren Willens.

Dieses Argument von der Authentizität des Willens, die gerade durch dessen neuronale Determination verbürgt werde, steht ersichtlich nicht nur Identitätstheoretikern offen. Plausibel beanspruchen kann es auch, wer lediglich einen stabilen, neuronal erzeugten und naturgegebenen Regularien folgenden Zusammenhang zwischen Geist und Gehirn annimmt, ohne doch deren Identität zu behaupten. Auf der Hand liegt freilich, dass allein damit die Frage nach Verantwortlichkeit und Schuld nicht zu beantworten ist. Ich komme darauf zurück.

Doch selbst, wenn man sich mit einem so explizierten Freiheitsbegriff zufrieden gibt, ist damit von den beiden oben genannten Grundproblemen allenfalls das erste, das der autonomen Urheberschaft, gelöst. Identitätstheoretiker behaupten nun, ihre Grundthese bringe auch das zweite, das der mentalen Verursachung, zum Verschwinden. Denn wenn das Mentale mit seiner neuronalen Grundlage identisch sei, dann sei die vermeintlich mentale Verursachung, deren Möglichkeit so unbegreiflich erscheine, in Wahrheit eine neuronale Verursachung.

Und an *deren* Möglichkeit als der eines Kausalzusammenhangs zwischen zwei physischen Ereignissen zweifelt niemand; die spezifische Ursächlichkeit neuronaler Erregungsmuster für Körperbewegungen stehe empirisch ohnehin außer Zweifel. Darüber hinaus sprächen für die Identitätstheorie, so behaupten ihre Anhänger, weitere entscheidende Gründe, die von ihrem Potential zur Lösung der beiden genannten Grundfragen unabhängig seien. Insbesondere passe sie von allen verfügbaren Konzeptionen am besten zu unserem wissenschaftlichen Weltbild. So wie die mittlere Molekularbewegung in einer Flüssigkeit deren Temperatur einfach *sei* und diese nicht erst auf geheimnisvollem Wege erzeuge, oder wie ein Blitz eine elektrische Entladung sei und nichts darüber hinaus, und wie ganz allgemein makrophysikalische Eigenschaften eines Gegenstands in dessen atomarer Struktur unmittelbar realisiert seien und nicht erst (zusätzlich) von ihr erzeugt würden – ganz genauso seien mentale Ereignisse unmittelbar identisch mit den sie realisierenden neuronalen Ereignissen.

### 1.1.1 *Mentale Verursachung?*

Stimmt das? Für und gegen die Identitätstheorie sind in den vergangenen Jahrzehnten zahllose Argumente in Stellung gebracht worden, die hier auf sich beruhen müssen.<sup>131</sup> Zwei aus der Gruppe der Gegenargumente erscheinen mir aber durchschlagend. Das erste bestreitet die oben skizzierte Lösung zum Problem der mentalen Verursachung; das zweite verwirft aus einem prinzipiellen Grund die Identitätsannahme im ganzen. Zum letzteren, prinzipiellen Argument komme ich im nächsten Abschnitt (1.1.2). Der Einwand gegen die behauptete Lösung des Kausalitätsarguments lautet so: Ereignisse verursachen andere Ereignisse nicht einfach schlechthin und *tout court*, sondern jeweils vermöge bestimmter Eigenschaften. Diesen komme daher im Rahmen des Verursachungsvorgangs die eigentliche kausale Relevanz zu,

131 Gute Übersicht bei Kim, *Philosophy of Mind*, 1996, S. 47 ff. („Mind as Brain“); ausgezeichnet Walter, *Mentale Verursachung*. Eine Einführung, 2006; instruktiv auch Heil/Mele (eds.), *Mental Causation*, 1993.

während andere Eigenschaften des ursächlichen Ereignisses für dessen Kausalwirkung bedeutungslos sein könnten. Nehmen wir an, die Sopranistin S singt vor einem Fenster im hohen C und in einer bestimmten Lautstärke das Wort „Zerspring!“, woraufhin die Scheibe kaputtgeht.<sup>132</sup> Es ist völlig korrekt zu sagen, das Ereignis des Singens dieses Wortes habe die Scheibe zerspringen lassen. Aber für eine gründliche Beschreibung ist es zu ungenau. Dafür müsste es lauten: Nicht vermöge der Bedeutung des gesungenen Wortes, sondern vermöge der dabei angeschlagenen Tonhöhe und Lautstärke ist das Singen für die Zerstörung der Scheibe kausal geworden; hätte die Sängerin in gleicher Höhe und Lautstärke „Bleib heil!“ gesungen, so wäre die Scheibe genauso kaputtgegangen.

