# Final-mentale Kausalität
## Eine neurophilosophische Untersuchung

Inaugural-Dissertation zur Erlangung des
Doktorgrades der Philosophie (Dr. phil.)
durch die Philosophische Fakultät der
Heinrich-Heine-Universität Düsseldorf

vorgelegt von
Kai-Nils Eicke
aus Düsseldorf

Betreuer:
Priv.-Doz. Dr. Georg Northoff

Düsseldorf 2009

**D61**

Diese Dissertation wurde durch den Deutschen Akademischen Austauschdienst unterstützt und gefördert. Hierfür möchte ich mich ausdrücklich bedanken. Meinem Doktorvater, PD Dr. Georg Northoff, möchte ich für die Möglichkeit danken, diese Dissertation unter seiner Betreuung zu realisieren. Die damit verbundenen Gespräche waren für mich stets hilfreich. Ich danke Dr. Felix ermpohl für seine engagierte Beratung und Diskussion hinsichtlich neurowissenschaftlicher Fragen. Für ihren unermüdlichen Einsatz im Lektorat danke ich Thomas Fischer und Dr. Wolfgang Wietzker. In technischen Fragen war mir Tobias Heide eine wichtige Hilfe, auch hierfür mein Dank. Meinen Eltern und meiner Schwester danke ich für ihre bedeutende moralische Unterstützung. Für ihr aufrichtiges Interesse möchte ich den vielen Freunden, Bekannten und Kollegen danken, deren Zuspruch mich während meiner Doktorarbeit begleitet hat. Mein besonderer Dank gilt meiner Freundin Esther Wegener für ihre verständnisvolle und geduldige Unterstützung.

*Anders eilt das Geschoss, anders der Geist. Der Geist aber eilt, auch wenn er behutsam zu Werke geht und wenn er seine Betrachtung von allen Seiten macht, nichtsdestoweniger geradeaus und auf sein Ziel los.*

*(Marc Aurel: Selbstbetrachtungen. Achtes Buch, 60)*

**Inhalt**


**Kapitel 1**

**Thematische und methodische Einführung** 1
1.1 Einordnung des Begriffs der mentalen Kausalität 1
1.2 Neurophilosophie als interdisziplinärer Lösungsansatz 6
1.3 Entwicklung der neurophilosophischen Methodik 10
1.4 Aufbau und Vorgehensweise der Arbeit 19

**Kapitel 2**

**Die Entwicklung der zeitgenössischen philosophischen Diskussion** 24
2.1 Identitätstheorie und mentale Kausalität 28
2.1.1 Bestimmung der psychophysischen Identität 28
2.1.2 Das kausale Argument für psychophysische Identität 31
2.1.3 Typenidentität in der Kritik 35
2.2 Funktionalismus und mentale Kausalität 38
2.2.1 Die funktionale Rolle als eine kausale Rolle 39
2.2.2 Die Realisierung funktionaler Zustände 41
2.2.3 Die funktionale Definition für kausale Rollen 43
2.2.4 Funktionalismus und Physikalismus 45
2.2.5 Funktionalismus in der Kritik 47
2.3 Anomaler Monismus und mentale Kausalität 51
2.3.1 Entwicklung der Ereignisontologie 52
2.3.2 Tokenidentität als Konsequenz 54
2.3.3 Tokenidentität in der Kritik 58
2.4 Supervenienztheorie und mentale Kausalität 63
2.4.1 Kovarianz und starke Supervenienz 64
2.4.2 Abhängigkeit und Nicht-Reduzierbarkeit 66
2.4.3 Kausaler Ausschluss mentaler Eigenschaften 68
2.4.4 Kims Version der Supervenienz in der Kritik 71

**Kapitel 3**

**Voraussetzungen für mentale Kausalität** 76
3.1 Das ontologisches Grundmodell 78
3.1.1 Ereignisse als ontologische Grundeinheit 80
3.1.2 Eigenschaften und mentale Kausalität 82
3.1.3 Die Realisierung von Eigenschaften im Rahmen des Schichten-Modells 84
3.2 Die Unterscheidung des mentalen und des physischen Bereichs 87
3.2.1 Merkmale des mentalen Bereichs 89
3.2.2 Merkmale des physischen Bereichs 91
3.2.3 Das Gehirn als mental-physisches Organ 93
3.2.4 Die Erklärungslücke 95
3.3 Die kausale Geschlossenheit der Physik 98
3.3.1 Theorienreduktion und induktive Generalisierung 100
3.3.2 Die physikalische Geschlossenheit in der Kritik 102
3.4 Das Argument des kausalen Ausschlusses 105

3.4.1 Begründung des Ausschlussarguments 105
3.4.2 Ablehnung systematischer Mehrfachverursachung 108
3.4.3 Das Ausschlussargument in der Kritik 110
3.5 Die Voraussetzung der strikten Gesetze 114
3.5.1 Geltungsbereich von Rationalitätsprinzipien 114
3.5.2 Nicht-strikte Gesetze als Alternative 116
3.6 Möglichkeiten für mentale Kausalität 121
3.6.1 Die Gefahr des Epiphänomenalismus 121
3.6.2 Die Voraussetzungen in der Kritik 125
3.6.3 Homogenität als implizite Voraussetzung 129

**Kapitel 4**

**Kriterien für mentale Kausalität** 134
4.1 Das Kriterium der Regularität 140
4.1.1 Vertikale Regularität 141
4.1.2 Horizontale Regularität und mentale Kausalität 144
4.2 Das Kriterium der Multikausalität 148
4.2.1 Das Prinzip der Multirealisierbarkeit 148
4.2.2 Multikausale Relationen und mentale Kausalität 153
4.3 Das Kriterium des Inhalts 156
4.3.1 Die Language of Thought Hypothese 158
4.3.2 Internalismus versus Externalismus 161
4.3.3 Inhalt und mentale Kausalität 164
4.4 Das Kriterium der Relationalität 168
4.4.1 Intrinsische versus extrinsische Eigenschaften 169
4.4.2 Relationalität und mentale Kausalität 172
4.5 Das Kriterium der Intentionalität 177
4.5.1 Intentionalität unter teleologischen Gesichtspunkten 179
4.5.2 Intentionalität und mentale Kausalität 183
4.6 Das Kriterium der Proportionalität 189
4.6.1 Das Prinzip der Proportionalität 190
4.6.2 Proportionalität und mentale Kausalität 195

**Kapitel 5**

**Neuronale Prozessierung und Repräsentation** 200
5.1 Das Paradigma der emotionalen Aufmerksamkeit 204
5.1.1 Untersuchungsmethode und Datensammlung 206
5.1.2 Abgrenzung emotionaler Aufmerksamkeit 209
5.1.3 Diskussion der Ergebnisse 213
5.2 Das Paradigma der emotionalen Beurteilung 219
5.2.1 Untersuchungsmethode und Datensammlung 220
5.2.2 Unterscheidung neuronaler Repräsentationen 224
5.2.3 Diskussion der Ergebnisse 227

**Kapitel 6**

**Vergleich der Kriterien und Paradigmen** 234
6.1 Emotionale Aufmerksamkeit und Regularität 235
6.1.1 Gesetze für neuronale Aktivität 236
6.1.2 Gesetze für psychische Funktionen 239

6.1.3 Korrelationsgesetze 240
6.1.4 Realisierungsrelation als Alternative 243
6.2 Emotionale Aufmerksamkeit und Multikausalität 246
6.2.1 Multikausalität auf der psychologischen Ebene 247
6.2.2 Multikausalität auf der neuronalen Ebene 249
6.2.3 Multikausalität innerhalb der Korrelation 252
6.3 Emotionale Aufmerksamkeit und Inhalt 258
6.3.1 Grenzen repräsentationaler Theorien 258
6.3.2 Empirische Validität und enger Inhalt 262
6.3.3 Naturalisierung ohne Vereinfachung 265
6.4 Emotionale Aufmerksamkeit und Relationalität 270
6.4.1 Extrinsische Eigenschaften der subjektiven Ebene 271
6.4.2 Intrinsische Eigenschaften der neuronalen Ebene 272
6.4.3 Binnendifferenzierung neuronaler Eigenschaften 274
6.4.4 Wechselseitige Bedingung intrinsischer und extrinsischer Eigenschaften 276
6.5 Emotionale Aufmerksamkeit und Intentionalität 279
6.5.1 Intentionale Verursachung als bidirektionales Kausalitätsmodell 280
6.5.2 Normativität als Herausforderung des neuronalen Bereichs 284
6.5.3 Dispositionen als Ausdruck neuronaler Intentionalität 287
6.6 Emotionale Aufmerksamkeit und Proportionalität 291
6.6.1 Korrelation zweier Erklärungsbereiche 291
6.6.2 Ausdifferenzierung des Abstraktionsgrades 293
6.6.3 Übertragungsmöglichkeiten für das Paradigma 295
6.6.4 Proportionale Kausalität und Eigenschaftspluralismus 299

**Kapitel 7**

**Entwicklung der neurophilosophischen Hypothese** 304
7.1 Die aristotelische Ursachenlehre 308
7.1.1 Hylëmorphismus als ontologischer Hintergrund 309
7.1.2 Vier Ursachen als einheitliches Phänomen 311
7.1.3 Möglichkeiten für mentale Kausalität 314
7.2 Das Modell der finalen Kausalität 319
7.2.1 Finale Kausalität als Ereignisontologie 324
7.2.2 Finale Kausalität und Regularität 330
7.2.3 Finale Kausalität und Multikausalität 334
7.2.4 Finale Kausalität und Inhalt 339
7.2.5 Finale Kausalität und Relationalität 345
7.2.6 Finale Kausalität und Intentionalität 349
7.2.7 Finale Kausalität und Proportionalität 354

**Kapitel 8**

**Schlussbemerkung und Ausblick** 361

**Literaturverzeichnis** 375

## Kapitel 1
## Thematische und methodische Einführung

### 1.1 Einordnung des Begriffs der mentalen Kausalität

Versucht der philosophisch Vorgebildete beispielsweise auf einer Cocktailparty mit einem Menschen mit gesundem Menschenverstand die Frage der mentalen Kausalität zu erörtern, befindet er sich schnell in einer schwierigen Situation. Denn er kann den philosophisch Unbelasteten nicht mit den gängigen Begriffen der philosophischen Diskussion konfrontieren, sondern muss die gewohnte fachwissenschaftliche Terminologie verlassen. Erste Schwierigkeiten ergeben sich beispielsweise in der Differenzierung zwischen „mentalen" und „physischen Eigenschaften" und ihren „Instanziierungen". Ganz zu schweigen von subtilen theoretischen Überlegungen, die z. B. mit Fragen der „physikalischen Geschlossenheit", der „Mehrfachverursachung", der „Multirealisierbarkeit", der „Intentionalität" oder – nicht zuletzt – der „externen Inhalte" verbunden sind. Solche Begriffe lassen sich voraussetzungslos kaum erklären. Die Frage der mentalen Kausalität eignet sich schlecht für den Smalltalk auf Cocktailpartys.

Überhaupt muss der unvoreingenommene und wenig informierte, wenngleich interessierte Zuhörer meist erst einmal für die Problemlage sensibilisiert werden. Dem Laien ist die epistemische wie ontologische Kluft zwischen Leib und Seele bzw. Gehirn und Bewusstsein im alltäglichen Erleben wenig oder selten präsent. Er erfährt sich durchweg als Subjekt seiner Handlung innerhalb seiner Erste-Person-Perspektive und erlebt als solches täglich, dass seine Gedanken, Wünsche oder Pläne tatsächlich sein Verhalten beeinflussen. Er erlebt sich insofern als eine Einheit, denn Denken und Handeln entsprechen einander anscheinend. Fehlentscheidungen oder Missgeschicke werden oft in diesem Sinne verstanden, auch wenn hier vielleicht ein erster Riss durch die selbstverständliche Subjektivität verläuft. Dass dieser Selbstwahrnehmung ein „Fehlschluss aus Abneigung" (englisch: „antipathetic fallacy"; vgl. Papineau 2004, 169ff.) zugrunde liegt, würde dem Menschen mit gesundem Menschenverstand von daher nicht ohne weiteres einfallen und wäre ihm so zunächst wenig plausibel.

Wer allerdings öffentliche Beiträge zur Gehirnforschung beispielsweise in Zeitschriften oder Zeitungen verfolgt, dem ist der Gedanke wenig befremdlich, dass das Gehirn die Grundlage für unser Denken, unsere Gedanken bildet. Implizit geht der „normale Mensch" vermutlich – grob gesagt – von einer Identität zwischen Geist und Gehirn aus, vielleicht auch verbunden mit einer emergentistischen Idee, dass das Gehirn Bewusstsein produziert. Zumindest scheinen ihm die Widersprüche, Fallstricke und ausgefochtenen Sträuße um ein tragfähiges Modell einer Körper-Geist-Identität verborgen zu sein. Am Ende einer solchen Diskussion mit einem „normalen Mann mit gesundem Menschenverstand" kann der Philosoph schon zweifeln, ob die Frage der

mentalen Kausalität ein echtes Problem oder lediglich ein philosophisches Scheinproblem ist (vgl. Rorty 1979). Hier beginnt die eigentliche philosophische Arbeit: Zu bestimmen, was sich begrifflich hinter der Frage der mentalen Kausalität genau verbirgt, die Voraussetzungen, unter denen diese Frage diskutiert wird, offenzulegen und zu prüfen. Den Begriff der „mentale Kausalität" darzustellen, ist insofern, auch im Rahmen dieser Arbeit, zunächst eine definitorische Aufgabe. Dabei soll auch herausgefunden werden, was eine Kausalität des Geistes speziell leisten kann.

Traditionell nimmt auch die zeitgenössische Diskussion auf Descartes' Substanzdualismus als Ausgangs- und Reibungspunkt für die Frage der mentalen Kausalität Bezug. Der französische Philosoph hatte sich bereits im 17. Jahrhundert mit der Frage auseinander gesetzt, wie ein Einwirken der Seele auf den Körper möglich ist. Obwohl Descartes von einem mechanisierten Weltbild ausgeht, gesteht er dem Geist volle kausale Kraft auf den Körper und dem Körper kausale Kraft auf den Geist zu. Zwischen Geist und Körper unterscheidet er als res cogitans und res extensa (vgl. Descartes 1977, 48ff.). Zwar charakterisiert der Geist den Menschen als denkendes Wesen, doch ist die geistige Substanz räumlich unausgedehnt, immateriell. Der Körper dagegen ist nicht denkend, völlig materiell und funktioniert gewissermaßen wie eine Maschine. Geist und Körper sind nach Descartes zwei verschiedene Substanzen, die durch kein Attribut miteinander verbunden sind (vgl. ebd., 141f.). Dennoch nimmt er an, dass eine Wechselwirkung zwischen Körper und Geist möglich ist. Als Ort des kausalen Austauschs bestimmt er die Zirbeldrüse, der als „Sitz der Seele" im Gehirn gewissermaßen eine Relaisfunktion bei der Leib-Seele-Interaktion zukommt (vgl. Descartes 1911). Seiner Vorstellung nach wird die Zirbeldrüse von animalischen Geistern bewegt, so dass sowohl die Seele Zugriff auf den Körper erhält als auch über entsprechende Empfindungen und Wahrnehmungen informiert wird:

> Sodann bemerke ich, dass der Geist nicht von allen Körperteilen unmittelbar beeinflusst wird, sondern nur vom Gehirn, oder vielleicht sogar nur von einem ganz winzigen Teile desselben [...]. Werden z.B. die Fußnerven heftig und ungewohnt bewegt, so gibt jene durch das Rückenmark bis zu den inneren Gehirnteilen dringende Bewegung dort dem Geiste ein Zeichen, etwas zu empfinden, nämlich einen Schmerz und diesen so, als sei er im Fuße; dieser veranlasst den Geist, seine Ursache, als etwas für den Fuß gefährliches, nach Möglichkeit zu entfernen. (Descartes 1977, 155f.)

Den Bemühungen des französischen Philosophen zum Trotz bleibt das Problem des Austauschs zwischen Körper und Geist bestehen. Die ausgedehnte Substanz wird der unausgedehnten Substanz unverbunden gegenübergestellt, letztendlich bleibt Descartes' postulierter Substanzdualismus unaufgelöst. Wie aber soll etwas Immaterielles mit einem materiellen Objekt interagieren können? Das Konzept der Wechselwirkung passt in Descartes' ansonsten mechanistische Theorie nicht hinein. Die Vorstellung einer solchen Interaktion bereitet Schwierigkeiten. Und auch aus heutiger Sicht erscheint dieses Modell unplausibel. Wie sollte ein nicht-materieller Geist in die

elektrochemischen Prozesse im Gehirn einwirken, um Nervensignale zu aktivieren? Um die Bewegungen unserer Gliedmaßen zu erklären, scheint eine substanzialisierte Seele nicht notwendig oder hilfreich.

Konsequenterweise behandelt die zeitgenössische philosophische Diskussion nicht mehr die Frage, wie zwei verschiedene Substanzen miteinander interagieren können. Das vorherrschende physikalische Weltbild lässt im Grunde keine eigenständige geistige Substanz mehr zu. Das Problem einer kausalen Interaktion zweier unterschiedlicher Substanzen hat sich hin zu der Überlegung verschoben, inwiefern Bewusstsein existiert oder durch das Gehirn realisiert wird (vgl. Heil 1998, 4). Die alltäglich wahrgenommene Differenz zwischen der inneren mentalen Welt und der äußeren materiellen bleibt allerdings bestehen:

> Minds are non-material entities: entities with properties not possessed by any material object. Minds bear intimate relations to material objects, perhaps, and especially intimate relations to brains. (ebd., 5)

Ob „minds" tatsächlich keine materiellen Eigenschaften besitzen und somit aus der physischen Welt herausfallen, ist auf der theoretischen Ebene allerdings umstritten. Für die Frage der mentalen Kausalität ist daher die Relation zwischen Bewusstsein und Gehirn von Interesse. Konkret: Wie verhalten sich mentale und physische Eigenschaften bzw. mentale und physische Ereignisse zueinander? Auf der epistemischen Ebene gibt es eine Erklärungslücke zwischen dem mentalen und dem physischen Bereich, die jeweils unterschiedlichen Beschreibungen unterliegen (vgl. Levine 1983, 1993). Beide Bereiche sind durch ihnen eigene Merkmale gekennzeichnet. Deshalb gibt es eine Lücke, wenn beispielsweise erklärt werden soll, wie aus physiologischer Hirnaktivität subjektive Empfindungen resultieren oder wie Gedanken eine körperliche Bewegung verursachen können. Ob der begriffliche Dualismus auf eine ontologische Kluft verweist, ist dabei umstritten. Je nach Position werden mentale und physische Eigenschaften als identisch aufgefasst (vgl. z. B. Feigl 1958), werden erstere auf letztere reduziert (vgl. z. B. Block 1990; Kim 1989; Ryle 1949) oder gänzlich eliminiert (vgl. z. B. Churchland 1981; Rorty 1965, 1979). Dualistische Positionen werden inzwischen von einer Minderheit vertreten (vgl. z. B. Eccles 1994; Meixner 2004; Popper/Eccles 1977; Swinburne 1994). [1] Einen weiteren Standpunkt vertreten die so genannten Nicht-Reduktivisten. Sie stellen mentale Phänomene in einen materialistischen Kontext, beharren aber darauf, dass mentale Eigenschaften sich nicht vollständig auf physische reduzieren lassen.

[1] Sicherlich kann der interaktionistische Dualismus von John Eccles als bekannteste dieser Positionen betrachtet werden. Innerhalb seines Modells ist die Annahme eines externen selbstbewussten Geistes sogar grundlegend dafür, dass die mikrobiologischen bzw. -chemischen Prozesse im Gehirn nicht rein mechanistisch willkürlich, sondern zumindest bedingt sinnvoll im Interesse des Individuums verlaufen. An seiner Position wird u. a. kritisiert, dass jener „selbstbewusste Geist" letztlich sehr zufällig und nur auf der Basis von quantenmechanischen Prozessen in den Körper und menschliches Handeln eingreifen kann (vgl. Eccles 1994).

Insofern versucht diese Position einen Mittelweg zwischen der kausalen Geschlossenheit der physikalischen Welt und der alltagspsychologischen Sichtweise zu beschreiten:

> We confront a dilemma. Either we concede that purposive reasongiving explanations of behavior have only a pragmatic standing, or we abandon our conception of the physical domain as causally autonomous. (Heil/Mele 1999, v)

Damit verbunden ist die Frage, ob der intentionalistische bzw. psychologische Diskurs von dem physikalischen zu unterscheiden oder doch nur ein mythischer Restbestand ist. Wenn wir in der psychologischen Terminologie über dieselben Dinge reden wie in der physikalischen, könnte damit auch eine Ersetzbarkeit der Begriffe geboten sein. Sollten wir, um ein typisches Beispiel zu verwenden, in Zukunft anstatt „Ich habe Schmerzen." sachlich formulieren „Meine C-Fasern feuern." (vgl. Rorty 1981, 98)? Und stellen wir dann anstatt des üblichen Hungers lediglich die Unterversorgung des Blutzuckers fest?

Tatsächlich wird das Problem der mentalen Verursachung vor allem in den unterschiedlichen und teilweise sogar widersprüchlichen Positionen der Alltagsbeschreibung und der wissenschaftlichen Theorie deutlich. Von neurowissenschaftlicher Seite wird unter anderem der Vorwurf erhoben, dass im Alltagsverständnis die physischen Tatsachen in der Welt nicht zutreffend beschrieben werden. Den Vorstellungen verschiedener physikalischer Positionen zufolge gibt es nur physische Ereignisse; die psychologisierenden Beschreibungen sind das Ergebnis einer fehlgeleiteten Sprache, die uns eine falsche Vorstellung von der Welt vermittelt. Sämtliche Ausdrücke, die mentale Phänomene beschreiben, ließen sich prinzipiell auf eine physikalische Ausdrucksweise reduzieren, so dass darüber hinaus die falsche Vorstellung eines mentalen Bereichs völlig verschwände (vgl. z. B. P. M. Churchland 1981; Putnam 1981). Dagegen wird jedoch betont, dass eine Welt ohne Überzeugungen, Absichten oder Erwartungen schwer vorstellbar ist. Es wäre beispielsweise überflüssig, moralische oder juristische Urteile zu fällen, Geistes- und Sozialwissenschaften verlören ihren Sinn, der Mensch verhielte sich gewissermaßen interesselos zu seiner Umwelt. Das Fehlen sämtlicher intentionaler Handlung würde einen sozialen Umgang nicht nur unverständlich, sondern sogar unmöglich machen (vgl. Baker 1987, 130ff.; Fodor 1987, xii).

Einige Autoren (z. B. Chalmers 1997, Jackson 1998, Levine 2001) behaupten, die Frage der mentalen Kausalität sei prinzipiell gelöst und könne gänzlich der Neurowissenschaft überantwortet werden. Danach sei das eigentliche Problem der Philosophie des Geistes die Frage der Erklärungslücke zwischen Qualia und physischen Eigenschaften. Auch wenn die Vertreter dieser Position zugestehen, dass mentale Kausalität ein Kernthema dieses Bereichs ist.

Insbesondere Chalmers unterscheidet aufgrund der Differenzierung zwischen phänomenalen und psychologischen Konzepten bzw. Eigenschaften ein leichtes und ein schweres Körper-Geist-Problem.[2] Soweit psychologische Eigenschaften als funktionale Eigenschaften über ihre kausale Rolle definiert werden, kann die Frage, wie ein Zustand in einem physischen System eine bestimmte kausale Rolle spielen kann, seiner Ansicht nach an die Naturwissenschaften bzw. die Wissenschaften über physische Systeme übertragen werden (vgl. Chalmers 1997, 24). Es müsse lediglich geklärt werden, wie ein physisches System auf Umweltstimuli mit geeignetem Verhalten reagieren kann. Abgesehen von „technischen" Schwierigkeiten hält er das Körper-Geist-Problem für psychologische Eigenschaften weitestgehend gelöst und erkennt auch keine tiefergehenden metaphysischen Rätsel (ebd.).

Dennoch nennt Chalmers zumindest einige weiterhin bestehende Schwierigkeiten zur Frage der mentalen Kausalität: Nach wie vor ist nicht definitiv entschieden, wie psychologische bzw. mentale Eigenschaften im Allgemeinen und intentionale oder semantische Eigenschaften im Besonderen konzeptionell gefasst werden können. Weitere Schwierigkeiten ergeben sich aus der Frage, wie z. B. höherstufige Eigenschaften, die weder intrinsisch sind noch strikten Gesetzen unterliegen, eine echte kausale Rolle für die Verursachung von Verhalten spielen können (ebd.).[3] Diese genannten Schwierigkeiten sind jedoch symptomatisch dafür, dass nach wie vor ein grundsätzliches Problem der mentalen Kausalität besteht, das nicht nur „technisch" ist, sondern darüber hinaus metaphysische Relevanz besitzt. So wird im Rahmen der folgenden Diskussion deutlich, dass die physikalischen Voraussetzungen – beispielsweise die Geschlossenheit der Physik, die Ablehnung systematischer Mehrfachdeterminierung, die Forderung nach strikten Gesetzen – unter denen das Problem diskutiert wird, nach wie vor umstritten und nicht eindeutig entschieden sind. Ebenso muss die Frage nach dem Realisierungsverhältnis zwischen mentalen und physischen Eigenschaften sowie mentalen und physischen Ereignissen nach wie vor als nicht entschiedene metaphysische Frage betrachtet werden. Die unterschiedlichen Positionen haben sich in der zeitgenössischen Diskussion inzwischen weitestgehend etabliert und stehen sich mehr oder weniger parallel als alternative Erklärungsansätze gegenüber – je nachdem, ob man eher zu einer

[2] Chalmers unterscheidet zwischen zwei Konzepten bzw. Begriffen (englisch: „concept") des Geistes, einem phänomenalen und einem psychologischen. Während das phänomenale Konzept das bewusste Erleben mentaler Zustände und so genannte qualitative Zustände beschreibt, betrachtet das psychologische Konzept den Geist als kausale und explanatorische Basis für Verhalten. Ein Zustand ist dann mental, wenn er die passende kausale Rolle für die Entstehung von Handlung oder zumindest eine angemessene Erklärung für Verhalten bietet (vgl. Chalmers 1997, 11). Wer sich für die Frage der mentalen Kausalität interessiert, muss sich vor allem mit der Rolle der psychologischen Eigenschaften auseinandersetzen. Wer sich für das bewusste Erleben mentaler Zustände interessiert, muss sich mit phänomenalen Eigenschaften auseinandersetzen (vgl. ebd., 16).

[3] Darüber hinaus kann festgehalten werden, dass phänomenale und psychologische Eigenschaften auch korrelieren – einige mentale Zustände, wie z. B. Schmerz oder Wahrnehmung, verfügen über beide Eigenschaften (vgl. Chalmers 1997, 17). Eine eindeutige Trennung zwischen beiden Eigenschaften ist somit begrifflich stets problematisch. Insofern Chalmers den Umstand als Tatsache anerkennt, dass jede Instanziierung einer phänomenalen Eigenschaft mit der korrespondierenden Instanziierung einer psychologischen Eigenschaft einhergeht, kann vor allem hinsichtlich der Frage der mentalen Kausalität der Schwerpunkt der Untersuchung auf die psychologischen Eigenschaften gelegt werden (vgl. ebd., 22).

reduktionistischen oder nicht-reduktionistischen Position tendiert. Folglich lässt sich auch aus Uneinigkeit der zeitgenössischen Debatte um mentale Kausalität ersehen, dass diese Frage, dieser Begriff eben nicht eindeutig geklärt ist und beantwortet werden kann. Dieses bestehende Dilemma macht auch Papineaus Selbsteinschätzung deutlich, dass er sich je nach Tagesform nicht sicher ist, ob höhere Zustände kausal wirksam sind oder nicht:

> The issue here hinges on whether we can seriously allow that higher states *cause* what their realizers cause. I am not sure what to say about this. […] At other times I feel less fussy about causation. In particular, I sometimes worry that we will be left with precious few causes, if we are going to hold that higher states are pre-empted as causes whenever they have realizers in virtue of which they cause. (Papineau 2004, 34f.)

Dieses lakonische Geständnis eines überzeugten Physikalisten verdeutlicht die Spannung, die noch immer mit der Frage der mentalen Kausalität verbunden ist. Kim unterscheidet deshalb drei Versionen des Problems der mentalen Kausalität: die erste entsteht durch Davidsons Argument des anomalen Monismus, die zweite durch den semantischen Externalismus, die dritte durch das Prinzip des kausalen bzw. explanatorischen Ausschlusses (vgl. Kim 1998). Alle drei Versionen sind unterschiedliche Aspekte des generellen Problems der kausalen Relevanz mentaler Eigenschaften. Ferner betont Kim, dass das Problem der mentalen Kausalität in seiner aktuellen Form gerade aus der Struktur der physikalistischen Argumentation heraus entsteht und mit ihr verbunden ist (vgl. Kim 2001, 272). Ihm zufolge gibt es zwei wesentliche Gründe für mentale Kausalität: erstens die Möglichkeit genuiner menschlicher Handlung einschließlich der moralischen Dimension Willen. Für willentliche Handlungen müssen unsere Glaubensannahmen, Wünsche, Intentionen oder Entscheidungen unsere Gliedmaßen „kausal aktivieren" bzw. „verursachen", damit diese innerhalb der Umwelt die erforderliche Bewegung o. ä. ausführen. Und zweitens um menschliches Wissen, einschließlich z. B. Wahrnehmung, logischem Denken oder Erinnerung, plausibel zu erklären (ebd.). Diese Formen menschlichen Wissens sind letztlich die Basis, um mit der Welt, wie sie uns umgibt, in Kontakt zu treten. Letztendlich betreffen diese Fragestellungen unser menschliches Selbstverständnis. Es stellt sich auch die Frage, ob unsere spezifische Humanität vor der übergreifenden Naturalisierung zu retten ist. In Ergänzung müsste aber ebenso gefragt werden, ob die Humanität durch die Naturalisierung überhaupt bedroht ist.

### 1.2 Neurophilosophie als interdisziplinärer Lösungsansatz

Die oben dargestellten Dilemmata machen deutlich, dass die Frage der mentalen Kausalität noch immer nicht vollständig und abschließend beantwortet ist. Warum gibt es nach wie vor Einwände

und Gegner der identitätstheoretischen oder der reduktionistischen Position? So existieren noch immer viele verschiedene Ansätze, die die genaue Relation zwischen mentalen und physischen Eigenschaften erklären und sich über die kausale Wirksamkeit bzw. Relevanz des Mentalen streiten. Wäre die Frage der mentalen Kausalität hinreichend beantwortet, wäre die unterschiedliche Beantwortung durch verschiedene Funktionalismusvarianten, Supervenienztheorien, Identitätstheorien, reduktionistische wie nicht-reduktionistische Positionen oder schließlich eliminativistische Positionen nicht weiter plausibel. Aus der Diskussion wird aber deutlich, dass ein Konsens ebenso aussteht, wie sich keine Position unangefochten durchsetzen kann.

Ein möglicher Lösungsweg ist, sich verstärkt der empirischen Seite zuzuwenden, dem Bereich, über den wir zumindest vorläufig sichere Aussagen treffen können. So fordert Roth (2001, 205) explizit eine „Neurobiologie des Geistes" in enger interdisziplinärer Zusammenarbeit von z. B. experimenteller und theoretischer Neurobiologie, Neurologie oder Kognitionspsychologie. Zudem ist seit einiger Zeit eine Tendenz erkennbar, wonach Naturwissenschaftler sich vermehrt ursprünglich klassisch philosophischen oder kulturwissenschaftlichen Fragestellungen zuwenden. Dabei wird deutlich, dass die Grenzen der verschiedenen wissenschaftlichen Disziplinen in bestimmten Bereichen aufbrechen. Insbesondere die Verbindung mentaler Phänomene der Erste-Person-Perspektive mit materiellen Vorgängen im Gehirn aus der Dritte-Person-Perspektive, also der Versuch, die Grenzen zwischen den Beschreibungssystemen für neuronale und psychische Prozesse zu überbrücken, wird mitunter als Dammbruch angesehen (vgl. Singer 2002e, 178). Diese Entwicklung schreitet als Syntheseprozess voran, wie auch die Doppelnamen neuer Forschungsrichtungen zeigen: Neuropsychologie, Neurolinguistik, Biopsychologie, Soziobiologie, evolutionäre Psychologie oder schließlich Neurophilosophie (vgl. ebd.). Dabei stehen sich beide Beschreibungssysteme, das philosophische wie das neurowissenschaftliche, zunächst gleichberechtigt gegenüber. Auch die Phänomene, die sie beschreiben, müssen als Teil der Weltbeschreibung gleichermaßen berücksichtigt werden – Phänomene der Erste-Person-Perspektive ebenso wie Phänomene der Dritte-Person-Perspektive.

Von neurowissenschaftlicher Seite her lässt sich an dieser Stelle ein Beispiel anbringen, warum die Neurophilosophie die Perspektive erweitert bzw. hilfreich für die philosophische wie neurowissenschaftliche Diskussion ist: Libets Untersuchung der spontanen Fingerbewegung (vgl. z. B. Libet 1982, 1985, 1993; Soon et al. 2008). Diese Versuche werden häufig im Zusammenhang mit der Frage der Willensfreiheit diskutiert. Dies war auch seinerzeit der intendierte Untersuchungsgegenstand von Libets Forschung. Dabei handelt es sich hier grundsätzlich auch um ein Beispiel für mentale Kausalität. Denn innerhalb der Versuchsanordnung konnte zumindest gezeigt werden, dass nach einer Aufgabenstellung als „mentalem Input" innerhalb der Gehirne der Probanden neuronale Aktivität im motorischen Kortex zu einer Handlung bzw. spontanen Fingerbewegung als physischen, behaviouralen Output führte. Insofern wird angenommen, dass

nach hinreichendem Aufbau von Bereitschaftspotenzial als Ursache die anschließende spontane Fingerbewegung bewirkt wird. Entscheidend ist, dass den Untersuchungen zufolge das Bereitschaftspotenzial eindeutig vor dem bewussten Erleben gemessen wurde. Problematisch ist jedoch die Interpretation der Ergebnisse, die, wie die anschließende Diskussion zeigte, sowohl als Beleg für interaktionistischen Dualismus (vgl. Eccles 1985) als auch für reduktionistische Positionen (vgl. Deecke 1987, 781; Roth 2003, 518ff.) dienen kann. Libet selbst stand vor dem zusätzlichen Problem, dass ihm von den Probanden über eine „Veto-Phase" innerhalb der Versuchssituation berichtet wurde: Die Probanden konnten sich anders entscheiden und das spontane Fingerheben abbrechen (vgl. Libet 1985, 537). Diese Veto-Phase ist jedoch als subjektives Phänomen der Erste-Person-Perspektive per se nicht messbar. Zwar gab es über die Berichte der Probanden hinaus zumindest Indizien im EEG bzw. EMG, diese halten einer wissenschaftlichen Prüfung jedoch nicht Stand. Es war Libet nicht möglich, ein angemessenes neurowissenschaftliches Design zu entwickeln, um dieses spontane Phänomen zu untersuchen. Jenseits der Diskussion um mögliche Willensfreiheit verdeutlichen seine Experimente jedoch zwei Aspekte: Erstens ist ein direkter Zugriff auf die Erste-Person-Perspektive über eine rein neurowissenschaftliche Methodik bzw. bildgebende Verfahren o. ä. nicht möglich. Als reale Erfahrung sollte dieser Bereich aber nicht ausgeklammert, sondern weiterhin untersucht werden. Deshalb ist die Verbindung mit der Philosophie für die Neurowissenschaft eine mögliche Perspektive. Zweitens wurde aufgrund der Versuchsergebnisse geschlossen, dass Willensfreiheit oder sogar eine bewusste mentale Wirkung auf physische Prozesse nicht möglich ist. Diese radikale Schlussfolgerung ist auch jenseits der Veto-Frage problematisch, da sich der Versuch nur auf eine relativ basale Handlung einer spontanen Fingerbewegung bzw. einer einfachen motorischen Handlung bezieht. Hiervon auf andere Bereiche wie komplexe Gedanken und Handlungen zu schließen, erscheint vorschnell, da diese schließlich auch in anderen Bereichen des Gehirns prozessiert werden. Andererseits muss es nicht erstaunen, dass neuronale Aktivität mit bewussten oder unbewussten Handlungen einhergeht. Hier stellt sich jedoch zusätzlich die Frage, welchen Stellenwert das gemessene Bereitschaftspotenzial oder auch andere neuronale Aktivität innerhalb mentaler Prozesse einnehmen (vgl. Pauen 2001, 111). Dies zeigt, wie wichtig eine theoretische Fundierung für neurowissenschaftliche Interpretationen ist. Dieses kann im Rahmen einer Neurophilosophie unter Rekurs auf philosophische Erkenntnisse und begriffliche Analyse geleistet werden.

Was haben die Neurowissenschaften zum Problem der mentalen Kausalität beizutragen? Welchen Bezug haben philosophische Theorien zu empirischer Forschung? Gibt es Anhaltspunkte und Hinweise, die verschiedene philosophische Auffassungen bestätigen oder verwerfen? Wegen der unterschiedlichen Geltungsansprüche, Bedingungen und Problemstellungen können neurowissenschaftliche Befunde und philosophische Theorien nicht direkt aufeinander bezogen werden. Auf der neurowissenschaftlichen Seite ist nach wie vor problematisch, dass empirische

Erkenntnisse trotz z. B. starker Korrelationen alternative, sogar gegensätzliche Interpretationen zulassen (vgl. Pauen 2001, 84). Durch einen empirischen Bezug wird philosophische Theorie zwar eingegrenzt, ermöglicht aber andererseits, Fehler in der empirischen Frage aufzuzeigen, Lösungsstrategien zu entwickeln und somit eine logisch, analytisch konsistente und plausible Interpretation der empirischen Ergebnisse zu fundieren (vgl. ebd., 101, 112f.). In Erwiderung auf Polemiken gegen die Philosophie des Geistes von neurowissenschaftlicher Seite erläutern Bennet und Hacker (2003), was Philosophie sowohl eigenständig als auch in komplementärer Ergänzung zur Neurowissenschaft leisten kann (vgl. Bennett/Hacker 2003, 396-409).[4] So wird betont, dass sich die analytische Philosophie auf konzeptionelle, begriffliche Untersuchungen konzentriert.

Zentrale Aufgabe ist es danach, im Rahmen einer Bedeutungsanalyse, philosophische Probleme zu lösen, Repräsentationen, Konzepte und Vorstellungen zu bestimmen sowie begriffliche Verwirrungen aufzulösen. In Abgrenzung zur empirischen Wissenschaft geht es nicht um die Ermittlung von Tatsachen auf der Basis von empirischer Wahrheit und Falschheit, sondern um die Prüfung sinnvoller Aussagen. Der Grenzbereich von Sinn bzw. der sinnvolle Gebrauch von Sprache wird z. B. durch die Prüfung kohärenter Aussagen, logischer Schlussfolgerungen oder eben die Kritik an falschen, sinnlosen Problemen oder Aussagen beschrieben (vgl. ebd., 399; ebd., 405). Die Kognitionswissenschaft gerät immer wieder in die Gefahr begrifflicher Unschärfe, weshalb eine begriffliche Klärung sowohl für die Identifizierung als auch für die Bestimmung der Probleme und schließlich für die konsistente Beschreibung und Bewertung der empirischen Forschungsergebnisse notwendig ist (vgl. ebd., 402). Neurowissenschaft kann den naturwissenschaftlich-empirischen Rahmen z. B. hinsichtlich der Gehirn-Bewusstsein-Relation untersuchen. Sie kann aber nicht epistemische Fragen begründen, die auf eine logische, begriffliche oder auch normative Verknüpfung verweisen. Dennoch ist den Autoren eine klare Trennung des Aufgabenbereichs der Neurowissenschaft und der Philosophie wichtig. Die Herangehensweise beider Disziplinen betrifft zwar teilweise dieselben Problemfelder, aber im Kern andere Fragestellungen. Die Philosophie treibt die begriffliche Schärfung des Untersuchungsgegenstands der mentalen Kausalität voran, während die Neurowissenschaft die empirischen bzw. neuronalen Bedingungen für mentale Kausalität untersucht. Insofern sind sie komplementär, aber keineswegs identisch oder kompetitiv bzw. sich wechselseitig ausschließend (ebd., 403). Der neurophilosophische Ansatz der vorliegenden Arbeit geht über diese Trennung insofern hinaus, als eine interdisziplinäre

[4] Einer der Vorwürfe von Neurowissenschaftlern gegenüber der Philosophie des Geistes lautet, dass letztere im Grunde nichts erreicht habe, um die Frage der Relation zwischen Körper und Geist abschließend zu klären, während die Methoden der Neurowissenschaft in vergleichbar kurzer Zeit viel mehr Ergebnisse geliefert hätten. Deshalb sei der Beitrag der Philosophie zu Fragen des Bewusstseins o. ä. irrelevant und wertlos. Diese starke Polemisierung ist letztlich auf ein falsches Verständnis des Aufgabenbereichs, der Zielsetzung und der Methodik von Philosophie zurückzuführen, wie Bennett und Hacker schlüssig darstellen (vgl. z. B. Crick 1995; Edelmann 1994; beide zitiert nach Bennett/Hacker 2003, 397f.). In einem weiteren Ansatz verwirft Bickle philosophische Argumentation zugunsten eines „rücksichtslos reduktiven Ansatzes", indem er sich insbesondere auf Erklärungen der zellularen/molekularen Neurobiologie konzentriert (vgl. Bickle 2003, 107ff.).

Verknüpfung erfolgt. Der Unterschied wird in der methodischen Perspektive deutlich: Philosophie und Neurowissenschaft schließen sich nicht aus, sondern ergänzen sich wechselseitig. Neurophilosophie ermöglicht aufgrund ihrer transdisziplinären Methodik dem Problemfeld der mentalen Kausalität einen Untersuchungsrahmen (vgl. Northoff 2000, 22).[5] Angestrebt wird dabei eine Annäherung und gegenseitige Einflussnahme zwischen den Disziplinen der Philosophie und Neurowissenschaft. Als Brückendisziplin zwischen subjektiver Erfahrung, philosophischer Theoriebildung und empirischer Forschung hat sie einen interdisziplinären Charakter. Auf der einen Seite führen neurowissenschaftliche Probleme zwangsläufig zu philosophischen Fragestellungen. Auf der anderen Seite sollen durch die Anwendung philosophischer Argumentationslinien schnelle Schlussfolgerungen und übermäßig naive Theorien vermieden werden (vgl. Walter 1999, 160f.).

Deshalb soll in der vorliegenden Arbeit auch untersucht werden, wie eine solche Verbindung von philosophischer und empirischer Theorie unter neurophilosophischen Gesichtspunkten aussehen kann. Eine neurophilosophische Perspektive kann Neues zu dieser Diskussion beitragen, gegebenenfalls einige bestehende Fragen beantworten oder aber darüber hinaus zu weiterer Untersuchung anregen. Dabei wird deutlich, dass die Beschäftigung mit „mentaler Kausalität" noch immer lohnt und keineswegs geklärt ist. So ist es wichtig, das Modell der Kausalität selbst zu betrachten und zu hinterfragen. Von besonderem Interesse ist dabei, mögliche Kriterien für mentale Kausalität zu entwickeln, die mit ihnen verbundenen Voraussetzungen kritisch zu diskutieren – auch in Hinblick auf ihre empirische Plausibilität. Die Diskussion der mentalen Kausalität unter neurophilosophischen Gesichtspunkten soll mögliche Irrwege und Scheinlösungen aufdecken und darüber hinaus in Form einer neurophilosophischen Hypothese neue Lösungswege aufzeigen.

### 1.3 Entwicklung der neurophilosophischen Methodik

Obwohl es inzwischen zahlreiche Aufsätze und Publikationen gibt, die unter den Ansatz einer Neurophilosophie gefasst werden, fehlt bislang eine einheitliche neurophilosophische Methodik. Northoff (2004) versucht deshalb, exakte methodische Prinzipien und Strategien für die neurophilosophische Disziplin zu entwickeln. Die transdisziplinäre Neurophilosophie soll methodisch philosophische Theorien und neurowissenschaftliche Hypothesen miteinander verbinden, um so

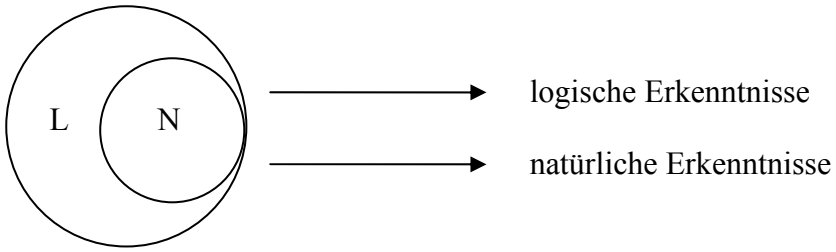
[5] Der Begriff der Neurophilosophie geht auf Churchland (1986) zurück. Die damit verbundene frühe eliminativistische Position unterscheidet sich allerdings deutlich von der in dieser Arbeit durchgeführten interdisziplinären Methodik, da sie letztlich eine Abschaffung der psychologischen Sprache verfolgt. Dem entgegen bleibt die psychologische Sprache und auch die philosophische Theorie in Northoffs Ansatz notwendiger und grundlegender Bestandteil für Epistemologie wie Ontologie – und damit eben auch für Neurophilosophie (vgl. Northoff 2004, 25ff.).

einen Austausch der jeweiligen Argumente und Standpunkte zu ermöglichen. Northoff betont, dass die Neurophilosophie als eigenständige Disziplin keine Übergangsphase hin zu einer reinen Neurowissenschaft darstellt, die Philosophie oder auch Psychologie ablöst. Dies wird allein dadurch deutlich, dass das Gehirn innerhalb der Neurophilosophie nicht nur als ein „physisches Gehirn", sondern auch als ein „mentales Gehirn" diskutiert wird. Erst durch die Verknüpfung bzw. Differenzierung der natürlichen und logischen Bedingungen ist überhaupt eine vollständige Untersuchung des Gehirns möglich, ein rein empirischer Ansatz kann dies jedoch nicht leisten (vgl. ebd., 56f.). Eine Besonderheit in dieser Verbindung ist dabei die „Selbstreferenzialität": Abhängig von den jeweiligen epistemisch-ontologischen Voraussetzungen ist der Geist bzw. das Gehirn selbst notwendige Bedingung für die epistemische Möglichkeit, philosophische Theorien und neurowissenschaftliche Hypothesen miteinander zu verbinden. Um logische Zirkularität in der Verbindung zwischen beiden Bereichen auszuschließen, bedarf es deshalb besonderer transdisziplinärer Prinzipien (vgl. Northoff 2004; 27). Auch wenn die von Northoff ausgearbeitete Methodik nicht den Anspruch erheben kann, für sämtliche neurophilosophischen Richtungen maßgeblich zu sein, liefert sie mindestens ein solides Instrumentarium, einen systematischen Vergleich zwischen philosophischer Theorie und neurowissenschaftlichen Hypothesen durchzuführen. Insofern orientiert sich diese Arbeit an den Prinzipien dieser transdisziplinären Methodik, die im Folgenden vorgestellt werden.

Innerhalb seiner speziellen neurophilosophischen Methodik entwickelt Northoff drei Prinzipien, durch die die Gleichrangigkeit sowie die Vergleichbarkeit von Philosophie und Neurowissenschaft gewährleistet werden soll: das „Prinzip der Asymmetrie", das „Prinzip der Bidirektionalität" und das „Prinzip der transdisziplinären Zirkularität".

Das Prinzip der Asymmetrie beschreibt ein epistemisches Ungleichgewicht zwischen den logischen und den natürlichen Bedingungen, die innerhalb der Neurophilosophie miteinander verknüpft werden. Denn während logische Bedingungen alle möglichen bzw. logisch denkbaren Welten umfassen, beziehen sich natürliche Bedingungen ausschließlich auf die tatsächliche, natürliche Welt mit ihren physischen bzw. biologischen Gesetzmäßigkeiten. Da innerhalb der logischen Bedingungen grundsätzlich auch natürliche Bedingungen enthalten sind, ist ihr epistemischer Rahmen weiter gefasst, weshalb das Verhältnis zwischen logischen und natürlichen Bedingungen als asymmetrisch charakterisiert werden kann (vgl. Northoff 2004, 27f.). Das Prinzip impliziert, dass eine direkte Schlussfolgerung von den logischen auf die natürlichen Bedingungen aufgrund des weiteren Rahmens der logischen Bedingungen nicht zulässig ist. Ebenso wenig ist eine direkte Schlussfolgerung von den natürlichen auf die logischen Bedingungen aufgrund des engeren Rahmens der natürlichen Bedingungen möglich. Konsequenz ist, dass das Geist-Gehirn-Problem in seiner Fragestellung unter logischen Bedingungen nicht, wie häufig angenommen,

vollständig durch neurowissenschaftliche Forschung bzw. Überlegungen unter natürlichen Bedingungen gelöst werden kann. Umgekehrt lassen sich Lösungsansätze aus der philosophischen Diskussion nicht ohne weiteres auf das tatsächliche Gehirn, wie es innerhalb der neurowissenschaftlichen Forschung untersucht wird, anwenden. [6] Jenseits der Asymmetrie der logischen und natürlichen Bedingungen kann jedoch eine Schnittmenge beider Bedingungen betrachtet werden – logische Bedingungen umfassen logische *und* natürliche Bedingungen. Eine Schnittmenge hinsichtlich der tatsächlichen, natürlichen Welt ist daher grundsätzlich gegeben. Deshalb werden eben auch Kriterien für die Unterscheidung zwischen den verschiedenen logischen Bedingungen und den daraus folgenden möglichen Verbindungen mit den natürlichen Bedingungen benötigt (vgl. Northoff 2004, 29f.).




Grafik 1: Das Prinzip der Asymmetrie impliziert, dass die logischen Bedingungen L die natürlichen Bedingungen N umfassen. Direktes Schließen allein von den logischen auf die natürlichen Bedingungen ist deshalb problematisch, da aufgrund des weiteren Rahmens die logischen und natürlichen Bedingungen nicht klar getrennt werden können. Umgekehrt ist der Rahmen der natürlichen Bedingungen zu klein, um direkt auf die logischen Bedingungen zu schließen. Beide Bedingungen sind aber wichtig, um das gesamte menschliche epistemische Potenzial auszuschöpfen und nicht aufgrund natürlicher wie logischer epistemischer Limitationen falsche ontologische Schlussfolgerungen zu treffen (vgl. Northoff 2004, 27ff., Figure 6).

Das Prinzip der Bidirektionalität betont die bidirektionale Verknüpfung zwischen philosophischen Theorien und neurowissenschaftlichen Hypothesen und damit zwischen logischen und natürlichen Bedingungen. Zum einen kann die empirische Konsistenz der philosophischen Theorien überprüft werden. Diese spiegelt sich in jenen logischen Bedingungen wieder, die mit den natürlichen Bedingungen übereinstimmen. In dem Fall einer empirischen Inkonsistenz, also einer Abweichung logischer und natürlicher Bedingungen kann entweder diese Differenz zwischen philosophischer Theorie und neurowissenschaftlicher Hypothese akzeptiert werden. Oder es wird die philosophische Theorie hinsichtlich der entsprechenden neurowissenschaftlichen Hypothese modifiziert, was einen „definitorischen Wandel“ und eine „konzeptionelle Weiterentwicklung“

[6] Ausgangspunkt für die von Northoff entwickelte neurophilosophische Methodik ist ein epistemischer Pluralismus, dem in methodischer Hinsicht gegenüber einem ontologischen Pluralismus ein Primat eingeräumt wird. Einerseits sind epistemische Aussagen ohne ontologische Fundierung „leer“, andererseits sind ontologische Annahmen notwendig von der Epistemologie abhängig und ohne epistemische Fähigkeiten „blind“: „If our brain were different […] we would probably have different ‚epistemological capacities'. Different ‚epistemological capacities' would provide us with a different access to ‚reality and existence' which consecutively would lead to different ontological assumptions about ‘reality and existence'.“ (Northoff 2004, 39). Das epistemische Primat ist notwendig, um die Verknüpfung zwischen diesen epistemischen Fähigkeiten und jenen ontologischen Aussagen zu untersuchen. Es ist insofern eine methodische Strategie. Mit der zusätzlichen Annahme des ontologischen Pluralismus wird dem Vorwurf der ontologischen Zirkularität begegnet (vgl. ebd., 43ff.).

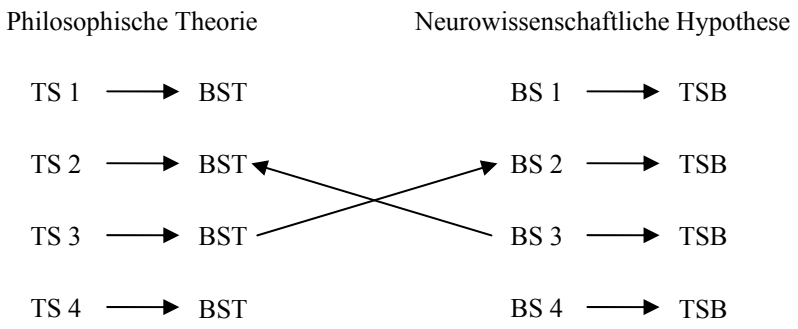
(englisch: re-clarification) ermöglicht. Zum anderen kann die logische Konsistenz der neurowissenschaftlichen Hypothese überprüft werden. Deren ontologische bzw. epistemische Voraussetzungen bestimmen die natürlichen Bedingungen, welche auf Konsistenz mit den logischen Bedingungen geprüft werden. Im Falle einer Inkonsistenz können die Voraussetzungen der natürlichen Bedingungen überprüft und gegebenenfalls modifiziert werden. Auf diese Weise schafft das Prinzip der Bidirektionalität die Grundlage für einen wechselseitigen Vergleich zwischen philosophischer Theorie und neurowissenschaftlichen Hypothesen bezüglich ihrer logischen wie natürlichen Bedingungen, durch den ein gemeinsamer Anwendungsrahmen hergestellt wird (vgl. Northoff 2004, 30f.).

Das Prinzip der transdisziplinären Zirkularität beschreibt die eigentliche Methode, den systematischen Prozess der wechselseitigen Übertragung und Kreuzvergleiche bzw. des Kreislaufs zwischen philosophischer Theorie und neurowissenschaftlichen Hypothesen bis hin zu der daraus resultierenden Bildung einer neurophilosophischen Hypothese. Aufgrund der Unterschiede zwischen logischen und natürlichen Bedingungen erscheint eine direkte Verknüpfung zwischen beiden Bereichen nicht möglich. Stattdessen wird von Northoff ein indirekter Vergleich vorgeschlagen, der in einem zirkulären Prozess durchgeführt wird:

> One may consider this ‚disciplined circularity' […] between philosophical theories and neuroscientific hypotheses as a linkage between ‚theoretical sentences and ‚observation sentences'. 'Theoretical sentences' refer to logical conditions and are thus independent from the actual world. They reflect ontological and epistemological assumptions which are discussed explicitly in philosophical theory […]. 'Observations sentences', in contrast, refer to natural conditions and empirical observations within the actual world [.] (Northoff 2004, 32f.)

In einem ersten Schritt sollen nun die „Explikationen" und „Implikationen" deutlich herausgestellt werden: Explikationen beziehen sich auf die impliziten ontologischen und epistemischen Voraussetzungen neurowissenschaftlicher Hypothesen; Implikationen beziehen sich auf mögliche empirische Konsequenzen philosophischer Theorien. Die Verknüpfung expliziter Beobachtungssätze mit impliziten Theoriesätzen wird als „theoretische Explikation", das Schließen auf mögliche Beobachtungssätze aus Theoriesätzen als „empirische Implikation" bestimmt. Durch den wechselseitigen Vergleich der theoretischen Explikationen und der empirischen Implikationen soll ein geeigneter Rahmen für die Verknüpfung philosophischer Theorie mit neurowissenschaftlichen Hypothesen bestimmt werden (vgl. Northoff 2004, 31ff.). In einem zweiten Schritt sollen zum einen die theoretischen Explikationen bzw. die ontologischen und epistemischen Voraussetzungen einer neurowissenschaftlichen Hypothese mit ontologischen und epistemischen Voraussetzungen philosophischer Theorien diskutiert werden: so kann die logische Konsistenz überprüft werden. Zum anderen sollen die empirischen Implikationen der philosophischen Theorien

mit empirischen Ergebnissen der neurowissenschaftlichen Hypothesen verglichen werden: so kann die empirische Konsistenz überprüft werden. Der wechselseitige Vergleich zwischen Theoriesätzen und Beobachtungssätzen unter dem Gesichtspunkt der logischen wie empirischen Konsistenz kann als notwendige Bedingung für den Vergleich zwischen philosophischen Theorien und neurowissenschaftlichen Hypothesen betrachtet werden (vgl. ebd. 33f.).




Grafik 2: Im Rahmen des Prinzips der transdisziplinären Zirkularität werden die Theoriesätze {TS 1, TS 2, …} der philosophischen Theorie und die Beobachtungssätze {BS 1, BS 2, …} der neurophilosophischen Hypothese bestimmt. Aus ihnen folgt die Entwicklung der empirischen Implikationen BST der Theoriesätze sowie der theoretischen Explikationen TSB der Beobachtungssätze. In einem ersten Kreuzvergleich werden die empirische Konsistenz der philosophischen Theorie und die logische Konsistenz der neurowissenschaftlichen Hypothese überprüft. Durch die Analogisierung und gegebenenfalls Modifizierung einer empirischen Implikation BST im Kontext der neurophilosophischen Hypothese sowie einer theoretischen Explikation TSB im Kontext einer philosophischen Theorie wird die Voraussetzung für die Verknüpfung von natürlichen und logischen Bedingungen geschaffen (vgl. Northoff 2004, 32f., Grafik 7).

In einem dritten Schritt werden sowohl die ontologischen bzw. epistemischen Annahmen sowie die empirischen Hypothesen der philosophischen Theorie und der neurowissenschaftlichen Hypothese analogisiert bzw. homogenisiert. Um hierfür logische bzw. empirische Konsistenz der neurowissenschaftlichen Hypothese bzw. der philosophischen Theorie zu gewährleisten, müssen deren theoretische Explikationen bzw. empirische Implikationen gegebenenfalls modifiziert werden. Die Voraussetzungen für einen Vergleich bzw. eine Verknüpfung zwischen Beobachtungssätzen und Theoriesätzen bzw. zwischen natürlichen und logischen Bedingungen sind nunmehr geschaffen (vgl. ebd., 34f.). In einem vierten Schritt kann ein inverser Kreuzvergleich zwischen den Disziplinen durchgeführt werden. Einerseits können die Konsequenzen der modifizierten theoretischen Explikationen auf die ontologischen und epistemischen Voraussetzungen der neurowissenschaftlichen Hypothese überprüft werden, andererseits können die Konsequenzen der modifizierten empirischen Implikationen auf die empirische Hypothese der philosophischen Theorie überprüft werden. Neurowissenschaftliche Hypothese wie philosophische Theorie bleiben möglicherweise unbeeinträchtigt durch diesen inversen Vergleich, müssen aber gegebenenfalls wiederum modifiziert werden. Der inverse Kreuzvergleich kann als notwendige Bedingung für die Offenlegung der direkten Interaktion zwischen ontologischen bzw. epistemischen An-

nahmen und empirischen Hypothesen innerhalb der philosophischen Theorie bzw. der neurowissenschaftlichen Hypothese betrachtet werden (vgl. ebd., 36).

Die Notwendigkeit für die Entwicklung aller drei Prinzipien ergibt sich nach Northoff aus Versuchen, logische Bedingungen auf natürliche Bedingungen zu reduzieren bzw. philosophische Theorien durch neurowissenschaftliche Hypothesen zu eliminieren (vgl. Churchland/Churchland 2001). Er betont, dass ein interdisziplinärer Vergleich unvollständig bleibt, solange die logischen Bedingungen nicht vollständig berücksichtigt werden. Außerdem argumentiert er von einem Standpunkt des epistemischen Pluralismus aus, der philosophischer Theorie wie neurowissenschaftlicher Hypothese zunächst ein gleiches Gewicht zuspricht (vgl. Northoff 2004, 30f.; ebd., 37ff.).

Unter Berücksichtigung der Ergebnisse aus dem wechselseitigen Vergleich der logischen und der empirischen Bedingungen soll schließlich eine eigenständige neurophilosophische Hypothese gebildet werden. Vor dem Hintergrund der neurophilosophischen methodischen Prinzipien, dem Prinzip der Asymmetrie, dem Prinzip der Bidirektionalität und dem Prinzip der transdisziplinären Zirkularität, soll die neurophilosophische Hypothese auch eine systematische Verknüpfung beider Bereiche beinhalten. Die neurophilosophische Hypothese muss insofern verschiedenen Ansprüchen, logischer Konsistenz, empirischer Konsistenz sowie transdisziplinärer Konsistenz gerecht werden. Ihre Signifikanz beruht schließlich auf der Verknüpfung der logischen und natürlichen Bedingungen (vgl. Northoff 2004, 49). Die neurophilosophische Hypothese reicht insofern über eine empirische Hypothese hinaus, als sie sich nicht nur auf empirische Konsistenz bezieht, sondern ausdrücklich ontologische wie epistemische Annahmen logischer Konsistenz mit einbezieht. Sie ist breiter angelegt als empirische Hypothesen. Sie reicht aber auch über eine philosophische Theorie hinaus, da sie sich nicht nur auf logische Konsistenz bezieht, sondern ausdrücklich empirische Hypothesen mit einbezieht. Die neurophilosophische Hypothese bestimmt ihre Begriffe nicht vor der nachfolgenden Untersuchung. Stattdessen wird die Begriffsbestimmung selbst Bestandteil eines „definitorischen Shiftings", da Begriffe auf empirischer Basis modifiziert und erneuert werden. Hieraus ergibt sich für die neurophilosophische These auch, dass sie zusätzlich einer transdisziplinären Falsifikation unterliegen kann, welche die Konsistenz der Verknüpfung prüft (vgl. ebd. 49f.). Andererseits ergibt sich aus der Verknüpfung natürlicher Bedingungen mit logischen Bedingungen, dass neurophilosophische Begriffe zumindest eingeschränkt vorstrukturiert sind und der Bereich möglicher ontologischer wie epistemischer Annahmen durchaus begrenzt ist (vgl. ebd., 51).

Ferner kann die neurophilosophische Hypothese implizite empirische Hypothesen innerhalb philosophischer Theorie untersuchen. Sie ermöglicht die wechselseitige Anpassung von philosophischer Theorie und empirischen Hypothesen nicht nur für die logische Analyse, sondern

auch für eine begriffliche Klärung. Die begriffliche Klärung oder Erneuerung kann die Untersuchung der Konsistenz der Verknüpfung als einen Test für die systematische Interaktion zwischen philosophischer Theorie und empirischer Hypothese ermöglichen. Damit besteht die zusätzliche Möglichkeit, eine neurophilosophische Hypothese zu entwickeln, die empirische Hypothese und philosophische Theorie systematisch und konsistent miteinander verbindet. Daraus folgt, dass ausschließlich analytische Urteile ebenso wie ausschließlich synthetische Urteile für eine neurophilosophische Hypothese ausgeschlossen sind. Beide Komponenten werden miteinander verknüpft. Einige neurophilosophische Hypothesen zeigen möglicherweise eine stärker analytische Komponente, während andere eine stärkere synthetische Komponente beinhalten. Trotzdem müssen beide Komponenten Bestandteil einer neurophilosophischen Hypothese sein, damit diese als solche gelten kann (vgl. ebd., 50f.). Insofern können neurophilosophische Hypothesen als gemischte Urteile gelten, die sowohl a priori- als auch a posteriori-Komponenten beinhalten. Zusätzlich können diese Hypothesen auch notwendige oder kontingente Urteile umfassen.[7] Epistemische, semantische und ontologische Unterscheidungen werden zwar aufgehoben bzw. relativiert. Aber damit ist eine Trennung bzw. sind neue, variable Kombinationen zwischen epistemischer, semantischer und ontologischer Charakterisierung für neurophilosophische Urteile möglich (vgl. ebd., 51f.).

Gegen die neurophilosophische Hypothese kann als ein mögliches Gegenargument aufgeführt werden, dass aufgrund ihrer Kontingenz durch die philosophisch-neurowissenschaftliche Verknüpfung ihre Notwendigkeit in ontologischer Hinsicht sowie ihre Gültigkeit in epistemischer Hinsicht bezweifelt werden kann. Northoff betont deshalb, dass das Zulassen von Kontingenz nicht zu einer vollständigen Eliminierung der Notwendigkeit führt. Im Rahmen der neurophilosophischen Hypothese sind logische Notwendigkeit und empirische Kontingenz nicht ausschließlich zu verstehen. Vielmehr werden empirische und logische Konsistenz eingefordert, wobei sich erstere auf die empirische Kontingenz und letztere auf die logische Notwendigkeit bezieht. Durch diese unterschiedliche Zuordnung ist es möglich, dass sie gemeinsam Geltung haben und sich nicht wechselseitig ausschließen (vgl. ebd., 54 f.).[8]

Die vorliegende Arbeit folgt der beschriebenen neurophilosophischen Methodik in wesentlichen Punkten, auch wenn nicht alle Schritte übernommen werden. Es erfolgt insofern eine Konzentration auf die theoretische Entwicklung innerhalb der analytischen Philosophie des Geistes, als diese grundsätzlich sehr geeignet für die Verknüpfung mit empirischen Ergebnissen der Neuro-

[7] Northoff spricht in allen Fällen davon, dass je nach Gewichtung eine Hypothese „mehr oder weniger" a priori, a posteriori, notwendig oder kontingent sein kann. Wie von Kripke betont, soll somit im Rahmen einer neurophilosophischen Hypothese im günstigen Fall eine Trennung zwischen a priori und notwendig einerseits sowie a posteriori und kontingent andererseits möglich sein (vgl. Kripke 1980).
[8] Für weitere Argumente gegen die neurophilosophische Methodik und ihre Diskussion s. Northoff 2004, 52 ff..

wissenschaft ist. Traditionell orientiert sich die analytische Philosophie des Geistes an wissenschaftlichen Tatsachen.[9] Ziel dieser Arbeit ist der Vergleich zwischen den logischen Bedingungen der philosophischen Theorie und den natürlichen Bedingungen der Neurowissenschaft für mentale Kausalität, um ein empirisch und theoretisch valides Modell zu entwickeln. So können das Prinzip der Asymmetrie und das Prinzip der Bidirektionalität zunächst als notwendige Voraussetzungen und Vorüberlegungen aufgefasst werden, auf denen mit dem Prinzip der transdisziplinären Zirkularität die eigentliche Methode fußt. Die in diesem dritten Prinzip vorgestellten vier Schritte werden aus verschiedenen Gründen, unter anderem der Praktikabilität halber, nicht vollständig durchgeführt. Das Vorgehen wird insofern modifiziert, als auf der theoretischen Ebene Kriterien bestimmt werden, die hinsichtlich der mentalen Kausalität einen spezifischeren und ausgewogeneren Vergleich ermöglichen (vgl. Kap. 4).

In Voruntersuchungen hat sich gezeigt, dass es äußerst problematisch ist, theoretische Explikationen und empirische Implikationen zu bestimmen – diese sind teilweise deutlich interpretationsabhängig, grob oder vage. Es ist genauso spekulativ, (alle) impliziten theoretischen Voraussetzungen der neurowissenschaftlichen Hypothese darzustellen, wie (alle) impliziten empirischen Konsequenzen der philosophischen Theorie. Beispielsweise können theoretische Explikationen der gleichen neurowissenschaftlichen Hypothese durchaus zu unterschiedlichen Positionen, wie der Identitätstheorie, dem anomalen Monismus, der Supervenienztheorie oder auch dualistischen Positionen führen. Diese Schwierigkeit der epistemischen bzw. ontologischen Interpretation empirischer Daten wird bereits im Rahmen der Diskussion der Ergebnisse der Versuche Libets deutlich, die sowohl dualistische als auch eine reduktive Beweisführung eingenommen wurden (vgl. Libet 1985; Eccles 1985; Roth 2003). Zudem sind, was wenig überrascht, die empirischen Implikationen der verschiedenen philosophischen Theorien in der aktuellen Diskussion der analytischen Philosophie insgesamt stark empirisch orientiert und unterscheiden sich nur subtil in wichtigen Detailfragen. Eine mögliche empirische Implikation der mentalen Kausalität, wonach neuronale Prozesse ursächlich für Handlung sind, würden weder Reduktivisten noch Nicht-Reduktivisten bestreiten. Deshalb erscheint beispielsweise die Gegenüberstellung zweier Positionen

1. „Es gibt mentale Kausalität“ versus „Neuronen verursachen Handlung“

sowie

2. „Es gibt keine mentale Kausalität“ versus „Neuronen verursachen Handlung“

[9] So ist die Diskussion um repräsentationale Eigenschaften des Gehirns bzw. des Geistes und die Entwicklung der „Language of Thought“-Theorie ein konkretes, etabliertes Beispiel für die Verknüpfung von philosophischer Theorie und neurowissenschaftlichen Hypothesen (vgl. Fodor 1975; vgl. auch Kap. 4.3.1. Inhalt). Auch die Diskussionsbeiträge und -erwiderungen zur Position des eliminativen Materialismus (vgl. z. B. Churchland 1986) oder auch der interaktionistische Dualismus des Neurowissenschaftlers Eccles (1994) verweisen auf die empirische Orientierung der theoretischen Diskussion innerhalb der Philosophie des Geistes.

wenig zielführend, da die Differenz zwischen den Theorien eine andere ist. Die Tatsache, dass neuronale Aktivität mit der Verursachung von Handlung verbunden ist, wird von keiner Position bestritten. Mentale Kausalität ist kein Gegensatz zu neuronaler Aktivität. Allerdings muss bestimmt werden, was das spezifisch mentale dieser Kausalität ist. Auch eine einfache Kontrastierung zwischen Eigenschaften und Ereignissen mit Neuronenaktivität oder zwischen mentalen und physischen Eigenschaften reicht nicht aus, der Frage der mentalen Kausalität angemessen zu begegnen, da eine solche Beschränkung die Komplexität des Phänomens verfehlen würde. Außerdem kann zwar zwischen den empirischen Implikationen der physischen Eigenschaften und den empirischen Beobachtungssätzen sehr leicht ein Bezug hergestellt werden, aber auch für mentale Eigenschaften sind empirische Implikationen denkbar, die dieses leisten (vgl. Kap. 6). Wer diese Möglichkeit negiert, tappt wechselseitig in den von Northoff beschriebenen „konditionalen Fehlschluss" (vgl. Northoff 2004, 28f.).[10] Dafür erscheinen Implikationen wie Explikationen insgesamt zu vage, um sie beispielsweise auf eine einzige Aussage bzw. einen einzigen Theorie- bzw. Beobachtungssatz zu beschränken. Und damit könnte das Phänomen der mentalen Kausalität in seiner gesamten Dimension und in seinen Verknüpfungen nicht angemessen untersucht werden.

Das abweichende methodische Vorgehen begründet sich auch darin, dass „mentale Kausalität" im Grunde ein hybrider Begriff ist, der bereits empirisch und theoretisch unterschiedliche Positionen verbindet. Schließlich wird der Bereich des Mentalen nicht zwangsläufig als physisch bzw. kausal betrachtet. Der Begriff der Kausalität, so wie er heute verwendet wird, verweist auf einen empirischen physischen Prozess. Hieraus entsteht ja gerade das Problem der mentalen Kausalität, da beide Begriffe bzw. ihre dahinter stehenden Voraussetzungen nicht ohne weiteres miteinander vereinbar scheinen. Nicht grundlos wird von einer explanatorischen Lücke gesprochen. Diese Lücke oder auch Kluft ist bereits in dem Begriff der „mentalen Kausalität" selbst angelegt. In der Diskussion innerhalb der analytischen Philosophie des Geistes wird aber deutlich, dass es gute Gründe dafür gibt, den physischen Begriff der Kausalität auf den mentalen Bereich anzuwenden und an jenem Begriff insgesamt festzuhalten (vgl. Kap. 3.6). Insofern ist der empirische Bereich innerhalb einer sehr differenzierten philosophischen Diskussion in den logischen Bedingungen stark vertreten. Damit bietet sich die Frage der mentalen Kausalität für eine neurophilosophische Untersuchung geradezu an. Andererseits ist diese Diskussion derart vielschichtig und hat über die Jahre an Komplexität zugenommen, dass eine Gegenüberstellung beispielsweise der Theoriesätze und der Beobachtungssätze nicht ausreicht. Der Begriff der men-

[10] Eine Konjunktionsanalyse kann nur die Schnittmenge zwischen natürlichen Bedingungen und logischen Bedingungen ermitteln, also nur jene logischen Theoriesätze, die „deckungsgleich" mit den Beobachtungssätzen der empirischen Bedingungen sind. Diejenigen Theoriesätze, die sich außerhalb der empirischen Bedingungen befinden, sind insofern relevant, als sie unter der Annahme des epistemischen Primats gleichwertig sind. Anders formuliert: auch die Beobachtungssätze unterliegen epistemischen Limitationen und können keine vollständige Auskunft über epistemische wie ontologische Sachverhalte geben.

talen Kausalität ist inhaltlich bzw. epistemisch wie ontologisch inzwischen stark aufgeladen und kann nur schwer aus dem jeweilig diskutierten Kontext herausgerissen werden. Deshalb ist Northoffs Vorschlag, die jeweiligen Theorie- bzw. Beobachtungssätze wieder in ihren jeweiligen Kontext zurückzuführen, formal durchaus sinnvoll. Allerdings stellt sich die Frage, ob jene sinnvoll aus ihrem jeweiligen Kontext herausgelöst werden können.

Letztlich kann man anmerken, dass eine strikte Anwendung der neurophilosophischen Methode Northoffs in ihrem mehrfachen wechselseitigen Austausch bzw. Kreuzvergleich der Theorie- und Beobachtungssätze Gefahr läuft, zu stark zu formalisieren und damit den eigentlichen Inhalt der Untersuchung zu verlieren. Andererseits werden mit den aufgestellten Prinzipien der Asymmetrie, der Bidirektionalität und der transdisziplinären Zirkularität sowie der Forderung nach Konsistenzprüfung und Analogisierung auf empirischer wie theoretischer Ebene wichtige Voraussetzungen für die systematische Verknüpfung von empirischen und logischen Bedingungen bzw. neurowissenschaftlicher Hypothese und philosophischer Theorie geschaffen. An ihnen orientiert sich deshalb auch das Vorgehen der vorliegenden Arbeit.

### 1.4 Aufbau und Vorgehensweise der Arbeit

In der vorliegenden Arbeit soll mentale Kausalität unter neurophilosophischen Gesichtspunkten betrachtet werden. Das bedeutet, dass nicht allein theoretische Fragen der Bewusstseinsphilosophie, sondern auch neurowissenschaftliche Forschung mit einbezogen wird. Ziel ist es dabei, mögliche neue Lösungsansätze für mentale Kausalität aufzuzeigen und zu entwickeln. Neben den gerade gestellten Fragen soll untersucht werden, inwiefern philosophische Überlegungen zu mentaler Kausalität mit empirischen Beobachtungen kompatibel sind. Im Gegenzug muss untersucht werden, ob empirische Hypothesen theoretischen Anforderungen genügen.

Im Verlauf dieser Arbeit wird zunächst ein theoretischer Überblick über die Entwicklung der Diskussion der Frage der mentalen Kausalität innerhalb der analytischen Philosophie des Geistes gegeben. Im Schwerpunkt werden dabei die Identitätstheorie, funktionalistische Positionen, die Position des anomalen Monismus sowie supervenienztheoretische Überlegungen hinsichtlich mentaler Kausalität vorgestellt und diskutiert (vgl. Kap. 2). Dabei werden bereits wesentliche Voraussetzungen und Merkmale deutlich, die maßgeblich für ein systematisches Problemverständnis mentaler Kausalität sind. Ebenso wird aufgrund der Kritik, die an den verschiedenen Positionen geübt wird, deutlich, dass die Frage der mentalen Kausalität nicht zugunsten einer Position als vollständig abgeschlossen gedeutet werden kann, auch wenn anscheinend die Argumente eines (starken) Physikalismus überwiegen (vgl. Kap. 2.1.3; 2.2.5; 2.3.3; 2.4.4). Die

philosophische Diskussion befindet sich somit in einem Dilemma, da einerseits mentale Kausalität erklärt werden soll, dies andererseits aufgrund der Voraussetzungen nicht möglich ist.

Insofern wird deshalb die Entwicklung der theoretischen Diskussion beschrieben, als hieraus wesentliche ontologische wie epistemische Voraussetzungen für mentale Kausalität abgeleitet und definiert werden können (vgl. Kap. 3). Auf diese Weise wird der Merkmalsraum bzw. der theoretische Rahmen, unter dem mentale Kausalität diskutiert wird, bestimmt und auf Konsistenz hin überprüft. Dieser theoretische Bezugsrahmen findet im weiteren Verlauf der Diskussion seine notwendige Berücksichtigung. Abhängig von der jeweiligen ontologischen Grundposition, der damit verbundenen Bestimmung von Eigenschaften, Ereignissen sowie deren Einbettung in ein z. B. mereologisches Weltbild wird die Problematik unterschiedlich definiert und werden unterschiedliche Lösungsansätze entwickelt (vgl. Kap. 3.1). Zentral ist dabei die Unterscheidung zwischen dem mentalen und dem physischen Bereich, aus dem per definitionem die Erklärungslücke für mentale Kausalität folgt (vgl. Kap. 3.2). Insofern der traditionelle Dualismus implizit bzw. allein begrifflich aufrechterhalten wird, scheint unter den gegebenen Voraussetzungen der kausalen Geschlossenheit der Physik, des Ausschlussarguments sowie der Anforderung strikter Gesetze eine eigenständige kausale Bedeutung für den mentalen Bereich nicht möglich zu sein. So laufen z. B. mentale Eigenschaften Gefahr, lediglich epiphänomenal und damit wirkungslose Anhängsel physischer Eigenschaften zu sein (vgl. Kap. 3.3; 3.4; 3.5; 3.6.1). Dabei wird allerdings ein grundsätzliches Dilemma deutlich, da der mentale Bereich einseitig, beinahe deduktiv ausgeschlossen wird, aber dennoch Teil des alltäglichen Lebens oder auch wissenschaftlicher Forschung ist. So geraten zunehmend die Voraussetzungen selbst in die Kritik. Es stellt sich dabei schließlich die Frage, inwiefern der Begriff der physischen Kausalität auf den mentalen Bereich übertragbar ist (vgl. Kap. 3.6; 3.3.2; 3.4.3; 3.5.2; 3.6.3).

Deshalb wird in einem ersten Kernkapitel versucht, den Begriff „mentale Kausalität“ genauer zu bestimmen. Hierzu werden sechs unterschiedliche Kriterien entwickelt, auf mentale Kausalität bezogen und diskutiert: das Kriterium der Regularität, das Kriterium der Multikausalität, das Kriterium des Inhalts, das Kriterium der Relationalität, das Kriterium der Intentionalität sowie das Kriterium der Proportionalität (vgl. Kap. 4). An dieser Stelle wird bereits deutlich, dass der Begriff der mentalen Kausalität, richtig verstanden, den herkömmlichen Rahmen der physischen Kausalität überschreitet und als Begriff modifiziert werden muss. Ferner erfolgt durch die Entwicklung der Kriterien für mentale Kausalität auch eine Zuspitzung des theoretischen Bereichs bzw. der Theoriesätze. Somit soll ein systematischer Vergleich zwischen Theoriesätzen und Beobachtungssätzen bzw. zwischen empirischen Implikationen und theoretischen Explikationen im Rahmen der vorgestellten neurophilosophischen Methodik durchgeführt werden. Das Herausarbeiten der Kriterien kann dabei durchaus als Bestimmung detaillierter Theoriesätze mit Bezug zur mentalen Kausalität betrachtet werden. Entsprechend der zeitgenössischen Diskussion

sowie der theoretischen Voraussetzungen werden die Kriterien für mentale Kausalität zunächst im Rahmen der logischen Bedingungen, der auch empirische Bedingungen umfasst, diskutiert. Dabei werden auch unterschiedliche Positionen und Widersprüche innerhalb der philosophischen Diskussion herausgearbeitet, um so einerseits die Kriterien auf ihre Relevanz für mentale Kausalität hin zu bestimmen und andererseits mentale Kausalität zu spezifizieren.

In einem weiteren Schritt werden den entwickelten Kriterien für mentale Kausalität die empirischen Ergebnisse bzw. Beobachtungssätze zweier neurowissenschaftlicher Studien gegenüber gestellt (vgl. Kap. 5). Diese beiden Studien wurden dahingehend ausgewählt, dass sich die verschiedene Kriterien, die mit der Frage der mentalen Kausalität verbunden sind, in einem weiteren Schritt auf diese beziehen lassen und einen spezifischen Vergleich ermöglichen. Insofern folgt die Arbeit der Methodik, wonach der philosophischen Theorie zunächst die neurowissenschaftliche Hypothese gegenübersteht bzw. Theoriesätze unter ihren logischen Bedingungen mit den Beobachtungssätzen unter natürlichen Bedingungen systematisch konfrontiert werden müssen. Um dem Phänomen der mentalen Kausalität auf empirischer Ebene gerecht zu werden, das heißt, eine Vergleichbarkeit zwischen möglichen theoretischen Explikationen und den philosophischen Theoriesätzen sowie möglichen empirischen Implikationen und den neurowissenschaftlichen Beobachtungssätzen zu erreichen, wurden Studien ausgewählt, die verschiedene Aspekte der mentalen Kausalität zunächst implizit berücksichtigen und sich inhaltlich ergänzen (vgl. Kap. 5). Die Studie „Dissociable Networks for the Expectancy and Perception of Emotional Stimuli in the Human Brain" legt ihren Schwerpunkt auf die Differenzierung antizipatorischer und perzeptiver emotionaler Stimulus-Prozessierung bzw. die Wirkung von Aufmerksamkeit und Wahrnehmung auf den Faktor Emotion (vgl. Kap. 5.1). Die Studie „Segregated neuronal representation of distinct emotion dimensions in the prefrontal cortex – an fMRI study" legt ihren Schwerpunkt auf die Differenzierung emotionaler und kognitiver Prozessierung sowie die neuronale Prozessierung emotionaler Dimensionen (vgl. Kap. 5.2). Aus beiden Studien lassen sich als theoretische Explikationen unter anderem intra-mentale (kausale) Prozesse, mögliche kausale Wechselwirkungen zwischen Gehirn und Umwelt sowie mental-kausale Handlungsumsetzungen ableiten. Ferner lassen sich die für mentale Kausalität definierten Kriterien unter Einbeziehung der theoretischen Voraussetzung prüfen und diskutieren (vgl. Kap. 6).[11]

Dieser Vergleich bzw. die Anwendung der Kriterien auf die empirische Ebene erfolgt in einem weiteren Kernkapitel dieser Arbeit (vgl. Kap. 6). In der Diskussion werden die Forderung nach Konsistenz, Analogisierung und möglicher Modifizierung der Begriffe in wechselseitiger

[11] In der Anwendung wird deutlich, dass die Vorauswahl der philosophischen Frage einerseits sowie der entsprechenden neurowissenschaftlichen Hypothese bzw. Studie andererseits bereits Teil der neurophilosophischen Methodik ist. Es müssen bei diesem Schritt implizite wie explizite Übereinstimmungen wie auch Differenzen auf der theoretischen wie empirischen Ebene beachtet werden. Die Entwicklung der Kriterien ist hierbei von Nutzen, da eine feinkörnige, subtile Gegenüberstellung und Untersuchung beider Bereiche hinsichtlich mentaler Kausalität möglich ist.

Orientierung berücksichtigt, wenn auch nicht in der von Northoff entwickelten systematischen Strenge durchgeführt. Dabei werden durchaus die mit den Kriterien für mentale Kausalität verbundenen Theoriesätze den Beobachtungssätzen der neurowissenschaftlichen Studien gegenübergestellt und wechselseitig über den Weg der „empirischen Implikation" wie „theoretischen Explikation" miteinander verglichen. Die Tendenz aus der näheren Bestimmung der Kriterien für mentale Kausalität bestätigt sich insofern, als tatsächlich der herkömmliche Begriff der mentalen Kausalität redefiniert und modifiziert werden muss. Beispielsweise impliziert im Rahmen des Kriteriums der Regularität die psychophysische Korrelation auf der empirischen Ebene einen veränderten Gesetzesbegriff über strikte Gesetze hinaus. Ebenso muss die Relation von einer dualistischen Beschreibung hin zu einer dreifachen Relation geführt werden (vgl. Kap. 6.1). Für das Kriterium der Multikausalität zeichnet sich im Vergleich ein Dilemma ab: Auf verschiedenen Ebenen gibt es Argumente für und gegen dieses Kriterium. Einen möglichen Ausweg bietet die weitere Differenzierung des Begriffs zwischen den verschiedenen Ebenen (vgl. Kap. 6.2).

Dem Vergleich des Kriteriums „Inhalt" kommt insofern eine besondere Bedeutung zu, als sich hieran verschiedene zentrale Problemfelder mentaler Kausalität anschließen. Schließlich steht eine Naturalisierung z. B. propositionaler Eigenschaften bislang aus. Die Verknüpfung von externem, abstraktem, mentalem Inhalt mit internen, konkreten, physischen Neuronen bleibt problematisch. Ein monokausales Modell scheint hierfür jedoch nicht ausreichend zu sein. Zusätzlich wird durch den Vergleich die Bedeutung sowohl kognitiver als auch phänomenaler Eigenschaften hinsichtlich mentaler Kausalität deutlich (vgl. Kap. 6.3). In Anknüpfung an das Kriterium des Inhalts macht der Vergleich des Kriteriums der Relationalität die Notwendigkeit begrifflicher Redefinition besonders deutlich. Er führt zu dem Ergebnis, dass weder die subjektive noch die neuronale Ebene als ausschließlich intrinsisch oder extrinsisch einzustufen ist. Auch aufgrund der neuronalen Binnendifferenzierung wurde deutlich, dass das Gehirn vielmehr als „intrinsisch-relationales" Organ betrachtet werden kann (vgl. Kap. 6.4). Die Anwendung des Kriteriums der Intentionalität auf die Beobachtungssätze der neurowissenschaftlichen Studien betont die wechselseitige Herausforderung des Normativen für den neuronalen Bereich, u. a. am Beispiel des Phänomens der Disposition. Im Ergebnis darf mentale Kausalität nicht von Normativität getrennt werden, da ansonsten keine echte Intentionalität dargestellt wird (vgl. Kap. 6.5). Schließlich ist auch für den Vergleich des Kriteriums der Proportionalität zunächst eine weitergehende Differenzierung des physischen wie des mentalen Erklärungsbereichs notwendig, um das Prinzip auf der empirischen wie epistemischen Ebene angemessen zu beschreiben. Für die ontologische Ebene konnte in der weiteren Diskussion festgestellt werden, dass adäquate Kausalität keine Eigenschaften auslassen darf, da diese sich wechselseitig bedingen und komplementär ergänzen (vgl. Kap. 6.6).

Unter Berücksichtigung der Ergebnisse aus dem wechselseitigen Vergleich der logischen und der empirischen Bedingungen wird schließlich eine neurophilosophische Hypothese zur Frage der mentalen Kausalität entwickelt (vgl. Kap. 7). Unter Rekurs auf die aristotelische Ursachenlehre wird als neurophilosophische Hypothese das Modell einer finalen Kausalität entwickelt. Dieses Modell wird auf die Ergebnisse des systematischen Vergleichs zwischen den sechs Kriterien für mentale Kausalität und den Beobachtungssätzen der neurowissenschaftlichen Studien bezogen. Bei Vermeidung theoretischer und empirischer Widersprüche soll dieses Modell sowohl in theoretischer als auch in empirischer Hinsicht valide sein. Deshalb wird über das epistemische Modell der finalen Kausalität auf der ontologischen Ebene eine spezifische Ereignisontologie entwickelt, die dem erweiterten Rahmen zwischen Gehirn und Umwelt bzw. zwischen mentalen Eigenschaften, physischen Eigenschaften sowie Umwelteigenschaften bzw. -ereignissen entspricht.

Abschließend werden die Ergebnisse dieser umfassenden neurophilosophischen Untersuchung, dem systematischen Vergleich wie auch der anschließenden neurophilosophischen Hypothese zusammengefasst (vgl. Kap. 8). Der Begriff „mentale Kausalität“ erscheint zumindest in Teilaspekten in einem anderen Licht, es ergeben sich weitere Fragen und Untersuchungsansätze. Dennoch kann diese Arbeit bereits als ein Plädoyer für eine „echte“ mentale Kausalität mit mentalen, propositionalen wie abstrakten Begriffen verstanden werden. Im Ausblick wird deutlich, dass weiterhin Forschungsbedarf hinsichtlich der Spezifizierung des Begriffs besteht.

## Kapitel 2

## Die Entwicklung der zeitgenössischen philosophischen Diskussion

Das Problem der mentalen Kausalität unterscheidet sich in der zeitgenössischen Diskussion der analytischen Philosophie fundamental von der traditionellen cartesianischen Version. Der cartesianische Substanzdualismus wurde von den analytischen Philosophen nahezu vollständig zurückgewiesen, heute werden andere ontologische und auch methodologische Wege beschritten. Als Ausgangspunkt der Debatte der analytischen Philosophie mag der Konflikt zwischen Positivisten und Historisten zu Beginn des 20. Jahrhunderts betrachtet werden (vgl. Marras 2003, 243). Während die Positivisten forderten, dass im Rahmen einer universalen Wissenschaftstheorie auch die sogenannten Humanwissenschaften das deduktiv-nomologische Modell der wissenschaftlichen Erklärung übernehmen sollten, beharrten die Historisten unter dem Hinweis auf die Singularität historischer Ereignisse und menschlichen Verhaltens auf der Methode des empathischen Verstehens und der hermeneutischen Interpretation (vgl. z. B. Colingwood 1946; Dray 1957; Hempel 1942, 1963; alle zitiert nach: Marras 2003, 243f.).

Darüber hinaus wurde die Diskussion um die Frage der mentalen Kausalität zunächst sehr stark aus einer analytisch-behaviouristischen Perspektive betrachtet. Danach können nach einer speziellen reduktionistischen Variante mentale Phänomene mehr oder weniger mit Verhaltensdispositionen identifiziert werden. Vergleichbar mit empirischen Dispositionen, wonach beispielsweise Salz die Disposition besitzt, sich in Wasser aufzulösen, oder Glas die Disposition besitzt, zerbrechlich zu sein, sollen für Menschen Verhaltensdispositionen gelten. Unter bestimmten Umständen würde sich eine Person auf eine ganz bestimmte Art und Weise verhalten; „etwas zu wünschen“ entspräche der Disposition „zu versuchen etwas zu erreichen, wenn die Möglichkeit dazu besteht“ (vgl. Ryle 1949). Über mentale Phänomene zu sprechen, bedeutet insofern, über Verhalten bzw. Verhaltensdispositionen zu sprechen. Wenn z. B. Carl eine Cola trinken möchte, wird er, unter gleichen Voraussetzungen, eine Cola trinken. Sowohl Ryles als auch Wittgensteins Kritik richtete sich gegen die Annahme, dass eine wie auch immer geartete Psychologie über verborgene geistige Zustände Verhalten und Handlungen erklärt bzw. es nicht öffentlich beobachtbare Vorgänge – einen „Geist in der Maschine“ – gibt (vgl. Ryle 1949, 325). In seinem Privatsprachenargument formulierte Wittgenstein u. a. die Position, dass mentale Ausdrücke ebenso wenig wie andere Ausdrücke privat oder geheim sein können – es kann keine Privatsprache geben. Für alle Ausdrücke muss es öffentlich zugängliche Kriterien geben: „An inner process stands in need of outward criteria[.]“ (Wittgenstein 1953, PU § 580) Nur dadurch ist letztlich eine relative Form der Gewissheit zu erreichen, ob Ausdrücke richtig oder falsch verwendet werden.

Eine mögliche Konsequenz aus der Übertragung des deduktiv-nomologischen Modells auf die Erklärung mentaler Phänomene wäre die Legitimierung eines mentalen Bereichs mit Ereignissen, Zuständen oder Prozessen, die als eigenständige, zählbare, episodische Entitäten in Abgrenzung zueinander und zu dem aus ihnen resultierenden Verhalten zu bestimmen wären, obwohl sie nicht als vollständig physisch gelten würden. Dagegen argumentiert insbesondere Ryle, es würde ein Kategorienfehler begangen, Gründe und Ursachen miteinander zu vermischen (vgl. Ryle 1949, 13f.). Der Geist kann nicht kausal auf den Körper einwirken, mentale Ereignisse bzw. Gründe können nicht verursachen, da Willensakte oder Handlungen nicht unter den Bereich des kausalen Systems fallen. Geistige Phänomene werden gewissermaßen außerhalb einer physischen Kausalität angesiedelt. Gründe können ferner deshalb keine Ursachen sein, da sie begrifflich mit dem Verhalten verknüpft sind, das sie verursachen sollen. Eine echte Ursache muss aber, nach Hume, eigenständig und unabhängig von ihrer Wirkung sein (vgl. Hume 1748). Solange jedoch ein Wunsch und das Objekt des Wunsches begrifflich miteinander verbunden sind, kann diese Beziehung nicht zugleich eine empirische und kontingente Relation von Ursache und Wirkung sein, wie das „Logical Connection-Argument" zum Ausdruck bringt (vgl. Melden 1961; Stoutland 1970; zitiert nach Marras 2003, 244).