Im gleichen Sinne kausal irrelevant, so lautet nun der Einwand gegen die Identitätstheorie, ist aber (zumindest möglicherweise) das mentale Ereignis des Willens oder Entschlusses zu einer Handlung, auch wenn es mit den neuronalen Vorgängen, durch die es realisiert wird, identisch sein sollte – so wie in unserem Beispiel das Ereignis des Singens des Wortes „Zerspring!“ identisch ist mit dem Ereignis des Singens eines Tons in bestimmter Höhe und Lautstärke, und dennoch nichts Kausales zum Zerspringen der Scheibe beiträgt. Kurz: nicht vermöge seiner mentalen, sondern nur vermöge seiner physischen Eigenschaften könne das psycho-physische Gesamtereignis „Willensentschluss“ kausale Wirkungen in der physischen Welt hervorrufen, z.B. die nachfolgende Handlung einleiten. Daher löse die Identitätstheorie entgegen ihrem Anspruch das Problem der mentalen Verursachung nicht.

Ich halte den Einwand für richtig. Die Frage, wie es möglich ist, dass etwas Mentales, also *per definitionem* Nichtkörperliches, wie der Wille kausale Wirkungen in der Körperwelt hat, erledigt sich nicht dadurch, dass man es einfach mit einem physischen (kausalitätsfähigen) Ereignis identifiziert. Freilich bleibt dann das Problem der mentalen Verursachung in seiner ganzen irritierenden Rätselhaftigkeit

132 Vielzitiertes Veranschaulichungsbeispiel von *Dretske*, *Explaining Behaviour: Reasons in a World of Causes*, 1988, S. 79; sowie *ders.*, *Reasons and Causes*, in: *Philosophical Perspectives* 3 (1989), 1 f.

insgesamt ungelöst zurück. Kapitulierte man vor ihm und erkennt nur physische Verursacher bzw. nur ein Verursachen vermöge physischer Eigenschaften an, dann scheint das gesamte mentale Erleben des Menschen, also das, was jedes individuelle Leben für seinen Inhaber erst lebenswert macht, zur bloßen Begleitmusik des Physischen zu werden: zum Epiphänomen der neuronalen Vorgänge in seinem Gehirn, von diesen erzeugt, aber ohne jede kausale Rückwirkung auf sie (ja auf *irgendetwas* in der Welt).<sup>133</sup> Das ist nicht nur in hohem Maße kontraintuitiv, sondern außerdem ein Gedanke, den viele Philosophen schwer erträglich finden. Wenn Mentales niemals für physische Vorgänge kausal werden kann, dann kann auch kein physischer Vorgang mit etwas Mentalem erklärt werden – also z.B. keine Handlung mit dem Willen des Handelnden. Das erscheint vielen absurd, und müssten sie es dennoch zugestehen, erschiene es ihnen katastrophal. Der amerikanische Philosoph Jerry Fodor hat der entsprechenden Antipathie plastischen Ausdruck gegeben:

„Wenn es nicht buchstäblich wahr ist, dass ich nach etwas greife, *weil* ich es haben möchte, dass ich mich kratze, *weil* es mich juckt, und dass ich etwas sage, *weil* ich es glaube..., wenn nichts davon buchstäblich wahr ist, dann ist praktisch alles falsch, was ich über irgendetwas glaube, und das ist das Ende der Welt.“<sup>134</sup>

Vielleicht ein wenig übertrieben, denn unser subjektives Erleben müsste sich ja allein durch die Einsicht in seinen epiphänomenalen Charakter nicht unbedingt selber verändern. Aber die Mischung aus unwirschem und beschwörendem Tonfall in Fodors Satz steht exemplarisch für viele Abhandlungen, die in den vergangenen Jahrzehnten

133 Einer der ersten Exponenten des Epiphänomenalismus war *Thomas Huxley*, der die kausale Wirkungskraft des Mentalen mit dem Pfeifen einer Dampflokomotive verglich, das von deren Betrieb erzeugt werde, aber keine Rückwirkung auf ihn habe; s. *Huxley*, *On the hypothesis that animals are automata, and its history* (1874), in: *Collected Essays*: Vol I, 1893, S. 199 ff., (240); Huxley begründete seine These für Tiere, hielt sie aber auch für den Menschen für gültig (S. 244); gegen Huxley *W. James*, *Are we automata?*, in: *Mind* 4 (1879), 1 ff.