Gegen den analytischen Behaviourismus wurden verschiedene Argumente aufgeführt. Unter anderem wurde darauf hingewiesen, dass aufgrund der starken behaviouralen Perspektive keine Unterscheidung zwischen Menschen mit geistigen Zuständen und Zombies ohne geistige Zustände möglich wäre, solange die Zombies wie Menschen aussähen und sich wie Menschen verhielten. In einem Gedankenexperiment wurde über die Gruppe der „Super-Spartaner" nachgedacht, die aufgrund von Training und Disziplin jeden noch so entsetzlichen Schmerz klaglos ertragen, unfähig, über ihre innere Befindlichkeit in ihrer psychologischen Bandbreite zu reden – auch wenn sie behaviouralen Überlegungen entgegen über innere subjektive Zustände verfügen (vgl. Putnam 1975b). Daneben wurde auch auf das grundsätzliche Problem nicht nur des analytischen Behaviourismus hingewiesen, mentale Zustände bzw. Handlung ohne mentale Begriffe zu erklären (vgl. Chisholm 1957). Schließlich argumentierte Davidson in seinem fundamentalen Aufsatz „Actions, Reasons, and Causes" erfolgreich gegen das Logical Connection-Argument und konnte zeigen, dass Gründe, die eine Handlung erklären, sehr wohl zugleich Ursachen sein können. Danach bezieht sich das Argument lediglich auf eine bestimmte Beschreibung von Gründen und Handlungen, während eine kausale Verbindung zwischen Ereignissen besteht, die die Gründe und Handlungen konstituieren. Die Beziehung zwischen Ereignissen ist sowohl empirisch als auch kontingent, wie von Hume gefordert. Ferner argumentierte Davidson, dass es nicht einsichtig sei, wie Gründe eine Handlung rationalisieren könnten, wenn sie nicht zugleich eine echte kausale Erklärung liefern würden. Damit wird implizit gefordert, dass echte Erklärungen von Handlungen kausale Erklärungen sein müssen. Nach Davidsons Position muss eine Erklärung

einer Handlung die beiden notwendigen Bedingungen erfüllen, dass erstens die Glaubensannahmen bzw. Wünsche o. ä. des Handelnden relativ zu einer nachvollziehbaren Handlung spezifiziert werden und dass zweitens die Glaubensannahmen oder Wünsche auch tatsächlich die Handlung verursachen (vgl. Davidson 1963).

Doch auch wenn Davidson weithin akzeptiert zeigen konnte, dass entgegen dem Logical Connection-Argument Gründe durchaus Ursachen sein können, bleibt die Frage bestehen, *wie* es möglich ist, dass Gründe Ursachen sind: Wie können mentale Zustände bzw. Ereignisse als Glaubensannahmen, Wünsche oder Intentionen die kausale Kraft haben, intentionales Verhalten, Körperbewegungen bzw. physische Ereignisse zu verursachen? Der Begriff der mentalen Kausalität wurde und wird auch innerhalb der im Folgenden diskutierten Positionen der analytischen Philosophie des Geistes weiter differenziert. Dabei hat die Frage der mentalen Kausalität in der analytischen Diskussion bereits eine Entwicklung erfahren, wie auch die gesamte Leib-Seele-Diskussion sich weiterentwickelt hat. Diese Entwicklung führt von der identitätstheoretischen Position hin zu funktionalistischen bzw. supervenienztheoretischen Überlegungen.[12] Auch wenn die genannten Positionen teilweise in Widerstreit zueinander stehen, haben sie alle zur Ausdifferenzierung der Fragestellung und ihres metaphysischen Gerüsts Wertvolles und Wesentliches geleistet. So ist es nur konsequent, dass diese Theorien auch heute noch vertreten werden und inhaltlich zur Diskussion beitragen.[13] Um die Entwicklung und Veränderung der Fragestellung sowie des metaphysischen Gerüsts im Hinblick auf zentrale Kerngedanken aufzuzeigen, sollen die Positionen der Identitätstheorie, des Funktionalismus, des anomalen Monismus und der Supervenienztheorie im Folgenden in einem Überblick mit ihrem Bezug zur mentalen Kausalität dargestellt werden. Dieser Überblick dient dazu, die für mentale Kausalität notwendigen Voraussetzungen und Kriterien in ihrem theoretischen Bezugsrahmen darzustellen, damit diese in einem weiteren Schritt herausgearbeitet werden können (vgl. Kap. 3; Kap. 4).

Dass die Diskussion in der analytischen Philosophie um die Frage der mentalen Kausalität in ihrer Entwicklung auch ein Ringen um die bestmöglichen Erklärungsmodelle ist, wird allein dadurch deutlich, dass sich zum einen die verschiedenen Positionen zeitlich überschneiden und mit wechselseitigem Bezug zueinander diskutiert werden. Zum anderen modifizieren aber sogar verschiedene der Protagonisten, wie z. B. Putnam, Lewis, Armstrong, Fodor, Block oder auch Davidson und Kim ihre Positionen aufgrund neuer Argumente und neuer Überzeugungen. Dabei

[12] Streng genommen beginnt die Entwicklung mit der Position des analytischen Behaviourismus (vgl. z. B. Ryle 1949; Wittgenstein 1953). Da diese Position selbst jedoch mentale Zustände aus dem kausalen Geschehen ausklammert, heute kaum mehr diskutiert wird und zudem wesentliche Vorzüge, aber auch Nachteile z. T. in den Funktionalismus mit eingeflossen sind, wird diese Position in dieser Arbeit nachstehend nicht weiter erörtert.

[13] Auch wenn die Position des Anomalen Monismus stark an Davidsons Person geknüpft war und dieser selbst seinen Standpunkt hin zur globalen Supervenienz modifizierte, sind die Kerngedanken, wonach es z. B. keine psychophysischen strikten Gesetze gibt, nach wie vor fester Bestandteil der zeitgenössischen Diskussion (vgl. Kap. 3.5).

werden verschiedene Themenbereiche, Vorraussetzungen sowie Kriterien für mentale Kausalität diskutiert. Ihnen gemeinsam ist, dass sie sich sozusagen als Negativfolie auf eine weitere Position, die des Epiphänomenalismus mentaler Eigenschaften, beziehen. Dass mentale Eigenschaften lediglich epiphänomenal sind, zieht sich wie eine Drohung oder aber als ausdrücklich akzeptierte Konsequenz durch die gesamte Diskussion um die Frage der mentalen Kausalität (vgl. z. B. Chalmers 1997, 150ff.; Fodor 1989).

Für die Identitätstheorie bzw. Typenidentität ist dabei von Bedeutung, dass die Geist-Gehirn-Korrelation als gesetzmäßig und die Identität von physischen und mentalen Eigenschaften als eine empirische Wahrheit aufgefasst wird (vgl. Kap. 2.1). In ihrer Tendenz ist diese Position in Anknüpfung an das Wissenschaftsverständnis der Positivisten reduktiv – die physische Welt ist in sich kausal geschlossen. Andererseits soll über die Identität mentaler und physischer Eigenschaften die kausale Wirksamkeit mentaler Eigenschaften gewährleistet werden. Aus dem Prinzip der Multirealisierbarkeit werden in der Kritik zur Position der Typenidentität verschiedene funktionalistische Positionen entwickelt (vgl. Kap. 2.2). Üblicherweise soll diese Realisierungsrelation die Nicht-Reduzierbarkeit der mentalen Eigenschaften gewährleisten. Mentale Eigenschaften erfüllen als funktionale Eigenschaften eine kausale Rolle, über die sie bestimmt werden. In ihrer extremen physikalistischen Ausprägung lässt der so genannte Computerfunktionalismus (vgl. Kap. 2.2.2) allerdings wenig Raum für eigene kausale Kraft funktionaler Eigenschaften, denen sogar eher ein epiphänomenaler Charakter zukommt. Insofern lässt sich die Position des Anomalen Monismus bzw. der Tokenidentität sowohl als Reaktion auf Typenidentität als auch auf (nicht-) reduktiven Funktionalismus interpretieren (vgl. Kap. 2.3).

Davidson knüpft in den 1970er-Jahren an seine kausale Handlungstheorie an und macht sich erneut dafür stark, mentale Ereignisse als kausale Ereignisse zu bestimmen. Dies ist letztlich vor dem Hintergrund seiner speziellen Ereignisontologie möglich, die ähnlich seiner Verteidigung von Gründen als Ursachen beschreibungsunabhängig und dadurch kausal neutral ist. Ironischerweise wird ausgerechnet seine deutliche Ablehnung psychophysischer Gesetze aufgrund nicht vorhandener strikter Gesetze für mentale Eigenschaften in der Reaktion auf seine Position als Argument gegen die kausale Kraft mentaler Eigenschaften ausgelegt. Dies ist auch auf eine Verschiebung der ontologischen Perspektive zurückzuführen, wodurch die Eigenschaften gegenüber den Ereignissen ontologisch gestärkt werden. So soll die Supervenienzrelation aus dem anomalen Monismus heraus zwar zunächst die kausale Kraft mentaler Eigenschaften sichern und stärken. Doch entpuppen sich mentale Eigenschaften unter supervenienter Kausalität anscheinend nur als schwache Erben der eigentlichen kausal wirksamen physischen Eigenschaften (vgl. Kap. 2.4). In dem Spannungsverhältnis zwischen ontologischer sowie kausaler Reduzierbarkeit bzw. Nicht-Reduzierbarkeit findet dabei die Unterscheidung zwischen Gründen und Ursachen ein neues

Gewand: echte mentale Kausalität soll nicht nur als epistemische, sondern vor allem als ontologische kausale Erklärung begründet werden.

## 2.1 Identitätstheorie und mentale Kausalität

Seit den späten 1950er- und bis in die 1970er-Jahre hinein war die Identitätstheorie mit dem Modell des Typenphysikalismus bzw. der Typenidentität die herrschende Lehre in der Debatte um die Körper-Geist-Relation.[14] Sie geht von der plausiblen Annahme aus, dass jegliche mentale Prozesse, wie z. B. Glaubensannahmen, Wünsche oder Empfindungen einen Gehirnzustand bzw. einen neuronalen Prozess als physische Basis haben. Diese Geist-Gehirn-Korrelation wird als gesetzesartig aufgefasst. Wenn ein bestimmter Schmerz verspürt wird und zugleich bestimmte Neuronen aktiv sind, ereignet sich diese Korrelation nicht zufällig, sondern als eine gesetzesmäßige Regularität (vgl. Kap. 3.5; Kap. 4.1). Jegliche Änderung im so genannten mentalen Bereich ist verbunden mit einer Veränderung im physischen Bereich.

Die Identitätstheorie erklärt diese systematische Korrelation von mentalen und physischen Eigenschaften, Ereignissen oder Zuständen dadurch, dass z. B. Ereignistypen miteinander identisch sind. Genauso wie Blitzschläge und elektrische Entladungen sich auf dasselbe Phänomen beziehen, beziehen sich Schmerzen und C-Faser-Aktivierungen auf denselben neuronalen bzw. physischen Prozess – letztlich sind die mentalen Zustände und Ereignisse mit diesem identisch.

### 2.1.1 Bestimmung der psychophysischen Identität

Ein wesentlicher Grund, die Identitätstheorie zu etablieren, war das Prinzip der Einfachheit bzw. der ontologischen Sparsamkeit. Mit Verweis auf „Ockhams Rasiermesser" argumentiert Smart[15] (1959), dass es wenig plausibel ist, sämtliche Phänomene der Wissenschaft als zunehmend komplexe Anordnungen physischer Bestandteile zu beschreiben, das Bewusstsein aber eine zusätzliche Beschreibung benötigt, die über das Physische hinausgeht. Smart vertritt somit ein stark intuitiv gefärbtes Argument, wenn er aus Sparsamkeitsgründen ablehnt, dass für eine vollständige Beschreibung des inneren Zustands eines Menschen nicht nur die physischen Prozesse, sondern auch Sinneseindrücke und andere Bewusstseinszustände herangezogen werden müssen. Es ist

[14] Ein alternatives identitätstheoretisches Modell ist die Tokenidentität, welche insbesondere Davidson mit der Theorie des anomalen Monismus formuliert hat. Diese Position wird später diskutiert (vgl. Kap. 2.3).

[15] Smarts Artikel „Sensations and Brain Processes" (1959, 141-156) kann ebenso wie Places Aufsatz „Is Consciousness a Brain Process?" (1956, 44-50) als grundlegende Publikation für die moderne Formulierung der Identitätstheorie angesehen werden.

ihm nicht einsichtig, warum Sinneseindrücke bzw. Bewusstseinszustände aus dem physikalistischen Weltbild außen vor gelassen bleiben sollen. Er kann aus verschiedenen Gründen einfach nicht glauben, dass alles durch Rekurs auf die Physik explizierbar sein soll, außer dem Vorkommen von Sinneseindrücken:

> That everything should be explicable in terms of physics […] except the occurence of sensations seems to be frankly unbelievable. Such sensations would be 'nomological danglers', to use Feigl's expression. […] I cannot believe that ultimate laws of nature could relate simple constituents to configurations consisting of perhaps billions of neurons. […] Such laws would be like nothing so far known in science. They have a queer 'smell' to them. (Smart 1962, 161-162; vgl. auch Smart 1959, 142)

Nach dem Prinzip der ontologischen Sparsamkeit gibt es keine Schmerzen zusätzlich zur C-Faser-Aktivierung. Durch die Identität der mentalen und physischen Zustände wird zudem die psychophysische Korrelation begründet und muss nicht durch komplizierte Gesetze begründet werden. Denn im Rahmen einer psychophysischen Korrelation laufen mentale Zustände Gefahr, zu „nomologischen Anhängseln" (vgl. Feigl 1958) zu werden, die nicht weiter erklärt werden können. Die Behauptung, dass Schmerzen und C-Faser-Aktivierungen miteinander identisch sind, identifiziert zwei Typen miteinander. Insofern wird hier von einer Typenidentität gesprochen. Die Identität der genannten Paare wird in der Regel über identische Eigenschaften bestimmt. Nach dem sogenannten Leibnizschen Gesetz gilt beispielsweise (vgl. Leibniz 1985, 17f.; Guttenplan 1995, 431):

> Wenn X und Y all ihre Eigenschaften gemeinsam haben, dann sind X und Y identisch.

Eine Eigenschaft kann also nicht *nur* X aber nicht Y haben, damit von einer ontologischen Identität gesprochen werden kann. Andererseits wird deutlich, dass die Identitätsbehauptung falsifiziert werden kann, indem eine einzige Eigenschaft aufgeführt wird, die X besitzt, aber nicht Y. Im Rahmen des psychophysischen Identitätsmodells wird davon gesprochen, dass z. B. die mentale Eigenschaft des Ereignisses, ein Schmerz zu sein, mit der physischen Eigenschaft dieses Ereignisses identisch ist. Ein Ereignis wird deshalb als die Instanziierung einer Eigenschaft zu einem bestimmten Zeitpunkt aufgefasst. Ein mentales Ereignis ist die Instanziierung einer mentalen Eigenschaft und ein physisches Ereignis entsprechend die Instanziierung einer physischen Eigenschaft (vgl. Kap. 3.1.1; Kap. 3.1.2). Damit ein mentales und ein physisches Ereignis als identisch gelten können, müssen sie also dieselben Eigenschaften instanziieren (vgl. Kim 1976, 1993b). Da alle kausal wirksamen Eigenschaften physische Eigenschaften sein sollen, wird dem Primat der physischen Eigenschaften Rechnung getragen. Insofern die These besagt, dass menta-

le Eigenschaften physische Eigenschaften sind, läuft die Typenidentität letztlich auf einen Reduktionismus hinaus, da es keine mentalen Eigenschaften über physische Eigenschaften hinaus gibt.

Die Typenidentitätstheorie kann folglich als eine starke reduktive physikalistische Theorie betrachtet werden. Die Reduktion mentaler Eigenschaften auf physische Eigenschaften ist durch so genannte Brückengesetze möglich (vgl. Kap. 3.3.1). Dabei ist grundsätzlich eine Theorie auf einer höheren Ebene unter Anwendung jener Brückengesetze von der Theorie auf der niedrigeren Ebene ableitbar, beispielsweise lässt sich eine Theorie über Temperatur auf eine Theorie über kinetische Energie zurückführen. In diesem Sinne sollen mentale bzw. psychologische Gesetze auf fundamentalere Gesetze zurückgeführt werden, die weniger komplexe physische Strukturen als das Gehirn betreffen. Dahinter steckt die Idee von der Einheit und der Gleichförmigkeit der Natur. Allerdings ist die Geschlossenheit der Physik zwar selbst kein physikalisches Gesetz, aber eine hochwertige empirische Generalisierung über die Praxis und den Inhalt physikalischer Wissenschaft (vgl. Kap. 3.3). Demnach ermöglicht die Vereinigung der wissenschaftlichen Theorien einen wahren und erschöpfenden Zugang zu allen physischen Phänomenen:

> [T]here is some unified body of scientific theories of the sort we now accept, which together provide a true and exhaustive account of all physical phenomena. They are unified in that they are cumulative: the theory governing phenomena out of which that phenomenon is composed and by the way it is composed out of them. The same is true of the latter phenomena, and so on down to fundamental particles or fields governed by a few simple laws, more or less as conceived in present-day theoretical physics. (Lewis 1983a, 105)

Die Theorien sind in einem kumulativen Sinne vereinigt, da eine Theorie, die ein bestimmtes physisches Phänomen herleitet, durch Theorien erklärt wird, die jene Phänomene herleiten, aus denen dieses Phänomen sich zusammensetzt. Selbiges gilt für diese letzteren Phänomene bis hin zu den fundamentalen Partikeln bzw. Feldern, die durch einige wenige Gesetze hergeleitet werden (vgl. ebd.). Einerseits werden Entitäten nach dieser Annahme vollständig durch physische Teile zusammengesetzt. Andererseits werden die Theorien, die das Gesamte erklären, selbst durch jene Theorien erklärt, die die Teile erklären, aus denen sich das Gesamte zusammensetzt.

Konsequenterweise wurde und wird die Identitätsbeziehung mentaler und physischer Zustände von den Befürwortern dieser Position immer unter Hinweis auf und im Vergleich zu anderen Identitäten, die wir gewöhnlich leichter akzeptieren, diskutiert. Schließlich kann die psychophysische Identität nicht a priori gezeigt werden. Ebenso verhält es sich aber beispielsweise mit der Identität von DNA-Molekül und Gen. Die Entdeckung dieser Identität beruht auf empirischer wie theoretischer Forschung. So wird deutlich, dass die kausale Rolle bzw. Funktion des Gens, die seinen Begriff bestimmt, auch von dem DNA-Molekül erfüllt wird. Das DNA-Molekül ist kausal für die Übertragung von Erbinformationen verantwortlich. Aus der gleichen kausalen

Funktion von DNA und Gen wird nun deren Identität geschlossen. Bei der Identitätsbestimmung mentaler und physischer Zustände wird nach demselben Muster vorgegangen. Mentale Eigenschaften werden in einer ersten Prämisse durch ihre kausalen Rollen charakterisiert – der Begriff eines mentalen Zustands ist der geeignete Begriff des Zustands der Person, um ein bestimmtes Verhalten hervorzubringen (vgl. Armstrong 1968). Die nächste Prämisse besagt, dass letztlich nur die Physik erklärungsadäquat ist:[16]

> The definitive characteristic of any (sort of) experience as such is its causal role, its syndrome of most typical causes and effects. But we materialists believe that these causal roles which belong by analytic necessity to experiences, belong in fact to certain physical states. Since those physical states possess the definitive characteristics of experience, they must be the experiences. (Lewis 1983a, 100)

Mentale Eigenschaften müssen somit in physikalischen Begriffen erklärbar sein, damit ihnen eine physische Wirkung zugesprochen werden kann. Das a priori-Verständnis des Schmerzbegriffs wird mit der Funktion der C-Faser-Aktivität abgeglichen. Sollte die empirische Forschung nachweisen, dass die C-Faser-Funktion dieselbe Funktion wie der Schmerzbegriff erfüllt, z. B. Zusammenzucken und Stöhnen bei einem Gewebeschaden, kann die Identität festgestellt werden (vgl. Armstrong 1968).[17] Die psychophysische Identität ist somit eine empirische Wahrheit und kann nur a posteriori bestätigt werden. Sie sei vergleichbar mit der Identität von Wasser und $H_2O$, Hitze und Molekularbewegung sowie Licht und elektromagnetischer Strahlung.

### 2.1.2 Das kausale Argument für psychophysische Identität

Ein wesentlicher Vorteil der Identitätstheorie liegt darin, dass aufgrund der Identität mentale Zustände oder Ereignisse nicht wirkungslos sind. Damit wird die Identitätstheorie der alltäglichen Erfahrung von Handlung und bewussten kausalen bzw. intentionalen Prozessen gerecht. Wird beispielsweise eine Nadel in eine Hand gestochen und diese daraufhin zurückgezogen, lässt sich der gesamte Prozess plausibel physisch erklären: über Nervenbahnen werden Signale an das Gehirn gesendet, die C-Fasern werden aktiviert, woraufhin im motorischen Cortex das Zurückziehen der Hand initiiert wird. Solange im Rahmen einer Korrelation die Schmerzen eine eigene Ursache für das Zurückziehen der Hand sein sollen, scheint ein Fall der kausalen Überdetermi-

[16] Jede physische Wirkung muss in rein physischen Begriffen erklärbar sein. Hintergrund dieser Prämisse bildet u. a. Lewis Modell einer „unified theory", einer Einheitswissenschaft (vgl. Kap. 3.3).

[17] Eine offensichtliche Schwierigkeit dieser Argumentation liegt darin, dass der Begriff „Schmerz" nicht so eindeutig definiert werden kann, wie für eine Identitätsbestimmung erforderlich ist. So verfügen nicht unbedingt alle schmerzfähigen Lebewesen über die so genannten C-Fasern. Insbesondere das Argument der Multirealisierbarkeit hat dieses Problem in der Argumentation verdeutlicht (vgl. Putnam 1967; Kap. 4.2.1).

nierung vorzuliegen (vgl. Kap. 3.4). Dass sämtliche Wirkungen im mentalen Bereich systematisch überdeterminiert sind, ist jedoch höchst unwahrscheinlich.[18] Die C-Fasern scheinen die eigentliche Ursache der Bewegung zu sein, entsprechend sind die Schmerzen kausal überflüssig, reine Epiphänomene. Durch die Identität von C-Faser-Aktivierung und Schmerzen wird schließlich die kausale Wirksamkeit der letzteren gerettet.

Papineau fasst die identitätstheoretischen Überlegungen zur kausalen Wirksamkeit mentaler Eigenschaften in einem kausalen Argument zusammen. Für ihn ist dieses Argument zudem das maßgebliche, um die Identitätstheorie selbst zu begründen.[19] Er formalisiert das kausale Argument über drei Prämissen (vgl. Papineau 2004, 17f.):[20]

1. Conscious mental occurrences have physical effects
2. All physical effects are fully caused by purely physical prior histories
3. The physical effects or conscious causes aren't always overdetermined by distinct causes.

Aus diesen drei Prämissen folgt für Papineau als Konklusion die Gültigkeit des Materialismus für den Bereich des Mentalen. Die erste Prämisse beschreibt Fälle, in denen bewusste Gefühle oder andere mentale Zustände Verhalten verursachen (vgl. Kap. 3.6). Die zweite Prämisse beschreibt die Geschlossenheit der Physik (vgl. Kap. 3.3). Hiernach ist jedes physische Verhalten vollständig von z. B. physischer Aktivität in den Muskeln, Nerven und schließlich dem motorischen Kortex und dem sensorischen Kortex verursacht. Zusammengefasst besagen die Prämissen 1 und 2 somit, dass einige physische Wirkungen eine bewusste und eine physische Ursache haben. Prämisse 3 besagt dagegen ausdrücklich, dass Wirkungen nicht systematisch verschiedene Ursachen haben. Daraus folgt die Konklusion, dass die bewussten Ereignisse identisch mit den physischen Ursachen sein müssen; die Gefahr der Überdeterminierung wird vermieden, da keine verschiedenen Ursachen vorliegen (vgl. Kap. 3.4).

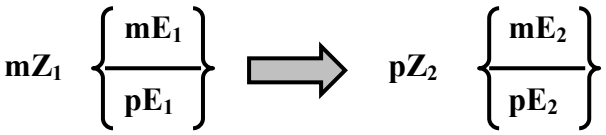
Ein mögliches Beispiel für durch Bewusstsein verursachtes körperliches Verhalten ist die Situation, sich (bewusst) durstig zu fühlen und deshalb zum Kühlschrank zu gehen, um sich ein Getränk zu holen. Nach der zeitgenössischen physikalischen Annahme wird die Bewegung durch physische Prozesse im Gehirn und in Nerven verursacht. Mentale Kausalität wird in diesem Sin-

[18] Zum Problem der Mehrfachdetermination bzw. des kausalen Ausschlusses vgl. insbesondere Kap. 3.4.

[19] In dem Aufsatz „Mind the Gap" (1998) nennt Papineau noch zwei andere Argumente für die Position des Materialismus: das Argument einer a priori Intuition und das Argument der a priori kausalen Rollen für Bewusstseinszustände. Allerdings hält Papineau beide Argumente für zu schwach und nicht ausreichend, um einen starken Materialismus zu begründen (vgl. auch Papineau 2004, 36-40).

[20] Papineau ist zwar ein „aktueller" Vertreter der Identitätstheorie, seine Argumentationslinie kann aber als repräsentativ verstanden werden. Allerdings legt Papineau selbst sich nicht fest, welche Form der Identität er favorisiert. So unterstützt das kausale Argument sowohl strikte als auch weniger strikte Identitäten (vgl. Papineau 2004, 34f.). Er selbst bezeichnet sich mit Block als einen „Inflationisten". Inflationistischer Materialismus soll mentale Konzepte in ihrer Besonderheit berücksichtigen – im Gegensatz zu so genannten „Deflationisten", die lediglich die Referenz zwischen materiellen Konzepten und Bewusstseinszuständen akzeptieren. Als Deflationisten nennt Papineau David Lewis und Daniel Dennett (vgl. ebd., 49).

ne als die Verursachung von Körperbewegung durch mentale bzw. neuronale Prozesse beschrieben. Die nahe liegende Konklusion ist die Identität von bewusstem Durst mit den physischen Prozessen (vgl. Papineau 2004, 17f.).




Grafik 3: Vor dem Hintergrund der kausalen Geschlossenheit der Physik kann eine physische Wirkung $pZ_2$ immer nur eine physische Ursache haben. Deshalb hat nur die physische Eigenschaft $pE_1$ des mentalen Zustands $mZ_1$ und nicht seine mentale Eigenschaft $mE_1$ kausale Kraft. Es besteht eine systematische Korrelation zwischen mentalen und physischen Eigenschaftstypen. Brückengesetze ermöglichen die Reduktion der mentalen Eigenschaften auf ihre Basis, die physischen Eigenschaften. In diesem Sinne sind sie miteinander identisch.

Die grundsätzliche Argumentation der Identitätstheorie orientiert sich also an der kausalen Struktur der physischen Welt. Zentral ist dabei die These der Geschlossenheit der physischen Welt. Diese These bezeichnet Papineau als die entscheidende Prämisse, um in Verbindung mit dem kausalen Argument die Position der psychophysischen Identitätstheorie stark zu machen:

> The crucial empirical premise is the completeness of physics, by which I mean that all physical effects are due to physical causes. And the argument is then simply that, if all physical effects are due to physical causes, then anything that has a physical effect must itself be physical. (Papineau 2001, 7; vgl. auch ders. 2004, 2)

Nach dem Physikalismus ist alles, auch prima facie nicht-physische Dinge, physisch. Es wird deutlich, dass man nicht den physischen Bereich verlassen muss, um für eine physische Wirkung eine völlig hinreichende Ursache auszumachen. Wenn also die Position der Geschlossenheit der Physik zutrifft, ist kein Raum übrig für nicht-physische Dinge als Ursache für physische Wirkungen – jegliche Ursache muss selbst physisch sein. In der 2. Prämisse wird das Prinzip der physikalischen Geschlossenheit in Verbindung mit ihrer Kausalstruktur formuliert:

> The thought behind premiss 2 is that such physical behaviour will always be fully caused by physical contractions in your muscles, in turn caused by electrical messages travelling down your nerves, themselves due to physical activity in your motor cortex, in turn caused by physical activity in your sensory cortex, and so on. (ders. 2001, 18)

Papineau liefert somit ein sehr detailliertes und plastisches Bild, wie er sich die kausalen Prozesse im Gehirn vorstellt, die wir Menschen als Empfindungen oder sogar Entscheidungen wahrnehmen und die schließlich die körperliche Bewegung verursachen – ganz im Sinne eines starken

Physikalismus.[21] Die Konzentration auf das Prinzip der Geschlossenheit der Physik sowie die Ablehnung von Mehrfachverursachung ist dabei paradigmatisch für die Position des Physikalismus (vgl. Kap. 3.3; Kap. 3.4).

Allerdings stellt die Frage der Mehrfach- bzw. Überdetermination die Identitätstheorie grundsätzlich vor ein schwieriges Problem. Mit welchem guten Grund kann eine Mehrfachdeterminierung ausgeschlossen werden? Schließlich sind auch Situationen denkbar, in denen unterschiedliche physische Ereignisse zu dem gleichen Resultat führen, ohne dass entscheidbar ist, dass ein Ereignis die eigentliche Ursache ist. Beispiel für mehrfach- bzw. überdeterminierte Ereignisse ist z. B. der Tod eines Mannes durch einen Schuss und einen zeitgleichen Blitzschlag. Das scheint aber das falsche Modell für mentale Kausalität zu sein, da impliziert wird, das Ergebnis hätte auch in Abwesenheit einer der beiden Ursachen stattgefunden. Papineau bezweifelt aber, dass der Satz „Ich gehe zum Kühlschrank, da meine Neuronen feuern, fühle aber keinen Durst." zutrifft (vgl. ebd.). Deshalb lehnt er kausale Mehrfachdetermination ab. Die Tatsache, dass eine Person ihren Arm hochhebt, wird nicht von zwei unterschiedlichen Ursachen – einer physischen und einer mentalen – gleichzeitig bewirkt:

> Thus, if some conscious desire causes my arm to move, and this movement in turn has such inanimate effects as a stone flying through the air, […] then my conscious desire itself will cause these inanimate effects. (Papineau 2004, 44)

Bewusste Systeme können also kausal wirksam sein, sofern ihre bewussten Eigenschaften – aufgrund ihrer Identität mit physischen Eigenschaften – unbelebte oder nicht-mentale Wirkung haben. Insofern macht sich Papineau ausdrücklich dafür stark, dass bewusste mentale Eigenschaften (englisch: „conscious mental occurences") physische Wirkung haben. Die Prämisse 1 seines kausalen Arguments ist unverzichtbar, andernfalls liefen bewusste Eigenschaften Gefahr, lediglich Epiphänomene bzw. wirkungslose Anhängsel der physischen Eigenschaften zu werden.

Aufgrund verschiedener anti-materialistischer Argumente favorisieren einige zeitgenössische Philosophen eine epiphänomenalistische Position und verwerfen Prämisse 1, wonach bewusste Ereignisse eine physische Wirkung haben (vgl. Kap. 3.6.1). Andere Bedenken gegenüber der ersten Prämisse betreffen bewusste repräsentationale Zustände und ihre kausale Wirksamkeit. Hierbei stellt sich die Frage, ob repräsentationale Zustände „weiten" oder „engen Inhalt" haben (vgl. Kap. 4.3.2). Weiter repräsentationaler Inhalt hängt von Zuständen außerhalb der Personen bzw. ihrer Gehirne ab. Beispielsweise hängt der repräsentationale Zustand „über Wasser nachdenken" von der tatsächlichen natürlichen Art Wasser in einer bestimmten Umwelt ab (vgl. Put-

[21] Theoretische und empirische Perspektive verhalten sich nach Papineau komplementär zueinander. Die Theorie kann keine genauen physischen Eigenschaften identifizieren, während die empirische Forschung über die Korrelation bestimmter Paarungen keine Identität etablieren kann. "The abstract claim is important, then, since it is needed to license the move from detailed empirical correlations to property identifications." (Papineau 2004, 21)

nam 1975c), oder der repräsentationale Zustand „über Arthritis nachdenken“ wird von den Mitgliedern einer Gemeinschaft bestimmt (vgl. Burge 1979; 1982). Nun stellt sich aber die entscheidende Frage, wie Zustände, die von Tatsachen außerhalb des Gehirns abhängen, eine kausale Wirkung hinsichtlich z. B. Körperbewegung haben. Unbestritten scheint die Annahme, dass Körperbewegung von Tatsachen innerhalb des Körpers abhängt (vgl. Papineau 2004, 25f.). Die Diskussion, ob es repräsentationalen weiten Inhalt gibt, und ob er kausalen Einfluss hat, scheint noch immer ergebnisoffen zu sein. Weiter Inhalt ist jedoch ein wichtiges Gegenargument gegen die Typenidentität mentaler und physischer Eigenschaften.[22]

### 2.1.3 Typenidentität in der Kritik

Gegen die Identitätstheorie in ihrer reduktionistischen Tendenz gibt es verschiedene Einwände. Beispielsweise wird darauf hingewiesen, dass sich das Mentale und das Physische zumindest in ihren epistemischen Eigenschaften unterscheiden (vgl. Kap. 3.2). Zudem muss anscheinend das Vorhandensein phänomenaler Eigenschaften anerkannt werden, um empirisch die psychophysische Identität zu ermitteln. Ein anderer, traditioneller Einwand besagt, dass mentale Zustände bzw. Gedanken nicht genau im Raum lokalisierbar sind und deshalb keine Gehirnzustände sein können. Ein Beispiel für diesen Einwand sind so genannte „Phantomschmerzen“, bei denen Schmerzen von amputierten Körperteilen empfunden werden (vgl. Shaffer 1963). Ferner kann kritisiert werden, dass hinsichtlich psychophysischer Kausalität Verhalten vornehmlich als physisches Verhalten verstanden wird. Damit schafft man – unter der Prämisse der Geschlossenheit der Physik – bereits in der Terminologie Voraussetzungen, die nur durch physische Eigenschaften zu erfüllen sind. Die psychologische Komponente von Verhaltensweisen wird dagegen ausgeklammert. Die Hauptargumente gegen die Typenidentität sind das Argument der notwendigen Wahrheit sowie das Argument der Multirealisierbarkeit (vgl. Kap. 4.2.1).

Da psychophysische Identität im Typen-Physikalismus nur aufgrund von Erfahrung bzw. a posteriori gezeigt werden kann, ist diese Form der Identität lediglich kontingent und nicht notwendig wahr. Dies wäre zunächst unproblematisch, da Identität a posteriori grundsätzlich kontingent und wahr sein konnte (vgl. Smart 1959). Kripke zeigt dagegen, dass es nicht nur a priori notwendige Wahrheiten und a posteriori kontingente Wahrheiten gab, sondern auch a posteriori notwendige und a priori kontingente Wahrheiten. Seiner Auffassung nach war psychophysische Identität weder a priori noch a posteriori notwendig und somit auch nicht wahr (vgl. Papineau 2004, 76). Ausdrücke, die in verschiedenen möglichen Welten verschiedene Dinge benennen

[22] In Kapitel 4.3.2. wird das Problem des „weiten Inhalts“ nochmals aufgegriffen und vertieft diskutiert werden, weshalb an dieser Stelle die Diskussion lediglich in aller Kürze skizziert wird.

können, wurden als nicht-starre Bezeichner definiert (vgl. Kripke 1980, Vorlesung 3). Sie werden abgegrenzt von Ausdrücken, wie z. B. Eigennamen, die in allen möglichen Welten dasselbe Objekt bezeichnen. „Georg W. Bush" wäre nach Kripke ein starrer Bezeichner, während „der dreiundvierzigste Präsident der Vereinigten Staaten" ein nicht-starrer Bezeichner ist. In verschiedenen möglichen Welten kann dieser Ausdruck sich auf verschiedene Personen beziehen. Eine *kontingente* Identität setzt also voraus, dass mindestens einer der beiden Ausdrücke ein nicht-starrer Bezeichner ist. Die Identität „Schmerzen = C-Faser-Aktivierung" kann somit nur dann kontingent sein, wenn mindestens einer der Bezeichner „Schmerzen" oder „C-Faser-Aktivierung" kontingent ist. Dies ist jedoch nach Kripke nicht der Fall, weshalb die psychophysische Identität als kontingente Identität in Bedrängnis kommt.[23] Da aber auch nicht gezeigt werden kann, dass beispielsweise Schmerzen und C-Faser-Aktivierungen a priori notwendig sind, wird somit die psychophysische Typenidentität grundsätzlich in Frage gestellt. Zombies sind in diesem Sinne kohärente materielle Doppelgänger ohne Bewusstseinseigenschaften. Dagegen sind Geister die konverse Möglichkeit – als Wesen, die all unsere Bewusstseinseigenschaften teilen, aber nicht die materiellen (vgl. Papineau 2004). Zombies wie Geister sind metaphysisch kohärent. Wenn sie denkbar möglich sind, können phänomenale Eigenschaften nicht identisch mit materiellen sein. Dass in der aktualen Welt einige bewusste und materielle Eigenschaften gleichzeitig auftreten, ist letztlich eine Korrelation, keine Identität.

Das Argument der Multirealisierbarkeit zielt in eine ähnliche Stoßrichtung, allerdings mit einer anderen Pointe. Sie ist letztlich das Hauptargument gegen Identität von mentalen und physischen Eigenschaften bzw. gegen Reduzierbarkeit. Insbesondere dieses Argument der Multirealisierbarkeit zwang die Typenidentitätstheorie zum Rückzug und etablierte die Position des nicht-reduktiven Physikalismus.[24] Diese Position ist seit den siebziger Jahren vor allem durch die Theorie des Funktionalismus einflussreich und weitreichend akzeptiert (vgl. z. B. Fodor 1974; Block/Fodor 1972). Auch dieses Argument zielt auf die empirische Kontingenz sowie die Variabilität der Realisierung in verschiedenen möglichen Welten. Danach sind mentale Eigenschaften und physische Eigenschaften nicht nomologisch koextensiv, weshalb erstere auch nicht z. B. über Brückengesetze auf letztere reduziert werden können. Entscheidend ist dabei einerseits der modale Anspruch, der an die Identitätsbeziehung gestellt wird. „Schmerzen" und „C-Faser-

[23] Üblicherweise werden zur Festlegung eines starren Bezeichners nur Eigenschaften verwendet, die der Gegenstand kontingenterweise besitzt. Wasser ist z. B. durchsichtig, geruchlos, in Flüssen, Seen etc.. Die Eigenschaften von „C-Faser-Aktivierung" oder „Schmerzen" sind jedoch nicht kontingent. „C-Faser-Aktivierung" muss ein starrer Ausdruck sein, da dieses Ereignis per definitionem sonst nicht in der entsprechenden möglichen Welt existiert. Ebenso soll auch der Ausdruck „Schmerzen" als Ereignis für alle möglichen Welten gelten, insofern sich die gleichen Schmerzen in verschiedenen Welten gleich anfühlen sollen (vgl. Kripke 1971, 213f.).

[24] Daneben spielte sicherlich auch Davidsons Theorie des anomalen Monismus eine wichtige Rolle gegen den reduktiven Physikalismus. Schließlich zeigte er, dass mentale Ereignisse sich nicht ohne weiteres auf physische reduzieren lassen, da sie „anomal" sind bzw. keinen strikten Gesetzen unterliegen (vgl. Davidson 1970, 1974; Kap. 2.3; Kap. 3.5).

Aktivierung“ sollen in allen möglichen Welten identisch sein. Dagegen wird aber angenommen, dass die Eigenschaften sowohl in verschiedenen möglichen Welten, als auch für verschiedene Lebewesen multirealisierbar sind. Es wird betont, dass nicht nur Menschen mit C-Fasern Schmerzen haben können, sondern auch (in dieser möglichen Welt) andere Organismen, wie der Octopus, dessen Gehirn sich von dem menschlichen unterscheidet. Biologische Systeme scheinen grundsätzlich zu weitreichend zu sein, so dass sie nicht auf eine einzige physische Basis reduziert werden können. Zudem ist auch vorstellbar, dass auch außerirdische Lebewesen mit einer vollkommen anderen Physiologie oder intelligente elektromechanische Roboter mentale Zustände haben (vgl. Putnam 1967a; Kim 1993a, 313). Andererseits scheinen mentale Funktionen auch innerhalb derselben Person bzw. eines einzelnen Systems sich über einen Zeitraum hin zu verändern, beispielsweise aufgrund von Lernprozessen. All diese Beispiele sollen zeigen, dass mentale Zustände mehrfach, variabel und auf unterschiedliche Weise in unterschiedlichen physischen Systemen realisierbar sind. Deshalb scheint es nach der These der Multirealisierbarkeit nicht möglich, einen mentalen Zustand mit einem physischen Zustand zu identifizieren. Mentale Zustandstypen werden nicht allein durch einen physischen Zustandstyp, sondern durch verschiedene physische Zustandstypen realisiert. Das Argument der Multirealisierbarkeit beruht vor allem auf der Annahme, dass mentale Zustände mit kausalen Rollen identifiziert werden. Mentale Eigenschaften werden insofern von physischen realisiert, als sie eine bestimmte kausale Rolle darstellen (vgl. Putnam 1967a, 1967b).

***Quintessenz Typenidentität***

- *Identität von mentalen und physischen Ereignistypen*
    - *Identität von Ereignissen entspricht Identität ihrer Eigenschaften*
- *empirische Generalisierung: kausale Geschlossenheit der Physik*
- *Bestimmung von Verhalten als physisches Verhalten: Reduktion der psychologischen Ebene*
    - *Gefahr Epiphänomenalismus: mentale Eigenschaften als nomologische Anhängsel physischer Eigenschaften*
- *keine Sicherung der mentalen Kausalität durch Typenidentität*
    - *Wirksamkeit extramentaler Tatsachen problematisch*
- *starke Gegenargumente gegen Typenidentität*
    - *Multirealisierbarkeit von Eigenschaften*
    - *Identität lediglich a posteriori kontingent wahr*
- *keine metaphysische Begründung der Typenidentität über epistemische Korrelation hinaus möglich*

Darüber hinaus hat Fodor in seinem Aufsatz „Special Sciences“ (1974) gezeigt, dass nicht jedem einzelwissenschaftlichen Artbegriff F genau ein physikalischer Artbegriff P so zugeordnet werden kann, dass sich aus jedem einzelwissenschaftlichen Gesetz ein physikalisches Gesetz ableiten lässt. Das ist für die Frage der mentalen Kausalität interessant, da es Realisierungen $F_a$ geben kann, die wider Erwarten keine Realisierung von $F'_a$ verursachen. Das ist der systematische Grund dafür, dass einzelwissenschaftliche Gesetze im Allgemeinen nicht ausnahmslos gelten – im Gegensatz zu physikalischen Gesetzen (vgl. Kap. 3.5; Kap. 4.1). Das ist allerdings auch ein Grund, weshalb die kausale Wirksamkeit mentaler Eigenschaften in Zweifel gezogen wird.

### 2.2 Funktionalismus und mentale Kausalität

Der bereits angesprochene Aufsatz „Psychological Predicates“ von Putnam leitete 1967 eine Wende in der philosophischen Debatte des Geistes ein und begründete die Theorie des Funktionalismus. In der Folge gaben während der 1970er-Jahre viele Philosophen die reduktionistische Position des Typen-Physikalismus auf; die Nichtreduzierbarkeit mentaler Eigenschaften und anderer so genannter höherstufiger Eigenschaften wurde zur akzeptierten Auffassung.[25] Funktionalismus verband insofern die Vorzüge der Identitätstheorie und des Behaviourismus, als mentale Zustände sich durch ihre kausalen Relationen zueinander und zu sensorischen Inputs sowie behaviouralen Outputs konstituieren. Das Gehirn wurde zunächst mit einer Rechenmaschine verglichen, wobei mentale Zustände als relationale Zustände mit einem wesentlichen Bezug zum Verhalten gekennzeichnet wurden.[26] Mentale Zustände lassen sich gewissermaßen über äußere Faktoren, innere Zustände und Verhalten definieren. Inzwischen gibt es eine unüberschaubare Vielzahl an Funktionalismusversionen.[27] In den Hauptströmungen kann zwischen „role-functionalism“ und „realizer-functionalism“ unterschieden werden, davon abhängig, ob funktionale Zustände über die Rolle, die sie spielen, definiert werden, oder über die Realisierer der Rolle im Einzelfall. Rollen-Funktionalismus vertritt das Prinzip der Multirealisierbarkeit, während Rea-

[25] Vgl. hierzu beispielsweise Block/Fodor 1980, die eine Reihe ihrer Meinung nach entscheidender auch empirischer Argumente anführen, wie z. B. die Lashleyans Lehre der neurologischen Equipotenzialität oder Darwins Lehre der Konvergenz. So heißt es hinsichtlich ähnlicher Verhaltensweisen unterschiedlicher Arten: „[O]rganisms of widely differing phylogeny and morphology may nevertheless come to exhibit superficial behavioral similarities in response to convergent environmental pressures. […] Psychological similarities across species may often reflect convergent environmental selection rather than underlying physiological similarities.“ (ebd., 238)

[26] Auch der ontologische Behaviourismus beschrieb mentale Zustände in gewisser Weise als relational, da die so genannten Verhaltensdispositionen die Beziehung zwischen Reizen und Verhalten charakterisieren. Allerdings betrachtete diese Position die mentalen Zustände eben lediglich als Verhaltensdispositionen und nicht als Ursachen von Verhalten (vgl. Schröder 2004, 86).

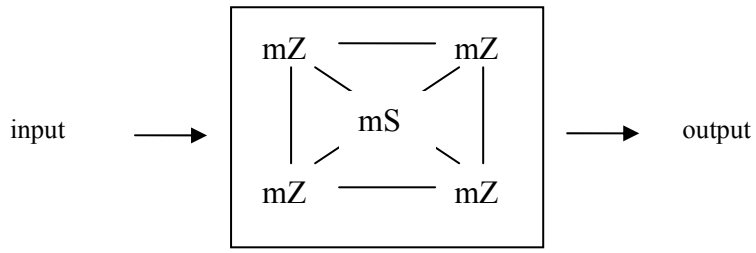
[27] Braddon-Mitchell und Jackson geben hierzu eine gute Übersicht, in der sie sieben Hauptarten unterscheiden (vgl. Braddon-Mitchell/Jackson 1996, 88f.). Die vorliegende Arbeit wird sich auf die Grundzüge der Hauptlinien des Funktionalimus beschränken.

lisierer-Funktionalismus wie die Typenidentitätstheorie zulässt, dass verschiedene Organismen verschiedene Schmerzzustände haben, möglicherweise sogar verschiedene Schmerztypen, weil die Realisierer der verschiedenen Organismen eben verschieden sind. Andererseits wird aus der Perspektive des Realisierer-Funktionalismus der Rollen-Funktionalist wichtigen Unterschieden im Detail nicht gerecht, während aus der Perspektive des Rollen-Funktionalismus der Realisierer-Funktionalist gerade aufgrund dieser Aufmerksamkeit gegenüber den Realisierern nützliche und sinnvolle Generalisierungen nicht zulässt.

### 2.2.1 Die funktionale Rolle als eine kausale Rolle

Zentral für den Funktionalismus ist die These der Multirealisierbarkeit mentaler Eigenschaften, die bereits im Zusammenhang mit der Kritik der Identitätstheorie beschrieben wurde (vgl. Kap. 2.1.3). Die These der Multirealisierbarkeit bildet gewissermaßen die Grundlage für eine neue Konzeption des Mentalen (vgl. z. B. Block 1980b, 173). Mentale Zustände können – wie alle funktionalen Zustände – auf die unterschiedlichste Weise realisiert werden. Es ist nicht möglich, einen einzigen physischen Zustand zu bestimmen, der die funktionale Rolle für alle tatsächlichen und möglichen Fälle von Schmerzen spielt. Funktionalisten unterscheiden deshalb häufig mentale Eigenschaften als „höherstufige" Eigenschaften (englisch: „higher-level properties") von ihren physischen Realisierern, den „niedrigstufigen" Eigenschaften (englisch: „lower-level properties"). Objekte oder Subjekte haben erstere aufgrund der Tatsache, dass sie die geeigneten letzteren Eigenschaften besitzen (vgl. Kap. 3.1).