134 *Fodor*, *Making Mind Matter More*, in: *ders.*, *A Theory of Content and Other Essays*, 1990, S. 137 ff. (156) (Hervorh. von mir); s. auch *McLaughlin*, *Epiphenomenalism*, in: *Guttenplan* (Anm. 129), S. 277 ff.

zum Problem des Epiphänomenalismus erschienen sind: Zumeist wurde dessen Abwegigkeit fraglos vorausgesetzt und nur ein entsprechender Nachweis gesucht. Gleichwohl blieb das Problem gegenwärtig und renitent. Denn wie eine mentale Verursachung denkbar sei, das war und ist trotz der Aura des Absurden, die den Epiphenomenalismus umgibt, so rätselhaft wie eh und je. Eine rundum überzeugende Lösung des Problems ist auch heute, nach unzähligen Erklärungsversuchen, nicht in Sicht. Anfang 2006 erschien im renommierten „Journal of Consciousness Studies“ ein spezielles Themenheft unter dem Titel „Epiphenomenalism“, dem die Herausgeber in ihrem Vorwort die sprechende Ergänzung „Dead End or Way Out?“ beifügten. Das Meinungsspektrum der acht Beiträge spiegelt genau diese Frage als aktuelles und ungelöstes Problem der Philosophie des Geistes wider.<sup>135</sup> Und damit zugleich das fortbestehende Rätsel seines Zwillingproblems, der Frage nach einer Möglichkeit mentaler Verursachung.

Vor diesem Hintergrund darf man es erstaunlich finden, dass im Strafrecht das Problem der mentalen Verursachung bislang so gut wie vollständig ignoriert worden ist.<sup>136</sup> Ein Freiheitsproblem sieht man allenfalls in der Entstehung, nicht aber in der möglichen Wirkung des Willens. Danach gefragt wird deshalb regelmäßig erst im Rahmen der Schuldprüfung, nicht dagegen bereits beim Handlungsmerkmal im Tatbestand. Aber schon dort gibt es eines: das des „willensgesteuerten“ Verhaltens. Dessen Vorliegen, Grundbedingung einer Zuständigkeit des Strafrechts überhaupt, wird für den Normalfall bewussten menschlichen Handelns regelmäßig ohne weiteres bejaht. Das könnte sich als voreilig erweisen. Gibt es so etwas überhaupt? Was hieße es, wenn die Frage verneint werden müsste? Wäre ein *nicht* handlungssteuernder Wille oder eine nicht willensgesteuerte Handlung „frei“?

135 *Pauen/Staudacher/Walter* (eds.), Epiphenomenalism, *Journal of Consciousness Studies* 13, No 1-2, 2006; zur Offenheit des Rätsels (nach 250 Seiten gründlicher Analyse) auch *Walter* (Anm. 131), S. 258 f.

136 Eine Ausnahme ist *Dettefsen* (Anm. 10), S. 309 ff., 324 f., die (allerdings ohne Blick auf die philosophische Debatte) das Problem der mentalen Verursachung deutlich sieht, deren Möglichkeit – neben einer vorausgesetzten neuronalen Determination – skeptisch beurteilt, gleichwohl aber den Epiphenomenalismus ablehnt.



Und bejaht man Handlungssteuerung durch den Willen dennoch: wie genau macht er das? Wie kann er als etwas Immaterielles materielle Wirkungen erzeugen ohne den Energieerhaltungssatz zu verletzen? Und wenn er's könnte: warum braucht er gleichwohl stets ein funktionierendes Gehirn, um eine körperliche Handlung auszulösen? Wie „steuert“ er diese dann, wenn doch das Funktionieren des unmittelbar auslösenden Gehirns naturgesetzlichen Regularien folgt? Und warum kann mein Geist, wenn er wirklich *so etwas* kann, immer nur auf mein Gehirn und auf kein anderes einwirken?