Es wurde früh die implizite Voraussetzung der Typenidentität diskutiert (vgl. Putnam 1975b; Fodor 1968), wonach mentale Typen in allen möglichen Welten die ihnen zugerechneten neurologischen Typen haben müssen. Aus der Identität von Schmerzen und C-Faser-Aktivierung folgt beispielsweise, dass jedes Lebewesen nur dann Schmerzen haben kann, wenn es C-Fasern bzw. eine C-Faser-Aktivierung hat. Dieser Zusammenhang ist jedoch biologisch schwer vertretbar. Es wurde bereits darauf hingewiesen, dass nicht für jeden Organismus, der Schmerz empfindet, die gleiche bio-chemische Struktur angenommen werden kann. Letztlich scheint der identitätstheoretische Ansatz zu stark auf die menschliche Biologie fokussiert und verfällt so in einen artenspezifischen Chauvinismus. Um diese Gefahr zu vermeiden, verweist der Funktionalismus darauf, dass es nicht die C-Faser-Aktivierung selbst, sondern ihre funktionale bzw. kausale Rolle ist, die mentale Zustände bestimmt. Mentale Zustände werden als Schaltstellen in einem kausalen Netzwerk aufgefasst, das durch den Empfang von sensorischen Inputs und dem Senden von behaviouralen Outputs zu kausaler Transaktion mit der äußeren Welt fähig ist.




Grafik 4: Funktionalismus und mentale Kausalität. Die mentale Eigenschaft wird nicht nur durch die input-/output-Relation, sondern auch durch die Relation zu weiteren inneren Eigenschaften bestimmt. Anders formuliert: Jeder einzelne mentale Zustand (mZ) ist von seiner Rolle innerhalb des Netzwerkes (internes Kriterium) und zusätzlich von seiner Verhaltensrolle (externes Kriterium) determiniert (vgl. Block 1980b).

Danach werden beispielsweise Schmerzen durch eine Verletzung der Haut verursacht, woraufhin sie wiederum ein Jammern verursachen können, sowie die Versorgung der Schmerzstelle und eine Fokussierung der Aufmerksamkeit mit dem Wunsch, den Schmerz zu beseitigen. Funktionalistisch betrachtet ist Schmerz der Zustand eines Wesens, der durch diese kausale Rolle charakterisiert ist, bzw. ein Wesen hat genau dann Schmerzen, wenn es in einem Zustand ist, der diese kausale Rolle innehat:

> Mental states are, according to functionalists, internal states within us, but we identify and name them by the effect the world has on them, the effect they have on one another, and the effect they have on the world by causing our behaviour. (Braddon-Mitchell/Jackson 1996, 41)

Die Grundthese des Funktionalismus besagt also, dass mentale Zustände ihrer Natur nach funktionale Zustände bzw. mentale Eigenschaften demnach funktionale Eigenschaften sind (vgl. Putnam 1967b; Fodor 1968, 1975; Block 1980b; Heil 1998, Kap. 4; Kap. 3.2.1). Funktionale Eigenschaften sind die Eigenschaften eines Systems, die allein durch ihre kausale Rolle charakterisiert werden. Die kausale Rolle ergibt sich aus:

1. den (äußeren )Inputs, die die Eigenschaften verursachen,
2. den Outputs, die die Eigenschaften außerhalb des Systems selbst verursachen,
3. ihrer kausalen Relation zu anderen Systemzuständen desselben Typs.

Schmerzen zu haben, bedeutet damit, sich in einem bestimmten Zustand zu befinden, der eine funktionale kausale Rolle erfüllt, wie die C-Faser-Aktivierung bei Menschen. Wir können zwar annehmen, dass bestimmte Schmerzen zu einem bestimmten Zeitpunkt bei einem bestimmten Subjekt identisch mit den bestimmten neurophysiologischen Zuständen dieses Subjekts sind, aber Schmerzen selbst als Typ können nur abstrakt über ihre funktionale bzw. kausale Rolle bestimmt werden. Zugleich scheint die Idee des Funktionalismus ihrem Kern nach auch ihre Schwäche zu sein. Denn der Schmerz wird als abstrakte kausale Eigenschaft aufgefasst, die nur aufgrund ihrer

direkten und indirekten Relationen sowie Input- und Outputrelationen mit der Welt verbunden ist. Damit ist zugleich vorstellbar, dass etwas zwar die kausale Eigenschaft besitzt, aber trotzdem keine Schmerzen hat (Block 1980b, 175; Shoemaker 1980a).

### 2.2.2 Die Realisierung funktionaler Zustände

Putnam verglich mentale Zustände mit den logischen bzw. funktionalen Zuständen von Computern – das psychologische Programm wurde von verschiedenen Organismen durch verschiedene bio-chemische Zusammensetzungen realisiert, und deshalb können verschiedene physiologische Zustände unterschiedlicher Organismen denselben mentalen Zustandstypen realisieren. Nach Putnams so genannten „Maschinenfunktionalismus" bedeutet es für ein System S, in einem Zustand M zu sein, in einem physiologischen Zustand o. ä. zu sein, der die Rolle R in dem relevanten Computerprogramm spielt. Dieser Zustand spielt insofern die Rolle R, als er die Relation physischer Inputs, Outputs und anderer innerer Zustände umfasst, die zu der abstrakten input-output-Relation mit logischen Zuständen eines Computerprogramms passt.[28]

Der Maschinen- bzw. Computerfunktionalismus vertritt die These, dass das Mentale tatsächlich wie ein Computer funktioniert und als solches charakterisierbar ist. Allgemein ist die Position des Funktionalismus jedoch nicht auf das Computermodell festgelegt, sondern lediglich auf die Auffassung, dass mentale Zustände bzw. Eigenschaften funktionale Zustände bzw. Eigenschaften sind und dass es daher für mentale Phänomene eine eigene funktionale Beschreibungsebene gibt, die von der Ebene der physischen Realisierung eingeschränkt unabhängig ist. Daraus lässt sich ableiten, dass die Einzelwissenschaften, die sich mit der Realisierung mentaler Phänomene beschäftigen, als eigenständige Wissenschaften weitgehend unabhängig sind (vgl. Fodor 1974).

(Maschinen-)Funktionalismus bietet drei Beschreibungsebenen, eine neurophysiologische Beschreibung auf der neurologischen Ebene, eine funktionale Beschreibung auf der Programmierungsebene sowie eine mentale Beschreibung im Sinne einer alltäglichen Zuschreibung mentaler Zustände. Es ist also nicht nur wichtig, funktionale Zustände zu erklären. Es gilt zu beschreiben, was es heißt, dass funktionale Zustände durch physische Zustände realisiert sind. Anders ausgedrückt: Es muss eine Beschreibung möglich sein, die ohne psychologisches oder mentales Voka-

[28] Die so genannte Turing-Maschine (Turing 1936) hat den Begriff der algorithmischen Zeichenkette bzw. der berechenbaren Funktion zunächst auf der Basis einer Operation mit Papier und Bleistift präzisiert. Turing konnte zeigen, dass Operationen, die schrittweise nach formalen Regeln von Zahlenzeichen in andere Zahlenzeichen überführen, auch von Maschinen mit minimalen Grundoperationen ausgeführt werden können. So kann ein Zusammenhang zwischen Programmen und funktionalen Zuständen gezeigt werden: Zu jedem Programm P gibt es eine Menge von durch entsprechende Verhaltensgesetze charakterisierten funktionalen Zuständen, für die gilt, dass jede Maschine, die diese funktionalen Zustände annehmen kann, genau das Programm P bearbeitet. Somit sind auch Computer funktional definierte Systeme und daher multirealisierbar.

bular auskommt. Schließlich soll ein Erklärungsregress – mentale Zustände mit mentalen Zuständen zu erklären – vermieden werden. Da jedoch mentale Zustände per definitionem durch Input- bzw. Outputrelationen und den Bezug zu anderen mentalen Zuständen sozusagen in einem holistischen System gekennzeichnet sind, kann eine Erklärung zunächst nicht auf mentale Zustände verzichten. Mit der Methode der Ramseyfizierung hat Lewis (1972) eine formal korrekte und explizite Definition funktionaler Zustände ermöglicht.[29] Es werden mentale Ausdrücke mit Variablen versehen und an Existenzquantoren gebunden. Die impliziten Verhaltensgesetze werden in einem einzigen Satz zusammengefasst, indem sie jeweils mit einem „und" verbunden werden. Beispielsweise wird aus der psychologischen Beschreibung (vgl. auch Block 1980b, 173f.)[30]

> *(T)* Für jedes *x* gilt: Wenn *x einen Gewebeschaden* erleidet und **bei normalem Bewusstsein** ist, hat *x* Schmerzen: wenn *x wach* ist, ist er wahrscheinlich **bei normalem Bewusstsein**; wenn *x* Schmerzen hat, *windet sich x* und *stöhnt* und **wechselt in einen Zustand des Sich-nicht-Wohlfühlens**; und wenn *x* **nicht bei normalem Bewusstsein ist oder sich nicht wohl fühlt**, *macht x in der Regel mehr Tippfehler.*

über die Ramseyfizierung

> *($T_R$)* Es gibt die Zustände $M_1$, $M_2$ und $M_3$, sodass für jedes *x* gilt: Wenn *x einen Gewebeschaden erleidet* und im Zustand $M_1$ ist, hat *x* Schmerzen; wenn *x wach* ist, ist er wahrscheinlich in $M_2$; wenn *x* Schmerzen hat, *windet sich x* und *stöhnt* und wechselt in $M_3$; und wenn *x* nicht in $M_2$ oder in $M_3$ ist, *macht x in der Regel mehr Tippfehler.*

der abgekürzt repräsentierte Satz

$$\text{x hat Schmerzen} := \quad \sum M_1, M_2, M_3, [\text{T}(M_1, M_2, M_3) \,\&\, x \text{ ist in } M_2],$$

wobei die Variable $M_2$ den Ausdruck „hat Schmerzen" in der Theorie T ersetzt. Diese Definition liefert einen Begriff von Schmerzen durch Rekurs auf deren kausale bzw. nomologische Relation, wobei sich unter deren Ursachen und Wirkungen weitere „mentale Zustände" befinden. Die Gefahr der Zirkularität wird umgangen, es sind nur „gewisse Zustände" möglich, und die drei mentalen Begriffe $M_1, M_2, M_3$ bedingen sich wechselseitig.

---

[29] Ausgehend von einer geeigneten Theorie lassen sich mentale Zustände also mithilfe dieses Verfahrens explizit definieren (vgl. Lewis 1972). Damit wird die Anforderung des Behaviourismus erfüllt, die Bedeutung mentaler Ausdrücke ohne jeden Rückgriff auf mentales Vokabular zu spezifizieren.

[30] Vgl. hierfür Kim 1998, 117. Die Thesen, die T ausmachen sind gesetzesartige Regularitäten/Kausalrelationen. Kursiv steht für nicht-mentale Prädikate (physisch, biologisch, behavioural), fett steht für psychologische Ausdrücke. Um unsere Muster-Definition handhabbar zu machen, kürzen wir TR wie folgt ab: ∑ M1, M2, M3 [T(M1, M2, M3)]. Dabei steht das Symbol ∑ für den Existenzquantor, der als „Es gibt" zu lesen ist.

### 2.2.3 Die funktionale Definition für kausale Rollen

Mit Hhilfe des Lewisschen Verfahrens lässt sich der Begriff der Realisierung zwar präziser fassen. Um die funktionalen Zustände eines Systems definieren zu können, benötigt man jedoch zusätzlich eine Theorie, in der die kausalen Rollen dieser Zustände exakt formuliert sind. Um die kausalen Rollen der einzelnen mentalen Zustände im Detail zu beschreiben, wird zwischen dem so genannten Common-Sense-Funktionalismus, der von der Alltagspsychologie ausgeht, und dem so genannten Psychofunktionalismus unterschieden, der von der wissenschaftlichen Psychologie ausgeht. Beide sollen eine realistischere und umfassendere funktionale Definition psychologischer Begriffe ermöglichen, die alle psychologischen Eigenschaften einschließt, hinreichend Information über sie hat (Input-Bedingungen, Verhaltens-Outputs, nomologische Verknüpfung der psychologischen Eigenschaften) und welche zusätzlich die psychologischen Eigenschaften eng genug umschreiben, damit diese identifiziert werden können.

> Common-sense functionalism, as we saw, characterizes these clauses in the terms familiar to common-sense: the inputs are described as tigers and cups of coffee, the outputs in terms of the way movements affect how we relate to tigers and cups of coffee. Empirical functionalism, by contrast, typically fix on the inputs to and the outputs from the central nervous system. (Braddon-Mitchell/Jackson 1996, 82)

Der Common-Sense-Funktionalismus fordert, dass die kausalen Relationen zwischen den mentalen Zuständen durch Begriffe der Alltagspsychologie beschrieben werden. Damit sollen nur Relationen und Begriffe, die jeder kennt, wie z. B. „Schmerzen durch spitze Gegenstände“ oder „Handlungsvollzug bei Gelegenheit und entsprechender Handlungsabsicht“, zugelassen werden. Allerdings ist die Alltagspsychologie als solche nicht ausbuchstabiert – wir haben lediglich implizite Kenntnis über sie, insofern sie zu unserem gemeinsamen Wissen gehört. Eine funktionale Definition müsste jene psychologischen Begriffe ergeben, die eine Sprachgemeinschaft teilt. Zwar sind die Alltagsbegriffe des Mentalen kollektiv definiert, aber somit unvollständig und fehlerhaft.

Die wissenschaftliche Psychologie dagegen scheint in diesem Zusammenhang die beste Theorie derjenigen kausalen bzw. nomologischen Relationen zu sein, in denen mentale Zustände eine Rolle spielen. Eine solche Beschreibung mentaler Zustände in den kausalen bzw. nomologischen Relationen selbst sowie mit physischen und behaviouralen Prozessen wird schließlich gefordert. Mögliche empirisch beschriebene Regelmäßigkeiten, die mentale Zustände charakterisieren, sind z. B., dass Frustration zu Aggression oder die subjektiv erfahrene Einschränkung des Handlungsspielraums zu unkooperativem Verhalten führt (vgl. Dollard et al. 1939; Brehm 1966; beide zitiert nach Schröder 2004). Allerdings sind die Grundlagen der wissenschaftlichen Psychologie teilweise umstritten, und es gibt keinen ähnlich robusten Kern an Gesetzen wie in der Phy-

sik. Allerdings ergibt sich ein gewichtiges Problem: Wenn die zugrunde liegende Theorie falsch ist, wird sich wahrscheinlich herausstellen, dass alle auf ihrer Basis und mittels der Ramsey-Lewis-Methode definierten Begriffe die gleiche leere Extension haben. Beide Theorie-Möglichkeiten können keine eindeutigen und letztgültigen Definitionen liefern. Es gibt zahlreiche Ausnahmen der Regularitäten, Begriffe mit unscharfen Rändern und unechten Wahrheiten (vgl. Kap. 3.2.1; Kap. 4.1).[31]

(Common-Sense-)Funktionalismus hängt also stark von einer guten Referenztheorie, wie z. B. teleologischen Theorien ab (vgl. Kap. 4.5.1). Es geht in beiden Fällen um die Beziehung zwischen Welt und Wörtern. Die kausale Referenztheorie[32] besagt, dass Begriffe kausal auf tatsächliche Dinge referieren. Beispielsweise werden alle Wasser-Stereotypen aus demselben fundamentalen Stoff – $H_2O$ – gemacht, weshalb wir Menschen sie aus dem direkten *kausalen* Kontakt heraus alle „Wasser" nennen. Diese Annahme geht dem Wissen, dass sie aus $H_2O$ bestehen, voraus. Wir benutzen das Wort Wasser, wenn wir kausal mit der natürlichen Art (natural kind) konfrontiert werden, z. B. um zu trinken oder zu schwimmen in einem „Wasser-Stereotypen" (vgl. Braddon-Mitchell/Jackson 1996, 68):

> When we called the exemplars of water ‚water', we were responding to a huge range of factors that came together to cause us to use the word ‚water', including: a certain pattern of retinal stimulation, a kind of wave packet reflecting from the exemplars, the weather conditions that caused the water to be in front of us to begin with, the position of our heads, the factors that led us to use ‚water' and not some other word, and the ambient lighting. (ebd.)

Auf ähnliche Weise soll die Referenz mentaler Begriffe bzw. ihrer physischen Repräsentationen durch Kausalbeziehungen zwischen mentalen Zuständen und Begriffen vermittelt werden (vgl. Kap. 4.3; Kap. 4.5). Eine kausale Referenztheorie ist für den Funktionalismus u. a. aufgrund der Möglichkeit wichtig, dass ein Begriff durch Merkmale bestimmt wird und nicht ein Begriff ist über etwas, das diese Merkmale hat. „In particular, it serves up the possibility that we should

[31] In dem Aufsatz „Wants as Explanations of Actions" (1963) geben Richard Brandt und Jaegwon Kim ein Beispiel für sechs Gesetze der Alltagspsychologie. Diese lauten u. a.: a.) Wenn x Freude empfindet, falls sie p nicht erwartet hat, aber nun plötzlich zu der Überzeugung gelangt, dass p der Fall sein werde, dann wünscht x p. oder: b.) Wenn x p wünscht, dann gilt unter günstigen Bedingungen: Falls x annimmt, dass der Vollzug von H vermutlich zu p führen wird und dass die Unterlassung von H vermutlich zu nicht-p führen wird, so wird x eine Regung verspüren, H zu tun. – Solche Gesetze sind allerdings höchstens ein Teil der Theorie, durch die für den Funktionalisten mentale Zustände charakterisiert sind. Denn sie enthalten nur Gesetze, die kausale Relationen zwischen mentalen Zuständen wiedergeben. Gesetze, in denen mentale Zustände mit möglichen Inputs und Outputs in kausale Beziehungen gesetzt werden, fehlen jedoch.

[32] Eine Alternative zur kausalen Referenztheorie ist die Beschreibungstheorie (description theory). Ihr zufolge ergibt sich die Referenz eines theoretischen Begriffs aus seinen charakteristischen Eigenschaften. So referiert z. B. der Begriff „Wasser" auf eine farblose Flüssigkeit, die vom Himmel fällt oder in Seen, Flüssen zu finden ist (vgl. Braddon-Mitchell/Jackson 1996, 61f.; Kripke 1980).

think of the common-sense roles of common-sense functionalism as fixing the reference rather than giving the meaning of the mental state terms." (ebd., 73)

### 2.2.4 Funktionalismus und Physikalismus

In einem gewissen Sinne ist das Mentale vom Physischen zu unterscheiden, da beide Bereiche eine andere Charakterisierung haben. Die Eigenschaft, Schmerz zu haben, ist nicht mit der Eigenschaft C-Faser-Aktivierung identisch (vgl. Kap. 3.2). Der Funktionalismus behauptet zunächst nicht, dass Bewusstsein physisch ist, vielmehr erlaubt er sogar theoretisch, dass das Mentale nicht-physisch ist, solange das Bewusstsein durch relevante Programme realisiert wird (vgl. Putnam1975b; Fodor 1974):

> The functionalist answer to "What are mental states?" is simply that mental states are functional states. […] Most functionalists are willing to allow that each particular pain is a physical state or event, and indeed that for each type of pain-feeling organism, there is (perhaps) a single type of physical state that realizes pain in that type of organism. Where functionalists differ with physicalists, however, is with respect to the question of what is common to all pains in virtue of which they are pains. The functionalist says the something in common is functional, while the physicalist says it is physical […]. (Block, 1980b, 172)

Das Verhältnis des Funktionalismus zum Physikalismus ist unklar, wie auch, ob der Funktionalismus überhaupt Aussagen über den ontologischen Status mentaler Eigenschaften macht (vgl. Kap. 3.1.2). Die Auskunft „Mentale Zustände sind ihrer Natur nach funktionale Zustände" ist ontologisch gesehen zunächst neutral. Allerdings wird deutlich gemacht, dass die Hypothese funktionaler Zustände mit dem Dualismus nicht inkompatibel ist (vgl. Putnam 1967b, 130). Beispielsweise wird der mentale Zustand Zorn durch seine kausale Rolle definiert: Eine Person ist genau dann zornig, wenn sie sich in einem Zustand befindet, der durch bestimmte äußere Ereignisse verursacht wird, der seinerseits bestimmte physische Reaktionen und ein bestimmtes Verhalten verursacht, und der in bestimmten kausalen Relationen zu anderen Zuständen dieser Person steht. Über die Art des Zustands ist damit nichts gesagt. Es kann sich um einen Gehirnzustand handeln, aber auch um einen nicht-physischen Zustand dieser Person oder vielleicht sogar um einen Zustand einer immateriellen Seele. Zu einer Version des Physikalismus wird der Funktionalismus erst durch die zusätzliche These, dass alle mentalen Zustände durch physische Zustände realisiert sind (vgl. Block/Fodor 1980).[33] Der Funktionalismus fordert zunächst ledig-

[33] Lewis dagegen argumentiert von einem Standpunkt des Common-Sense-Funktionalismus aus: Danach zeigt der Funktionalismus, dass Physikalismus wahr ist (vgl. Lewis 1980; Rosenthal 1991). Auch Kim argumentiert, dass die Konsequenz der Multirealisierbarkeit mentaler Zustände den Physikalismus stützt und sogar in einen

lich, dass es innere Zustände von S gibt, die untereinander sowie mit den Inputs und Outputs wie durch die Psychologie von S spezifiziert zusammenhängen, ohne dass über die Natur dieser inneren Zustände irgendetwas gesagt wird. Um eine Auswahl zu treffen, bietet sich eine Ergänzung durch physikalische Bedingung an. Ontologisch werden dann die von einer korrekten psychologischen Theorie postulierten inneren Zustände als physische Zustände aufgefasst. Eine psychologische Theorie scheint nur geeignet, wenn Zustände als innere physische Zustände des jeweiligen Organismus betrachtet werden: „Die von einer wahren psychologischen Theorie postulierten psychologischen Fähigkeiten und Mechanismen müssen real sein, und die einzige Realität, an die man appellieren kann, ist die physikalische Realität." (vgl. Kim 1998, 125) Aus dieser Argumentation heraus impliziert der Funktionalismus den Physikalismus, da eine kausale Rolle nur durch etwas ausgefüllt werden kann, das selbst physisch ist, weil es in einer Relation zu anderen physischen Zuständen bzw. deren Eigenschaften steht (vgl. Salmon 1984; Kistler 1999, beide zitiert nach Schröder 2004, 90).

Trotzdem gibt es nach der funktionalistischen Position z. B. intentionale Zustände, da innere physische Zustände bzw. Eigenschaften, die propositionale Eigenschaften realisieren, tatsächliche und mögliche Zustände repräsentieren (vgl. Kap. 4.3; Kap. 4.5). Und was sie repräsentieren wird zumindest teilweise durch ihre funktionale Rolle bestimmt. Allerdings ist es nicht möglich, genau zu bestimmen, wie der repräsentationale Inhalt physischer Dinge bestimmt werden kann bzw. wie ein neurologischer Zustand propositionalen Inhalt repräsentieren kann (vgl. Fodor 1987). In der Folge wird deshalb innerhalb der Diskussion zwischen „weitem" und „engem Inhalt" bzw. „weiten" und „engen Eigenschaften" unterschieden. Diese sind durch die intrinsischen physischen Zustände des Subjekts, jene Eigenschaften durch extrinsische Eigenschaften bestimmt. Insofern ist repräsentationaler Inhalt „weit". Wenn funktionale Rollen vorgeblich „eng" sind, scheiden propositionale Eigenschaften als Kandidaten für funktionale Rollen aus (vgl. Burge 1986, 1989; Loar 1991; Stalnaker 1991; Goldman 1993; Kap. 4.3.2).

Das Problem der mentalen Kausalität wird nun einerseits unter der Voraussetzung eines Funktionalismus mit weiten oder engen mentalen Eigenschaften diskutiert und andererseits unter der Voraussetzung eines begründbaren starken Externalismus über intentionalen Inhalt (vgl. Jackson 1998, 96).[34] Häufig wird gerade den Eigenschaften, *wie* Dinge sind, kausale Verantwortung zu geschrieben. Das Problem ist nun, welche Arten von Eigenschaften Ursachen sein können – was macht es wahr, dass ein bestimmter Gehirnzustand ein mentaler Zustand ist. Aufgrund des Arguments der Multirealisierbarkeit ist ein Gehirnzustand nicht ein mentaler Zustand, weil er ein neuronaler Typ ist, sondern aufgrund seiner funktionalen Rolle. Es gibt aber gute

---

Reduktionismus mündet (vgl. Kim 1992; 1999b). Vgl. auch Braddon-Mitchell/Jackson (1996), die vertreten, dass Funktionalismus die Identitätstheorie auf sicheren Grund stellt (vgl. ebd., 91f.).

[34] Eine erste Skepsis, dass Zustände allein aufgrund ihrer funktionalen Rollen keine kausale Kraft haben, findet sich bei Fodor (1987).

Gründe, zu bezweifeln, dass funktionale Eigenschaften echte Ursachen sein können: Angenommen, es gibt ein zerbrechliches Glas mit der bestimmten Art „Bindung B" zwischen den Glasmolekülen, die verantwortlich für das Brechen des fallenden Glases sind. Jackson meint nun, dass es „Bindung B" zusammen mit dem Fallen ist, was das Zerbrechen verursacht, für die Dispositionen bleibt nichts übrig.

> Either way, it is the categorical bases of functional properties, and not functional properties themselves, that cause: the transitions between mental states, the transitions from inputs to mental states, and the transitions from mental states to outputs will be governed by the neurological or syntactical nature of the mental states […] and not their functional properties. (Jackson 1998, 97)

In der Konsequenz sind funktionale bzw. mentale Eigenschaften epiphänomenal (vgl. Kap. 3.6.1). Es scheint, dass die kausale Ursächlichkeit allein auf die „niedrigstufigen" Eigenschaften zurückgeht, während die „höherstufigen" funktionalen Eigenschaften kausal ausgeschlossen sind. Die enge Anbindung an den Physikalismus lässt mentale Eigenschaften, wie auch bei der Typenidentität, eigentlich wirkungslos werden. Das hat aber auch weitgehende Folgen für biologische, neuronale oder chemische Eigenschaften als höherstufige Eigenschaften (vgl. Kap. 3.1.3).

### 2.2.5 Funktionalismus in der Kritik

Obwohl der Funktionalismus vor allem die Vorteile des Behaviourismus und der Identitätstheorie bietet, besitzt er zugleich einige Schwächen beider Positionen. Diese lassen sich einerseits in Probleme mit phänomenalen Zuständen bzw. Qualia[35] und andererseits in Probleme mit Intentionalität unterteilen.

Es ist eine verbreitete Strategie, auf Fälle zu verweisen, die funktional äquivalente Zustände zu den menschlichen haben, denen wir intuitiv jedoch keine mentalen Zustände zuschreiben würden. Das Argument der invertierten bzw. vertauschten Qualia soll zeigen, dass mentale Zustände nicht einfach funktionale Zustände sein können, da sie Aspekte haben, die sich nicht auf ihre kausale Rolle zurückführen lassen. Für die kausale Rolle einer Empfindung ist es egal, wie es sich anfühlt, diese Empfindung zu haben oder ob sie mit einem qualitativen Eindruck verbunden ist. In dem Gedankenexperiment der vertauschten Qualia werden jeweils rote und grüne Farbeindrücke und Dinge miteinander vertauscht. Die Grün-Empfindungen einer derart gestalteten Person A haben aber dieselbe kausale Rolle wie die Rot-Empfindungen der übrigen Menschen –

[35] Als Qualia werden gemeinhin die qualitativen Empfindungen eines mentalen Zustands bezeichnet, also der introspektive phänomenale Charakter, wie es sich „anfühlt" in einem bestimmten mentalen Zustand zu sein. Vgl. hierzu Nagels berühmten Aufsatz „What is it like to be a bat?" (1974).

damit gehören sie funktionalistisch zum selben Typ mentaler Zustände. Nach der Definition des Funktionalismus wären diese Grün-Qualia sogar streng genommen Rot-Qualia, für ihn ist es belanglos, ob ein mentaler Zustand überhaupt mit einem Eindruck verbunden ist. Diese Konsequenz ist aber absurd und entspricht nicht der Natur mentaler Zustände; somit wird gefolgert, dass kausale Rollen nicht das entscheidende Merkmal mentaler Zustände sein können. Gegen dieses Argument ließe sich einwenden, dass es nicht intuitiv selbstverständlich ist, zwei mentalen Zuständen dieselbe kausale Rolle zuzuschreiben, obwohl sie mit verschiedenen Qualia verbunden sind. Empfindungen, die mit unterschiedlichen Eindrücken verbunden sind, besitzen in der Regel auch verschiedene kausale Rollen.[36]

In seinem berühmten China-Gedankenexperiment (1980c) argumentiert Block gegen den Maschinenfunktionalismus, es sei unplausibel, dass jedes denkbare System mentale Zustände hat. Das Gedankenexperiment sieht, kurz gesagt, eine Simulation vor, in der eine Milliarde Chinesen jeweils ein spezielles Funkgerät mit einem input-Lämpchen und einem output-Knopf erhalten, welches in Kontakt zu einem Androiden und einer Satellitenanzeigetafel steht. Erhält der Androide einen Sinnesreiz, gibt das Funkgerät einen Signalton von sich. Auf der Anzeigetafel werden durch Zahlenreihen die funktionalen Zustände des Systems beschrieben. Immer wenn der ein Signal erhält, soll der jeweilige Chinese mit dem Funkgerät die Satellitenanzeige und somit auch das motorische System des Androiden beeinflussen. Das Volk der Chinesen vermittelt dem Androiden die Sinnesreize und die Bewegung, so dass ein Außenstehender den Eindruck eines „normalen Verhaltens" des Androiden hätte. Nach Block sind die Rollenspieler der funktionalen Zustände die Anzeigen auf dem Satelliten; sie haben die kausale Rolle. Er würde jedoch intuitiv aufgrund der Systemkonstruktion ablehnen, dass die Abfolge von Symbolen auf einer Satellitenanzeige einer Abfolge bewusster Gedanken oder Empfindungen entspricht (vgl. ebd., 277-281). Allerdings lässt sich der intuitive Einwand insofern entkräften, als in einem weiteren, abstrakten Sinn die Satellitenanzeige und alle weiteren gegebenen Bedingungen zusammen den funktionalen Rollenspieler bilden. Der Zustand eines Systems wird erst dann zu einem funktionalen Zustand, wenn er in kausaler Relation zu anderen Zuständen steht. Das China-Gedankenexperiment würde demnach einen ausgelagerten Bewusstseinszustand abbilden.

[36] Mit der Annahme kausal gleicher mentaler Zustände bei unterschiedlichen Qualia sind zudem epistemische Probleme verbunden. Wie soll eine Person erkannt werden bzw. sich selbst erkennen als Person mit Grün- ohne Rotempfindung, wenn sie sich wie eine rot empfindende Person verhält? Aufgrund solcher Überlegungen haben die Churchlands sowie David Lewis Qualia nicht mit kausalen Rollen, sondern mit physischen Merkmalen der jeweiligen Rollenträger identifiziert (vgl. Churchland/Churchland 1981; Lewis 1980). Auch die Möglichkeit philosophischer Zombies (die über genau dieselben funktionalen Zustände verfügen wie wir, aber überhaupt nichts fühlen, weil keiner ihrer Zustände einen qualitativen Charakter aufweist) erscheint unplausibel. Epistemisch betrachtet, gibt es keine Möglichkeit, zwischen zwei Wesen hinsichtlich des qualitativen Charakters ihrer mentalen Zustände zu unterscheiden, wenn sie nicht nur in ihrem Verhalten, sondern auch in allen Überzeugungen, ihrer Introspektion übereinstimmen.

Auf Probleme der funktionalistischen Position mit Intentionalität machte besonders das so genannte Chinesisches-Zimmer-Argument von Searle (1980) aufmerksam und richtet sich im Kern ebenfalls gegen den Maschinenfunktionalismus. In einem Gedankenexperiment ist eine Person in ein Zimmer gesperrt. Sie spricht und liest kein Chinesisch, ist aber in der Lage, allein mit einer Menge von Transformationsregeln Symbolketten – chinesische Ausdrücke – in andere Symbolketten umzuwandeln. Die Person wird mit der Zeit im Hantieren mit diesen chinesischen Ausdrücken in Übereinstimmung mit den ihr gegebenen Regeln sehr geschickt und sendet jedes Mal, wenn eine Kette von chinesischen Schriftzeichen von außen in den Raum hinein gegeben wird, eine entsprechende Kette von chinesischen Zeichen heraus. Aus der Perspektive außerhalb des Zimmers ist aufgrund der Input-Output-Relationen nicht zu überprüfen, ob die Person in dem Zimmer die chinesische Sprache beherrscht oder nicht, während die Person tatsächlich nicht des Chinesischen mächtig ist. Was im Zimmer vor sich geht, ist nichts weiter als ein Hantieren mit Symbolen auf der Basis ihrer Formen oder ihrer „Syntax": Echtes Verstehen beinhaltet aber nach Searle semantisches Verständnis – die Kenntnis dessen, was diese Symbole repräsentieren. Und echtes Denken umfasst für Searle bewusste intentionale Zustände, also Zustände, die sich auf etwas beziehen und die einen repräsentationalen Inhalt haben (vgl. Kap. 4.5.2).

Einen Geist haben, ist für Searle insofern mehr, als sich in syntaktischen, binären Prozessen eines Computers zu befinden. Mentale Prozesse werden von repräsentationalen Inhalten oder von Bedeutungen angetrieben; diese können aber nur von komplexen biologischen Systemen, wie z. B. dem menschlichen Gehirn, realisiert werden. Dagegen ließe sich jedoch einwenden, dass die gleichen neurobiologischen Kausalprozesse ablaufen werden, unabhängig davon, was die involvierten neuronalen Zustände bezüglich der Welt repräsentieren oder ob sie überhaupt etwas repräsentieren. Neuronale Prozesse scheinen mit Bedeutungen und repräsentationalen Inhalten genauso wenig zusammenzuhängen, wie binäre Prozesse. Wie Bedeutung und Verstehen aus Molekülen und Zellen entstehen können, ist nicht weniger mysteriös, als wie sie aus Ketten von aus Nullen und Einsen entstehen könnten. Außerdem ließe sich, ähnlich wie bei Block, intuitiv einwenden, dass das ganze System Bewusstsein haben könnte.

Auch wenn die genannten Einwände den Funktionalismus per se nicht entkräften, wird deutlich, dass auch dieser zumindest angreifbar ist. Zudem verbirgt sich in den Argumenten implizit das Problem, die Ebene zu bestimmen, von der mentale Zustände bzw. Eigenschaften abhängen. Handelt es sich lediglich um eine Kausalrelation auf der Makroebene oder ist der Stoff bzw. die Materie, die den funktionalen Eigenschaften zugrunde liegt, entscheidend? Der Funktionalismus beantwortet nicht, inwiefern die höherstufigen oder die niedrigstufigen Eigenschaften für funktionale Zustände und damit auch für deren kausale Wirksamkeit relevant sind (vgl. Schröder 2004, 113).

Wo sind die Grenzen für funktionale Zustände bzw. ein funktionales System und seine Inputs und Outputs zu ziehen? Werden die Inputs bzw. Outputs ausschließlich als elektrochemische Signale aufgefasst, die das Gehirn erhält bzw. versendet, wird das Gehirn – und nicht eine Person – implizit als logisches Subjekt mentaler Zustände anerkannt. Außerdem führt diese Auffassung zu jenem Speziezismus bzw. Chauvinismus, den der Funktionalismus eigentlich vermeiden möchte, da Wesen mit anders funktionierenden Sinnesorganen oder Effektoren aus dem Kreis der mentalen Wesen von vornherein ausgeschlossen werden.[37] Werden die Inputs dagegen weiter gefasst als verschiedene Umweltsituationen, in denen wir uns befinden, und die Outputs als Veränderungen in unserer Umwelt, die wir durch unsere Handlungen hervorrufen, steht diese Einschätzung in Kontrast zu der Erfahrung, wonach die Bedingungen, unter denen eine Umweltsituation mentale Zustände hervorruft, nicht beliebig sind. Zudem käme der Reizverarbeitung ein sehr starkes Gewicht zu – in der Folge könnten Wesen mit unterschiedlichen Sinnesorganen nicht dieselben Überzeugungen haben.

***Quintessenz Funktionalismus***

- *Kennzeichnung mentaler Zustände durch Input-Output-Relation sowie kausale Relation innerhalb mentaler Zustände*
    - *Bestimmung mentaler Eigenschaften allein als funktionale Eigenschaften durch kausale Rolle und kausale Relation innerhalb des Systems*
- *Herausforderung der Multirealisierbarkeit als Generalisierung für mögliche Welten gegenüber konkretem Artenchauvinismus für tatsächliche Welt*
- *für mentale Kausalrelationen keine sinnvolle Erklärung ohne mentale Zustände möglich*
    - *Common-Sense-Funktionalismus: vage Begriffe*
    - *wissenschaftlicher Psychofunktionalismus: empirische Regelmäßigkeit*
- *ontologisch neutral: Ergänzung durch Physikalismusthese als Herausforderung für repräsentationale Modelle*
    - *Gefahr Epiphänomenalismus: kausale Kraft mentaler Eigenschaften nicht durch funktionale Rolle gesichert*
    - *keine adäquate epistemische Beschreibung durch Konzentration auf Realisierer*
- *Beschränkung mentaler Zustände auf funktionale Rolle unzureichend*
    - *Notwendigkeit von semantischem Verständnis für genuines Denken*
- *zentraler Stellenwert der Umwelt für Denkinhalte und -prozesse*

[37] Außerdem steht diese Auffassung vor dem konversen Problem: Für Inputs und Outputs als elektrochemische Signale ist es egal, wie sie zustande kommen – es kommen die Argumente der invertierten Qualia bzw. der rot-grün-Vertauschung zum Tragen. Zudem impliziert diese Möglichkeit über einen unplausiblen Speziesismus einen Normalismus; aufgrund der eingeschränkten Wahrnehmung müssten bereits Taube und Blinde vom Kreis der mentalen Wesen ausgeschlossen werden.

In seinem inzwischen berühmten Zwillingserde-Beispiel hat insbesondere Putnam gezeigt, dass propositionale Eigenschaften nicht allein in den inneren Zuständen des Subjekts bzw. seines Gehirns anzusiedeln sind, sondern auch durch ihre Umwelt beeinflusst werden (vgl. Putnam 1967b; Kap. 4.4.1). In dem Beispiel werden zwei molekular völlig identische Menschen einander gegenübergestellt, die sich dennoch in ihren Glaubensannahmen und Wünschen, aufgrund verschiedener Faktoren in ihrer räumlichen und lebensgeschichtlichen Umgebung unterscheiden. Das Zwillingserde-Gedankenexperiment ist nach wie vor eines der zentralen Gegenargumente gegen den Funktionalismus, da somit die Wirksamkeit rein intrinsischer Eigenschaften mit dem Repräsentationsproblem kontrastiert wird. Wenn Inhalt tatsächlich relational und extrinsisch ist, und wenn nur intrinsische Eigenschaften kausale Kraft besitzen, wie können Inhaltseigenschaften dann wirksam sein? Diesen Widerspruch gilt es aufzulösen, wenn man zeigen will, dass mentale (inhaltliche) Eigenschaften kausal wirksam sind.

### 2.3 Anomaler Monismus und mentale Kausalität

Die Position des anomalen Monismus ist eine weitere Alternative zur Typenidentität. Ähnlich wie der Funktionalismus wurde sie bereits Anfang der 1970er Jahre als Reaktion auf diese Position entwickelt und insbesondere von Davidson (1970) vertreten. Während der Funktionalismus sich auf die kausale Rolle mentaler Zustände konzentrierte, untersuchte Davidson, wie sich mentale und physische Prädikate, die Ereignisse beschreiben, zueinander verhalten (vgl. Kap. 3.2.1; Kap. 3.2.2). Er verfolgte somit zunächst eine begriffliche Unterscheidung. Mit dem anomalen Monismus versuchte er, eine Identitätstheorie zu begründen, die sowohl dem materiellen Weltbild als auch dem intentionalen Alltagsdiskurs gerecht wird. Davidsons anomaler Monismus (griech.: monas, „Einheit“) kann nur bedingt als eine Einheitslehre verstanden werden, da es zwar auf der ontologischen Ebene nur Ereignisse gibt, aber auf einer epistemischen Ebene verschiedene Beschreibungen möglich sind. Ähnlich der dritten kantischen Antinomie (vgl. Kant 1998, 548f.) sieht Davidson Schwierigkeiten, die Grundüberzeugungen einer deterministischen Welt mit der Nichtreduzierbarkeit des Mentalen in Einklang zu bringen:

> Die [...] gesetzte Grenze ist also nicht naturgegeben, sondern stammt von uns, sobald wir entscheiden, die Menschen als rationale handelnde Wesen mit Zielen und Zwecken zu sehen, als Wesen, die auch einer moralischen Bewertung unterliegen. (Davidson 1998, 335)

Als rationale Wesen, als die wir uns begreifen, können wir seiner Ansicht nach zum gegenseitigen Verständnis nicht zugunsten eines rein physikalischen Diskurses auf unser interpretatives mentales Vokabular verzichten. Die kausale Abhängigkeit und die Anomalie geistiger Ereignisse stehen

in einem scheinbaren Widerspruch zueinander, der sich aus drei Prinzipien ergibt: Erstens das Prinzip der kausalen Wechselwirkung, wonach wenigstens einige geistige Ereignisse kausal mit physischen Ereignissen interagieren. Das zweite Prinzip betrifft den nomologischen Charakter der Kausalität. So fallen Ereignisse, die in einer kausalen Relation zueinander stehen, unter strikte deterministische Gesetze. Das dritte Prinzip betrifft die Anomalie des Geistigen. Hiernach gibt es eben keine strikten deterministischen Gesetze, auf deren Grundlage geistige Ereignisse prognostiziert und erklärt werden können (vgl. ebd., 292f.; vgl. Kap. 3.5; Kap. 4.1). In der Konsequenz scheint es nicht möglich, dass mentale Ereignisse in kausaler Wechselwirkung mit physischen Ereignissen stehen – das erste Prinzip steht in offenem Widerspruch zu den beiden anderen.

### 2.3.1 Entwicklung der Ereignisontologie

Grundlegend für seine Konzeption der mentalen Ereignisse ist Davidsons Sprach- bzw. Handlungstheorie, in der er untersucht, was es bedeutet, sprachliche Äußerungen bzw. andere Menschen zu verstehen. Es geht also um eine theoretische Bestimmung linguistischer Kompetenz. In seinem Modell der radikalen Interpretation werden die Aussagen und Überzeugungen einer Person von einem Zuhörer oder Interpreten[38] auf ihre innere Konsistenz hin überprüft und ebenso nach außen hin mit den übrigen Überzeugungen der Person abgestimmt. Davidson fasst diese Bestimmung im „Principle of Charity" zusammen (vgl. Davidson 1986a, 316). Zum einen dient es dem Interpreten als Handlungsanweisung, die ihm mit den Prinzipien der Wahrheits- und der Konsistenzunterstellung das Verstehen des Sprechers ermöglichen soll. Zum anderen werden diese Prinzipien dem Sprecher einfach unterstellt, damit er interpretierbar wird. Er muss im Wesentlichen glauben, was er sagt und worauf er sich bezieht und dabei möglichst wenige widersprüchliche Überzeugungen vertreten. Insofern können intentionale Begriffe mitunter etwas „freestyle" (ders. 1987, 447) sein, eben sehr flexibel, kontextabhängig und nicht festgelegt. Geistige Ereignisse sind gekennzeichnet durch propositionale Einstellungen, wie z. B. zu glauben, zu wünschen oder zu hoffen. Sie werden als ein „Haben von Gründen" durch ihren intuitiven Gebrauch bestimmt (vgl. Kap. 3.2.1). Diese Definition ist unbefriedigend, doch Davidson legt sich nicht weiter fest: „[E]s ist uns offenbar misslungen, den intuitiven Begriff des Geistigen in

[38] Mit dem Modell der radikalen Interpretation sowie dem Principle of Charity bezieht sich Davidson auf Quine (vgl. Quine 1960, Kap. 2). In einem Gedankenexperiment wird ein Feldlinguist vor die Aufgabe gestellt, ein ihm völlig unbekanntes Sprachsystem, beispielsweise eines Indianerstammes im Dschungel, zu übersetzen und zu verstehen. Die Grundthese, die Davidson dabei entwickelt, besagt, dass alles Verstehen letztlich Teil einer Strategie der Verhaltenserklärung ist, und die intentionalen wie semantischen Begrifflichkeiten nur in ihrem Beitrag zu dieser Strategie verständlich sind. In dem Interpretationsmodell, das Davidson schließlich entwirft, sind die intentionalen Einstellungen eingebunden in die kausale Struktur der Welt. Es besteht eine Verbindung zwischen den Überzeugungen eines Sprechers und der Welt. Auf dieser Grundlage ist es dem Interpreten möglich, mithilfe von Kausalhypothesen Verständnis zu erreichen (vgl. Davidson 1990, 76).

den Griff zu bekommen.“ (Davidson 1998, 297) Ein mentales Ereignis zu sein, bedeutet, eine intentionale Beschreibung zu haben. Gemeint ist eine Beschreibung, die mindestens ein Verb essentiell enthält, das eine propositionale Einstellung ausdrückt (vgl. ebd., 298). Immerhin wird die Unbestimmtheit, sprich die intensionale Bezugnahme auf Ereignisse, als wesentliches Merkmal mentaler Ereignisse angegeben (vgl. ders. 1982, 320).

Die Bestimmung physischer Ereignisse ist demgegenüber ungleich einfacher. Zwar liefert Davidson auch hier kaum eine konkrete Bestimmung, allenfalls wird erklärt, physische Ereignisse hätten eine physikalische Beschreibung. So gelten alle Ereignisse in der materialistischen Welt als physische Ereignisse, sie sind mittels der exakten Naturwissenschaften bestimmbar (vgl. ders. 1990, 18; Kap. 3.2.2). Schließlich kann als weitere, sozusagen negative Definition gesagt werden, dass physische Ereignisse alle Ereignisse sind, die keine mentalen Ereignisse sind.