Auf keine dieser Fragen gibt es eine überzeugende Antwort *pro libertate*. Nicht nur gibt es keine Spur eines empirischen Indizes für unmittelbare Wirkungen des Geistes auf das Gehirn, es gibt im Gegenteil gute Gründe für die Annahme, dass solche Wirkungen unmöglich sind. Das schließt es ja keineswegs aus, das Strafrecht dennoch für tatbestandliches Handeln zuständig zu machen. Man wird aber fragen dürfen, ob die Strafrechtsdogmatik ihre schlafwandlerische Selbstsicherheit im Umgang mit dem Grundbegriff der „willensgesteuerten Handlung“ nicht doch vorübergehend gegen eine aufgewecktere Skepsis eintauschen sollte. Nicht jeder dogmatische Begriff, der in seiner Buchstäblichkeit keinen fassbaren Inhalt (mehr) hat, muss ja sofort abgeschafft werden, und der der „Willenssteuerung“ wäre bei weitem nicht der erste und einzige dieser Art im geläufigen Inventar des Strafrechts. Was aber nötig wäre, ist eine größere Klarheit darüber, was daran sachlich glaubigster Sinn und was bloß noch überlieferte Metaphorik ist.

### 1.1.2 *Qualia, oder: Was Mary nicht wissen konnte*

Der zweite Einwand gegen die Identitätstheorie ist radikaler: Ein reduktiver (identitärer) Physikalismus, der alle Erscheinungen der Welt für (im Prinzip) erklärbar hält durch ihre analytische Rückführung auf ein letztes, regelmäßig molekulares oder atomares Fundament des Physischen, sei falsch. Wohl könne man mittlerweile die meisten der traditionell rätselhaften Phänomene, z.B. die Entstehung von Leben aus „toten“ materiellen Bausteinen, physikalistisch-reduktiv erklären, nicht anders als den Druck eines Gases durch dessen mittlere Mole-

külgeschwindigkeit.<sup>137</sup> Aber eines dieser Probleme bleibe für jeden denkbaren wissenschaftlichen Physikalismus unerreichbar: die Frage nach dem Zusammenhang von Geist und Gehirn. Das lasse sich zeigen: „Ein Baum im Wald stürzt um. Ist jemand in der Nähe und hört dies, dann gibt es das Geräusch eines stürzenden Baumes; andernfalls gibt es nur Schallwellen. Wenn nun jemand den Baum stürzen hört, wo genau steckt dann dieser Unterschied – das subjektive Erfahren des Geräusches? Öffne den Kopf des Hörenden, untersuche sein Gehirn in jeder einzelnen Windung; du wirst weder das Geräusch noch die Erfahrung des Geräuschhörens finden.“<sup>138</sup>

Hier wird die Idee plastisch. Nicht auf Physi(kali)sches reduzierbar sei der spezifisch subjektive Charakter des inneren, mentalen Erlebens. Der starke (reduktive) Physikalismus sei daher falsch.<sup>139</sup> Eine Beschreibung des gesamten Universums müsste als ein irreduzibles Grundelement im Bau der Welt das subjektive Erleben des Mentalen enthalten. Anders gewendet: eine rein physikalische Beschreibung des Inventars der Welt wäre notwendig unvollständig.

Traditionell unterscheidet man zwei verschiedene Typen von mentalen Zuständen: intentionale und phänomenale.<sup>140</sup> Zum ersteren Typus gehören innere Zustände „über etwas“ oder „gerichtet auf etwas“, wie nachdenken, wünschen, hoffen, lieben, hassen, etc.; zu ihnen gehört auch, leicht erkennbar, der Wille oder Entschluss zu einem bestimm-

137 Dazu *McLaughlin*, *The Rise and Fall of British Emergentism*, in: *Beckermann/Flohr/Kim* (eds.), *Emergence or Reduction?*, 1992, S. 49 ff.; instruktiv auch *Papineau* (Anm. 41). – Die manchmal in diesem Kontext genannten Inhalte des „objektiven Geistes“ oder *Karl Poppers* „dritter Welt“ (s. *ders.*, *Objektive Erkenntnis*, 1973, S. 88 f., *passim*), also etwa Normen, Religionen, Kulturwerte etc., sind übrigens ohne weiteres reduktiv-physikalistisch erklärbar. Was sich mit einer solchen Erklärung anfangen lässt (und v.a. was nicht), ist eine andere Frage.

138 *Glymour*, *A Mind is a Terrible Thing to Waste – Critical Notice*: Jaegwon Kim, *Mind in a Physical World*, in: *Philosophy of Science* 66 (1999), 455 ff.

139 *Thomas Nagel*, der mit seinem berühmten Aufsatz „What Is It Like to Be a Bat?“ (*Philosophical Review* 83 [1974], 435 ff.) das physikalistisch Unerklärbare des subjektiven Erlebens zum Thema gemacht hat, hielt freilich am Physikalismus fest, weil er eine physikalistisch-reduktive Erklärung anhand künftiger wissenschaftlicher Einsichten nicht ausschließen wollte.

140 S. bereits oben, zu und in Anm. 54.

ten Handeln. Der zweite Typus umfasst u. a. alle Sinneseindrücke: Empfindungen von Hitze, Kälte, Schmerz, Eindrücke beim Sehen, Hören, Riechen, Schmecken, etc. Die Begründungsstrategien gegen den reduktiven Physikalismus stützen sich meist allein (oder hauptsächlich) auf das Irreduzible der phänomenalen, nicht aber der intentionalen Zustände. Berühmt geworden ist ein entsprechendes Gedankenexperiment des australischen Philosophen Frank Jackson aus dem Jahr 1982:

Mary, eine exzellente Neurowissenschaftlerin, ist von Geburt an gezwungen, subjektiv in einer schwarz-weißen Welt zu leben. Denn sie hat einen Augendefekt, der jedes farbliche Sehen ausschließt. Nach langen Studien verfügt sie – so sei unterstellt – über *alle* physikalischen und neurobiologischen Informationen über *sämtliche* Vorgänge, die in einem Menschenkopf beim Sehen der Farbe Rot stattfinden: von der Reaktion der Retina auf das eintreffende Licht bis zu den neuronal-molekularen Vorgängen und den „feuernden“ Synapsen bei der Verarbeitung im Gehirn. Alle diese Vorgänge kennt sie auch aus züglicher Beobachtung in Experimenten. Irgendwann kann Marys Augendefekt medizinisch behoben werden. Nun *sieht* sie zum erstenmal selbst die Farbe Rot.<sup>141</sup>

Jacksons Pointe ist, dass Mary nun etwas prinzipiell Neues lernt, etwas, das ihr kein noch so vollständiges Wissen über alle neurophysiologischen Vorgänge des Rotsehens vorher habe vermitteln können. Dieses kategorial Neue ist das Subjektiv-Qualitative oder Phänomenale des Sinneseindrucks (technisch: „Quale“, im Plural „Qualia“). Kenntnis davon sei nur mit dem Haben des mentalen Zustands selber, aber mit keinem Wissen über dessen neuronale Grundlagen zu erlangen. Der reduktive bzw. identitäre Physikalismus sei daher falsch.

Jacksons Gedankenexperiment hat eine Flut von Aufsätzen, zustimmenden wie ablehnenden, zum Problem der Qualia und des Physikalismus ausgelöst.<sup>142</sup> Diese Diskussion, die profunde Einsichten in das

141 Jackson, Epiphenomenal Qualia, in: *Philosophical Quarterly* 32 (1982), 127 ff.; ich habe das Szenario des Gedankenexperiments etwas verändert; das Jacksonsche ist deutlich bizarrer als das hier gewählte (eine solche vollständige Farbenblindheit, den sog. Achromatismus, gibt es tatsächlich); die gedankliche Pointe ist identisch.

142 Gute Dokumentation der Debatte in Ludlow/Nagasawa/Stoljar, *There's Something*

Problem erbracht, es aber nicht annähernd konsensfähig gelöst hat, muss hier auf sich beruhen. Zumindest in ihrem Kern halte ich Jacksons Überlegung für richtig.<sup>143</sup> Und in eben diesem Ausmaß erscheint der reduktive Physikalismus daher falsch.

Was bedeutet das für unsere Frage nach der Freiheit des Willens? Jedenfalls dies: Wir haben zumindest einige gute Gründe, Willensakte nicht für identisch mit ihren neuronalen Grundlagen zu halten. Zwar demonstriert Marys Beispiel dies unmittelbar nur für phänomenale Zustände (Qualia), nicht aber für intentionale (wie den Entschluss zu einer Handlung). Doch ist schon zweifelhaft, ob Qualia und intentionale Zustände stets klar voneinander unterschieden werden können. Jedenfalls erscheint es nicht unvernünftig, aus dem Umstand, dass Qualia nicht auf ihre neurophysiologischen Grundlagen reduzierbar sind, den Schluss zu ziehen, dass *kein* subjektiv-mentaler Zustand dies ist.<sup>144</sup> Der reduktive Physikalismus gibt uns daher auf unsere beiden Grundfragen, die nach einem hinreichend autonomen Ursprung des Willens und die nach dessen kausaler Wirkung auf Einleitung und Kontrolle von Handlungen, keine überzeugende Antwort. Und ebenso