> Ein physikalisches Vokabular zu charakterisieren, ist weniger wichtig, weil es sich bei der Feststellung, ob eine Kennzeichnung mental oder physikalisch ist, relativ zum mentalen Vokabular sozusagen rezessiv verhält. (ders. 1998, 296)

Die Unterscheidung, die Davidson zwischen mentalen und physischen Ereignissen trifft, ist allerdings keine ontologische. Man muss auch hier seine spezielle Ereignisontologie im Blick behalten, der zufolge Ereignissen ein beschreibungsunabhängiger Status zufällt, wobei die Ereignisse in unterschiedlicher Terminologie, also entweder mit mentalen oder mit physischen Prädikaten beschrieben werden können. Und manchmal beziehen sie sich auf dieselben Ereignisse. In ihrer Beschreibung sind die Eigenschaften jedoch verschieden, denn im Unterschied zu den Ereignissen fällt ihren Eigenschaften kein genuin ontologischer Status zu. Die Eigenschaft eines Gegenstandes zeigt für Davidson lediglich, dass ein bestimmtes Prädikat auf diesen Gegenstand zutrifft. Sinnvolle Identitätsaussagen lassen sich dagegen nur über Einzeldinge, und das sind eben Einzelereignisse, machen:

> [I]f causal relations and causal powers inhere in particular events and objects, then the way those events and objects are described, and the properties we happen to employ to pick them out or characterize them, cannot affect what they cause. Naming the American invasion of Panama “Operation Just Cause” does not alter the consequences of the event. (ders. 1999, 8)

Ihre kausale Kraft haben Ereignisse unabhängig von der Terminologie, mit der sie beschrieben werden. Allerdings stellt sich die Frage nach dem ontologischen Status von Ereignissen. In ihrer Beschreibungsunabhängigkeit bestimmen sie schließlich das kausale Netzwerk und sollen zugleich, als mentale Ereignisse, den Beschreibungsanspruch des psychologischen Diskurses gewährleisten. Als Ereignisse versteht Davidson unwiederholbare, datierte Einzeldinge, wie z. B.

historische Begebenheiten oder Äußerungen. In ihrem ontologischen Status ähneln sie Gegenständen. Als Beispiele werden „Der Tod von Walter Scott" = „Der Tod des Verfassers von Waverly", „Die Ermordung des Erzherzogs Ferdinand" = „Das Ereignis, das den Ersten Weltkrieg auslöste", „Der Ausbruch des Vesuvs im Jahre 79 n. Chr." = „Die Ursache der Zerstörung von Pompeji" genannt (vgl. ders. 1970, 294; Kap. 3.1.1). Die gewählten Beispiele formulieren für Davidson bereits Identitätsbehauptungen.

Aber auch der Begriff des Ereignisses ist mehrdeutig und schwer einzugrenzen. So ist es bereits problematisch, Ort und Zeitpunkt eines Ereignisses genau zu bestimmen. Es ist möglich für das Auftreten physischer Ereignisse spezielle Raum-Zeit-Regionen anzunehmen, doch übertragen auf mentale Ereignisse erscheint diese Einteilung zunehmend vage (vgl. McLaughlin 1985, 344). Doch auch die Extension der Ereignisse ist schwer zu bestimmen. Wenn wir im alltäglichen Gebrauch davon sprechen, dass es beispielsweise Paul schlecht geht und er Magenschmerzen hat, kennzeichnet diese intentionale Beschreibung auf der Makroebene andere Ereignisse, als eine molekularbiologische Beschreibung von Enzymen und Säuren. Die Ursache-Wirkung Relation, die wir auf der makroskopischen Ebene den Ereignissen zuordnen, entspricht unserem alltäglichen Sprachgebrauch. Intentionale Gründe wie Wünsche und Absichten gelten demnach als ursächlich für physische Wirkungen. Dem gegenüber steht die Mikroebene der nomologischen Relation. Ereignisse, die unter Gesetze fallen, werden mikroskopisch durch physikalische Termini bestimmt. Problematisch ist nun die Zuordnung der Ereignisse von der makroskopischen auf die mikroskopische Beschreibungsebene und natürlich umgekehrt. Inwiefern ist eine solche Zuordnung möglich und überhaupt gültig? Wenn Mikroereignisse nicht identisch mit alltäglichen Ereignissen sind, benötigen wir auf der Makroebene eine andere Antwort als auf der Mikroebene. So ist es beispielsweise aussichtslos, die physikalische Beschreibung einer Wellenbewegung im Wasser mit einer bestimmten Menge an Molekülbewegungen identifizieren zu wollen. Damit wird deutlich, dass die Bestimmung von Ereignissen von der Terminologie des jeweiligen Diskurses abhängt. Diese beschreibt die Ereignisse aber nicht nur, sondern konstituiert sie auch. Die Annahme, dass es gänzlich beschreibungsunabhängige Ereignisse gibt, kann demnach zurückgewiesen werden (vgl. Haugeland 1982, 101). Für den anomalen Monismus hat das zur Folge, dass auch hier ein triftiges Argument für die Identität von mentalen und physischen Ereignissen fehlt.

### 2.3.2 Tokenidentität als Konsequenz

Wieso spricht Davidson aber von einem anomalen Monismus? Anstatt von seiner Ereignisontologie ausgehend zu argumentieren, begründet er zunächst die Anomalie des Mentalen. In seiner Handlungs- und Sprachtheorie hat Davidson gezeigt, dass der Bereich der Intentionalität bzw. die

psychologische Terminologie anderen Kriterien unterliegt als die physikalischen Diskurse (vgl. Davidson 1998, 324). Er ist insofern anomal, als es für ihn keine strikten Gesetze gibt, auf denen sich intentionale Diskurse begründen (vgl. Kap. 3.5; Kap. 4.1). Funktionierende Generalisierungen sind zwar auch für den mentalen Bereich möglich, es handelt sich bei diesen allerdings eher um Daumenregeln des Denkens, die von Ausnahmen bestimmt werden. Davidson verweist auf den holistischen Charakter des kognitiven Bereichs, was bedeutet, dass Überzeugungen stets in ihrem Gesamtsystem betrachtet und bewertet werden müssen. Intentionale Diskurse unterliegen dabei den Kriterien der Kohärenz, Rationalität und Widerspruchsfreiheit (vgl. Davidson 1998, 324). Verallgemeinerungen oder Regeln, die womöglich solche Aussagen vorhersehbar machen sollen, sind lediglich ungefähr wahr, durch großzügige Auswegklauseln gesichert und daher heteronom (vgl. ebd., 308). Der Interpret muss sich flexibel auf den Sprecher und die Umwelt einstellen, um diesen zu verstehen. Und hierbei helfen ihm eben keine strikten Gesetzte wie in der idealen Physik.

Das zentrale Argument gegen die Verbindung von physischen und mentalen Ereignissen ist insofern das von Davidson aufgestellte „Principle of Charity". Dessen für die intentionalistischen Diskurse konstitutiven Prinzipien sind nicht vereinbar mit den Grundsätzen physikalischer Diskurse. Mentale Ereignisse werden im Gegensatz zu physischen Ereignissen relational und normativ bestimmt (vgl. Kap. 4.5). Wir wissen gewissermaßen zu viel über menschliches Verhalten als dass wir exakten und universellen Aussagen Vertrauen schenken könnten. Denn Überzeugungen und Wünsche werden stets von weiteren Überzeugungen und Wünschen beeinflusst und modifiziert. „Dieser Holismus des geistigen Bereichs ist offenbar ein Fingerzeig sowohl auf die Autonomie als auch den anomalen Charakter des Geistigen." (Davidson 1998, 305f.) Wir sind dazu gezwungen, die Einstellungen und Handlungen unserer Mitmenschen als plausibel und kohärent anzuerkennen, da ein zu hohes Maß an Widersprüchen zu Konfusion und Unverständnis führen würde (vgl. ders. 1970). [39] So neigen wir dazu, Inkonsistenzen als scheinbar auszulegen und Überzeugungen und Wünsche grundsätzlich positiv zu bewerten, indem wir ihnen ein hohes Maß an Rationalität und Widerspruchsfreiheit zubilligen (vgl. ders. 1998, 332).

Innerhalb des intentionalistischen Diskurses sind strikte Gesetze somit ausgeschlossen. Davidson untersucht, ob es solche Gesetze zwischen mentalen und physischen Ereignissen geben kann. Diese Gesetze sind aber nur möglich, wenn sich beide Terminologien auf die physikalische reduzieren lassen. Da für Davidson fest steht, dass sich beide Terminologien über Ereignisse quantifizieren, muss lediglich geklärt werden, in welchem Verhältnis die unterschiedlichen Beschreibungsweisen zueinander stehen. Eine mögliche Bestimmung des Verhältnisses wäre das der

[39] In diesem Punkt wird einmal mehr deutlich, dass Davidson vor dem Hintergrund seiner Handlungstheorie argumentiert. Handlungen werden dabei als Unterarten von Ereignissen aufgefasst; Fragen, die sich auf Handlungen beziehen, können auch auf Ereignisse übertragen werden. So sind auch Handlungen Ereignisse in der Welt materieller Objekte und als solche unterschiedlich beschreibbar (vgl. Beckermann 1999, 187).

Typenidentität. Das würde bedeuten, dass es für jedes mentale Ereignis bzw. sein intentionales Prädikat M ein physikalisches Prädikat P gäbe, wobei beide Prädikate M und P extensionsgleich wären. Davidson spricht sich jedoch gegen eine Typenidentität aus. Intentionalistische und physikalische Diskurse sind seiner Meinung nach inkommensurabel. Die für eine Reduktion erforderlichen sogenannten Brückengesetze[40] können nicht aufgestellt werden (vgl. Kap. 3.3.1). Aufgrund der unterschiedlichen Extension ihrer Prädikate ist eine nomologische Identität der Typen M und P ausgeschlossen:

> Durch nomologische Aussagen werden Prädikate zusammengebracht, von denen wir a priori wissen, dass sie füreinander gemacht sind, d. h. von denen wir dies unabhängig von unserer Kenntnis, ob die Belege für einen Zusammenhang zwischen ihnen sprechen, wissen. [...] Die Richtung, in die unsere Erörterung nun zu gehen scheint, ist die: dass mentale und physikalische Prädikate nicht füreinander gemacht sind. (Davidson 1998, 307; vgl. auch ebd., 349)

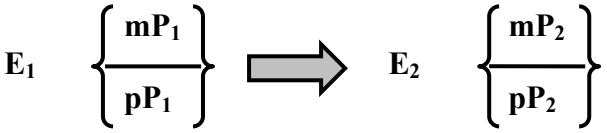
Insofern käme die Reduktion der mentalen auf die physikalischen Prädikate einem Themenwechsel gleich. Aufgrund dieser Normen wird der intentionalistische Diskurs nach Prinzipien strukturiert, die nicht in den physikalistischen Diskurs übersetzbar sind. Somit ist eine nomologische Reduktion der mentalen Ereignisbeschreibungen auf physikalische Ereignisbeschreibungen nicht möglich, schließlich werden beispielsweise die Ereignistoken $E_1$ und $E_2$ in einem mentalen Sinne anders sortiert als in einem physikalischen. Aus physikalistischer Perspektive sind die Ereignistoken der mentalen Prädikate zwar zugänglich, aber in der neuen Beschreibung würde ihre speziell mentale Kennung wegfallen. Ein sehr plakatives Beispiel für die Richtigkeit seiner Behauptung ist für Davidson der Misserfolg des definitorischen Behaviorismus (vgl. ebd., 303f.).

Da eine Typenidentität nicht möglich scheint, konzentriert sich Davidson auf die Identität von Token.[41] Diese stützt sich auf die ersten beiden der oben genannten Prämissen, dem Prinzip der kausalen Interaktion und dem Prinzip des nomologischen Charakters der Kausalität. Dem erstgenannten Prinzip zufolge interagieren einige mentale Ereignisse mit physischen Ereignissen. Davidson hält dies sogar für eine „unbestreitbare Tatsache". Unserem Verständnis nach sind solche kausalen Verknüpfungen alltäglich beobachtbar. Hinter der Benutzung eines Zuges anstel-

[40] Brückengesetze ermöglichen nach Nagel erst Theorienreduktion. So können beispielsweise alle Gesetze einer Theorie $T_1$ – eventuell mithilfe geeigneter Brückengesetze – auf Gesetzen der Theorie $T_2$ abgeleitet werden (vgl. Nagel 1961; Hooker 1981). Eine Theorienreduktion ist möglich aufgrund derselben kausalen Rolle, die beispielsweise die Prädikate F und G in den Theorien $T_1$ und $T_2$ einnehmen. Notwendige Vorraussetzung ist allerdings, dass F und G nomologisch koextensional sind, was Davidson bestreitet.

[41] Im Unterschied zur Typenidentität, die eine Identität zwischen Ereignisarten propagiert, geht Davidson lediglich von der Identität einzelner Ereignisse aus. Demnach umfassen Ereignistypen verschiedene Ereignistoken. Deshalb ist eine Reduktion des mentalen auf den physischen Bereich ausgeschlossen – jene sind nicht durch Brückengesetze reduzierbar. Die Identität einzelner Ereignisse kann – wörtlich genommen – nur jenseits der Beschreibung gefunden werden. Wenn man „token" (englisch) als Zeichen übersetzt, erhält der Begriff der Tokenidentität die zusätzliche Bedeutung, dass die Identität mentaler und physischer Zeichen nur auf der beschreibungsunabhängigen Ereignisebene zu finden ist.

le des Autos als Verkehrsmittel stehen beispielsweise eine ganze Reihe von Überlegungen, Meinungen und Entscheidungen. Banalere Beispiele sind das Kochen eines Kaffees oder die Wahl, eine Banane zu essen anstelle eines Apfels. „Hier kommen wir [...] nicht um die Feststellung herum, dass Kausalität im Spiel ist, d. h. dass der Wunsch und die Überzeugung kausale Bedingungen der Handlung waren." (ebd., 325) Das Prinzip des nomologischen Charakters der Kausalität besagt nun, dass Ereignisse einer Ursachen-Folge Relation unter strikte Gesetze fallen.




Grafik 5: Auf der ontologischen Ebene gibt es ein kausales Netzwerk von beschreibungsunabhängigen Ereignissen. Die Ursache-Wirkungs-Relation ist somit weder vorrangig mental noch physisch. Mentale und physische Prädikate gehören zwei unterschiedlichen Beschreibungssystemen an. Da mentale Prädikate bzw. die durch sie beschriebenen psychischen Eigenschaften keinen strikten Gesetzen unterliegen, ist keine Reduktion auf physische Eigenschaften möglich. Es ist lediglich eine Identität dieser Eigenschaften bei einzelnen Ereignistoken möglich. Allerdings muss für diese Identität letztlich ein Perspektivwechsel von der epistemischen auf die ontologische Ebene erfolgen.

Betrachten wir die Konsequenz, die sich aus den beiden Prinzipien für mentale Ereignisse ergibt, nämlich den Beweis der Identität. Wenn nach dem ersten Prinzip ein mentales Ereignis M ein physisches Ereignis P verursacht, muss es nach dem zweiten Prinzip ein striktes Gesetz geben, unter das M und P fallen. Da es, wie wir oben gesehen haben, weder strikte psychische noch psychophysische Gesetze gibt, muss dieses Gesetz ein physikalisches sein. Daraus ergibt sich wiederum, dass das mentale Ereignis M physikalisch beschreibbar sein muss, was letztlich bedeutet, dass M auch ein physisches Ereignis sein muss. „Also ist jedes geistige Ereignis, das in kausaler Beziehung zu einem physischen Ereignis steht, ein physisches Ereignis." (ebd., 315) Den beiden Prinzipien nach besteht also Identität zwischen mentalen und physischen Ereignissen.

Der entscheidende Punkt ist aber, dass dabei keine Identität zwischen den Ereignis*typen*, sondern zwischen den Ereignis*token* besteht. Davidson hält eine systematische Zuordnung zwischen physikalischen und psychologischen Beschreibungen für ausgeschlossen. Offene Klassen von physikalisch beschriebenen Ereignissen können nicht mit Klassen von psychologisch beschriebenen Ereignissen korrelieren; lediglich zwischen gegebenen, endlichen Klassen ist eine solche Zuordnung möglich (vgl. ebd., 348). Davidson behauptet nämlich, dass für zeitlich bestimmte psychische Einzelereignisse eine physikalische Beschreibung geleistet werden kann. Insofern ist es notwendig, „[...] zeitlich bestimmte Einzelereignisse von Arten von Ereignissen zu unterscheiden [...]" (ebd., 354). Reduktion ist für Davidson demnach ausgeschlossen. Schließlich sind Kausalität und Identität Beziehungen zwischen individuellen Ereignissen. Diese werden mental oder, sofern sie unter Gesetze fallen sollen, physikalisch beschrieben. Gesetze beschreiben also – und das ist entscheidend – einen sprachlichen Bereich. Es ergeben sich damit zwei

Merkmale für mentale Ereignisse in Verbindung mit physischen – eine kausale Abhängigkeit auf der ontologischen Ebene und eine nomologische Unabhängigkeit auf der Beschreibungsebene.

Unter Einbeziehung von Supervenienzüberlegungen hat Davidson schließlich versucht, die Asymmetrie der Tokenidentität zugunsten der physischen Ereignisse aufzuheben (vgl. ders. 1999, 5; vgl. Kap. 2.4). Seine frühe Formulierung des Supervenienzgedankens, wonach „[...] geistige Merkmale in gewissem Sinne von physischen Merkmalen abhängig sind oder zu diesen hinzukommen" (Davidson 1998, 301), beziehen sich zunächst auf Ereignisse oder Gegenstände und sind insofern stark auslegungsbedürftig.[42] Schließlich hat Davidson nicht genau gezeigt, wie mentale Ereignisse zu definieren sind, von daher ist eine mögliche Unterscheidung „in einem mentalen Aspekt" nicht genau bestimmt. Damit ist auch unklar, welche mentalen Ereignisse M auf physischen Ereignissen P supervenieren. Ist es zulässig, davon zu sprechen, dass beispielsweise Wünsche auf neuronalen Zuständen supervenieren, oder sind die Klassen der Ereignisse kleiner oder größer bestimmt?[43] In einer neueren Version bestimmt Davidson die Tokenidentität als eine superveniente Relation zwischen Prädikaten: „[…] a predicate *p* is supervenient on a set of predicates *S* if and only if *p* does not distinguish any entities that cannot be distinguished by *S*" (ders. 1999, 4).[44] Allerdings wird in der Kritik deutlich, dass durch diese implizite Aufwertung der Eigenschaften bzw. Prädikate auch in kausaler Hinsicht Konsequenzen erwartet werden. Schließlich wird die Supervenienzbeziehung als ontologische und nicht als rein epistemische Relation verstanden.

### 2.3.3 Tokenidentität in der Kritik

Gegen Davidsons Annahme, dass die Gesetze der Physik maßgeblich seien und es keine psychophysischen Gesetze geben könne, gibt es allerdings verschiedene Einwände: Beispielsweise besäßen alle Einzelwissenschaften ebenso wenig Erklärungskraft wie die Psychologie (vgl. Burge 1989, 317). Zudem ist es durchaus möglich, dass es im Rahmen des wissenschaftlichen Fortschritts psychologische Gesetzmäßigkeiten gibt (vgl. Fodor 1989, 75f.). Schließlich ist es möglich, dass es mentale Ereignisse gibt, die außerhalb der Physik beschrieben werden, aber dennoch mit

[42] Unter einer Supervenienzrelation (lat. supervenio, „sich über etwas werfen") wird allgemein verstanden, dass eine supervenierende Eigenschaft M sozusagen zu einer subvenienten Eigenschaft P hinzukommt und von dieser „getragen" wird (vgl. Kim 1984, 1990; vgl. Kap. 2.4).

[43] Das Zwillingserde-Beispiel hat gezeigt, dass Überzeugungen extrinsisch individuiert werden, weshalb ihre Supervenienz nicht allein von den neuronalen Zuständen abhängt. Davidson akzeptiert das Argument, verwirft jedoch nicht die Identität von mentalen und physischen Ereignistokens (vgl. Davidson 1989, 12f.).

[44] Davidson vertritt das Prinzip einer schwachen Supervenienz, da er weiterhin auf der Differenz zwischen physischen und mentalen Ereignisbeschreibungen besteht. So interessieren im Zusammenhang mit mentalen und physischen Relationen vor allem die Fälle, in denen p jeglicher Form von Reduktion widersteht. Insofern impliziert Supervenienz zwar Monismus, jedoch nicht definitorische oder nomologische Reduzierbarkeit. Deshalb ist sie interessant für Davidson (vgl. Davidson 1999, 5).

physischen Ereignissen kausal interagieren; damit wäre die Geschlossenheit der Physik nicht mehr absolut (vgl. McLaughlin 1985, 343). Wenn es also beobachtbare psychophysische Korrelationen gibt – und solche nimmt Davidson ja mit Verweis auf die Supervenienz an – unterliegen sie entsprechenden empirischen Gesetzen. Und damit wären intentionale Prädikate auch in die physikalische Terminologie übersetzbar:

> But to say that observed psycho-physical correlations, or for that matter observed psycho-psycho correlations, can be projected is just to say that they are empirically lawful. And if there are lawful psycho-physical generalizations, then there are locally sufficient physical conditions for the application of intentional predicates, in violation of the thesis of the anomalism of the mental. (Antony 1994, 243)

Gerade diese letzte Feststellung, die in dem Zitat angedeutet wird, zeigt die weitreichenden Folgen für das Verhältnis zwischen mentalen und physischen Ereignissen. Schließlich würde sich aus einer wie auch immer gearteten gesetzmäßigen Relation auch die Möglichkeit zur Reduktion ergeben. Insofern trifft Kim es ganz passend, wenn er formuliert: „I think „non-strict laws" are bad news for anomalous monists. In embracing them they may end up losing anomalism from anomalous monism." (Kim 1999a, 26) Seiner Auffassung nach würden nicht-strikte Gesetze den Anti-Reduktionismus in Erklärungsschwierigkeiten bringen.

Diese reduktionistische Tendenz kennzeichnet einen weiteren Punkt der Kritik an Davidsons anomalem Monismus. Gesetzmäßige Verbindungen, egal ob strikt oder nicht-strikt, sind evident für die Konsequenzen zwischen mentalen und physischen Ereignissen. „Evidential promiscuity is a direct and unavoidable consequence of the discovery of any nomologically reliable link between the mental and the physical, and it is this consequence Davidson wants to block." (Antony 1994, 236)

Donald Davidsons Theorie des anomalen Monismus hat sicherlich großen Einfluss auf die Debatte innerhalb der Philosophie des Geistes ausgeübt. Seine Position gegen die Möglichkeit psychophysikalischer Gesetze kann ebenso wenig ignoriert werden, wie seine Argumente bezüglich der kausalen Relation mentaler Ereignisse. Die Diskussion des anomalen Monismus innerhalb der letzten dreißig Jahre hat allerdings auch auf einige Schwächen dieser Theorie hingewiesen. Davidson selbst unterteilt die Kritiker in zwei Gruppen: Jene, die seine Argumentation gegen die Existenz psychophysikalischer Gesetze obskur finden und jene, die die drei grundlegenden Prämissen des anomalen Monimus als inkonsistent ablehnen:

> The criticisms are connected: if AM [Anomalous Monism, der Verf.] makes the mental causally inert, then AM apparently implies the falsity of the first premiss and hence the

> inconsistency of the three premisses. The third premises seems to many critics the relevant offender, so they urge that it should be dropped. (Davidson 1999, 3)

Letztendlich vereinigen sich alle Kritiken in der Aussage, dass der anomale Monismus das Mentale kausal unwirksam lasse. Es sei lediglich eine Begleiterscheinung des Physikalischen, ohne eigene Kraft oder Wirkung.[45] Mitschuld an dieser Sorge trägt Davidsons Überlegung, die Beziehung zwischen mentalen und physischen Ereignissen auf der Basis von Supervenienzüberlegungen zu beschreiben (vgl. ebd., 5). Zwar vertritt Davidson die Version der schwachen Supervenienz, um somit Reduktion zu vermeiden. Zugleich fehlt aber die modale Stärke, die die mentalen Ereignisse an die physischen binden soll (vgl. Kap. 3.1.3; Kap. 4.1.1). Unter Einbeziehung aller möglichen Welten wäre es somit möglich, dass in einer Welt $W_2$, die unserer aktuellen Welt $W_1$ detailgetreu gleicht, die psychologischen Eigenschaften ganz anders verteilt sind. Die gewünschte Abhängigkeit der supervenierenden von der subvenierenden Ebene ist nicht erreicht, vielmehr handelt es sich um einen scheinbar willkürlichen Parallelismus. Folgerichtig bemerkt Kim bezüglich der schwachen Supervenienzversion:

> For anomalous monism entails this: the very same network of causal relations would obtain in Davidson's world if you were to redistribute mental properties over its events any way you like; you would not disturb a single causal relation if you randomly and arbitrarily reassigned mental properties to events, or even removed mentality entirely from the world. (Kim 1989, 34f.)

Die Beseitigung aller mentalen Eigenschaften von Ereignissen hätte in dieser Welt keine Konsequenz hinsichtlich der Verteilung der physischen Eigenschaften. Damit wird die Relevanz einer Tokenidentität recht fragwürdig. Welchen Sinn macht es, von Identität zwischen mentalen und physischen Token zu sprechen, wenn die mentalen völlig austauschbar sind, und sich dennoch kein Unterschied für die physischen Token ergibt? Darüber hinaus entzieht eine solche Argumentation den mentalen Eigenschaften jegliche kausale Kraft.[46]

Zudem wird unter der Voraussetzung einer gegebenen Supervenienz mit der grundlegenden Prämisse gebrochen, dass ausschließlich physikalische Gesetze singuläre Kausalereignisse beschreiben können. Schließlich beschreiben Supervenienzen die nomologische Verbindung zwischen mentalen und physischen Eigenschaften. Damit wird deutlich, dass auch für singuläre To-

[45] Diese nicht unberechtigte Angst vor einem Epiphänomenalismus hat Fodor treffend als „epiphobia" charakterisiert (vgl. Fodor 1989, 59).

[46] Auch der starke Supervenienzbegriff bringt den anomalen Monismus in Schwierigkeiten. Zwar liefert er anstatt einer zufälligen Kovarianz die gewünschte modale Stärke für das Abhängigkeitsverhältnis, doch verbunden sind damit – und das ist der springende Punkt – Gesetze, die zwischen supervenienten und subvenienten Eigenschaften bestehen: „whenever anything has mental property M there is some physical property Q such that it has Q and everything that has Q has M." (Vgl. Kim 1999a, 22f.) Gefordert wird damit eine Gleichheit in physischen und mentalen Eigenschaften für alle möglichen Welten. Wenn demnach zwei Personen sich beispielsweise zur gleichen Zeit in demselben neuronalen Zustand befinden, ist es nicht möglich, dass nur eine von beiden Personen Schmerzen empfindet und die andere nicht.

kenereignisse psychophysische Gesetze erforderlich sind, um eine kausale Relation zwischen den mentalen und physischen Ereignissen zu sichern.

> Does this not therefore undermine the very argument Davidson offers for anomalous monism breaking it at the point where it is held that only physical laws are available to back singular causal statements? (Johnston 1985, 417)

Das Prinzip der Supervenienz verhält sich widersprüchlich zum anomalen Monismus. Zum einen treffen hier zwei unterschiedliche Gesetzesbegriffe aufeinander, die nicht vereinbar sind. Zum anderen ist die Reduzierbarkeit der mentalen Eigenschaften auf die physischen in einer Supervenienzrelation nicht völlig ausgeschlossen.

Wem genau soll aber kausale Wirkung zugeschrieben werden? Die meisten physikalistischen Theorien gehen davon aus, dass Ereignisse oder Zustände kraft ihrer Eigenschaften kausale Wirkung haben (vgl. Kap. 3.1.2; Kap. 3.6). Davidson muss eine solche Zuordnung ablehnen. In seinem Universum gibt es nur Ereignisse, deren Eigenschaften bereits Bestandteil einer spezifischen Terminologie sind. Seiner Ansicht nach ist es unerheblich, in welcher Terminologie Ereignisse beschrieben werden, sie haben ihre kausale Kraft unabhängig von den verschiedenen Möglichkeiten, in denen sie beschrieben werden (vgl. Davidson 1999, 8). Es wurde aber gezeigt, dass gerade diese Voraussetzung problematisch und nicht begründbar ist. Mitunter sind Eigenschaften konstituierend für Ereignisse, so dass mentale Ereignisse von mentalen Eigenschaften und physische Ereignisse von physischen Eigenschaften bestimmt werden. Da aber mentale und physische Eigenschaften unmöglich miteinander identisch sein können, ergibt sich erneut ein Problem für die von jenen konstituierten Ereignisse, denn auch sie können somit nicht identisch sein (vgl. McLaughlin 1985, 340).

Darüber hinaus ist aber auch die Abgrenzung von mentalen und physischen Ereignissen bzw. Eigenschaften ungenau. Davidson legt sich nicht eindeutig fest, wie die Klasse der Prädikate hinsichtlich mentaler und physischer Prädikate verteilt ist (vgl. ebd., 350). Es ergeben sich somit auch Probleme hinsichtlich der Zuordnung. So bleibt offen, ob das Erlebnis „Grün Sehen" dasselbe Ereignis beschreibt wie die entsprechende neurologische Messung im Gehirn bzw. die Aktivität bestimmter Neuronen. Entsprechend ist es schwer vorstellbar, wie ein sehr komplexes mentales Ereignis wie ein Wunsch, an den sich vermutlich zahlreiche Dispositionen anschließen, mit einem entsprechenden physischen Ereignis, z. B. einem Neuronen-Ensemble, identifiziert werden kann. Damit hängt wiederum der Vorwurf zusammen, dass die Ereignisontologie die mentalen Eigenschaften wirkungslos belässt und somit den Epiphänomenalismus stärkt. Jene

sind allenfalls relevant, aber nicht wirksam; den meisten Philosophen, die pro mentale Verursachung argumentieren, ist das aber zu wenig (vgl. Kim 1999a, 24; Kap. 3.6.1).[47]

***Quintessenz anomaler Monismus***

- *Widerspruch zwischen kausaler Wechselwirkung mentaler und physischer Ereignisse und fehlenden psychophysischen Gesetzen*
    - *Geltung von Rationalitätsprinzipien für intentionale Begriffe: holistischer kognitiver Bereich (Kohärenz, Rationalität, Widerspruchsfreiheit)*
    - *Geltung exakter, strikter Naturgesetze für physische Begriffe*
- *anstelle von Typenidentität lediglich Identität von Ereignistoken*
    - *mentale Ereignisse physisch und mental beschreibbar*
    - *beschreibungsunabhängige Ereignisontologie*
    - *kausale Kraft für Ereignisse im kausalen Netzwerk*
- *Gefahr Epiphänomenalismus: mentale Eigenschaften kausal wirkungslos*
    - *Eigenschaften konstituierend für Ereignisse*
    - *psychophysische Gesetze auch für singuläre Tokenereignisse notwendig*
- *ceteris-paribus-Gesetze als Alternative zu strikten Gesetzen*

In der Tat erscheint die Annahme der Kausalrelation auf der Ebene der Ereignisse als eine Verlagerung des Problems mentaler Kausalität. Mit diesem Kunstgriff soll die Erklärungskraft des intentionalen Diskurses zwar gewahrt bleiben, eine wirklich befriedigende Lösung wird jedoch nicht angeboten. Mit dem Ausweichen auf eine Ereignisontologie hat der anomale Monismus letztlich das Ziel verfehlt, die kausale Wirkung mentaler Ereignisse zu beschreiben. Die Tokenidentität soll zwar gewährleisten, dass der anomale Monismus ein ontologischer Monismus mit einem eigenständigen intentionalistischen Diskurs ist. Andererseits behauptet er, dass mentale Ereignisse auch als physische beschreibbar sind. Davidson legt Wert darauf, dass die „[...] mentalistischen Begriffe in keiner Weise zweitklassig oder eliminierbar [...]“ (Glüer 1993, 154) sind. Aber wenn die physikalischen Diskurse den Rahmen oder das kausale Netz liefern, ist nicht einsichtig, inwiefern mentale Ereignisse kausal effizient für physische Ereignisse sein können. Somit wird deutlich, dass Davidsons scheinbar beschreibungsunabhängige Ereignisontologie nicht frei

[47] Ein Beispiel hierfür ist das Ereignis eines Schusses, mit dem ein Mörder einen Menschen erschießt. Für das Resultat des Todes des Opfers spielt es keine Rolle, ob der Schuss laut war oder leise. Demnach sind in dem gewählten Beispiel ausschließlich die physischen Eigenschaften des Ereignisses ausschlaggebend für den Todesfall. Sie stellen die notwendigen und hinreichenden Bedingungen für die kausale Wirkung des Ereignisses. Sosa (1999) räumt aber im Gegensatz zu Davidson den Eigenschaften einen eigenen ontologischen Status zu. Seiner Ansicht nach muss der anomale Monismus Raum für Eigenschaften zulassen, um kausale Relationen zu erklären (vgl. Sosa 1999, 46f.).

von Spannungen ist. Die Gefahr, dass mentale Ereignisse lediglich epiphänomenal sind, bleibt bestehen. Auch der anomale Monismus scheint letztlich nicht stabil genug, um eine echte kausale Kraft mentaler Ereignisse in einer physischen Welt zu begründen.

### 2.4 Supervenienztheorie und mentale Kausalität

Innerhalb der Philosophie des Geistes hat sich vor allem Jaegwon Kim mit der Supervenienztheorie auseinandergesetzt, die Theorie maßgeblich ausformuliert und in der Diskussion etabliert.[48] In diversen Aufsätzen beschreibt er die verschiedenen Formen der Supervenienz, die so genannte schwache, globale und starke Supervenienz (vgl. z. B. Kim 1984, 1987, 1990). Er konzentriert sich dabei zunehmend auf die Supervenienz von Eigenschaften, da diese die meisten übrigen Entitäten, beispielsweise Sätze, Zustände, Ereignisse oder Propositionen miteinbezieht: „It will become plausible, I believe, that property supervenience is fundamental, and that supervenience for most other entities can be explained in terms of it." (ders. 1984, 55; vgl. Kap. 3.1) Die Supervenienztheorie wurde einerseits von Kim auf der Basis von Davidsons anomalem Monismus heraus entwickelt, kann aber andererseits auch als Ergänzung zur funktionalistischen Position betrachtet werden. So soll die Relation zwischen mentalen und physischen Eigenschaften durch die Supervenienz genauer bestimmt werden und somit – zumindest nach Kim – eine starke physische Basis erhalten.

Die Supervenienztheorie liefert zudem für die nicht-reduktionistische Position einen wichtigen theoretischen Beitrag, die Beziehung zwischen mentalen und physischen Eigenschaften zu beschreiben, ohne erstere auf letztere zu reduzieren. Das Supervenieren mentaler Eigenschaften auf physischen Basen ist vergleichbar mit einer Beethoven-Symphonie, die von verschiedenen Orchestern, der New Yorker Philharmonie oder der Berliner Philharmonie, gespielt werden kann. Die Melodie wird dabei von unterschiedlichen Instrumenten getragen, die personell auswechselbar sind, und kann in diesem Sinne vielfach realisiert werden. Mit der Supervenienztheorie sind vor allem drei Ideen verbunden, die der Kovarianz, der Abhängigkeit und der Nicht-Reduzierbarkeit von Eigenschaften.

[48] Brian McLaughlin geht sogar so weit, dass der Begriff „Supervenienz" dank Kim inzwischen zum traditionellen Vokabular der analytischen Philosophie gehört: „It is fair to say that he has done more than anyone else to identify types of supervenience, to examine their pairwise logical relationships, and to explore their usefulness for various philosophical purposes." (McLaughlin 1997, 209)

### 2.4.1 Kovarianz und starke Supervenienz

Kovarianz bedeutet, dass Gegenstände oder Ereignisse, die in ihren subvenienten Eigenschaften nicht unterscheidbar sind, auch in ihren supervenienten Eigenschaften nicht unterscheidbar sind. Als Beispiel kann auf die Eigenschaft „ein guter Mensch zu sein" verwiesen werden. Wer sich durch bestimmte Taten auszeichnet, dem werden entsprechende moralische Eigenschaften zugeschrieben.[49] Hinsichtlich des Leib-Seele-Problems verbirgt sich dahinter die Frage, ob ein Replikant oder ein Klon, der seinem Original physisch vollständig gleicht, auch in den psychologischen Eigenschaften identisch ist. Anders formuliert: Sind physische Duplikate notwendig auch psychologische Duplikate (vgl. ders. 1995, 577)? Die verschiedenen Formen der Kovarianz beschreiben unterschiedlich starke Abhängigkeiten zwischen mentalen und physischen Eigenschaften im Rahmen logisch möglicher Welten (vgl. Kap. 3.1.3). So bezieht sich die schwache Supervenienz lediglich auf eine Welt; für verschiedene mögliche Welten ist die Determination zwischen den mentalen und den physischen Eigenschaften nicht mehr gegeben. Somit kann die schwache Supervenienz nicht die geforderte starke Abhängigkeit der mentalen Eigenschaften von den physischen garantieren (vgl. ders. 1984, 57ff.).[50]

Die Variante der globalen Supervenienz bezieht sich dagegen auf verschiedene logisch mögliche Welten, in denen die Kovarianz mentaler und physischer Eigenschaften in gewünschter Stärke gegeben ist. Die Überlegung der Ununterscheidbarkeit mentaler Eigenschaften hinsichtlich ihrer physischen Korrelate bezieht sich auf ganze Welten, eben „global", wodurch im Gegenzug ermöglicht wird, dass sich die Welten in ihren physischen Eigenschaften unterscheiden, obwohl sie mental ununterscheidbar sind (vgl. ebd., 68f.).[51] Kim selbst misst allerdings der Variante der starken Supervenienz am meisten Bedeutung bei, da sie nicht nur die geforderte modale Stärke gewährleistet und innerhalb einer sowie für verschiedene logisch mögliche Welten gilt, sondern

[49] Ursprünglich kommt der Begriff der Supervenienz aus der Ethik und beschreibt das Abhängigkeitsverhältnis zwischen moralischen und nicht-moralischen Eigenschaften. Beispielsweise werden zwei Personen, die sich in derselben Weise verhalten, auch dieselben moralischen Eigenschaften zugeschrieben. Es gibt keinen Unterschied moralischer Zuschreibung ohne einen Unterschied der nicht-moralischen Eigenschaften (vgl. Hare 1952, 145; vgl. Moore 1922).

[50] Eine mögliche Formel der schwachen Supervenienz lautet: „A *weakly supervenes* on B if and only if necessarily for any x and y if x and y share all properties in B then x and y share all properties in A – that is, indiscernibility with respect to B entails indiscernibility with respect to A." (Kim 1984, 58) Das Problem der schwachen Supervenienz ist, dass sie nur innerhalb einer Welt gilt. Innerhalb ihres theoretischen Rahmens ist es dagegen durchaus möglich, dass sich zwei physisch identische Ereignisse in ihren mentalen Eigenschaften unterscheiden.

[51] Eine mögliche Formel der globalen Supervenienz lautet: „A *globally supervenes* on B just in case worlds that are indiscernible with respect to B („B-indiscernible", for short) are also A-indiscernible." (Kim 1984, 68) Obwohl Kim aufgrund lokaler Anwendung die Variante der starken Supervenienz bevorzugt, ist nicht entschieden, ob diese tatsächlich mehr modale Stärke liefert als die globale Variante, beziehungsweise letztere nicht erstere sogar enthält. So wurde immerhin gezeigt, dass globale und starke Supervenienz gleichwertig sind, wenn sie sich auf intrinsische Eigenschaften beziehen (vgl. ders. 1995, 578).

darüber hinaus für Individuen, sozusagen lokal, die Abhängigkeit der mentalen von den physischen Eigenschaften sichert. Mögliche Formulierungen der starken Supervenienz lauten:

> (SI) For any individuals x and y, and any worlds wj and wk, if x in wj is *P*-indiscernible from y in wk (that is, x has in wj exactly the same properties in *P* that y has in wk), then x in wj is *M*-indiscernible from y in wk. (ders. 1995, 578)

oder:

> (SII) A *strongly supervenes* on B just in case, necessarily, for each x and each property F in A, if x has F, then there is a property G in B such that x has G, and *necessarily* if any y has G, it has F. (ders. 1984, 65)

Bei der starken Supervenienz handelt es sich um eine asymmetrische Beziehung. Sie soll die geeignete Abhängigkeitsrelation mentaler und physischer Eigenschaften beschreiben und sich somit als physisch ausweisen. Demnach ist jede mentale Eigenschaft vollständig von einer entsprechenden physischen Eigenschaft determiniert. Auf der anderen Seite soll sie auch die nomologische Nichtreduzierbarkeit des Mentalen auf das Physische garantieren – eine Forderung, die mit der starken Variante der Supervenienz allerdings nur schwer zu erfüllen ist, wie sich zeigen wird. Die Formulierung (SI) macht deutlich, dass sich die starke Supervenienzthese eben insbesondere auf Einzelereignisse oder -zustände bezieht, während die Formulierung (SII) die Notwendigkeit in den Blickpunkt rückt, die für die Korrelation zentral ist. Allerdings ist es eine Frage der Interpretation, ob psychophysische Supervenienzen als metaphysisch, logisch oder nomologisch notwendig bestimmt werden. Kim lässt diese Frage zunächst offen (vgl. ebd., 66).

Beispiele für alltägliche Supervenienzen finden sich in der Beziehung von Beschleunigung, Geschwindigkeit und Lage eines Objekts im Raum. Denken wir an ein Kreuzschiff, das die Passage von Amsterdam nach Casablanca fährt. Ohne Veränderung der Geschwindigkeit kann es nicht beschleunigen, ebenso kann das Schiff nicht die Geschwindigkeit erhöhen, ohne den Standort zu wechseln. Somit supervenieren die Tatsachen der Beschleunigung eines Objekts bzw. Schiffs auf den Tatsachen seiner Geschwindigkeit und die Tatsachen der Geschwindigkeit wiederum supervenieren auf den Tatsachen bezüglich der Position des Objekts bzw. Schiffs (vgl. Mandik 2008). Das Beispiel verdeutlicht den wichtigen Punkt von Supervenienzbeziehungen: Makroeigenschaften hängen von Mikroeigenschaften ab, ohne identisch mit ihnen zu sein. Die Beschleunigung des Schiffs ist ohne eine Veränderung der Geschwindigkeit nicht möglich, dennoch ist beides voneinander verschieden. Auf den Bereich der Bewusstseinsphilosophie übertragen, ergibt sich eine Bindung zwischen mentalen und physischen Eigenschaften, die über die bloße Korrelation hinausgeht.

Zur Erläuterung der starken Supervenienz soll noch einmal das Beispiel der „guten Menschen" herangezogen werden. Angenommen, ein guter Mensch zu sein, definiert sich über die drei Charaktereigenschaften Wohltätigkeit, Tapferkeit und Ehrlichkeit. So nehmen wir weiter an, dass der Philosoph Sokrates ein ehrlicher und tapferer Mensch ist. Dem heiligen Franziskus dagegen werden vor allem die Eigenschaften der Wohltätigkeit und Ehrlichkeit zugeschrieben. Von beiden Männern ließe sich aber sagen, dass es sich um gute Menschen handelt. Es wird deutlich, dass es alternative subveniente Eigenschaften gibt, die für sich hinreichende Basen für die supervenienten Eigenschaften sind (vgl. Kim 1984, 65). Übertragen auf psychophysische Supervenienz bedeutet das, dass es für jede mentale Eigenschaft M eine Menge an physischen Korrelaten $P_1$, $P_2$, ..., $P_n$ gibt. Ein Schmerz ist demnach in der Welt $W_1$ für einen Menschen durch das Feuern von C-Fasern realisiert, während in der Welt $W_2$ bei einem Marsianer, der unter Schmerzen leidet, die M-Fasern feuern. Die Version der starken Supervenienz hat also den Vorteil, dass sie auch dem Prinzip der Multirealisierbarkeit Rechnung trägt (vgl. Kap. 4.2.1).

### 2.4.2 Abhängigkeit und Nicht-Reduzierbarkeit

Aus den Ansprüchen, die an die Korrelation von mentalen und physischen Eigenschaften herangetragen werden, wird deutlich, dass Supervenienzen in einem Spannungsfeld zwischen Abhängigkeit und Nichtreduzierbarkeit stehen. Die Nichtreduzierbarkeit von Eigenschaften höherer Level bezeichnet Kim als zentrales Argument für den nichtreduktiven Physikalismus. Dass mentale Ereignisse physische Basen benötigen, um realisiert zu werden, verweist allerdings auf ein großes Problem hinsichtlich der Nicht-Reduzierbarkeit. Da die physischen Basen darüber hinaus auch hinreichend für das Erscheinen der Eigenschaften der oberen Level sind, kommt die eigenständige Existenz mentaler Eigenschaften weiter in Bedrängnis. Natürlich versteht sich die Supervenienztheorie, zumindest in Kims Ansatz, als physikalistische Theorie und muss damit in die physikalistische Weltkonzeption eingeordnet werden. Wie bereits Davidson versteht auch Kim die Welt als ein Netzwerk kausaler Relationen, wobei die Verbundenheit der Dinge untereinander am Besten mit dem Begriff der Abhängigkeit oder sogar Determination beschrieben wird (vgl. Kap. 3.1.1):

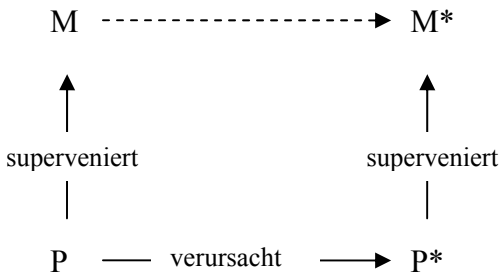
> [T]hings are connected with one another in that whether something exists, or what properties it has, is *dependent on*, or *determined by*, what other things exist and what kinds of things they are. (Kim 1984, 53)

Aufgrund dieser Abhängigkeitsverhältnisse ist es möglich, die Welt einsehbar und verständlich zu machen, Erklärungen und Prognosen zu liefern. Die Überlegung, wie nun mentale Eigenschaften

bestimmbar sein sollen, wird insofern durch die Abhängigkeit zu ihren physischen Eigenschaften bestimmt. Superveniente Eigenschaften sind abhängig oder determiniert von subvenienten Basen. Dieses Abhängigkeitsverhältnis führt aber nicht zwangsläufig zu Reduktionismus. Kim lehnt sich an ein Modell von Hellman und Thompson (1975) an, wonach physische Tatsachen lediglich alle Tatsachen der Welt, also auch mentale Tatsachen, determinieren.[52] Die Forderung der nichtreduktiven Physikalisten nach einer Abhängigkeit ohne Reduktion scheint durch die Supervenienztheorie somit erfüllt zu sein. Doch Kim bezweifelt, dass diese Einschätzung wirklich zutrifft. So wird für die starke psychophysische Supervenienz angenommen, dass jede mentale Eigenschaft M eine Serie physischer Eigenschaften $P_1$, $P_2$, ..., $P_n$ hat, von der jede einzelne hinreichend für M ist. Kims Überlegung ist nun, dass die Gesamtheit der physischen Eigenschaften $\cup P_i$ als notwendige Koextension von M auch als Basis zur Reduktion dienen könnte. Seiner Meinung nach zeigt die Existenz von $\cup P_i$ für M und ähnliche physische Extensionen für alle anderen mentalen Eigenschaften zumindest die metaphysische Reduzierbarkeit mentaler Eigenschaften. „Obviously, it is not the business of a philosophical argument to generate actual reductions of scientific theories or properties, but only to show their metaphysical possibility." (Kim 1995, 580) Wieso aber ist unter Supervenienzaspekten eine Reduktion möglich? Handelt es sich bei dem Abhängigkeitsverhältnis zwischen mentalen und physischen Eigenschaften nicht in erster Linie um eine bloße Korrelation? Tatsächlich räumt Kim ein, dass die Theorie nicht konkret angibt, welche Art von Abhängigkeit für Körper-Geist-Supervenienzen vorliegt. Insofern kann psychophysische Supervenienz auch kaum als eigenständige Theorie des Geistes bezeichnet werden kann, vielmehr ist sie eine Bestandsaufnahme des Körper-Geist-Problems. Für eine darüber hinausgehende Lösung muss das Abhängigkeitsverhältnis jedoch noch weiter spezifiziert werden (vgl. ebd., 582).

Kim muss seine theoretische Position also über den Supervenienzbegriff hinaus erweitern bzw. diesen mit theoretischem Inhalt unterfüttern. Dies geschieht im Rahmen der physikalistischen Konzeption mit dem Modell einer Schichtenwelt (vgl. Kap. 3.1.3). Innerhalb der Hierarchie dieses Modells nehmen die mentalen Eigenschaften einen Platz auf der Ebene der makrophysischen Entitäten ein und die physischen Eigenschaften einen Platz auf der Ebene der mikrophysischen Entitäten. Kim ergänzt das Prinzip der Supervenienz somit durch das Prinzip der physischen Realisation. Beide Prinzipien beschreiben die Abhängigkeitsrelation eigenständig, sie schließen sich jedoch nicht aus und können sogar miteinander verbunden werden: „[T]he claim that mental states are physically realized arguably entails the claim that they are physically supervenient." (ebd., 195)

[52] Kim verweist in diesem Zusammenhang auch auf den Emergentismus, der als Vorläufer der Supervenienztheorie betrachtet werden kann. Dem Emergentismus zufolge werden Eigenschaften höherer Level, wie Bewusstsein, von Eigenschaften niedriger Level erzeugt, jene sind folglich von diesen determiniert. Die emergierten Eigenschaften werden als genuin neu und nicht-reduzierbar betrachtet (vgl. Kim 1995).




Grafik 6: Die dargestellte Supervenienz zwischen M und P sowie zwischen M* und P* entspricht der vertikalen Abhängigkeit, während die Relation zwischen P und P* die horizontale Kausalität darstellt (vgl. Kim 2003, 159 und 166; vgl. auch Block 1990). Die offene Frage für eine „superveniente Kausalität" ist, ob Eigenschaft M auch M* verursacht oder lediglich Eigenschaft P Ursache von P* ist. Dabei muss berücksichtigt werden, dass nicht die Eigenschaften M oder P, sondern im Grunde ihre Instanziierung, also das Ereignis, auf der horizontalen Ebene Ursache für die nachfolgende Instanziierung ist.

Kim schließt an das Konzept von LePore und Loewer an, wonach die Eigenschaft P von E die Eigenschaft F von E realisiert, wenn E zugleich P ist und E zugleich F ist und zwischen P und F eine starke Verbindung besteht. Diese Verbindung wird als notwendig und explanatorisch aufgefasst und soll zudem stärker sein als die Annahme, wenn P dann M sei physisch (vgl. LePore/Loewer 1989, 179; Kap. 4.1; Kap. 4.2). Kim akzeptiert dieses Modell der Realisierung allerdings nur unter Vorbehalt. Es ist ihm wichtig, darauf hinzuweisen, dass P nur dann eine Erklärung für M sein kann, wenn damit eine metaphysische Relation beschrieben wird.[53] Außerdem ergänzt er die Annahme, dass wenn P dann M mit nomologischer Notwendigkeit gilt, durch die zusätzlichen Annahme, dass M von einem (in)finitem Set physischer Eigenschaften $P_1$, $P_2$, usw. realisiert wird, und diese in Form einer disjunkten Eigenschaft $P_1 \vee P_2 \vee ... \vee P_n$ eine nomologische Koextension von M bilden (vgl. Kim 1999b, 197). An dieser Stelle wird bereits deutlich, dass Kim die Position der Nichtreduktivisten in seinem speziellen Sinn auslegt. Schließlich will er zeigen, dass diese Form der Realisierung eine Reduktion der mentalen auf die physischen Eigenschaften von Ereignissen oder Zuständen ermöglicht.

### 2.4.3 Kausaler Ausschluss mentaler Eigenschaften

Kritik an der Supervenienztheorie kommt erstaunlicherweise von ihrem eigenen Begründer. Kim selbst ist es, der die Position als solche zwar nicht verwirft, aber gegen einen zentralen Bestand-

[53] Für Kim muss hinter einer Erklärung von M durch P eine metaphysische Relation stehen, damit diese zutrifft. Insofern kann P nur als Mikroeigenschaft in einem System s einen kausalen Mechanismus für die Realisierung von M liefern. P ist dabei als Teil eines Netzwerkes von nomologisch verbundenen Mikroeigenschaften zu betrachten, welche allesamt mikro-kausale Mechanismen für die nomologische Verbindung eines Systems mentaler Eigenschaften liefern, in dem wiederum M ein Element ist. Eine solche Erklärungsstrategie ist jedoch von der eigentlichen Realisierungsrelation zu unterscheiden, welche auch metaphysisch standhalten muss (vgl. Kim 1999b, 197).

teil, die Nichtreduzierbarkeit, argumentiert und somit die ursprüngliche Zielsetzung dieser Position aufhebt. Kim hat die Supervenienztheorie letztlich in den Dienst einer reduktionistischen Argumentation gestellt. „Kim's conceptualized functionalism turns out to be a somewhat surprising brand of reductionism – a reductionism with some eliminativist cut-outs and, possibly, some dualist leftovers." (David 1997, 133) Anhand der Frage der mentalen Kausalität möchte Kim exemplarisch das Scheitern des nichtreduktiven Physikalismus zeigen. Seiner Ansicht nach ist die Existenz von Eigenschaften an ihre kausale Kraft gebunden. „To be real is to have causal powers." (Kim 1999b, 202; vgl. Kap. 4.4.2)[54] Unterschiedliche Eigenschaften oder Arten, unter die sie fallen, grenzen sich somit auch über ihre kausale Wirksamkeit ab:

> Kinds in science are individuated on the basis of causal powers; that is, objects and events fall under a kind, or share in a property, insofar as they have similar causal powers. (ders. 1993, 326)

Diese grundlegende Annahme wird von den meisten nicht-reduktiven Physikalisten geteilt. Mentale Eigenschaften sollen über ihre physischen Basen hinaus eigene kausale Kraft besitzen. Denn nur insofern ist der Anspruch einer autonomen psychologischen Theorie begründet, da sie eine eigene kausale Erklärung für mentale Phänomene leisten kann. Eine mögliche kausale Wirkungslosigkeit mentaler Eigenschaften hätte auf der anderen Seite weitreichende Konsequenzen für deren Existenz (vgl. Kap. 3.6.1).

In seiner Argumentation zeigt Kim, dass mentale Kausalität nicht als „same-level" sondern nur als „cross-level"-Kausalität möglich ist. Das bedeutet, dass ein mentales Ereignis M nicht direkt ein anderes mentales Ereignis M* kausal beeinflussen kann. Vielmehr ist ein Umweg über dessen physische Basis P* notwendig, von der das mentale Ereignis M* realisiert wird. Demnach beinhaltet mental-mentale Kausalität mental-physische Kausalität. Letztere wird als „downward causation" bezeichnet, da sie eine Kausalität von einem oberen zu einem unteren Level beschreibt. Sie ist jedoch ebenfalls äußerst problematisch, wie gezeigt werden wird. Doch warum führt beispielsweise der Wunsch, ein klassisches Konzert zu besuchen, nicht direkt zu der Absicht, sich bei einem Konzertveranstalter nach Karten zu erkundigen? Die Antwort ist ernüchternd einfach: aufgrund der hinreichenden Realisierung durch eine physische Basis. Die Annahme, dass eine mentale Eigenschaft M gestützt auf geeignete Gesetze kausal wirksam hinsichtlich der mentalen Eigenschaft M* ist, läuft der Annahme entgegen, dass jede mentale Eigenschaft, also auch M*, durch eine physische Eigenschaft, in diesem Fall P*, hinreichend rea-

[54] Kim bezeichnet diesen Grundsatz als „Alexander's dictum" in Anlehnung an den Emergentisten Samuel Alexander. Ihm zufolge nimmt der Epiphänomenalismus an, dass etwas, das ohne Sinn und Zweck in der Natur existiere, in seiner Arbeit von seinen untergeordneten Einheiten abhinge und aufgrund dessen überflüssig sei und abgeschafft werden könne (vgl. Alexander 1927, 8). Alexander hatte das Diktum formuliert, um sich vom Epiphänomenalismus abzugrenzen.

lisiert wird. Denn ohne die Eigenschaft P* wäre M* nicht instanziiert worden. So ist es für Kim außerdem nicht einsehbar, dass M und P* zusammen kausal verantwortlich für M* sind oder sogar zwei getrennte Ursachen bilden und M* überdeterminieren. Unter der Annahme, dass P* hinreichende Bedingung für M* ist, bleibt somit für die Eigenschaft M nur die Möglichkeit, M* zu instanziieren, indem sie P* kausal beeinflusst (vgl. Kim 1999b, 204f.; Kap. 3.4). Aus dieser „downward causation" ergibt sich aber ein weiteres Problem für die kausale Wirksamkeit von M. Schließlich wird auch M von einer physischen Basis P realisiert. Warum kann dementsprechend nicht einfach P kausale Ursache für P* sein, also eine einfache Kausalrelation auf demselben Level, die zudem im Sinne der Supervenienz Basis für M's Kausalität hinsichtlich M* ist? Das Schicksal der Eigenschaft M wäre somit, ein Epiphänomen zu sein. Dafür spräche als Grund vor allem das Prinzip des kausal-explanatorischen Ausschlusses: Wenn das Erscheinen von P* hinreichend durch die kausale Wirkung von P erklärt wird, ist auch unter Gesichtspunkten der Einfachheit fraglich, welche kausale Arbeit noch für M zu leisten ist. Unter neurobiologischen Gesichtspunkten müsste untersucht werden, wie die kausale Kette oder das Netzwerk zwischen Neuronen P und P* von einer mentalen Eigenschaft verändert werden soll (vgl. Kim 1999b, 205).

Vor dem Hintergrund dieser Überlegungen kommt Kim schließlich zu dem Schluss, dass innerhalb des Modells einer Schichten-Welt nur kausale Interaktion innerhalb des physischen Levels geschehen kann, während jegliche Kausalrelation auf den oberen Leveln von den unteren abgeleitet werden. In dieser Richtung wird auch die Supervenienz gedeutet, wonach kausale Beziehungen von Makro-Ereignissen als auf mikro-kausalen Prozessen supervenierend bestimmt werden. Mentale Ereignisse sind Ursachen eines anderen Ereignisses kraft der Tatsache, dass sie auf einem passenden physischen Ereignis supervenieren. Mentale Eigenschaften verdanken ihre kausale Kraft den physischen Eigenschaften, auf denen sie supervenieren (vgl. ders. 1989, 282f.). Mentale Kausalität ist für Kim also nur in einem übertragenen Sinn möglich, er formuliert das Prinzip der kausalen Erbschaft:

> If *M* is instantiated on a given occasion by being realized by *P*, then the causal powers of *this instance of M* are identical with (perhaps, a subset of) the causal powers of *P*. (ders. 1999b, 208)

Dieses Prinzip trägt der Annahme der kausalen Geschlossenheit der physischen Welt Rechnung (vgl. Kap. 3.3). Damit das Prinzip Gültigkeit erlangt, werden die mentalen Eigenschaften M mit ihren physischen Basen identifiziert. Dem entgegen steht das Prinzip der physischen Realisierung, genauer gesagt, dass Prinzip der Multirealisierbarkeit (vgl. Kap. 4.2.1). Allerdings interpretiert Kim Multirealisierbarkeit als Realisierung durch eine offene, disjunkte Gruppe physischer Eigenschaften $P_1$, $P_2$, …, $P_n$. Er verknüpft Multirealisierbarkeit deshalb mit dem Prinzip der Individua-

tion von Arten. Danach sind Instanziierungen M – durch dieselben physischen Eigenschaften realisiert – dieselben Eigenschaftsarten, hingegen sind Instanziierungen M – durch unterschiedliche physische Eigenschaften realisiert – auch unterschiedliche Eigenschaftsarten (vgl. ders. 1993, 323ff.). Aus diesem Blickwinkel betrachtet, verträgt sich Multirealisierbarkeit sehr wohl mit multipler lokaler Reduktion. Auch wenn die Strukturen eines Systems sehr fein individuiert sind und es unmöglich erscheint, jegliche neuronale Basen der mentalen Zustände praktisch zu bestimmen, muss deren metaphysische Reduzierbarkeit jedoch akzeptiert werden (vgl. ebd., 328f.). Kim gesteht der Psychologie in diesem Rahmen zwar eine Daseinsberechtigung zu, allerdings strikt basierend auf biologischen und physischen Prozessen. Kim gesteht einem Physikalisten lediglich zwei Optionen zu: Eliminativismus und Reduktionismus. Da er die Existenz von speziesbezogenen mentalen Eigenschaften einräumt, grenzt er selbst sich von eliminativistischen Überlegungen ab und vertritt konsequent eine reduktionistische Position (vgl. ebd., 267).

### 2.4.4 Kims Version der Supervenienz in der Kritik

Eine Kritik an Kims theoretischem Konzept ist nicht umfassend zu leisten, sondern muss eher punktuell erfolgen. Dies liegt vor allem daran, dass er – im Gegensatz beispielsweise zu Davidson – im Grunde kein Gesamtkonzept aufstellt, sondern vielmehr gegen die Nicht-Reduktivisten argumentiert. „So apart from attempting to articulate and defend his own position, as a disputationalist he recently has been directing his guns mainly against nonreductive materialists." (Horgan 1997, 167) Immer wieder betont er, dass Verbreitung und Popularität im Gegensatz zur Unvollkommenheit und konzeptionellen Unsicherheit des nicht-reduktiven Materialismus stünden. Zudem wendet sich Kims Argumentation aber auch gegen Typenidentität oder Eliminativismus (vgl. Kap. 2.1; Kap. 3.6.1). Kim hat sich also gewissermaßen zwischen alle Stühle gesetzt, ein Zustand, der sich in seiner Theorie widerspiegelt. Gerade in dieser Unentschiedenheit sind Angriffspunkte gegen Kim zu finden.

Kim selbst räumt ein, dass die Supervenienzthese im Vergleich zu den klassischen Alternativen innerhalb der Leib-Seele-Diskussion keine eigenständige Erklärung für die Abhängigkeit von mentalen und physischen Eigenschaften liefert. Supervenienz beschreibt keine spezielle Art von Abhängigkeit, sondern verschiedene Fälle, wie z. B. kausale Abhängigkeit, Bedeutungsabhängigkeit, mereologische Abhängigkeit u. a. (vgl. Kim 1995, 582; McLaughlin 1997, 210). Somit kann nicht abschließend begründet werden, ob mentale und physische Eigenschaften lediglich kovariieren, oder erstere etwa von letzteren kausal abhängig sind. Für Kim ist die starke Supervenienz das naheliegende Modell, welches der kausalen Geschlossenheit der physischen Welt Rechnung trägt. Bereits hier zeigt sich die Spannung in seiner Argumentation. Denn auf der einen

Seite hängen die kausalen Kräfte der mentalen Eigenschaften sozusagen „Huckepack" von denen der subvenienten physischen Eigenschaften ab. Auf der anderen Seite sollen die mentalen Eigenschaften nicht völlig eliminiert werden und eine eigene Erklärungskraft behalten. So zumindest die ursprüngliche Haltung Kims. Horgan beschreibt diese Position als kausalen Kompatibilismus, wonach mentale Kausalität mit nicht-physischen Eigenschaften mit physischer Kausalität trotz der kausalen Geschlossenheit des physischen Bereichs koexistieren kann (vgl. Horgan 1997, 166).

Kims Kompatibilismus ist nun insofern speziell, als er für die mentale Verursachung eine „downward causation", also die Verursachung der oberen zu den unteren Leveln setzt. Diese ist aber abhängig von den zugrundeliegenden physischen Kausalverhältnissen. In dieser Richtung läuft die Position auf einen Eigenschaftsdualismus hinaus, da die Identität von mentalen und physischen Eigenschaften abgelehnt wird. Da sich Kim aber auch für eine Theorienreduktion stark macht, im Sinne einer Tokenreduktion, die die psychologische Theorie auf die physikalische (möglicherweise korrigierte) Version zurückführt, argumentiert Kim wiederum gegen die Existenz und kausale Kraft mentaler Eigenschaften (vgl. Kap. 3.3.1; Kap. 4.4). Nach „Alexander's dictum" können mentale Eigenschaften nur real sein, wenn sie eine kausale Erklärung liefern. Andererseits soll die Psychologie, wenn es mentale Eigenschaften gibt, auf naturalisierte Wissenschaft reduzierbar sein (vgl. Kim 1993a, 325). Letztendlich entzieht Kim den mentalen Eigenschaften den Boden, den sie für eine wissenschaftliche Psychologie liefern sollen.

Kims Annahme, dass lediglich homogene physische Arten voraussagefähig sind, ist letztendlich eine Setzung. Regelmäßigkeiten auf der Makroebene müssen nicht ausschließlich auf Regelmäßigkeiten auf der Mikroebene (lokal) zurückgeführt werden (vgl. Kap. 4.1; Kap. 4.6). Angeblich sind sämtliche Zustände und Relationen in der Welt auf physische Beschreibung zurückzuführen, unvorstellbar kleine Moleküle auf den Mikroleveln bestimmen irgendwie das Geschehen auf den Makroleveln. Dieses „irgendwie", polemisiert Fodor, wirkt einfach mysteriös und stört Kim (vgl. Fodor 1997, 161). Dennoch folgt aus der Dominanz des physikalischen Bereichs nicht zwangsläufig die Belanglosigkeit der Spezialwissenschaften, wie der Psychologie, Biologie oder Geologie. Ihre Autonomie und Erklärungskraft wird dadurch nicht zwangsläufig berührt. Fodor bleibt dabei: „[…] autonomy implies 'real' (viz projectible) patterns without reduction" (ebd., 162; Kap. 4.2.1). Letztendlich betreibt Kim über seine reduktionistische Position hinaus einen Eliminativismus. Unsere Vorstellungen von Schmerz, Glaube, Wunsch etc. sind lediglich Vorstellungen, Begriffe (englisch: concepts) und repräsentieren keine tatsächlichen Zustände oder Eigenschaften. Damit weicht er von gängigen funktionalistischen Positionen ab. Block weist darauf hin, wie radikal Kims Position in diesem Punkt ist.[55] Gemeinhin wird angenommen, dass es zusätzlich zu

[55] Nach Marian David unterscheidet sich Kim vor allem in drei Punkten von so genannten „Hardcore"-Funktionalisten. Er erkennt erstens nicht die Existenz jener mentalen Zustände an, welche die Funktionalisten

einem Konzept von Schmerz auch die Eigenschaft Schmerz gibt, welche multirealisierbar von verschiedenen Lebewesen oder Kreaturen geteilt wird. Kim jedoch verneint eine gemeinsame Eigenschaft Schmerz für die verschiedenen Organismen, sie fallen lediglich unter dasselbe Konzept (vgl. Block 1997, 120). Anstelle einer Identität von mentalen und funktionalen Zuständen beschreibt er lediglich die Identität von Konzepten mentaler Zustände und funktionalen Konzepten. Zwar gibt es mentale Zustände, aber eben nur arten-spezifisch. Sie sind als solche lokal identisch mit den Zuständen erster Ordnung, also nicht multirealisierbar. Multirealisierbar ist lediglich das unbeschränkte Konzept von Schmerz. Als Konsequenz wird Schmerz als mentaler Zustand dadurch aber letztlich eliminiert (vgl. David 1997, 137).[56]

Ein weiterer zentraler Kritikpunkt ist Kims Argument des kausalen Ausschlusses (vgl. Kap. 3.4.3). Seiner Meinung nach ist es nicht einsichtig, warum eine mentale Eigenschaft $M_1$, die von der physikalischen Eigenschaft $P_1$ realisiert wird (oder auf $P_1$ supervenient), kausale Ursache für $P_2$ sein sollte. Voraussetzung ist natürlich, dass $P_2$ bereits hinreichend durch $P_1$ verursacht wird, $M_1$ ist somit anscheinend epiphänomenal und kausal vollkommen überflüssig (vgl. Kim 1993a, 354f.). Das Problem der Überdetermination besteht also unter der Annahme, dass Makroleveleigenschaften voraussagefähig und autonom sind. Denn dann würden auch sie kausale Relationen mit Gesetzen beschreiben und erklären. Es gäbe mehr als nur ein Gesetz, das kausale Relationen beschreiben würde (vgl. Fodor 1997, 163). Warum also diese Annahme nicht einfach zurückweisen, wie Kim es zunächst auch macht? Warum gibt es das Problem des kausalen Ausschlusses? Schließlich bezieht er sich auf eine kausal geschlossene physische Welt, und in der sind die einzigen kausalen Eigenschaften physische Eigenschaften. Problematisch ist für Kims Überlegungen, dass er nach wie vor an einem Eigenschaftsdualismus festhält, da er das Prinzip der Multirealisierbarkeit akzeptiert und Typenidentität ablehnt. „Unless disjunctive types are countenanced […] the standard type physicalism is precluded; that is, we cannot identify mental properties or kinds with physical ones […]." (Kim 1993a, 364) Darüber hinaus sind mentale Eigenschaften kausale Eigenschaften; zwar eingeschränkt und auf der Basis ihrer subvenienten Eigenschaften, aber „irgendwie" eigenständig:

---

mit funktionalen Zuständen identifizieren. Er identifiziert zweitens die mentalen Zustände, die er anerkennt, mit physischen Zuständen erster Ordnung und lässt deshalb keine Multirealisierbarkeit zu. Sein Funktionalismus ist drittens grundsätzlich ein physikalischer Reduktionismus (vgl. David 1997, 135f.).

[56] Daraus ergibt sich aber ein weiteres Problem für Kims Funktionalismus. Unabhängig davon, was diese Konzepte genau sind, handelt es sich sicherlich um mentale Phänomene, wenn Konzepte erfasst oder verwendet werden. Diese Phänomene wiederum können auch nur von verschiedenen physischen Zuständen realisierte Konzepte sein, das „Konzept des Erfassens des Konzeptes von Schmerz" – ein Regress von immer mehr miteinbezogenen Konzepten würde sich ergeben. Dennoch muss Kim diese Konstruktion übernehmen, was allerdings zur Folge hat, dass zwangsläufig einige Konzepte übrig bleiben, die sich Kims funktionalistischer Version entziehen und nicht behandelt werden können. Die andere Möglichkeit wäre das Haben oder Erfassen von Konzepten nicht als Konzepte, sondern als genuine mentale Zustände zu behandeln. Allerdings hätte man in diesem Fall ein Erklärungsproblem, welches Argument die Unterscheidung von Schmerz als Konzept und dem Erfassen von Schmerz als Zustand rechtfertigt (vgl. David 1997, 140).

> The answer is that a mental property is causally efficacious as long as, and to the extent that, each of its possible realizers is causally efficacious; and a particular instance of a mental property has exactly the causal efficacy of its realizer on that occasion. This confirms the causal efficacy of mental properties. (ebd., 363)

Kim schafft es nicht, die gegensätzlichen Prämissen einer physisch kausal geschlossenen Welt und der autonomen kausal wirksamen mentalen Eigenschaften miteinander zu versöhnen. Er möchte aber beides beibehalten, anstatt sich konsequent der identitätstheoretischen Position zu verschreiben. Er versucht vielmehr, seinen speziellen Eigenschaftsdualismus mit dem Argument des kausalen Ausschluss in Einklang zu bringen. Es ist gezeigt worden, dass die meisten Nichtreduktivisten das Argument einfach nicht akzeptieren, da nicht einsichtig ist, warum nicht zwei gleichzeitige Ereignisse kausal verantwortlich für ein drittes sein können. Kims eigene Begründung ist an dieser Stelle sehr schwach: „For we would be allowing two distinct sufficient causes, simultaneous with each other, of a single event. This makes the situation look like one of overdetermination, which is absurd." (ebd., 354) Die Möglichkeit einer Mehrfachdetermination als absurd zu disqualifizieren, kann nicht als starkes Argument beurteilt werden. Es wäre also aus Kims Perspektive konsequent, die Prämisse aufzugeben, dass mentale Eigenschaften nicht identisch mit physischen sind. Damit würde aber seine gesamte Eigenschaftsontologie – sein Entwurf der Supervenienzbeziehung bzw. Ansicht einer deterministischen Relation zwischen mentalen und physischen Eigenschaften – aufgehoben. Eine monistische Version des reduktiven Materialismus wäre die Position, die Kim zu vertreten hätte (vgl. Horgan 1997, 175ff.).

Auch das Prinzip der kausalen Erbschaft steht in der Kritik, wonach im Rahmen einer funktionalen Realisierungsrelation die Eigenschaften zweiter Ordnung die kausale Kraft ihrer Realisier erben (vgl. Kim 1997, 295; Menzies 2003, 221f.). Beispielsweise ist für Schröder diese Metapher der kausalen Erbschaft irreführend, da es höchstens einen Erblasser, jedoch keinen echten Erben geben kann. Dies liegt daran, dass funktionale Eigenschaften als Eigenschaften zweiter Ordnung gedacht werden und sich damit auf einer anderen Ebene bewegen als ihre Realisierer. Schließlich erbt das Amt des Polizisten auch nicht das gesamte Schicksal der Person Jones, lediglich weil Jones die Rolle des Polizisten ausfüllt (vgl. Schröder 2003, 175f.). Ebenso wenig kann eine mentale Eigenschaft als Eigenschaft zweiter Ordnung, die die funktionale Rolle der Eigenschaft erster Ordnung ausfüllt, deren kausale Kraft erben. Dies liegt einfach daran, dass eine funktionale Rolle bereits kausale Kraft ist. Damit würde die Eigenschaft zweiter Ordnung sich im Grunde selbst beerben, da sie nicht die kausale Kraft hat, sondern selbst ist. Schröder betont den Unterschied bzw. die Ambiguität zwischen der kausalen Kraft einer Eigenschaftsinstanziierung und einer Eigenschaft selbst. Eine Eigenschaftsinstanziierung kann mehrere Eigenschaften und deren kausale Kraft beinhalten. Beispielsweise ist die Eigenschaft „eine Fensterscheibe um zwölf Uhr treffen" nicht kausal verantwortlich für das Zerbrechen der Scheibe,

aber eine der kausalen Eigenschaften der Instanziierung der Eigenschaft „eine Fensterscheibe um zwölf Uhr treffen“ verfügt über die kausale Kraft, weshalb die Fensterscheibe zerbricht. Entscheidend ist darüber hinaus, dass eine Instanziierung M, die $P_1$ ist, zwar die kausale Kraft von $P_1$ erhält bzw. eine Instanziierung, die $P_2$ ist, zwar die kausale Kraft von $P_2$ erhält. Damit ist aber noch keine Aussage über die kausale Kraft der Eigenschaft M gemacht worden, und diese kann nach Schröder nicht ihre kausale Kraft von $P_1$ oder $P_2$ beziehen, da sie nichts anderes als diese kausale Kraft selbst ist. Bei der Supervenienzrelation ist die Lage jedoch eine andere, da mentale und physische Eigenschaften verschieden und nicht identisch sind. Hierbei müsste es aus Schröders Sichtweise noch fraglicher sein, inwiefern mentale Eigenschaften, die lediglich von physischen Eigenschaften abhängen, aber nicht von diesen realisiert werden, deren kausale Kraft erben können.

***Quintessenz Supervenienztheorie***

- *Kovarianz, Abhängigkeit und Nicht-Reduzierbarkeit zwischen supervenienten Eigenschaften und subvenienten Basen*
- *asymmetrische Relation: Makroeigenschaften abhängig von Mikroeigenschaften*
- *Eigenständigkeit mentaler Eigenschaften in Bedrängnis:*
    - *physische Basen notwendig für Realisierung*
    - *mögliche Reduktion im Rahmen des Schichtenwelt-Modells*
- *Gefahr Epiphänomenalismus: kausaler Ausschluss aufgrund Ablehnung von Mehrfachdetermination*
    - *Alexander's dictum: Abhängigkeit der Existenz von kausaler Kraft*
    - *kausale Kraft mentaler Eigenschaften beruht auf kausaler Kraft physischer Eigenschaften*

**Kapitel 3**

**Voraussetzungen für mentale Kausalität**

Mentale Kausalität wird in der analytischen Philosophie vor dem Hintergrund bestimmter ontologischer Voraussetzungen diskutiert. So ergibt sich die Problematik der mentalen Kausalität selbst erst aufgrund eines Spannungsverhältnisses zwischen diesen Voraussetzungen und dem Anspruch, dass beispielsweise mentale Eigenschaften kausal wirksam sind. Für Vertreter der Identitätstheorie bzw. einer Typenidentität besteht das Problem der mentalen Kausalität als solches zunächst nicht (vgl. Crane 1995, 211; vgl. auch Armstrong 1968; Feigel 1958, 1967; Smart 1959, 1962). Da mentale und physische Ereignisse bzw. Zustände identisch sind, gibt es hier kein „Überbrückungsproblem" zwischen zwei ontologisch verschiedenen Bereichen.

Auch für die Tokenidentität muss zumindest nach Davidsons Auffassung aufgrund der von ihm beschriebenen Ereignisontologie streng genommen kein Problem mit mentaler Kausalität bestehen (vgl. Crane 1995; Davidson 1999). Kausalität ist eine Relation zwischen Einzelereignissen, unabhängig von ihren Eigenschaften und unabhängig von einer jeweiligen Terminologie. Da die Ereignisse sowohl unter eine psychologische als auch physische Beschreibung fallen, ist im Rahmen seiner ontologischen Bedingungen eine mentale Verursachung grundsätzlich möglich:

> [I]t is events that have causes and effects. Given this extensionalist view of causal relations, it makes no literal sense […] to speak of an event causing something as mental, or by virtue of its mental properties, or as described in one way or another. (Davidson 1999, 13)

Davidson weist somit das Prinzip zurück, wonach Ursachen aufgrund ihrer Eigenschaften wirksam sind. Ebenso weist er die Vorstellung zurück, dass eine Ursache in einer bestimmten Weise „wirksamer" beschrieben werden kann. Fragen nach der Wirksamkeit der Eigenschaften einer Ursache würde er als Fragen nach dem Informationsgehalt einer kausalen Erklärung verstehen. Er macht somit keine Unterscheidung zwischen der Kausalität verschiedener Eigenschaften, sondern zwischen alternativen kausalen Erklärungen.

Die Frage der Erklärung berührt jedoch nicht unbedingt den Kern mentaler Kausalität, da mentale Kausalität als eine metaphysische Relation betrachtet werden muss. „Mental causation – the causation of what we do by what we think – becomes a problem against a particular metaphysical background." (Baker 1999, 75; Audi 1999). Kausale Beziehungen werden als metaphysische und nicht als epistemische Tatsachen aufgefasst. Mentale Kausalität wird vor allem durch die bestehenden metaphysischen Vorraussetzungen problematisch. Diese Vorraussetzungen können beschrieben werden in der Form, dass erstens die Realität auf physischen Ereignissen beruht, zweitens Kausalität einen objektiv realen Gehalt hat, unabhängig von psychologischen

Erklärungen und drittens Verhalten als Verbindung innerer Kausalität zu physischen Ereignissen verstanden wird (vgl. Baker 1999, 75f.).[57] Das Problem mentaler Kausalität entsteht aus der kausalen Geschlossenheit der physischen Welt und einer starken Materialismusthese, wonach jede Eigenschaft durch eine physische Eigenschaft instanziiert wird. Damit wäre die Frage nach mentaler Kausalität allein metaphysisch und unabhängig von Diskussionen um Intentionalität, Existenzfragen nach Glaubensannahmen und Wünschen, psychophysischen Gesetzen oder einer stimmigen Verhaltenstheorie zu lösen (vgl. ebd., 86f.). Unter den bestehenden metaphysischen Annahmen scheint eine Kausalität mentaler Eigenschaften jedoch nicht möglich zu sein.

Das Problem der mentalen Kausalität besteht vor allem für jene Physikalisten, die zwischen mentalen und physischen Eigenschaften unterscheiden, wobei erstere nicht auf letztere reduziert werden sollen. „The standard view in the current literature is that it is ‚non-reductive' physicalism which has this problem, and that it has it precisely because it is ‚non-reductive'." (Crane 1995, 213) Mentale Kausalität muss somit vor allem als ein Problem für Nicht-Reduktivisten unter funktionalistischen bzw. supervenienztheoretischen Aspekten diskutiert werden.[58]

In der Folge werden die Voraussetzungen, die den verschiedenen Positionen teilweise gemein sind, im Detail dargestellt. Auf diese Weise wird deutlich, dass die daraus resultierenden Argumente als solche zwar anerkannt, aber nicht unbedingt akzeptiert sind. Einerseits ist es für eine nicht-reduktionistische Position schwer, den Gegensatz zwischen der Voraussetzung der Geschlossenheit der Physik mit dem Anspruch einer mentalen Kausalität zu versöhnen. Dies scheint grundsätzlich nur möglich, wenn im Gegenzug eine der üblichen Voraussetzungen zurückgewiesen wird. Andererseits kann einer starken materialistischen Position ein Dogmatismus vorgeworfen werden, der dem eigentlichen Phänomen nicht gerecht wird. So werden die ontologischen Voraussetzungen nur auf Kosten der mentalen Phänomene bewahrt, die z. T. ausgeklammert, eliminiert oder ebenfalls nur unzureichend erklärt werden. Die Frage der mentalen Kausalität wird auf der Grundlage folgender Voraussetzungen diskutiert:

(1) dem ontologischen Grundmodell,
(2) der Unterscheidung zwischen mental und physisch,
(3) der kausalen Geschlossenheit der Physik,

[57] Baker will den Vorrang der Metaphysik zugunsten eines epistemischen Erklärungskonzeptes der Kausalität verwerfen. Ob eine solche Verschiebung der Prioritäten eine befriedigende Lösung des formulierten Problems bietet, bleibt allerdings offen (vgl. Baker 1999, 93f.).

[58] Crane spricht anstelle einer nicht-reduktivistischen Position von einem orthodoxen Physikalismus. Diesen Terminus bevorzugt er, da z. B. auch Davidson eine nicht-reduktivistische Position vertritt, aber, wie oben geschildert, aufgrund seiner Ereignisontologie kein Problem mit mentaler Kausalität hat. Orthodoxer Physikalismus behauptet die schwächere Relation einer „Konstitution" zwischen dem Mentalen und dem Physischen als die Identitätstheorien. Deshalb betrifft das Problem der mentalen Kausalität auch nur jene physikalistische Position (vgl. Crane 1995, 211).

(4) dem Ausschlussargument sowie

(5) der strikten Gesetze.

Bei den genannten Voraussetzungen handelt es sich um teilweise miteinander verknüpfte, aber auch miteinander konkurrierende Annahmen. So werden z. B. nach wie vor verschiedene Eigenschafts- bzw. Ereignisbegriffe verwendet (vgl. Kap. 3.1). In der gegenwärtigen Diskussion stehen vor allem die Voraussetzungen der „kausalen Geschlossenheit der Physik" sowie das „Ausschlussargument" der kausalen Wirksamkeit mentaler Eigenschaften entgegen (vgl. z. B. Van Gulick 1999).[59] Hierin scheint ein wesentliches Problem zu liegen. Unter der Annahme, dass bestimmte mentale Ereignisse bzw. Eigenschaften nicht identisch mit physischen Ereignissen bzw. Eigenschaften sind, ergibt sich in Verbindung mit diesen Voraussetzungen folgendes Dilemma: Wenn mentale Eigenschaften verantwortlich für physische Wirkungen sein sollen, und diese nicht mehrfachverursacht sein dürfen, ist eine mentale Kausalität mit der Geschlossenheit der Physik nicht vereinbar. Nicht-reduktiver Physikalismus scheint somit entweder inkonsistent zu sein oder endet in einem Epiphänomenalismus, gegen den ursprünglich argumentiert wird.

### 3.1 Das ontologische Grundmodell

Die Frage der mentalen Kausalität und die damit verbundenen Lösungsansätze hängen von der gewählten ontologischen Position ab. Substanzdualisten beispielsweise müssen erklären, wie der kausale Nexus zwischen so unterschiedlichen Substanzen wie dem Mentalen und dem Physischen möglich ist (vgl. Eccles 1994; Meixner 2004; Popper/Eccles 1977). Für die Vertreter der Identitätstheorie dagegen stellt die Frage der mentalen Kausalität zunächst kein Problem dar, da aufgrund der ontologischen Identität mentaler und physischer Eigenschaften, Zustände oder Ereignisse keine Energieübertragung zwischen verschiedenen Substanzen bzw. Entitäten geleistet werden muss (vgl. Kap. 2.1; Kap. 2.3; Kap. 3).

Die Schwierigkeiten für eine mentale Kausalität hängen eben auch davon ab, welche ontologischen „Grundeinheiten" vorausgesetzt werden. Die mentale Kausalität zwischen verschiedenen Substanzen, die Descartes beschäftigte (vgl. Descartes 1977; Kap. 1.1), ist eine andere, als die mentale Kausalität zwischen verschiedenen Ereignissen oder Eigenschaften. So wird in der aktuellen Debatte der analytischen Philosophie hauptsächlich zwischen Zuständen, Ereignissen und

[59] Crane nennt als Annahmen 1. The Anomalism of the Mental, 2. The Causal Closure of the Physical, 3. The Inefficacy of the Functional und 4. The Inefficacy of the Semantic (vgl. Crane 1992, 186). Robert Audi nennt vier Quellen des Zweifels bezüglich kausaler Kräfte des Mentalen: 1. Gründe sind keine Ursachen, 2. abstrakte Inhalte können keine kausale Rolle einnehmen, 3. im Falle der Supervenienz erklären allein physische Ereignisse die Wirkungen, 4. ohne Gesetze ist eine genuin kausale Rolle des Mentalen nicht möglich (vgl. Audi 1999, 53ff.).

Eigenschaften unterschieden. Wobei diese Trennung bei den verschiedenen Positionen unterschiedlich stark vorgenommen wird – Zustände und Ereignisse werden hinsichtlich ihrer kausalen Funktion häufig ähnlich beschrieben. Allerdings gibt es auch hier Unterschiede.

Beispielsweise unterscheidet Audi (1999) zwischen mentalen Ereignissen und mentalen Zuständen. Die „Formation" bzw. Anordnung von Glaubensannahmen oder Wünschen bezeichnet er als Ereignisse, die propositionalen Einstellungen als dispositionale mentale Zustände.[60] Es sind die geistigen Zustände, insbesondere Wünsche, Glaubensannahmen, Intentionen, und nicht abstrakte Gründe, die Ursache für Handlung sein können (vgl. Audi 1999, 54f.). Burge dagegen unterscheidet weniger trennscharf zwischen Ereignissen und Zuständen. Ihm zufolge erscheinen mentale Ereignisse ebenso wie mentale Zustände, wenn die zugrunde liegenden physischen Zustände oder Ereignisse erscheinen. Die Frage der mentalen Kausalität stellt sich auch für ihn vor allem hinsichtlich der Eigenschaften der Ereignisse oder Zustände: „The worry about epiphenomenalism arises by considering that some properties of events (or states) are relevant to their causal relations, while others are not." (Burge 1999, 98) Die kausale Kraft der physischen Ereignisse wird üblicherweise durch deren physische Eigenschaften bestimmt. „Bestimmt" bzw. „determiniert" ist insofern zunächst offen, als hiermit abhängig vom jeweiligen Modell eine Supervenienzrelation, eine Erklärungsabhängigkeit oder die Individuation gemeint sein kann (vgl. Burge 1999, 100; Kap. 2.4). Ebensowenig ist Burge zufolge klar definiert, was sich hinter der Bezeichnung „physische Eigenschaft" verbirgt. In einem weiten Sinn können hiermit auch Relationen beschrieben werden, bei denen physische Eigenschaften ihre Umgebung umfassen, die in nicht-mentalistischen, nicht-intentionalen Begriffen beschrieben wird.

Auch Baker (1999, 75) weist darauf hin, dass die Terminologie ein Problem innerhalb der Diskussion darstellt. Zwar geht auch sie davon aus, dass die Realität von der physischen Realität abhängt und diese aus einem Netzwerk von Ereignissen besteht.[61] Dieser Ereignisbegriff ist jedoch durch Davidson und Kim sehr unterschiedlich geprägt. Während Davidson ein Ereignis als konkretes Partikular an einem speziellen raum-zeitlichen Ort versteht, fasst Kim ein Ereignis als eine Instanziierung einer Eigenschaft zu einer bestimmten Zeit auf. In der Diskussion über Davidson benutzt Kim jedoch teilweise seinen eigenen Eigenschafts- bzw. Ereignisbegriff. Er spricht davon, dass innerhalb des anomalen Monismus die mentale Eigenschaft eines Ereignisses

[60] Dabei unterscheidet Audi weiter zwischen nachhaltigen und dynamischen auslösenden Ursachen. Für jeden dispositionalen mentalen Zustand, der als Handlungserklärung dient, kann ein kausales Ereignis bestimmt werden. Dispositionen bringen dabei nichts ohne ein Ereignis hervor. Diese Zustände sind nicht nur nachhaltige Ursachen, sondern operieren in ständiger Verbindung mit den dynamischen Ereignisursachen. „The best course here is to distinguish sustaining causation from event causation and treat the standard action-explaining generalizations, in a broad sense, as causal in virtue of expressing the former and, in a narrow sense, as non-causal in virtue of omitting the latter from their explicit scope." (Audi 1999, 69)

[61] Baker unterscheidet in der Folge mentale Ereignisse als „innere Ereignisse" mit physischen und mentalen Eigenschaften von „äußeren" rein physischen Ereignissen, die objektiv mess- und erfahrbar sind. Ein inneres Ereignis wäre demnach ein Gedanke oder ein Schmerz, während ein äußeres Ereignis der Knall einer Pistole oder auch die Umsetzung eines Fußballspiels wären (vgl. Baker 1999, 75ff.).

keinen kausalen Unterschied zwischen zwei mentalen Ereignissen macht (vgl. Kim 1999a, 20ff.). Aufgrund dieser Vermischung folgern die Kritiker Davidsons somit, dass mentale Eigenschaften epiphänomenal sind. Diese Vermischung der beiden Ereignisbegriffe ist jedoch problematisch, weil Davidson selbst nicht von Eigenschaften, sondern von der Beschreibung der Ereignisse sprechen würde (vgl. Kap. 2.3.1).

### 3.1.1 Ereignisse als ontologische Grundeinheit

Der Begriff „Ereignis" ist im Grunde ein Oberbegriff für verschiedenartige Geschehnisse. Davidson hat seine Ereignisontologie vor dem Hintergrund seiner Handlungstheorie entwickelt (vgl. Kap. 2.3.1). Erfolgreiche Kommunikation scheint nur möglich, wenn zwischen den Gesprächspartnern über bestimmte Grundstrukturen der Wirklichkeit ein gewisses Einverständnis herrscht. Reportern und Zuhörern sind sich beispielsweise darin einig, dass es nicht nur Fußballspieler und Musiker, sondern auch „Fußballspiele" und „Konzerte" gibt (vgl. Davidson 1986b, 1990). Der entscheidende Punkt von Davidsons Analyse ist, dass wir Ereignisse als eigene Entitäten für die Wahrheit unserer Aussagen über Ereignisse annehmen müssen. Wir verpflichten uns gewissermaßen ontologisch auf die Existenz von Ereignissen, wenn wir über Ereignisse sprechen. Die Identitätsbedingung von Ereignissen lässt sich durch ihre „kausale Rolle" bestimmen (Davidson 1998, 230f.). Ausgangspunkt für die Überlegung ist die Feststellung bzw. Setzung, dass Ereignisse Ursachen und Wirkungen haben. Jedem Ereignis ist die Summe seiner Ursachen und Wirkungen zueigen: Es hat im kausalen Gefüge eine ganz bestimmte Position.

Kritikern an der kausalen Rolle als notwendiger und hinreichender Identitätsbedingung von Ereignissen stellt sich die Frage, ob diese Identitätsbedingung tatsächlich die ontologische Konstitution von Ereignissen erklärt oder nicht in einem Zirkel endet. Dabei ist interessant, dass die Identitätsbedingungen – einmal für Ereignisse akzeptiert – ebenso für Zustände gelten müssten. Damit legt man sich jedoch auf eine schwerwiegende ontologische These fest, dass der Unterschied zwischen dynamischen Geschehnissen, wie Ereignissen, und statischen Zuständen kein kategorialer sein kann. Außerdem ist es fraglich, ob sich im Anschluss an die Annahme der kausalen Rolle als Identitätsbedingung auch Kriterien darlegen lassen, die Aufschluss darüber erlauben, ob es sich bei den Ereignissen um numerisch identische handelt oder nicht. Hinsichtlich unendlicher Bereiche muss dann allerdings für die Angabe der kausalen Rolle die Summe aller Ursachen und Wirkungen berücksichtigt werden. Die vollständige Angabe der kausalen Rolle ist aber ein aussichtsloses Unterfangen, da sich die Summe aller Ursachen und Wirkungen eines Ereignisses aus einer unabsehbar großen Anzahl anderer Zustände und Ereignisse zusammensetzt, für die

wiederum jeweils eine solche Summe von Ursachen und Wirkungen anzugeben wäre (vgl. Kap. 2.3.1; Davidson 1969, 1970).

Nach Kims Ereignisbegriff, der sich innerhalb der analytischen Diskussion des Geistes zunehmend durchgesetzt hat, vollziehen sich Ereignisse stets an etwas, einem Träger, wie Dingen oder Personen. Letztere sind Konstituenten von Ereignissen (vgl. Kim 1976, 160). Ereignisse haben eine Dauer und finden in einer bestimmten Zeitspanne statt. Der wesentliche Aspekt besteht jedoch darin, dass konkrete Ereignisse Instanziierungen, also partikuläre Vorkommnisse bestimmter Eigenschaften oder generischer Ereignisse sind. Das momentane Tippen am Computer ist demnach eine Instanziierung der Eigenschaft oder des generischen Ereignisses des Am-Computer-Tippens. Die Identitätsbedingung für Ereignisse besagt auch, dass Ereignisse genau dann identisch sind, wenn sämtliche ihrer konstitutiven Elemente identisch sind: die Träger der Ereignisse, die Eigenschaften, die sie instanziieren, sowie die Zeiten, zu denen sie stattfinden (vgl. ebd., 161).

Kritisch hinterfragt wird vor allem die Behauptung, Ereignisse seien Eigenschaftsinstanziierungen. Nach Kim werden durch Variationen in der Beschreibung von Ereignissen verschiedene Eigenschaften, folglich auch numerisch verschiedene Ereignisse eingeführt. Kim unterscheidet folgende Ereignisse:

1. Mein Tippen am Computer zur Zeit t.
2. Mein Tippen eines Textes am Computer zur Zeit t.
3. Mein schnelles Tippen eines Textes am Computer zur Zeit t.

Träger und Zeit sind zwar dieselben, aber die Eigenschaftsinstanziierung ist eine andere. Diese Auffassung wird in der Diskussion als eine extrem „feinkörnige" bezeichnet. Es scheint nicht plausibel zu sein, dass meine Tätigkeit durch jede Beschreibung, die neue Umstände berücksichtigt, einfach vervielfältigt wird (vgl. Ginet 1990; Runggaldier 1996, 51f.). Ontologisch problematisch ist an Kims Position zudem, dass weitere Entitäten vorausgesetzt werden, ohne diese genauer zu beschreiben. Dies läuft mindestens einer Sparsamkeitsargumentation zuwider. Ein Vorteil ist allerdings, dass Kims Ereignisbegriff die Veränderung betont. Änderungen sind immer Änderungen von etwas oder jemandem in einer Zeitspanne.