About Mary, 2004. Angegriffen hat man v.a. Jacksons „knowledge argument“, die Behauptung, Mary erwerbe im Moment ihres erstmaligen Rotsehens eine bislang nicht vorhandene *Kenntnis von etwas*: einer (neuen) Tatsache. Ein phänomenaler Sinneseindruck sei keine Tatsachenkenntnis; Mary lerne also nichts Neues über die Welt; sie werde nur mit einem veränderten Modus des Sehens vertraut; der reduktive Physikalismus sei daher nicht widerlegt. Mir erscheint der Einwand weder trifftig noch richtig: Mary lernt, dass *andere* Menschen beim Sehen von Rot etwas Ähnliches empfinden wie sie selbst im Moment (gewiss eine Tatsache). Überzeugend für ein (modifiziertes) „knowledge argument“ M. Nida-Rümelin, On Belief About Experiences. An Epistemological Distinction Applied to the Knowledge Argument..., in: Philosophy and Phenomenological Research 58 (1998), 51 ff.

- 143 Jackson selbst tut dies übrigens nicht mehr; er ist zum reduktiven Physikalismus zurückgekehrt, s. *ders.*, Postscript on Qualia, in: Ludlow et al. (Anm. 142), S. 417 ff., sowie *ders.*, Mind and Illusion, ebd., S. 421 ff.
- 144 Ist z.B. das innere Ringen um eine schwere Entscheidung nur als intentionaler Zustand fassbar, oder „fühlt es sich“ nicht doch auch „irgendwie an“? Dass man Qualia und intentionale Zustände trennen und jedenfalls die letzteren (z.B. den „Willen“) physikalistisch-reduktiv erklären könne, meint aber Kim, Mind in a Physical World, 1998, S. 118 ff.; *ders.*, The Causal Efficacy of Consciousness, in: *Velmans/Schneider* (Anm. 62), S. 406 ff.

wenig überzeugend ist sein Vorschlag, diese Fragen hinter immer genaueren physikalistischen Beschreibungen der neuronalen Vorgänge im Gehirn eines Handelnden gänzlich verschwinden zu lassen.

### 1.2 Andere Lösungen: Moderate Dualisten – Emergenz? Supervenienz?

Wer den vollständigen Reduktionismus, die Identität von Geist und Gehirn, verneint, kann dennoch Physikalist sein und daher die These einer souveränen Unabhängigkeit des Willens vom Gehirn ganz genauso verwerfen. Das tut der nicht-reduktive Physikalismus. Auch er hat verschiedene Spielarten.<sup>145</sup> Ihre gemeinsame Grundthese, dass sich das Mentale vom Neuronalen irreduzibel unterscheidet, führt freilich (erstens) zwingend zu irgendeiner Form von Dualismus zwischen „Leib und Seele“ und (zweitens) zu der Notwendigkeit, den postulierten besonderen Zusammenhang zwischen Gehirn und Geist wenigstens begrifflich plausibel zu machen.

Der Dualismus-Vorwurf ist zwar beliebt, aber langweilig.<sup>146</sup> Dass jemand, der einen Identitätsmonismus ablehnt, irgendeine Form von Dualismus akzeptieren muss, ist trivial und nicht im mindesten zu beanstanden. Viele setzen einen Monismus der Substanz (Gehirn) und einen Dualismus der Eigenschaften dieser Substanz (neuronale / mentale) voraus. An Benennungen hängt nicht viel; ein solcher Eigenschaftsdualismus ist jedenfalls kein unplausibler Ausgangspunkt für weitere Klärungen.

145 Ausgezeichnete Sammelbände dazu sind *Beckermann/Flohr/Kim* (Anm. 137); *Aler/Walter* (eds.), *Phenomenal Concepts and Phenomenal Knowledge. New Essays on Consciousness and Physicalism*, 2007; *Freeman* (ed.), *Consciousness and its Place in Nature*, 2006 (zu der verblüffenden Theorie *G. Strawsons*, dass Physikalisten das Mentale nur dann erklären könnten, wenn sie zugleich einen Panpsychismus annähmen).

146 Er knüpft an den diskreditierten Substanzdualismus von „res cogitans“ und „res extensa“ bei *Descartes* an (s. Anm. 30); einen solchen muss (sollte) aber kein Dualist heute noch vertreten. Für einen modernen, wissenschaftlich aufgeklärten Dualismus *Chalmers*, *Naturalistic Dualism*, in: *Velmans/Schneider* (Anm. 62), S. 359 ff.