Vor allem seine Charakterisierung der Ereignisse als Elemente von Kausalbeziehungen sowie als Objekte von Erklärungen ist hervorzuheben (vgl. Kim 1976, 169). Vor diesem Hintergrund macht auch die feinkörnige Deutung der Ereignisse Sinn, da die unterschiedlichen Abläufe 1-3 unterschiedliche Ursachen und Folgen haben können. Ebenso greifen wir in unserer Erklärung auf ein anderes Ereignis zurück, wenn wir 1 oder 2 oder 3 erklären wollen. Nach Kims erkenntnistheoretischen Überlegungen sollen beide Aspekte, die Erklärung der Wahrheit von

Behauptungen und die Erklärung von Ereignissen, zusammengebracht werden. Das kann nur geschehen, wenn es einen engen Zusammenhang zwischen der Wahrheit von Behauptungen und dem Vorkommen von Ereignissen gibt. Wenn immer wir die Wahrheit der Behauptung P erklären, müssen wir auch ein ganz bestimmtes Ereignis E als explicandum der Erklärung annehmen. Dabei ist jedes E einzigartig mit je einem P verbunden (vgl. Kim 1976; Lombard 1986, 50ff.).

### 3.1.2 Eigenschaften und mentale Kausalität

Wie aber sollen mentale Ereignisse, die anscheinend so schwer zu fassen sind, gegenüber anderen mentalen Ereignissen oder sogar physischen Ereignissen kausale Wirksamkeit haben? In den meisten Theorien wird angenommen, dass Eigenschaften kausal verantwortlich für verschiedene Ereignisse sind. Eine Ursache hat demnach kraft ihrer Eigenschaften (englisch: „in virtue of") eine Wirkung. Eigenschaften wie die Reichweite einer Explosion, das Timing eines Schmetterballs, die Höhe der Berge oder die Krümmung einer Kurve nennen wir verantwortlich für die Größe eines Kraters, den Gewinn des Tennisspiels, die Ermüdung der Bergsteiger oder die Häufigkeit von Unfällen (vgl. Jackson 1996, 391). Allerdings sind nicht alle Eigenschaften eines Ereignisses kausal verantwortlich für die Wirkung. Wenn ich mir die Hand an einem roten heißen Schürhaken verbrenne, ist es nicht die Farbe des Schürhakens, der die Verbrennung verursacht, sondern die Hitze (vgl. Crane 1992, 185). Wenn ein mentales Ereignis also eine Wirkung hat, müssen wir uns fragen, ob diese Wirkung aufgrund seiner mentalen Eigenschaften oder seiner nicht-mentalen Eigenschaften erreicht wird. An dieser Stelle ist es wichtig, genau darauf zu achten, dass mentale Ereignisse und mentale Eigenschaften voneinander unterschieden werden und als eigene Problemfelder zu betrachten sind. So ist es durchaus möglich, mentalen Ereignissen kausale Wirkung zuzuschreiben, während die kausale Wirksamkeit ihrer Eigenschaften ungeklärt bleibt.

> Mental events are effective, maybe, but not by way of their mental properties; any causal role that the latter might have hoped to play is occupied already by their physical rivals. (Yablo 1992, 249; vgl. auch Van Gulick 1999, 234)

Mentale Ereignisse können demnach durchaus kausal wirksam sein, nämlich über ihre physischen Eigenschaften. Die mentalen Eigenschaften scheinen dagegen wirkungslos zu sein und lediglich epiphänomenal. Es ist somit wichtig, mentale und physische Ereignisse sowie mentale und physische Eigenschaften zu unterscheiden und klar voneinander abzugrenzen.

Eigenschaften im engen Sinne umfassen Qualitäten, Attribute, Charakterzüge, Aspekte, Momente usw. Bestimmungen also, die die Dinge haben und verlieren können, welche ihnen kontingenterweise zukommen. Sie sind ungesättigt, unvollständig und können in Zeit und Raum nicht ohne einen Träger existieren; sie werden von Ereignissen instanziiert. Eigenschaften sind der Grund, warum manche Dinge einander ähneln oder sich qualitativ voneinander unterscheiden. Vor einem realistischen Hintergrund müssen Eigenschaften insofern allgemein und universal sein, als sie vielen verschiedenen Dingen zukommen. Damit sind sie niemals individuell, sie können aber als dieselbe Eigenschaft von verschiedenen Dingen instanziiert werden; die an den konkreten Dingen realisierten oder instanziierten Eigenschaften sind also individuell. Darin besteht das grundlegende ontologische Problem, zu klären, als was Eigenschaften existieren: als allgemeine oder als partikuläre, an den konkreten Dingen vorkommende Entitäten? Wenn Eigenschaften individuell und konkret realisiert sind, werden sie heute auch als „Tropen" bestimmt (vgl. Macdonald/Macdonald 1995). Tropen werden als ontologisch grundlegende Einheiten verstanden. Ein Ding bzw. Ereignis zeichnet sich durch eine Menge konkreter Eigenschaften in Raum und Zeit aus. Dabei können empiristisch grundsätzlich nur Einzelfälle von Eigenschaften bzw. Relationen – ein Nexus aktiv und passiv wirksamer Eigenschaften – erkannt werden. Tropen sind als konkrete Eigenschaften eines Dings bzw. eines Ereignisses kontingente Tatsachen.

Natürliche Eigenschaften sollen nahe legen, dass sie vorgegeben oder objektiv und nicht lediglich auf konventionelle Einteilungen der uns begegnenden Wirklichkeit zurückzuführen sind. Natürlich sind z. B. Masse oder Ladung. Sie werden vor einem realistischen Hintergrund als Erklärungsgrund für Ähnlichkeiten zwischen Dingen postuliert, oder auch, um die kausalen Wirkkräfte in der Natur zu erklären. Dinge und Ereignisse rufen unter bestimmten Umständen Wirkungen hervor, weil ihnen ganz bestimmte Eigenschaften zukommen (vgl. Armstrong 1968).

Dass etwas giftig ist, heißt demnach, dass es unter bestimmten Umständen die Wirkkraft hat, einen Organismus zu schädigen bzw. zu zerstören. Giftige Substanzen wirken aber kraft gewisser Eigenschaften, die ihnen zukommen. Es gibt somit eine Entsprechung zwischen den Wirkkräften (als Funktion von Umständen und Rahmenbedingungen auf kausale Folgen) und den Eigenschaften, von denen die Wirkkräfte abhängig sind. Nach dem Kausalgesetz wird angenommen, dass eine Eigenschaft nur dann eine natürliche bzw. echte Eigenschaft ist, wenn sie potenziell kausal wirksam sein kann (vgl. Kap. 2.1.2; Kap. 2.4.3; Heil 1998, 118). Damit stellt sich auch die Frage, ob mentale Eigenschaften natürliche Eigenschaften sind und Wirkkräfte haben. Sollte dies nicht zutreffen, stellt sich wiederum die Frage, was sie dann sind bzw. worauf die Wirkkräfte zurückgehen.

Deshalb erscheint es wichtig, die Beziehung zwischen mentalen und physischen Eigenschaften genauer zu betrachten. Gewöhnlich wird davon gesprochen, dass mentale Eigenschaften von physischen realisiert werden. Sie sollen zwar nicht identisch sein, aber ihre Beziehung soll

über eine bloße Korrelation hinausgehen. In gewisser Weise bleibt somit auch mit der physikalischen Konzeption der Welt eine Art Dualismus bestehen, allerdings ein Eigenschaftsdualismus mit physischen Eigenschaften auf einer niedrigeren Ebene und mentalen Eigenschaften auf einer höheren Ebene, bzw. mit physischen Mikro- und mentalen Makroeigenschaften.

### 3.1.3 Die Realisierung von Eigenschaften im Rahmen des Schichten-Modells

Was bedeutet es, dass eine Eigenschaft eine andere realisiert? LePore und Loewer formulieren ein Konzept, wonach zwischen zwei Eigenschaften eine starke Verbindung bestehen soll (vgl. auch Kap. 2.4.2)

> Exactly what is it for one of an event's properties to *realize* another? The usual conception is that *e's* being *P* realizes *e's* being *F* if *e* is *P* and *e* is *F* and there is a strong connection of some sort between *P* and *F*. We propose to understand this connection as a necessary connection which is explanatory. The existence of an explanatory connection between two properties is stronger than the claim that $P \rightarrow M$ is physically necessary since not every physically necessary connection is explanatory. (LePore/Loewer 1989, 179)

Im Grunde werden hier zwei verschiedene Voraussetzungen für Realisierung angegeben: (1) P erklärt M und (2) P $\rightarrow$ M ist nomologisch notwendig. Für die Voraussetzung (1) wird die Verbindung nicht nur als notwendig, sondern vor allem als explanatorisch aufgefasst, weshalb eine Realisierung besteht. Man kann also davon sprechen, dass die neurophysiologische Eigenschaft P des mentalen Zustands e die psychologische Eigenschaft F des Zustands e realisiert, da eine starke (nomologische) Verbindung zwischen P und F besteht und P das Vorhandensein von F erklärt. Levine betont dagegen, dass die Unterscheidung zwischen einer kausal-nomologischen Relation der Eigenschaften und einer Realisierungsrelation insofern bedeutend ist, als letztere Relation eine engere Bindung der Eigenschaften aneinander sichert, schließlich werden beide Eigenschaften zeitgleich instanziiert (vgl. Levine 2001, 14; Kap. 7.2). Dennoch ist die Relation keine der Identität, da mentale Eigenschaften multirealisierbar sind; es ist nicht möglich einen einzigen physischen Zustand zu bestimmen, der in allen möglichen Welten z. B. der Realisierer von „Schmerzen" ist (vgl. Kap. 2.2.2; Kap. 4.2.1).

Block (1990) unterscheidet zwischen Eigenschaften erster Ordnung, welche die Eigenschaften zweiter Ordnung hervorbringen. Als Beispiel dient das rote Tuch des Toreros. Angeblich reizt die Farbe rot den Stier, weshalb „rot" kausal relevant für die Wut des Stiers ist. Das Tuch hat somit auch die Eigenschaft zweiter Ordnung, „reizend" zu sein. Aber ist damit auch die Eigenschaft „Reizung" kausal relevant für die Wut des Stiers und reizt den Stier? Das würde man ablehnen, da der Stier letztlich auch zu dumm ist, die „Reizung" oder „Provokation" als solche

zu verstehen. Block beschreibt einen physischen Zustand als Zustand von etwas, das eine Eigenschaft erster Ordnung hat, die mit einer wahren physischen Theorie beschrieben werden kann. Eine Eigenschaft erster Ordnung wiederum bedarf nicht der Quantifizierung durch weitere Eigenschaften. Dagegen bedarf zwecks Bestimmung eine Eigenschaft zweiter Ordnung der Quantifizierung durch Eigenschaften erster Ordnung. Die physikalistische Position, gegen die der Funktionalismus argumentiert, besagt, dass mentale Eigenschaften identisch mit physischen Eigenschaften erster Ordnung sind. Funktionalisten müssen dagegen nicht verneinen, dass mentale Eigenschaften physische Eigenschaften zweiter Ordnung sind. Ihnen ist es gerade wichtig, diese Abhängigkeit zu betonen (vgl. Block 1980b, 182).

Den Hintergrund für die physikalistische Konzeption der Realisierung mentaler Eigenschaften bildet das Modell einer so genannten Schichten-Welt, welches insbesondere die funktionalistische Position beeinflusst hat. Nach diesem ontologischen Modell ist die Welt in verschiedene Ebenen, Klassen oder Schichten von Entitäten und deren charakteristische Eigenschaften eingeteilt und hierarchisch strukturiert. Mikrophysische Partikel, wie z. B. Elektronen, Neuronen oder Quarks, bilden auf einer Grundebene die kleinste Einheit. Sie sind mit fundamentalen Eigenschaften wie Masse oder Bewegung ausgestattet. Die Entitäten der oberen Ebenen basieren auf den unteren Ebenen. Sie haben aber eigene Eigenschaften, eigene Gesetze und werden in einer eigenen Taxonomie der jeweiligen Einzelwissenschaft beschrieben (vgl. Heil 1998, 114ff.). Insofern ist die hierarchische Struktur durch mereologische Überlegungen gesichert. Entitäten einer Ebene sind zusammengesetzt aus oder beinhalten die Entitäten der unteren Ebenen. Biologische Organismen sind ein Beispiel für diese Ebenenhierarchie. Der menschliche Körper setzt sich aus verschiedenen Formen von Biomasse zusammen; Muskeln, Knochen und Flüssigkeiten lassen sich weiter unterteilen in verschiedene Zellstrukturen und Moleküle. Das Schichtenmodell scheint die kausale Geschlossenheit der physischen Welt zu sichern (vgl. Kap. 3.3). Kausale Interaktion ereignet sich nur innerhalb der Ebenen, wobei die Autonomie der kausalen Prozesse der fundamentalen Ebene(n) berücksichtigt wird. Allerdings ist es möglich, dass kausale Beziehungen oberer Ebenen von den kausal-nomologischen Prozessen der unteren Ebenen abhängen (vgl. Kim 1999b, 192). Insofern ist in dem Schichtenmodell die Beziehung mentaler und physischer Eigenschaften als Abhängigkeitsrelation gekennzeichnet. Die subvenienten Basen bzw. die niedrigstufigen Eigenschaften werden bei den Prozessen auf der unteren Ebene angesiedelt, während die supervenierenden bzw. höherstufigen Eigenschaften bei den Prozessen auf der höheren Ebene angesiedelt werden. Die Idee der Instanziierung von Eigenschaften ist eng verbunden mit der Idee der Ebenenhierarchie. So hängen beispielsweise Inhaltseigenschaften nicht direkt von mikrophysischen Eigenschaften ab, sondern werden zunächst von neurophysiologischen Eigenschaften und deren Relation zu anderen, z. B. Umwelteigenschaften bestimmt. Neurophysiologische hängen selbst von molekularbiologischen Eigenschaften ab, die wiederum

von chemischen oder physischen Eigenschaften bestimmt werden – bis hin zur mikrophysischen Ebene. Diese Ebenen-Metapher hat sich im wissenschaftlichen Kontext als grundsätzlich sinnvoll erwiesen; allerdings ist zumindest in der philosophischen Debatte nicht eindeutig festgelegt, wie und in welcher Folge die verschiedenen Ebenen voneinander abhängen (vgl. Baker 1999, 81). Deshalb wird zumeist lediglich grob zwischen Makro-Eigenschaften und Mikro-Eigenschaften unterschieden, wobei Makro-Eigenschaften je nach Kontext mentale Eigenschaften oder aber neurophysiologische Eigenschaften bezeichnen können.

Vor diesem Hintergrund erscheint es schwierig, mentalen Eigenschaften eine eigene kausale Kraft zuzusprechen. Sie scheinen aufgrund ihrer Abhängigkeit von den physischen Eigenschaften allenfalls „kausal relevant" zu sein, wenn nicht epiphänomenal als bloße „nomologische Anhängsel" (vgl. Feigl 1958; Kap. 2.1). Innerhalb der Kausalrelation werden Ursachen ontologisch als Ereignisse bzw. Instanziierungen von Eigenschaften aufgefasst. Beispiele für Ursachen sind „sich in einem Schmerzzustand befinden" oder „aktivierte spezifische Neuronen haben". Aufgrund dieser Ontologie impliziert das Kausalargument der Identitätstheorie (vgl. Kap. 2.1.2), dass auch bewusste Eigenschaften wie z. B. „Durst haben" identisch mit physischen Eigenschaften wie „sich in einem bestimmten Gehirnzustand befinden" sind. Die Identität von bewussten Ursachen und physischen Ursachen impliziert wiederum die Identität von bewussten Ereignissen und physischen Ereignissen, was wiederum die Identität von deren Eigenschaften impliziert (vgl. Papineau 2004, 20). Allerdings besteht für diese Position eben die Schwierigkeit, dass hieraus ein theoretischer Chauvinismus hinsichtlich bewusster Zustände folgt. Werden bewusste Eigenschaften mit strikten physischen Eigenschaften – und nicht mit beispielsweise funktionalen Rollen – identifiziert, deren Instanziierungen wiederum physische Ursachen sind, ist es nicht möglich, dass Menschen oder Tintenfische jene bewussten Eigenschaften teilen (vgl. Kap. 2.1.3; Kap. 2.2).
Ziel des Funktionalismus war es u. a., diese Problematik zu umgehen. Da das Kausalargument der Identitätstheorie (vgl. Kap. 2.1.2) aus verschiedenen Gründen angreifbar ist, sowie aufgrund des Arguments der Multirealisierbarkeit die Identität physischer und mentaler Eigenschaften grundsätzlich auf dem Prüfstand steht, können mentale Eigenschaften bzw. Ereignisse eben nicht ohne weiteres als kausal wirksam betrachtet werden. Funktionalisten betonen im Rahmen des Schichten-Welt-Modells deshalb die Eigenständigkeit der einzelnen Ebenen bzw. der Einzelwissenschaften. Das System bzw. die Kategorien einer höherstufigen Taxonomie markieren letztlich auch die Grenzen gegenüber niedrigstufigen Taxonomien. Die Beobachtungen und Erkenntnisse einer solchen Einzelwissenschaft bleiben der Physik auf der niedrigstufigen Basisebene prinzipiell verschlossen:

> Hearts are made up of cells, and ultimately of quarks and electrons. But hearts are not just assemblages of electrons and quarks, and properties of hearts are not properties of electrons and quarks or assemblages of electrons and quarks. (Heil 1998, 118)

Deshalb ist auch eine Reduktion einer Wissenschaft auf einer höherstufigen auf eine niedrigstufige Ebene aus ausgeschlossen. Da die Einzelwissenschaften somit auch eigenständige Kausalgesetze formulieren können, müssen die Eigenschaften, die diese Kausalgesetze der Neurobiologie, der Psychologie etc. betreffen, echte Eigenschaften mit kausaler Wirkung sein (vgl. ebd., 116ff.).

***Quintessenz Ontologisches Grundmodell***

- *Modell der mentalen Kausalität abhängig von ontologischer Position*
- *kausale Rolle der Ereignisse als Eigenschaftsinstanziierungen innerhalb gesamtem Kausalgefüge*
- *Ereignisse und Eigenschaften mit unterschiedlicher kausaler Funktion*
    - *Bestimmung kausaler Kraft (physischer) Ereignisse durch kausale Kraft ihrer (physischen) Eigenschaften*
    - *Kontingenz von Eigenschaften für Ereignisse: als allgemeine Eigenschaften Instanziierung von verschiedenen Ereignissen*
- *zeitgleiche Instanziierung mentaler und physischer Eigenschaften bei wechselseitiger Abhängigkeit durch Realisierungsrelation*
- *hierarchische Struktur auf Basis des Schichtenwelt-Modells mit eigener Taxonomie für jeweilige Ebene*
    - *Abhängigkeit der Inhaltseigenschaften von neurophysiologischen Eigenschaften in Relation zu Umwelteigenschaften, nicht direkt von mikromikro Eigenschaften*
- *Anforderung echter Eigenschaften mit kausaler Wirkung für eigenständige Kausalgesezte*

### 3.2 Die Unterscheidung des mentalen und des physischen Bereichs

Eine weitere mögliche Voraussetzung für mentale Kausalität ist die Annahme, dass eine Beschreibung der Welt in verschiedener, konkurrierender Weise möglich ist – ohne die konkurrierenden psychologischen und physischen Beschreibungen gäbe es das Problem der mentalen Kausalität nicht. Wie eingangs gezeigt, unterscheidet der Substanzdualismus traditionell res extensa und res cogitans als zwei unterschiedliche ontologische Einheiten (vgl. Descartes 1977). Demnach wird das Physische als räumlich, ausgedehnt und zeitlich charakterisiert, während das Mentale zwar als zeitlich, aber eben nicht-räumlich, nicht-ausgedehnt, sondern als geistig beschrieben wird (vgl. Kutschera 2003, 11ff.). In der zeitgenössischen analytischen Diskussion wird

zwar nicht mehr zwischen zwei Substanzen unterschieden, dennoch wird ein Eigenschaftsdualismus zumindest auf der epistemischen Ebene aufrechterhalten.

Auf der ontologischen Ebene kann beispielsweise die Typenidentität keine Differenzierung vornehmen. Deshalb ist es nur konsequent, wenn Papineau in Anlehnung an Chalmers (vgl. Chalmers 1997, 11ff.) einen konzeptionellen bzw. begrifflichen Dualismus vorschlägt. Materielle Konzepte beschreiben Bewusstseinseigenschaften als Gegenstände der kausalen Welt aus der Dritte-Person-Perspektive, während phänomenale Konzepte sich auf die Erste-Person-Perspektive beziehen (vgl. Papineau 2004, 49).[62] Mit dem konzeptionellen Dualismus will Papineau allerdings keinen ontologischen Standpunkt vertreten, hier geht es vor allem um die epistemische Perspektive. So betont er: „[L]et me make clear that I do not take my conceptual ‚dualism' itself to demand any special non-material ontology. In my view, the deployment of phenomenal concepts depends on material processes in thinkers' brains, just as much as the deployment of any other concepts." (ebd.) Papineau räumt ein, dass es hinsichtlich der Identität mentaler und physischer Eigenschaften zunächst lediglich möglich ist, theoretische Aussagen zu machen. Jede Bewusstseinseigenschaft ist in dem Sinne mit einer noch zu identifizierenden physischen Eigenschaft identisch – um welche physische Eigenschaft es sich handelt, bleibt offen:

> It tells us that each conscious property must be identical with *some* physical property, but it doesn't tell us *which* specific physical property any given conscious property may be identical with. (Papineau 2004, 20)

Papineau betont, dass hierfür weitere empirische Forschung notwendig ist – die Geist-Gehirn-Identität ist nur a posteriori kontingent erkennbar. Die empirische Kontingenz der Relation mentaler und physischer Eigenschaften wird in dem Argument der Multirealisierbarkeit sowie dem Argument der starren Bezeichner stark gemacht (vgl. Kap. 2.1.3). Insbesondere Kripke (1980) hat darauf hingewiesen, dass auch die Geist-Gehirn-Identität notwendig sein muss, um wahr zu sein. Wenn nun die Identität bewusster und physischer Eigenschaften lediglich durch empirische Forschung validiert werden kann und nicht rein theoretisch a priori, muss in der Konsequenz

[62] Materielle Konzepte umfassen in der Regel „Rollenkonzepte", also Konzepte, die eine kausale Rolle beschreiben, wie z. B. die Rolle von „Schmerz" als Vermittler zwischen körperlichem Schaden und Vermeidungsverhalten. Papineau schließt auch direkte physikalische Konzepte als materielle Konzepte ein, insofern sie sich über intrinsische physikalische Konstitution wie Größe, Form, oder Masse auf ihre Referenten beziehen. Ein äußerer Reiz verursacht über die Nervenbahnen ein Neuronenfeuer. Das Feuern der Neuronen verursacht die körperliche Reaktion, wie z. B. das Zusammenkrümmen. Phänomenale Konzepte sind dagegen schwerer zu fassen. Wenn phänomenale Konzepte verwendet werden, werden üblicherweise mentale Eigenschaften nicht als Teil der materiellen Welt betrachtet, sondern unter dem Aspekt „wie es sich anfühlt", diese Eigenschaften zu haben. So werden Schmerzen nicht über Gehirnaktivität, sondern eher über die Vorstellung, wie sich ein bestimmter Schmerz anfühlt, beschrieben: „Ich habe Kopfschmerzen, weil es draußen laut ist", lautet eine mögliche phänomenale Beschreibung. Papineau macht allerdings deutlich, dass phänomenale Konzepte sich direkt auf materielle Eigenschaften beziehen – ebenso wie materielle Konzepte (vgl. Papineau 2004, 48).

überlegt werden, ob nicht ein anderes Modell die Relation zwischen den Eigenschaften besser beschreibt. Hierfür ist es zunächst einmal wichtig, beide Bereiche genauer zu bestimmen.

### 3.2.1 Merkmale des mentalen Bereichs

Im Rahmen des Schichten-Welt-Modells wurden mentale Eigenschaften bereits als „höherstufig" charakterisiert – eine Einschätzung, die zumindest unter funktionalistischen oder supervenienztheoretischen Gesichtspunkten getroffen werden kann (vgl. Kap. 3.1.3). Doch was zeichnet sie als höherstufige Eigenschaften über ihre möglicherweise „kausale Rolle" hinaus (vgl. Kap. 2.2.1; Kap. 2.2.3) aus? Über die verschiedenen theoretischen Positionen hinweg scheint die Bestimmung, was mental und was physisch ist, zunächst sehr ähnlich zu sein. So wird das Mentale einerseits in einer Descartes'schen Tradition noch immer negativ als Gegenbegriff zum Physischen definiert. Andererseits wird es positiv durch das Bewusstsein (vgl. Searle 1993), die Erste-Person-Perspektive (vgl. Nagel 1974), Qualia (vgl. Chalmers 1997; Jackson 1982) oder propositionale Einstellungen (Baker 1987) bestimmt. So werden mentale Ereignisse beispielsweise als bewusst gekennzeichnet, als nicht räumlich, intentional oder privat (vgl. Beckermann 1999, 9). Es gibt verschiedene Merkmale des Mentalen, die dieses teilweise erfassen, jedoch keines allumfassend.

Für eine Vielzahl von mentalen Phänomenen bzw. Ereignissen scheint von Bedeutung zu sein, dass sich eine Person einer Sache bewusst ist, sich also Gedanken über etwas machen kann. In propositionalen Einstellungen drückt sich die Haltung einer Person zu bestimmten Sachverhalten aus, z. B. die Überzeugung, Friede sei ein schützenswerter Zustand, der Glaube, Rotwein schmecke besser als Weißwein oder der Wunsch, die Mehrwertsteuer möge nicht erhöht werden. Auch so genannte voluntative Zustände gehören zu den propositionalen Einstellungen, wie z. B. die Absicht, ins Kino zu gehen. Da die Proposition den Inhalt der Einstellung bildet, werden diese Zustände auch intentionale Zustände genannt (vgl. Kap. 4.5). Propositionale Einstellungen oder intentionale Zustände sind für mentale Kausalität bedeutsam, da mit ihnen das abstrakte, nicht-physische Moment in eine Handlungskette gelangt. Es sind eben Absichten als abstrakte mentale Ereignisse, die Handlungen als konkrete physische Ereignisse verursachen sollen.

Als qualitative Zustände bezeichnen so genannte Qualia Empfindungen oder Gefühle, die eine Person bei einer Sinneswahrnehmung haben kann. Beispiele sind Schmerzen oder Farbempfinden. Mit diesen Eindrücken gehen charakteristische Eigenschaften einher, die sie uns unmittelbar als eben diese Zustände bewusst und erfahrbar machen. Als emotionaler Zustand wird die Rubrik der Gefühle zusammengefasst. Da sie sich auf Personen, Objekte oder Sachverhalte beziehen können, sind emotionale Zustände häufig zugleich propositionale Einstellungen (vgl.

Baker 1995, 488). Man freut sich beispielsweise auf eine Reise, ärgert sich über einen Kollegen oder trauert um einen Freund. Aus den genannten mentalen Ereignissen bzw. Zuständen lassen sich sicherlich zahlreiche weitere ableiten oder ergänzen. Qualitative Zustände sind für mentale Kausalität bedingt relevant. Zwar gibt es bei ihnen weniger einen Anspruch auf eine direkte Bewirkung von Ereignissen im Sinne der intentionalen Zustände. Allerdings müssen sie z. B. einerseits als Ergebnis kausaler Wahrnehmungsprozesse von extramentalen Stimuli o. ä. bewirkt werden. Andererseits bewirken auch qualitative Zustände in Verbindung mit z. B. intentionalen Zuständen möglicherweise emotional motivierte Handlungen. Ein extremes Beispiel wäre hier der sogenannte Mord im Affekt.

Gerade die Merkmale der Privatheit und das Erste-Person-Privileg machen es schwer, mentale Ereignisse vollständig zu eliminieren. Mit der Frage „What is it like to be a bat?" pointiert Thomas Nagel dieses Argument der Privatheit (vgl. Nagel 1974). Wir können eine Fledermaus zwar von Außen erforschen. Wie es sich anfühlt, eine Fledermaus zu sein, kann durch objektive Kriterien jedoch nicht erfahrbar gemacht werden. Ebenso können Intentionen, Emotionen und Qualia einer Person nur von ihr selbst direkt erlebt oder beobachtet werden. Eine außenstehende Person hat zwar aus der Dritte-Person-Perspektive die Möglichkeit zur Interpretation, aber keinen direkten Zugang zu den mentalen Zuständen, dem „inneren Erleben" ihres Gegenübers.[63] Privatheit und das Erste-Person-Privileg sind direkt an die Erfahrung propositionaler Einstellungen oder qualitativer Zustände gebunden. Personen erleben sich als Akteure und Verursacher ihrer Handlungen (vgl. Kutschera 2003, 74ff.).

So versteht Baker Handlungen bzw. Verhalten als eine Ereignisart mit speziellen inneren Ursachen. Innere Ereignisse, die Verhalten verursachen, haben verschiedene Arten von Eigenschaften, einige sind physisch und andere sind intentional. Intentionale Eigenschaften setzt Baker mit mentalen Eigenschaften gleich und bezeichnet sie als „content-properties" – Inhaltseigenschaften. Mentale Eigenschaften umfassen z. B. den Inhalt, ein Erfrischungsgetränk aus dem Kühlschrank zu holen oder den Inhalt, dass Brasilien die nächste Fußball-Weltmeisterschaft gewinnt. Die mentalen Eigenschaften werden bestimmt durch Glaubensannahmen oder Wünsche. Deshalb gelten sie im alltäglichen Gebrauch als Gründe für menschliches Verhalten:

> When we explain behaviour in terms of a person's reasons, what we cite in explanations are various content-properties – properties determined by what she believes, what she desires, and so on. In this picture, these content-properties are attributed to the person's internal events that cause the behaviour. (Baker 1999, 76)

[63] Dagegen scheinen Personen mit mentalen Zuständen über direktes Wissen um diesen Zustand zu verfügen. Dass Charly Chaplin traurig ist, vermutet der Kinobesucher zwar aufgrund seines leidenden Blicks; was Charly Chaplin dabei fühlt, ist jedoch seine Privatsache. Ebenso kann der Zahnarzt mit einer Röntgenaufnahme den medizinischen Grund für den Zahnschmerz objektiv feststellen, die Schmerzen selbst erlebt allerdings nur der Patient.

Die Basisfrage für mentale Kausalität – Wie können unsere Gedanken unsere Handlung beeinflussen? – wird vor diesem Hintergrund weiter differenziert. Sie lautet nun: Wie können Inhaltseigenschaften innerer Ereignisse kausal relevant für das Hervorbringen behaviouraler Ereignisse sein? Eine echte Erklärung dieser Frage muss somit die kausal relevante Eigenschaft des inneren Ereignisses benennen, die das Verhalten verursacht. In der Folge sind Inhaltseigenschaften nur erklärungsrelevant, wenn sie auch kausal relevant sind (vgl. ebd.).[64]

### 3.2.2 Merkmale des physischen Bereichs

Zu bestimmen, was denn genuin physisch ist bzw. physisch genug, dass es kausale Wirkung haben kann, scheint im Vergleich zur Bestimmung der Merkmale des Mentalen schwieriger zu sein. Zwar fallen letztlich sämtliche Dinge und Gegenstände der physischen Welt in diesen Bereich, also alle Objekte, die räumlich ausgedehnt, messbar oder empirisch überprüfbar sind. „It is widely believed that the most fundamental facts about our universe are physical facts, and that all other facts are dependent on these." (Chalmers 1997, 32) Dennoch scheint beispielsweise das Vorhaben nicht zu funktionieren, eine physische Sprache als Universalsprache der Wissenschaft auch auf mentale Phänomene anzuwenden (vgl. Carnap 1966). Insofern muss genauer bestimmt werden, welche physischen Phänomene gemeint sind, wenn die Rede davon ist, dass mentale Eigenschaften durch ihre physischen realisiert oder konstituiert sind, dass sie von den physischen Eigenschaften erzeugt werden oder auf ihnen supervenieren. Einem Stein oder einem Stück Holz werden keine mentalen Eigenschaften zugeschrieben, die von den physischen Eigenschaften des Steins realisiert werden können. Und auch wenn die Beantwortung dieser Frage zunächst simpel erscheint, legen sich viele Philosophen nicht eindeutig fest, was sie mit physischen Eigenschaften meinen, sondern bewegen sich in einer abstrakten Terminologie.[65] Die übliche Sichtweise ist, dass

[64] Baker selbst kommt zu dem Schluss, dass aufgrund der metaphysischen Voraussetzungen das Problem der mentalen Voraussetzungen nicht lösbar ist. Da Makro-Eigenschaften nicht als kausal relevant eingestuft werden können, befürchtet sie ferner, dass mit der starken Bindung von Erklärungen an echte Ursachen alle wissenschaftlichen und alltäglichen Erklärungen aufgehoben werden (vgl. Baker 1999, 77).

[65] Papineau versucht zu konkretisieren: Mentale Eigenschaften werden mit bestimmten als nicht-mental erkennbaren Eigenschaften identifiziert. Physische Eigenschaften werden als „erkennbar nicht-mental und nicht-biologisch" bzw. „unbelebt" charakterisiert. Grund ist, dass der Bereich des Unbelebten grundsätzlich als vollständig und geschlossen gilt. Welche Bereiche eine vollständige Physik umfasst, ist schwer zu definieren – darüber haben wir derzeit kein exaktes Wissen. So kann auch nur eine vorläufige Definition physischer Eigenschaften gegeben werden. (vgl. Papineau 2004, 41). „Conscious causes have inanimate effects. Inanimate effects always have full inanimate causes. So conscious properties must be identical with […] inanimate properties." (ebd., 43) Eine Eigenschaft ist insofern belebt, als sie in vor-theoretischen Begriffen wie Bewusstsein, Intelligenz oder auch Atmung und Verdauung erkennbar ist. Dagegen können unbelebte Eigenschaften über Größe, Form oder Masse identifiziert werden. Für Papineau ist wichtig, dass einige Eigenschaften beides sein können: Sowohl belebt als auch unbelebt $=_{df}$ physisch. Damit bleibt die Möglichkeit für Materialisten bestehen, dass biologische oder bewusste Systeme belebte und unbelebte Eigenschaften als identische Eigenschaften haben – mit der Konsequenz, dass bewusste Systeme kausal wirksam sind (vgl. auch Montero 2003, 182ff.).

mentale Zustände bestimmte innere Zustände des Gehirns einer Person sind – und damit physisch. Über unsere Sinnesorgane nehmen wir unsere Umwelt und uns selbst wahr. Unser Körper bewegt sich in verschiedenen Situationen. Die Eindrücke, die uns übermittelt werden, konstituieren dabei unsere Glaubensannahmen, Überzeugungen, Wünsche oder Absichten. Der empirische Ort, wo die entscheidenden Änderungen unserer Glaubensannahmen vollzogen werden, ist das Gehirn. „Thus, as an empirical matter of fact, the brain is where we should locate belief; and the same goes for mental states in general, as similar considerations apply across the board." (Jackson 1996, 387f.)

Aus diesem „engen" physischen Begriff ergibt sich eine definitorische Herausforderung, da auch die Einzelwissenschaften bzw. biologischen, chemischen und eben psychologischen Eigenschaften nicht physisch sind. Als höherstufige oder superveniente Eigenschaften hängen sie von den physischen Basen auf der Mikroebene ab. Andererseits sind auch neuronale Eigenschaften von mikrophysischen Prozessen abhängig.[66] Eine Gleichsetzung „neuronal" mit „physisch" scheint mit dem engen Begriff nicht vereinbar zu sein, da Neuronen und neuronale Prozesse zunächst eine biologische Größe sind. Dennoch ist es in der Praxis üblich, die neuronale Ebene als biologisch-physische Ebene und somit als Gegenpol zur mentalen Ebene zu betrachten. Von vielen Philosophen werden physische inzwischen mit neurophysiologischen Eigenschaften gleichgesetzt. Allerdings werden diese nicht im gleichen Maß der Anforderung der kausalen Geschlossenheit der Physik gerecht, wie die physischen Eigenschaften im strengen Sinne. Angenommen, jedes neurophysiologische Ereignis habe als vollständige Ursache zu einem bestimmten Zeitpunkt t eine neurophysiologische physische Ursache. Dann wird deutlich, dass diese Gleichsetzung nicht zutrifft. Phänomene auf einer niedrigeren Ebene unterbrechen oder fördern neurophysiologische Prozesse, die lediglich unter neurophysiologische ceteris-paribus-Gesetze fallen (vgl. Fodor 1974).[67] Ein System gilt aber nur als kausal geschlossen, wenn seine Elemente ausschließlich und vollständig von anderen Elementen innerhalb des Systems kausal beeinflusst werden. Ein Einfluss von außerhalb widerspricht somit dem Grundsatz der kausalen Geschlossenheit. Wenn dieser Grundsatz nur auf die niedrigste Ebene der Mikro-Physik zutrifft, kann physisch nur als mikro-physisch verstanden werden, um diesem Grundsatz gerecht zu werden. Dabei bezeichnet mikro-physisch als Grundeinheiten die physischen Basisteilchen bzw. Partikel und ihre Eigenschaften (vgl. Baker 1999, 79).

[66] Eccles beschreibt diese Prozesse detailliert und kenntnisreich. Für ihn ergibt sich als Konsequenz die Position des dualistischen Interaktionismus (vgl. Popper/Eccles 1977).

[67] Als ceteris-paribus-Gesetze werden Gesetze bezeichnet, die lediglich auf starken Regelmäßigkeiten beruhen und – im Gegensatz zu strikten Gesetzen – Ausnahmen zulassen.

### 3.2.3 Das Gehirn als mental-physisches Organ

An dieser Stelle ist zumindest ein grober Überblick über das menschliche Gehirn und dessen Aufbau wichtig. Zusammen mit dem Rückenmark und dem Hirnstamm, der sich am oberen Ende des Rückenmarks befindet, bildet das Gehirn das zentrale Nervensystem und ist für die Steuerung und Organisation des menschlichen Organismus die zentrale Schaltstelle. Den verschiedenen Hirnregionen, wie dem cerebralen Kortex, dem Thalamus oder dem Hypothalamus werden inzwischen unterschiedliche Funktionen zugeordnet, wie motorische, sensorische oder emotionale Aufgaben. Allerdings ist nach wie vor umstritten, inwiefern die verschiedenen geistigen Strukturen auf der Großhirnrinde zu lokalisieren sind (vgl. Crick 1994, 112ff.). Deshalb ist es wichtig festzuhalten, dass das Gehirn keine einheitliche Struktur hat. Die unterschiedlichen Bereiche haben zwar unterschiedliche Funktionen, sind aber eng miteinander verknüpft und interagieren. Die Funktion des Gehirns steht in enger Verbindung mit der Funktion seiner Grundeinheit, den Nervenzellen oder Neuronen. Die meisten Neuronentypen sind nach dem gleichen Schema aufgebaut: An ihren Zellkörper und sich verästelnden Dendriten sind Synapsen gekoppelt, die über das Axon – eine Art Kabel – in Verbindung zu anderen Neuronen stehen und elektrische Impulse austauschen (vgl. ebd., 122ff.). Es gibt auf diese elektrischen Entladungen hin drei mögliche Reaktionen für ein Neuron: Erregung, Hemmung oder Modulation. Damit ist bereits die Hauptaufgabe eines Neurons bestimmt:

> Es empfängt Informationen von vielen anderen Neuronen, gewöhnlich in der Form elektrischer Impulse. Es nimmt eine komplizierte zeitabhängige Addition dieser Inputs vor und sendet daraufhin auf seinem Axon Information (in der Form einer Abfolge elektrischer Impulse) an viele andere Neuronen. Obwohl ein Neuron auch Energie verbraucht, um seine Aktivitäten aufrechtzuerhalten und Moleküle zu synthetisieren, besteht seine Hauptfunktion darin, Signale zu empfangen und auszusenden – d. h. mit Informationen umzugehen. (ebd., 123)

Über die elektrochemischen Vorgänge teilen die Neuronen zunächst lediglich mit, wie sehr sie erregt sind. Allerdings steht diese Information in Zusammenhang mit der Außenwelt und Geschehnissen, die wir Menschen über die Sinne wahrnehmen (vgl. ebd., 137). Die Speicherung von Informationen geht einher mit einer Veränderung der Synapsen zwischen den Nervenzellen. Nervenzellen, die gleichzeitig elektrische Signale aussenden, stärken ihre Verbindung zueinander. Damit werden beispielsweise propositionale Einstellungen im Gehirn stabilisiert. Im Gehirn gibt es bestimmte Regeln, um beispielsweise Wahrnehmung oder Erinnerungen als elektrische Aktivität zu codieren. Einzelne Neuronenensembles speichern bestimmte Aspekte eines Gedankens. Sie werden mit weiteren Neuronenensembles nach einer hierarchischen, pyramidalen Struktur vernetzt: an der Basis werden allgemeine Informationen prozessiert, an der Spitze dagegen die

speziellen (vgl. Fried et al. 2007; Churchland 1997, 124f.; Singer 2002b, 134). Denken verschiedene Menschen an denselben Gegen-stand, werden in ihren Gehirnen ähnliche neuronale Signalmuster aktiviert. Aktuellen Studien zufolge geht jeder Gedanke mit einem eigenen Muster von Hirnaktivität einher (vgl. Haynes 2008; Shinkareva/Mason/Malave/Wang/Mitchell et al. 2008).

Daneben scheint sich zunehmend die Vorstellung eines „dynamischen Gehirns" zu etablieren, wonach neuronale Aktivität beispielsweise nach raum-zeitlichen Mustern erfolgt. Wahrnehmen, Erinnern und Denken entstehen durch eine Synchronisierung verschiedener Bereiche des Gehirns. Die Neuronen bzw. neuronale Ensembles werden auf einen Stimulus hin in einem rhythmischen Gleichtakt bzw. präzisen Sequenzen aktiviert. Neben der funktionalen Rolle der einzelnen Neuronen ist demnach ihr zeitliches Zusammenwirken relevant. Das Gehirn arbeitet sozusagen wie ein Orchester, das immer neue Variationen spielt (vgl. Singer 1999; Singer et al. 2004; Haenschel et al. 2007).

Aus dieser empirischen, neurowissenschaftlichen Perspektive scheinen naturgemäß physische Eigenschaften allein kausal relevant und hinreichend für die Wirksamkeit der mentalen Zustände zu sein. In seiner stärksten Form fordert deshalb der eliminative Materialismus, dass sämtliche mentale Zustände, wie Glaubensannahmen oder Wünschen, mit denen wir Voraussagen treffen oder Verhaltensweisen erklären, nicht existieren. Sie werden lediglich als Setzungen einer fehlerhaften und fiktiven Theorie dargestellt, die von richtigem Sprachgebrauch abgelöst werden können (vgl. Churchland 1981; Rorty 1979; Stich 1996, 3). Um mentale Kausalität zu erklären, reicht demnach die Klärung der intrinsischen kausalen Verbindung beispielsweise zwischen den physischen Eigenschaften $P_1$ und $P_2$ bzw. von Neuron zu Neuron.

Somit stellt sich die Frage, ob das Problem der mentalen Verursachung nicht einfach an die Neurowissenschaft übergeben werden kann. Wenn man sich allerdings die Bestimmung des Neurons und seine Funktion vor Augen führt, scheint es notwendig, genau zu hinterfragen, inwiefern mentale Zustände auf physischen supervenieren sollen oder durch diese realisiert werden. Ein Modell, welches einzelne Neuronen mit einzelnen physischen Zuständen gleichsetzt, erscheint unzureichend, um einen mentalen Zustand der Art „Heute abend gehe ich ins Kino" zu realisieren. Dies zeigte bereits der Einwand gegen die Identitätstheorie, wonach mentale und physische Eigenschaften nicht zwingend koextensiv sind (vgl. Kap. 2.1.3). Die Disposition ist zu komplex und scheint sich aufgrund ihres Inhalts einer Reduzierung zu entziehen. Die Aussage, dass verschiedene Neuronen elektrochemische Impulse aussenden, wird dem komplexen Plan des Kinobesuchs nicht gerecht. Vielversprechender wäre dagegen ein Modell, welches von neuronalen Ensembles ausgeht, die mentale Zustände realisieren. Hier ließe sich eher vorstellen, dass verschiedene Neuronenensembles in verschiedenen Hirnarealen mentale Zustände und natürlich kausale Verknüpfungen zwischen ihnen erzeugen. Es ist, wie gesagt, zumindest umstritten, ob

geistige Strukturen bzw. mentale Zustände auf der Großhirnrinde zu lokalisieren sind. Damit ist nicht anzunehmen, dass das Problem der mentalen Kausalität auf einer rein empirischen Ebene gelöst werden kann.

### 3.2.4 Die Erklärungslücke

Aus der Betrachtung der Merkmale des Mentalen sowie der Merkmale des Physischen bleibt auch unter Einbeziehung eines neurowissenschaftlichen Standpunkts die Schwierigkeit bestehen, mentale Eigenschaften auf physische Eigenschaften zu beziehen, vice versa. Wie soll es möglich sein, diese beiden so verschiedenen Erklärungsbereiche aufeinander zu beziehen bzw. innerhalb einer einzigen kausalen Erklärung zu beschreiben? Es gibt, so scheint es, eine „explanatorische Lücke" zwischen beiden Bereichen.[68]

Geht man von einer physischen Ursache aus und verfolgt die Kette ihrer Wirkungen, so bleibt man zunächst immer im physischen Bereich: wird z. B. Licht von einer roten Fläche reflektiert, trifft auf die Netzhaut eines Betrachters, wird dort als optischer Reiz in einem elektrochemischen Prozess in einen elektrischen Impuls umgewandelt und gelangt über die Nervenbahnen ins Gehirn, um dort weitere physiologische Wirkungen hervorzurufen. Die explanatorische Lücke manifestiert sich zunächst in der Frage: Wie kommt es von der physiologischen Hirnaktivität zur Rotempfindung des Betrachters?

Bei der Erklärungslücke der Identität phänomenaler und materieller Eigenschaften handelt es sich um einen Spezialfall. Im Unterschied zur Reduktion vorwissenschaftlicher Begriffe, wie z. B. Wasser auf $H_2O$ oder Temperatur auf kinetische Energie, lässt sich die Identität phänomenaler und materieller Eigenschaften womöglich weder a priori noch a posteriori ermitteln. Grund hierfür ist, dass phänomenale Konzepte nicht direkt mit Beschreibungen kausaler Rollen in Verbindung gebracht werden. Sie setzen mit der Beschreibung jenes *What is it like* eines Thomas Nagel (vgl. 1974) einen anderen Schwerpunkt als gängige Ursache-Wirkung-Relationen:

> If we are not thinking of pain as something with certain physical causes and effects, but as something that feels a certain way, then we find ourselves quite unable to offer any explanation of why brains yield pains. (Papineau 2004, 149)

[68] Das Problem der explanatorischen Lücke bzw. Erklärungslücke wurde ursprünglich von Joseph Levine (1983, 1993) eingeführt, wonach uns materialistische Erklärungen mit der Reduktion phänomenaler Zustände eine Lücke offen lassen. Egal wie kohärent eine Theorie ist, sie kann uns nicht erklären, warum die materielle Eigenschaft M zur phänomenalen Eigenschaft C führt. Sie kann nicht erklären, warum es sich so und nicht anders anfühlt, die materielle Eigenschaft M zu haben. Die Erklärungslücke betrifft beide Fälle, in denen z. B. Schmerzen a) direkt mit dem Feuern von spezifischen Neuronen im parietalen Kortex oder b) mit physisch realisierten höheren Eigenschaften identifiziert werden.

Letztlich handelt es sich bei der explanatorischen Lücke um eine epistemologische Lücke. Es ist uns einfach nicht möglich, über kausale Rollen a priori die Identität phänomenaler und materieller Eigenschaften herzuleiten. Selbst wenn wir eine kausale Rolle für phänomenale Eigenschaften identifizieren könnten, wären wir dazu nicht in der Lage. Evidenz für die Identität ergibt sich lediglich aus der Korrelation phänomenaler und materieller Eigenschaften in Verbindung mit dem kausalen Argument (vgl. Kap. 2.1.2), sowie dem analogen Vergleich zu anderen akzeptierten Identitäten, wie z. B. der Tatsache, dass Tullius und Cicero dieselbe Person ist. Es gibt für eine materialistische Position keine Alternative, als diese Identität als nackte Tatsache zu akzeptieren (vgl. Papineau 2004, 158).

Die explanatorische Lücke ergibt sich für Kutschera teilweise aus Unkenntnis einschlägiger Gesetze, teilweise aus einer falschen Analyse der Kausalbeziehungen (vgl. Kutschera 2003, 78ff.). Die Lücke entsteht, wenn man den physiologischen Vorgang des Sehens ohne geeignete psychophysische Gesetze analysiert und die neuronalen Vorgänge mit bewusstem Erleben verknüpft. Zwar sind verschiedene psychophysische Gesetze auf der Makroebene alltäglicher Betrachtungen bekannt, aber nur wenig informative Gesetze auf der Mikroebene. Aufgrund unserer Erfahrungen sind alltägliche Kausalbeziehungen intuitiv verständlich, Wechselwirkungen mit Vorgängen im zentralen Nervensystem dagegen unanschaulich. Das lediglich geringe Wissen über die Natur psychophysischer Korrelationen erschwert auch das Verständnis bzw. die Überwindung der explanatorischen Lücke. Zudem müssen psychophysische Korrelationen sehr komplex sein, da die psychologischen Unterscheidungen oft anderen Kriterien unterliegen, als die neurologischen Unterscheidungen. Allein die physische Spezifizierung eines biologischen Vorgangs, wie z. B. die Entfaltung einer Blüte, ist praktisch äußerst schwierig, wenn überhaupt möglich. Wesentlich problematischer muss dagegen die Spezifizierung neuronaler Korrelate für subjektive, psychische Zustände sein – zumal solche Korrelate nicht durchgängig angenommen werden können und darüber hinaus auch nicht exakt extensionsgleich sein müssen (vgl. ebd.).[69]

Um sich der begrifflichen Dichotomie zwischen mentalen und physischen Phänomenen und den damit verbundenen definitorischen Schwierigkeiten, wie letztlich auch der Erklärungslücke, zu entziehen, schlägt Searle im Unterschied zu den „Standardpositionen" (vgl. Searle 2001, 59ff.) des Dualismus und des Materialismus eine biologische Naturalisierung geistiger Phänomene bzw. des Bewusstseins vor. Dieses lässt sich nicht reduzieren oder ist metaphysisch verschieden, sondern muss als biologische Tatsache verstanden und akzeptiert werden. Folglich sollen sowohl der Dualismus als auch der Materialismus mit seinem immanenten Dualismus überwunden werden. Der

[69] Weiteres Beispiel für die explanatorische Lücke ist der Vorgang des„Armhebens". Hier unterscheidet Kutschera zwischen den physiologischen begleitenden Vorgängen, die – ebenso wie Absichten – keine echten Ursachen sind. Handlungen sind – im modallogischen Sinn – Erstursachen und haben selbst keine Ursachen. Handlungen haben lediglich Gründe (vgl. Kutschera 2003, 79; Kap. 7).

biologische Naturalismus kann insofern als eine monistische Position verstanden werden, als geistige Phänomene sich nicht auf physische Tatsachen reduzieren lassen, da sie selbst physische bzw. biologische Tatsachen sind. Searle will zeigen, dass das Vokabular (Dualismus, Monismus, Materialismus, Physikalismus) überholt ist. Wer von Reduzierbarkeit oder von Irreduzibilität spricht, befindet sich automatisch in diesem falschen Begriffssystem (vgl. Searle 1993, 16). Er will sich auf wirklich bewusste Zustände als innere, subjektive, qualitative, erstpersönliche Geisteszustände beziehen, als Begriff des Bewusstseins (vgl. ebd., 22).

Searle empfiehlt, sich an den biologischen Tatsachen zu orientieren. So ist das Gehirn ein biologisches Organ und das Bewusstsein ein biologischer Vorgang. Demnach werden Bewusstseinszustände zwar von Hirnvorgängen verursacht, diese Tatsache müssen wir anerkennen. Diese Kausalbeziehung zwingt uns aber nicht in einen Dualismus. Bewusstseinszustände haben spezielle Merkmale, sind aber dennoch ein eigenständiges und durchweg biologisches Phänomen:

> Bewusstsein wird von Hirnvorgängen verursacht und ist ein höherstufiges Merkmal des Gehirns. (vgl. Searle 2001, 70)
>
> Bewusstsein ist, kurz gesagt, ein biologisches Merkmal des Menschenhirns und des Hirns gewisser anderer Lebewesen. Es wird durch neurobiologische Vorgänge verursacht und ist ein Bestandteil der natürlichen biologischen Ordnung wie jedes andere biologische Merkmal (Photosynthese, Verdauung, Mitose). (Searle 1993, 109)

Dies beschreibt im Kern Searles Ansicht über die metaphysische Beziehung zwischen Bewusstsein und Gehirn – unter Einbeziehung der vorherigen Merkmale. Somit ist Bewusstsein nur bedingt vergleichbar mit anderen biologischen Phänomenen, wie z. B. Stoffwechsel. Sie unterscheiden sich allein dadurch, dass sich Bewusstsein nicht auf Mikro-Phänomene zurückführen lässt, da immer ein irreduzibles subjektives Element übrig bleibt. Die subjektiven Erlebnisse sind ein wesentliches Merkmal von Bewusstsein, können also per definitionem nicht ausgeschlossen werden. Trotz einer vollständigen Darstellung des Kausalgeschehens auf der neurobiologischen Basis bleibt das subjektive Element irreduzibel übrig.[70]

Mit dieser Position der Naturalisierung scheint der Dualismus zwischen mental und physisch zwar zunächst aufgehoben zu sein, allerdings führt Searle selbst die Begrifflichkeiten in ihrer Polarität durch seine gesamte Argumentation weiter fort. Außerdem verschiebt sich ähnlich wie bei der Typenidentität die Problematik, da nun, wenn das Mentale bzw. das Bewusstsein naturalisiert ist, immer noch eine Unterscheidung zwischen Ereignissen oder Eigenschaften, die nur physisch, und solchen, die auch mental sind, getroffen werden muss. Dies kann Searle jedoch

[70] Bewusstsein ist demnach kein Beobachtungsgegenstand wie andere Dinge, z. B. Stühle oder Tische, und lässt sich deshalb schwer beschreiben. Folglich lassen sich die subjektiven Erlebnisse des Bewusstseins nicht „herauskürzen", schließlich sind sie zentral für den Bewusstseinsbegriff und das Erste-Person-Phänomen und gerade deshalb Gegenstand des Interesses (vgl. Searle 2001, 74).

nicht leisten. Seine Position hat in der Folge zumindest ähnliche Begründungsschwierigkeiten wie die Position der Typenidentität. Letztlich manifestiert sich in der Erklärungslücke auch auf der epistemischen Ebene die Schwierigkeit der gängigen Position, die Bereiche des Mentalen und des Physischen in der kausalen Frage miteinander zu verknüpfen.

***Quintessenz Unterscheidung mental versus physisch***

- *konkurrierende Beschreibungen der Welt auf epistemischer Ebene*
    - *Kennzeichnung mentaler Kausalität auf abstrakter, inhaltlicher sowie subjektiver Ebene durch verschiedene Merkmale des Mentalen*
    - *Bestimmung des Gehirns als empirischen Ort für mentale Zustände durch verschiedene Merkmale des Physischen*
        - *Gleichsetzung von „neuronal" und „physisch" nicht zwangsläufig*
- *Erklärungslücke als epistemische Lücke vor dem Hintergrund fehlender psychophysischer Gesetze*
- *mentale Kausalität als konzeptionelle Frage: biologisches Phänomen umfasst beide Bereiche (mental; physisch); kein rein physisches Modell hinreichend*

### 3.3 Die kausale Geschlossenheit der Physik

Die gegenwärtige philosophische Diskussion bewegt sich im Rahmen eines erkenntnistheoretischen Realismus, der vom Naturalismus der Naturwissenschaften geprägt ist: „It is widely believed that the most fundamental facts about our universe are physical facts, and that all other facts are dependent on these." (Putnam 1975a, 32) Demnach wird keine ein physisches Ereignis involvierende Kausalkette je die Grenze des Physischen zum Nicht-Physischen bzw. Mentalen überschreiten. Zentrales Argument ist die Annahme einer Erklärungsvollständigkeit der physikalischen Theorie. Die physische Welt ist in sich geschlossen, und alle Fragen, die sich stellen, können innerhalb der physikalischen Terminologie beantwortet werden. Das Prinzip der kausalen Geschlossenheit gehört zur Kernargumentation der Typenidentität, wurde aber auch von den übrigen Positionen übernommen und weitergeführt (vgl. Kap. 2.1; Kap. 2.2; Kap. 2.3; Kap. 2.4).

Jede physische Wirkung hat also einen hinreichenden physischen Grund. Die Ursachen für den Vulkanausbruch des Vesuv sind mit verschiedenen geologischen Bedingungen erklärbar, unter anderem einer tektonischen Schwächezone der Erdkruste und dem Aufsteigen des Magma durch gespannte Gase. Rein theoretisch ist die kausale Verkettung lückenlos physisch überprüfbar, es bedarf nicht eines zornigen Gottes, um den Vulkanausbruch zu erklären. Diese metaphysische Position ist zugleich methodologische Voraussetzung der physischen Wissenschaften.

Das Prinzip der kausalen Geschlossenheit der Physik besagt also, dass alle physischen Wirkungen auf physische Ursachen zurückzuführen sind. Dieses Prinzip basiert auf der weniger starken These der Vollständigkeit der Physik, die zunächst eine reine Lehre über den physischen Bereich ist.[71] Danach kann jegliche physische Wirkung innerhalb des Bereiches der Physik hinreichend erklärt werden, ohne dass dieser Bereich überschritten werden muss:

> All physical effects are fully determined by law by prior physical occurences. (Papineau 2001, 8; vgl. auch ders. 1993, 16)[72]

Diese These besagt noch nicht, dass alles, einschließlich prima facie nicht-physische Dinge, physisch ist. Insofern kann die These der Vollständigkeit der Physik, im Sinne einer umfassenden Erklärung aller Dinge durch die Physik, von der These der kausalen Geschlossenheit der Physik unterschieden werden. Letztere These allein reicht zwar nicht aus, um den Physikalismus durchzusetzen. Allerdings ist die weitere Argumentation hierfür naheliegend, da, wenn alle physischen Wirkungen hinreichend auf physische Ursachen zurückzuführen sind, alles, was eine physische Wirkung hat, selbst physisch sein muss (vgl. Kim 1998, 40). Es ist kein Raum mehr übrig für nicht-physische Dinge oder Ereignisse, um einen Unterschied in der physischen Wirkung zu machen.[73]

Trotz des Primats der Physik wird in der aktuellen philosophischen Debatte versucht, die psychologische Fragestellung im physikalischen Kontext zu berücksichtigen. Gemeinhin wird als minimaler Physikalismus die Grundannahme akzeptiert, dass mentale Eigenschaften durch physische Basen realisiert sind:

> Why [...] this special deference to physical theory? […] The answer is not that everything worth saying can be translated into the technical vocabulary of physics; not even that all good science can be translated into that vocabulary. The answer is rather this: nothing happens in the world, not the flutter of an eyelid, not the flicker of a thought, without some redistribution of microphysical states. (Quine 1982, 98)

Mit der Forderung, dass das Mentale physisch beschreibbar ist, will der Physikalismus mentale Verursachung mit denselben Prinzipien erklären. Die Beantwortung dieser Frage stellt jedoch

---

[71] Das Prinzip variiert seine Form innerhalb der Debatte um mentale Kausalität. Yablo spricht von einem „physischen Determinismus", auch wenn die kausale Geschlossenheit von einem kausalen Determinismus unterschieden wird (vgl. Yablo 1992). So kann das Prinzip modifiziert werden, um mit einem Indeterminismus kompatibel zu sein. Anstelle von einer „hinreichenden physischen Ursache" kann z. B. von einer„physische[n] Ursache hinreichend für die objektive Wahrscheinlichkeit eines Ereignisses" gesprochen werden.

[72] Eine Auflistung aller Ansätze, die das Prinzip zu formulieren, findet sich in Lowe 2000.

[73] Papineau bezieht sich auch auf die Vertreter des Physikalismus in den 1950er und 1960er Jahren. Sie hätten implizit ähnlich argumentiert. So werden mentale Zustände von Armstrong und Lewis über ihre kausalen Rollen definiert und mit physischen Zuständen als identisch betrachtet, da letztere diese Rollen z. B. als Ursachen von Verhalten ausfüllen. Würde man auf die Annahme der Vollständigkeit der Physik verzichten, würden nicht die physischen Zustände, sondern die mentalen selbst diese Rollen ausfüllen (vgl. Papineau 2001, 8f.).

immer noch eines seiner Hauptprobleme dar. Dennoch ist es Aufgabe der Philosophie, die Fragen nach Dingen, Tatsachen und Eigenschaften zu untersuchen, da diese Fragen von den Wissenschaften unbeantwortet bleiben (vgl. Gadenne 1997). Allerdings scheint die Akzeptanz der Abgeschlossenheit des physischen Bereichs zwangsläufig zu einer Ablehnung der mentalen Verursachung zu führen. Selbst für die Position eines Eigenschaftsdualismus scheint die kausale Wirksamkeit von mentalen Eigenschaften ausgeschlossen zu sein, da diese eben in einem Widerspruch zu der Annahme der physischen Geschlossenheit steht (vgl. ebd.).

### 3.3.1 Theorienreduktion und induktive Generalisierung

Der geradlinigste physikalistische Ansatz in diesem Kontext diskutiert die Reduktion der psychologischen Theorie auf die physikalische. Demnach wird die menschliche Psychologie als ein Theoriesystem begriffen, das mikroreduzierbar auf die Neurowissenschaft ist. Man kann diesen Ansatz durchaus in den Kontext der so genannten „Einheitstheorie“ (englisch: unified theory; vgl. Kim 2003; Kap. 2.1.1) stellen und mit dieser vergleichen. Es sind die so genannten Brückengesetze, welche erst die Theorienreduktion ermöglichen. Eine Theorienreduktion ist beispielsweise möglich aufgrund derselben kausalen Rolle, welche die Prädikate F und G in den Theorien $T_1$ und $T_2$ einnehmen. Notwendige Vorraussetzung ist allerdings, dass F und G nomologisch koextensional sind (vgl. Nagel 1961; Kap. 2.3.2). Das übliche Beispiel für Theorienreduktion ist die klassische Thermodynamik, die mikro-reduzierbar auf die statistische Molekularmechanik ist. Innerhalb der Wissenschaftstheorie wird Mikroreduktion als Form der Erklärung aufgefasst. Die Theorie auf der höheren Ebene ist unter der Anwendung von Brückengesetzen von der Theorie auf der niedrigeren Ebene ableitbar – die nomische Koextension der theoretischen Eigenschaften der komplexen Entitäten auf der höheren Ebene (z. B. Temperatur) wird mit den theoretischen Eigenschaften auf der niedrigeren Ebene einschließlich ihrer Mikrobestandteile (z. B. kinetischer Energie) verrechnet. Eine psychologische Theorie kann demnach auf die physikalische Theorie der entsprechenden Spezies, wie z. B. Menschen, Reptilien oder Marsianer reduziert werden. Nach Kim kann diese Reduktion über lokale Brückengesetze erfolgen, die er verschiedenen empirischen Einwänden zum Trotz zumindest metaphysisch für möglich hält. Schließlich folgt aus der These physischer Realisierung, dass mentale Zustände eines Systems auf den physischen Zuständen beruhen müssen (vgl. Kim 1992; Kap. 2.4.2).

Es ist die gängige Meinung, dass die Brückengesetze, obwohl sie empirisch hergeleitet werden, die tatsächliche Identität der betroffenen Eigenschaften offenbaren. Beispielsweise ist die Temperatur von Gas lediglich molekulare kinetische Energie. Würden Brückengesetze nicht die zugrunde liegende Eigenschaftsidentität betreffen, wären sie nicht hinreichend für eine vollstän-

dige Reduktion. Und dies wird eben von Seiten nicht-reduktiver Physikalisten bezweifelt. So ist für Davidson aufgrund der unterschiedlichen Extension ihrer Prädikate eine nomologische Identität mentaler und physischer Eigenschaftstypen ausgeschlossen, die für eine Reduktion erforderlichen Brückengesetze können nicht aufgestellt werden (vgl. Kap. 2.1.3; Kap. 2.3.2). Argumentiert wird ferner über das Prinzip der Multirealisierbarkeit. Wenn mentale kausale Eigenschaften von verschiedenen physischen Eigenschaften vielfach realisierbar sind, kann die Psychologie nicht über Brückengesetze, welche die nomologische Koextension einer jeden mentalen Eigenschaft mit einer konkreten physischen bestimmen, auf physikalische Wissenschaft reduziert werden. Die kausale Bedeutung mentaler Eigenschaften bliebe somit erhalten (vgl. Kap. 4.2.1).

Papineau versteht vor diesem Hintergrund die starke Physikalismusthese als ein empirisch begründetes Argument, das sich aufgrund wissenschaftlicher Evidenz in einer wissenschaftshistorischen Entwicklung durchgesetzt hat (vgl. Papineau 2001, 7). [74] Um das Prinzip empirisch zu stärken, wird auch eine induktive Generalisierung formuliert, wonach in jedem Fall, in dem die Naturwissenschaft eine Ursache für ein physisches Ereignis entdeckt hat, jene eine hinreichende physische Ursache für das Ereignis war (vgl. auch Papineau 1993). Das Prinzip hat also eine induktive Basis insofern, als die Wissenschaftsgeschichte gezeigt hat, dass vorwissenschaftliche Erklärungen rein wissenschaftlichen Erklärungen weichen, die mit rein physischen Dingen oder Ereignissen operieren. So hat die wissenschaftliche Evolutionstheorie die vornehmlich religiös geprägte Schöpfungstheorie als Erklärung für die Anpassung und Abstimmung der Organismen an ihre Umwelt abgelöst (vgl. Darwin 1859; Dawkins 1987, Kap. 3).

Die Geschlossenheit der Physik ist zwar selbst kein physisches Gesetz, aber eine hochwertige empirische Generalisierung über die Praxis und den Inhalt physischer Wissenschaft. Demnach ermöglicht die Vereinigung der wissenschaftlichen Theorien einen wahren und erschöpfenden Zugang zu allen physischen Phänomenen. Die Theorien sind in einem kumulativen Sinne vereinigt. Eine Theorie, die ein bestimmtes physisches Phänomen herleitet, wird durch Theorien erklärt, die jene Phänomene herleiten, aus denen dieses Phänomen sich zusammensetzt. Selbiges gilt für diese letzteren Phänomene, bis hin zu den fundamentalen Partikeln bzw. Feldern, die durch einige wenige Gesetze hergeleitet werden (vgl. Lewis 1983a, 105). Einerseits werden Entitäten nach dieser Annahme vollständig durch physische Teile zusammengesetzt. Andererseits werden die Theorien, die das Gesamte erklären, selbst durch jene Theorien erklärt, die die Teile erklären, aus denen sich das Gesamte zusammensetzt. Nicht-Physikalisten haben wohlgemerkt

[74] Alternativ wird die Geschlossenheit der Physik als Arbeitshypothese der Naturwissenschaft betrachtet, die der Physikalismus voraussetzt (vgl. Lewis 1983a, 105; Kim 1993a, 290). Ferner weist Papineau darauf hin, dass zeitgenössischer Physikalismus als eine ontologische und nicht als eine methodologische Lehre verstanden werden sollte. Damit distanziert er sich von Forderungen nach einer Einheitswissenschaft, die von Vertretern des logischen Positivismus zu Beginn des 20. Jahrhunderts gefordert wurde. Ontologischer Physikalismus impliziert keine spezielle Methodologie. Dies wird auch durch seine verschiedenen Spielarten deutlich, wie z. B. physische Supervenienz, physische Realisation oder physische Token-Tokenidentität. „These more sophisticated doctrines leave plenty of room for different sciences to be studied in different ways". (Papineau 2001, 4.)

die Möglichkeit, diese Form der Theorienreduktion zu akzeptieren, ohne jedoch das Prinzip der kausalen Geschlossenheit akzeptieren zu müssen. Sie können ihre These so formulieren, dass physische Teile nicht-physische Eigenschaften haben und diese kausal verantwortlich sind. Auf diese Weise können sie weiterhin verneinen, dass alle Wirkungen ausschließlich physische Ursachen haben.

Allerdings ist mit dem Prinzip der Vollständigkeit der Physik und somit auch der kausalen Geschlossenheit ein Dilemma verbunden. Wird mit Physik der Fachbereich beschrieben, der an Universitäten gelehrt und erforscht sowie in Fachzeitschriften diskutiert wird, ist offensichtlich, dass die Physik derzeit nicht als vollständig bezeichnet werden kann. Viele Wissenschaftler würden heute einräumen, dass die Physik längst nicht alle Fragen beantworten kann. Es besteht immer noch eine Diskussion über fundamentale Fragen der Physik und es ist fraglich, inwiefern Physik zukünftig neue Kategorien für physische Ursachen ermitteln wird (vgl. Kuhn 1996; Montero 2003; Rescher 1985). Andererseits kann Physik als der Fachbereich solcher zukünftiger wissenschaftlicher Theorien verstanden werden. Dann hätten wir keinen Grund, deren Vollständigkeit zu bezweifeln, da wir diese derzeit nicht beurteilen können (vgl. Papineau 2001, 12). Diese terminologische Unterscheidung macht deutlich, dass es für das kausale Argument (vgl. Kap. 2.1.2) nicht notwendig ist, exakt zu bestimmen, was die Vollständigkeit der Physik beinhaltet, sondern es wichtiger ist, zu bestimmen, was sie nicht beinhaltet.

### 3.3.2 Die physikalische Geschlossenheit in der Kritik

Jemand, für den Physikalismus die Lösung des Körper-Geist- bzw. Gehirn-Bewusstsein-Problems beinhaltet, muss nicht durch das Argument der Geschlossenheit der Physik beeindruckt sein, Dafür ist das Argument zu naheliegend an der grundsätzlichen physikalistischen Argumentation, um zu überzeugen. Zudem müsste jemand, der nicht von der Wahrheit des Physikalismus hinsichtlich des Mentalen überzeugt ist, kaum die Behauptung akzeptieren, dass alle physischen Wirkungen physische Ursachen haben.

Die kausale Geschlossenheit der Physik folgt auch nicht aus philosophischen Behauptungen heraus, selbst wenn die Essenz der Kausalität ein Energiefluss wäre (vgl. Crane/Mellor 1990, 191f.; Hart 1988, Kap. 9). Wird physisch nicht trivial verwendet, scheint die Geschlossenheit eine substanzielle empirische Vermutung zu sein, der zumindest entscheidende Unterstützung fehlt. Vor diesem Hintergrund ist es verständlich, wenn ein Nicht-Physikalist die Geschlossenheit verneint und behauptet, dass einige physische Wirkungen nicht-reduzierbare mentale Ursachen hätten. Dies beinhaltet nicht zwangsläufig, dass Handlungen die Gesetze der Physik brechen.

Wenn es Gesetze gibt, die mentale und physische Phänomene miteinander verbinden, müssen diese Gesetze mit denen der Physik konsistent sein.

Grundsätzlich ist die Voraussetzung der kausalen Geschlossenheit der Physik nicht unumstößlich. Tatsächlich wird bezweifelt, ob das physikalische Weltbild mit all seinen Konsequenzen als maßgeblich zu akzeptieren ist oder nicht als ideales Modell gleichberechtigt neben anderen steht:

> [...] Many different world versions are of independent interest and importance, without any requirement or presumption of reducibility to a single base. The pluralist, far from being anti-scientific, accepts the sciences at full value. His typical adversary is the monopolistic materialist or physicalist who maintains that one system, physics, is pre-eminent and all-inclusive, such that every other version must eventually be reduced to it or rejected as false or meaningless. (Goodman 1978, 4)

Es ist fraglich, ob ein physikalisches Kausalitätsmodell einfach auf psychologische Phänomene übertragen werden kann. Schließlich ist zweifelhaft, dass die rein physikalische Erklärung den psychologischen Prozessen gerecht wird. „Brain sciences are concerned with patterns that are not as sensitive to individual-enviromental relations as the patterns that concern psychology." (Burge 1989, 310)

Zudem zeigen sich Schwierigkeiten, den Begriff „physisch" einzugrenzen (vgl. Kap. 3.2.2). Denn auch chemische und biologische Eigenschaften lassen sich als Makroeigenschaften auf eine mikrophysische Ebene reduzieren (vgl. Burge 1999, 103). In diesem Sinne lassen sich die neuronalen Eigenschaften weiter reduzieren. Das System läuft Gefahr, in einem Regress zu enden. Ebenso ist es nicht wirklich einsichtig, warum mentalen Eigenschaften auf einer psychologischen Ebene, auf der sie durchaus eine Erklärungskraft besitzen, weniger kausale Kraft zugestanden wird, als biologischen, chemischen oder makrophysikalischen Erklärungen. Intertheoretische Reduktion betrifft letztlich alle Wissenschaften, die als Einzelwissenschaften systematisch auf die Physik reduzierbar sind (vgl. Fodor 1997). Konsequenterweise ist jede Makroeigenschaft auf eine Mikroeigenschaft reduzierbar, also ist keine Form der Makrokausalität gültig. Eine solche metaphysische Annahme untergräbt ironischerweise nicht nur die Alltagspsychologie, sondern das gesamte wissenschaftliche Weltbild überhaupt (vgl. Baker 1999, 94). Zudem wurde bereits deutlich gemacht, dass eine Theorienreduktion problematisch ist, da die erforderlichen Brückengesetze aufgrund der unterschiedlichen Extension mentaler und physischer Eigenschaften bzw. der Multirealisierbarkeit der ersteren nicht möglich sind (vgl. Kap. 3.3.2).

Das Argument der kausalen Geschlossenheit der Physik wird gerade in Hinsicht auf die kausalen Beziehungen gestärkt. So liefert die Physik hinreichende Erklärungen für die Relation zwischen Ursache und Wirkung. Alle physischen Wirkungen werden nämlich – unter ein physisches Gesetz

fallend – von physischen Ursachen bestimmt. Dieses Argument ist die Basisannahme, auf der die übrigen physischen Argumente, vor allem die Zurückweisung einer Überdetermination kausaler Ereignisse und das Vorhandensein strikter Gesetze, ruhen. So begründet Chalmers die kausale Geschlossenheit der physischen Welt mit evidenten Ergebnissen der Naturwissenschaften: „The best evidence of contemporary science tells us that the physical world is more or less causally closed." (Chalmers 1997, 125) Dagegen lässt sich einwenden, dass sich Chalmers in diesem Satz selbst relativiert, da die Aussagen der Wissenschaft nur auf eine mehr oder weniger geschlossene physikalische Welt verweisen. Außerdem wird hier ein wichtiger Theoriebereich zeitgenössischer Wissenschaft ausgeklammert: „In der Quantenphysik geht man davon aus, dass jede Beobachtung eines mikrophysischen Geschehens einen Eingriff in dieses Geschehen darstellt. Sie betrachtet also gerade nicht geschlossene Systeme." (Kutschera 2003, 77) In der Konsequenz kann die Annahme einer kausalen Geschlossenheit der Physik auch nicht ein Prinzip dieser naturwissenschaftlichen Disziplin sein. Physikalische Theorien werden dagegen aufgrund neuer Erkenntnisse auch weiterhin modifiziert und erneuert. Verschiedene zeitgenössische Philosophen sind deshalb in einer radikalen Kritik besonders vorsichtig, was die Definition des Materialismus anbelangt. Einige sind sogar der Ansicht, dass die gesamte Idee „leer" sei und es deshalb keinen geeigneten Weg gebe, den physischen Bereich zu bestimmen (vgl. Crane/Mellor 1990; Crane 1991; Segal 2000; alle zitiert nach: Papineau 2004, 15).

***Quintessenz kausale Geschlossenheit der Physik***

- *Annahme der Erklärungsvollständigkeit aller physischen Ereignisse durch physische Ereignisse bzw. Eigenschaften (jede physische Wirkung hat eine physische Ursache)*
    - *Zusatzannahme, wonach alle Ereignisse durch physische Ereignisse bzw. Eigenschaften vollständig erklärbar*
- *Umgang mit psychologischen Fragestellungen als Herausforderung: Wie ist das Mentale physisch beschreibbar?*
    - *Reduktion mentaler auf physische Eigenschaften nicht möglich aufgrund unterschiedlicher Extension*
    - *mögliche Reduktion betrifft in letzter Konsequenz alle Wissenschaftsbereiche*
- *starke Physikalismusthese als induktive Generalisierung und hochwertige empirische Vermutung*
    - *ideales Modell; faktische Unvollständigkeit der Position*
    - *alternative Betrachtung nicht geschlossener Systeme durch Quantenphysik*
- *bestehende Schwierigkeit der Übertragung des physischen Kausalitätsmodells auf psychologische Kausalphänomene*

### 3.4 Das Argument des kausalen Ausschlusses

Mit dem Prinzip der kausalen Geschlossenheit der Physik eng verknüpft ist das Argument des kausalen Ausschlusses bzw. das so genannte Ausschlussargument. Dieses Argument ist letztlich eine Verschärfung der These, dass jede physische Wirkung einen hinreichenden physischen Grund hat. „[F]irst, that all events are physical events, and second, that the world of the physical constitutes a closed causal system." (Van Gulick 1999, 242) Wenn jede physische Eigenschaft eine andere physische Eigenschaft verursacht, scheinen weitere (mentale) Eigenschaften für die kausale Wirkung überflüssig zu sein, da die physische Wirkung hinreichend von einer physischen Ursache verursacht wird. Konsequenterweise wird die systematische Mehrfachverursachung bzw. Überdeterminierung ausgeschlossen – eine physische Eigenschaft als Ursache reicht aus, mentale Eigenschaften sind als zusätzliche Ursachen überflüssig.

#### 3.4.1 Begründung des Ausschlussarguments

Das Argument des kausalen Ausschlusses entstand einerseits in Zusammenhang mit der Kritik an Davidsons anomalem Monismus (vgl. Kap. 2.3.3), und andererseits innerhalb der funktionalistischen Diskussion aufgrund des Arguments der Mehrfachrealisierung (vgl. Kap.2.2.5). In der Konsequenz droht beiden Positionen, wie auch nicht-reduktiven Positionen unter Supervenienzbedingungen, dass mentale Ereignisse bzw. Eigenschaften lediglich epiphänomenal, also kausal wirkungslos sind.

Nach Davidsons Tokenidentitätstheorie ist jedes einzelne mentale (Token-) Ereignis nur identisch mit einem einzelnen physischen (Token-)Ereignis. In der Kritik an Davidsons anomalem Monismus wird eingewendet, dass deshalb nur physische Ereignisse bzw. Eigenschaften kausale Wirkung haben. Dies geht aus den beiden Prinzipien hervor, wonach alle Kausalrelationen strikten Gesetzen unterliegen und es keine strikten psychophysischen Gesetze gibt. Deshalb müssen alle Formen einer Körper-Geist-Kausalität unter physischen Gesetzen subsumiert werden (vgl. Kap. 2.3.2; Kap. 3.5.1). In der Folge muss nach Ansicht der Kritiker Davidsons erstes Prinzip abgelehnt werden – mentale Eigenschaften bzw. Ereignisse sind nur mehr epiphänomenal (vgl. McLaughlin 1999). Wie bereits gezeigt wurde, verneint Davidson die Annahme, wonach Ursachen aufgrund ihrer Eigenschaften kausale Wirkung haben. Als Nominalist bezieht sich Davidson überhaupt nicht auf Eigenschaften sondern lediglich auf kausale Ereignisse. Deshalb sind mentale Eigenschaften genauso wenig vom kausalen Geschehen ausgeschlossen, wie physische (vgl. Davidson 1999). Allerdings kann eingewendet werden, dass Davidson der aktuellen Diskussion nicht gerecht wird, wonach Eigenschaften in der Theorie der Kausalität eine zentrale Rolle

spielen (vgl. Kim 1976, 1989; Ehring 1997; Kap. 3.1.2). Ein berühmtes Beispiel ist das von der Sopranistin, die das hohe C singt, woraufhin Glas zerbricht. Hier würde man die Eigenschaft der hohen Frequenz und nicht z. B. die Bedeutungseigenschaft des Liedtextes als ursächlich für das Zerbrechen des Glases ansehen (vgl. Dretske 1988) – und somit auch nicht das Ereignis als solches.[75] Das Argument des kausalen Ausschlusses wird für die Position des anomalen Monismus erst relevant, wenn ein modifizierter Eigenschaftsbegriff – der beispielsweise der Position Kims folgt – verwendet wird. Dann aber wird deutlich, dass mentale Eigenschaften insbesondere aufgrund der Prämisse der strikten Gesetze kausal wirkungslos sind.

In der funktionalistischen Diskussion entsteht das Problem des kausalen Ausschlusses paradoxerweise aufgrund eines Kernarguments des Funktionalismus gegen die Reduktion mentaler Eigenschaften auf physische Eigenschaften. Gemeint ist das Argument der Multirealisierbarkeit (vgl. Kap. 2.2; Kap. 4.2.1). Hiernach kann die Eigenschaft, z. B. Schmerz zu haben, von einer Vielzahl physischer Eigenschaften realisiert werden. Jemand befindet sich in einem Schmerzzustand, da er sich in einem Zustand mit dem richtigen kausalen Profil, beispielsweise einem konkreten neuronalen Zustand befindet. Der Schmerzzustand und der neuronale Zustand sind jedoch nicht identisch, da Lebewesen mit anderen physischen Konditionen sich auch in dem Schmerzzustand befinden können (vgl. Kim 1992, Heil 1999).[76] Mentale Eigenschaften sind Eigenschaften einer multirealisierbaren höherstufigen Ebene von Gegen-ständen oder Objekten. Diese besitzen jene aufgrund ihrer Eigenschaften auf der niedrigstufigen Ebene – wobei die Eigenschaften auf der niedrigstufigen Ebene die Realisierer der Eigenschaften auf der höherstufigen Ebene sind (vgl. Kap. 3.1.3). Hinsichtlich des Ausschlussarguments stellt sich allerdings die Frage, inwiefern mentale Eigenschaften als Eigenschaften des höheren Levels eine kausale Wirkung haben können, wenn die eigentliche kausale Kraft bei ihren Realisierern liegt. Wenn ein Schmerzzustand durch die funktionale Eigenschaft F repräsentiert wird, F in einem Fall aber durch die neuronale Eigenschaft N realisiert wird und N für verschiedene Verhaltensweisen, wie z. B. das Einnehmen einer Schmerztablette, kausal relevant ist, scheint die funktionale Eigenschaft F kausal überflüssig zu sein (vgl. Kim 1989, 1998, 1999b, 2003).[77] Die physischen Eigen-

[75] In einer Erwiderung ließe sich aber fragen, ob hier nicht auf unterschiedlichen Ebenen diskutiert wird. So befindet sich eine Ereigniskausalität in Davidsons Sinn auf einer rein ontologischen Ebene, während die Kausalität der Eigenschaften zumindest teilweise in die epistemische Ebene hineinreicht – anders wäre eine erkennende Unterscheidung zwischen unterschiedlichen Eigenschaften nicht möglich (vgl. Kap. 7.2).

[76] Das Argument der Multirealisierbarkeit ist zunächst ein modales Argument, da die geforderte Identität von Eigenschaften nicht für alle möglichen Welten gilt. Es ist nicht möglich, einen einzigen physischen Zustand zu bestimmen, der z. B. der Realisierer von Schmerz in allen aktuellen und möglichen Welten ist. Nicht nur die neurologischen Zustände der Menschen unterscheiden sich sehr subtil in ihren physischen Konditionen, sondern auch die der Menschen von anderen Lebewesen, wie z. B. Hunden, Katzen oder Tintenfischen, ganz zu schweigen von Marsmenschen. Trotz dieser unterschiedlichen physischen Konditionen würde man aber immer noch von Schmerzen sprechen (vgl. Kap. 4.2.1).

[77] Hier ließe sich anmerken, dass die Formulierung, dass „Neuronen das Einnehmen von Schmerzmitteln verursachen" streng genommen ein Paradigmenfehler bzw. eine Regelverletzung im Wittgensteinschen Sinne ist, da der „physische" Sprachbereich der neuronalen Eigenschaften mit dem „behaviouralen" Sprachbereich – „ich

schaften scheinen die mentalen Eigenschaften, die durch jene realisiert werden, von jeglicher kausalen Relevanz auszuschließen.

Ähnlich der funktionalistischen Position stellt sich das Ausschlussproblem auch unter der Annahme, dass mentale Eigenschaften auf verschiedenen physischen Eigenschaften supervenieren können. Wie bereits erläutert wurde, respektiert die Supervenienztheorie das Primat des Physischen, da durch die Instanziierung physischer Eigenschaften das Mentale bestimmt wird (vgl. Kap. 2.4.2). Damit verbunden ist auch die Idee, dass mentale Ereignisse neuronale Korrelate haben. Demzufolge haben zwei Organismen, die dieselben neuronalen Ereignisse haben, auch dieselben mentalen Ereignisse. Diese bleiben jedoch von den neuronalen Ereignissen verschieden. Zwar können zwei beliebige Ereignisse, die sich in jeglicher physischer Hinsicht exakt gleichen, sich nicht in mentaler Hinsicht unterscheiden, doch gilt dies nicht auch umgekehrt. Denn wie das Prinzip der Multirealisierbarkeit besagt, können mentale Ereignisse, die sich gleichen, durchaus durch verschiedene physische Eigenschaften realisiert werden (vgl. Kim 1984; McLaughlin 1984). Hier wird abermals die Bedeutung des Eigenschaftsbegriffs für das Argument des kausalen Ausschlusses deutlich.

In Form eines Dilemmas argumentiert Kim (1998, 39ff.), dass mentale Kausalität nicht einsichtig ist. Er argumentiert, dass Körper-Geist-Supervenienz entweder zutrifft oder nicht. Wenn Körper-Geist-Supervenienz nicht zutrifft und wir der kausalen Geschlossenheit der Physik (als Materialisten) verpflichtet sind, ist es nicht sinnvoll möglich, von mentaler Kausalität zu sprechen. Wenn Körper-Geist-Supervenienz jedoch zutrifft, bleibt für mentale Eigenschaften lediglich die Wahl zwischen Epiphänomenalismus und Reduktionismus. Da Kim nicht gegen beispielsweise die Kognitionswissenschaft argumentieren kann, bleibt für ihn als einzig sinnvolle Option die Reduktion der mentalen Eigenschaften auf physische. Dies ist insofern bemerkenswert, als die Supervenienz ursprünglich die Nicht-Reduzierbarkeit jener Eigenschaften sichern sollte. Auf Basis der Supervenienz beschreiben auch verschiedene nicht-reduktivistische Physikalisten bzw. Nicht-Reduktivisten die Relation zwischen mentalen und physischen Phänomenen als eine asymmetrische Abhängigkeitsrelation. Aus den Behauptungen der kausalen Geschlossenheit und der mentalen Kausalität ergibt sich auch für sie das Dilemma, dass entweder mentale Ereignisse eigenständige Ursachen sind, woraus die Ablehnung der Voraussetzung der kausalen Geschlossenheit folgen würde. Oder aber mentale Ereignisse wären kausal wirkungslos, weshalb auch ein nicht-reduktiver Physikalismus in einen Epiphänomenalismus münden würde.

---

nehme ein Schmerzmittel, weil ich Schmerzen habe“ – vermischt wird. Ferner erscheint auch die in der funktionalistischen Diskussion oft implizit vorgenommene Gleichsetzung der neuronalen Eigenschaften mit physischen Eigenschaften insofern problematisch, als hierbei zwischen einer empirischen und einer ontologischen Ebene gewechselt wird.

### 3.4.2 Ablehnung systematischer Mehrfachverursachung

Dass physische Eigenschaften mentale Eigenschaften realisieren, bzw. letztere auf ersteren supervenieren, reicht nicht aus, um mentale Eigenschaften zwingend von kausaler Verursachung auszuschließen. Deshalb wird der kausale Ausschluss durch die Argumentation gegen systematische Mehrfachrealisierung verstärkt. Sowohl die physische als auch die mentale Eigenschaft eines Ereignisses scheint im Falle der Supervenienzbeziehung zunächst als geeignet, Ursache eines Folgeereignisses bzw. einer Folgeeigenschaft zu sein. Angenommen eine physische Eigenschaft E hat die mentale Eigenschaft M zur Ursache. Außerdem gilt, dass alle physischen Eigenschaften auch physische Ursachen haben, woraus folgt, dass auch E eine vollständige physische Ursache E hat. Somit verursachen M und P unabhängig voneinander die Eigenschaft E, letztere wäre von beiden mehrfachverursacht bzw. überdeterminiert. Eine Verursachung von E durch P kann dabei nicht als kausale Kette über die Eigenschaft M als kausalen Link erfolgen, da die Relation zwischen den Eigenschaften P und M keine kausale ist und beide zur selben Zeit instanziiert werden:

> No single event can have more than one sufficient cause occurring at any given time – unless it is a genuine case of causal overdetermination. (Kim 2003, 157)

Ein weiteres gängiges Beispiel ist der Schmerz, der auf den C-Fasern superveniert. Angenommen, ein ungeschickter Heimwerker schlägt sich aus Versehen mit dem Hammer auf den Finger, anstatt den Nagel zu treffen. Die wahrscheinliche Reaktion auf den folgenden Schmerz besteht in einem kurzen Aufschrei und dem Zusammenzucken des Heimwerkers. Der beschriebenen Situation nach befindet sich der Heimwerker in einem neuronalen Zustand N, worauf der mentale Zustand S „Schmerzen“ superveniert. Welcher Zustand verursacht nun das Zusammenzucken? Da N zeitgleich zu S erscheint, macht es keinen Sinn zu glauben, dass die Schmerzen von N verursacht werden und N über eine Kausalverbindung zu den Schmerzen das Zusammenzucken verursacht. Doch auch die Annahme, dass damit N und S zwei getrennte Ursachen für das Zusammenzucken sind, erscheint nicht plausibel, da in diesem Falle die Wirkung des Zusammenzuckens kausal überdeterminiert wäre. Schließlich scheint es viel sinnvoller anzunehmen, dass lediglich N das Zusammenzucken verursacht – und damit greift das Argument des kausalen Ausschlusses. Für jedes einzelne Ereignis bzw. jede einzelne Eigenschaft kann es nicht mehr als eine einzige hinreichende Ursache oder echte kausale Erklärung geben, es sei denn, es handelt sich um einen Fall kausaler Überdeterminierung (vgl. Kim 1993a, 259f.).

Die Formulierung des Prinzips macht deutlich, dass mentale Ereignisse als eigenständige Verursacher ausscheiden. Sie sind allenfalls in Verbindung mit ihren subvenierenden Basen kausal relevant. Denn obwohl der mentalen Eigenschaft in den aufgestellten Beispielen zunächst kausale

Wirksamkeit zufällt, wird diese unter Verweis auf die physischen Basiseigenschaften zurückgenommen. Um mentale Eigenschaften vollständig von einer kausalen Wirkung auszuschließen, wird also das Ausschlussargument üblicherweise durch die Prämisse ergänzt, wonach ein Ereignis o. ä. nicht kausal mehrfachdeterminiert sein darf. Solange es theoretisch zulässig ist, dass Ereignisse mehrfachdeterminiert werden, ist es auch möglich, dass ein Ereignis zumindest teilweise durch eine mentale Eigenschaft verursacht wird (vgl. Van Gulick 1999, 242f.). Physikalisten können nicht grundsätzlich ablehnen, dass Mehrfachdeterminierung möglich ist. Beispielsweise wird in Agatha Christies berühmtem Orient Express ein Mann von mehreren Fahrgästen auf verschiedene Weise ermordet. Sicherlich handelt es sich hierbei um eine rituelle Tat, der ein erster Mord vorausgeht. Doch es ist zumindest denkbar und logisch möglich, dass ein Mord gemäß dem Motto „sicher ist sicher" gleichzeitig durch zwei verschiedene Handlungen, beispielsweise verschiedene Gifte in Speis und Trank, vollzogen wird. Ein weiteres Beispiel ist die Hinrichtung durch ein Exekutionskommando. Es ist durchaus möglich, dass die Projektile zweier Schützen zeitgleich das Opfer an lebenswichtigen Organen treffen und unabhängig voneinander den Tod verursachen. Ebenso ist es denkbar, dass auf neuronaler Ebene zwei verschiedene Neuronenensembles $N_1$ und $N_2$ einen Gedanken oder eine Absicht hervorrufen. Warum sollte nicht auch die mentale Eigenschaft $M_1$ unabhängig von der physischen Eigenschaft $P_1$ kausale Ursache für die physische Eigenschaft $P_2$ sein?

Üblicherweise wird jedoch die Ausdehnung der Mehrfachverursachung bzw. systematische Mehrfachverursachung kritisiert. „[I]t is hard to believe that God is such a bad engineer […]." (Schiffer 1987, 148) Zudem lassen sich solche Standard- bzw. Einzelfälle von Mehrfachverursachung von einer systematisch notwendigen Mehrfachverursachung für mentale Kausalität abgrenzen. Wenn zwei Kugeln gleichzeitig das Herz des Opfers treffen oder ein Kurzschluss und eine umgefallene Laterne gleichzeitig einen Hausbrand auslösen, spielt jede Ursache eine eigene kausale Rolle; es sind zwei separate und unabhängige kausale Relationen, die zu einer gemeinsamen Wirkung führen. Aufgrund der Supervenienzrelation ist die Situation für mentale Kausalität eine andere, denn mentale Eigenschaften sind abhängig von ihrer physischen Basis (vgl. Kim 2003, 161).

Doch selbst wenn das Ausschlussproblem aufgrund der Generalisierung nicht zwangsläufig zu einem Epiphänomenalismus führt, muss untersucht werden, nicht ob, sondern wie mentale Eigenschaften kausal relevant sein können, wenn sie von ihren physischen Realisierern abgetrennt werden (vgl. Kim 1998, 61f., 77-80). Es gibt gute Gründe für die Annahme, jedes physische Ereignis habe eine physische Ursache. Mentale Eigenschaften können nur kausal relevant sein, wenn sie zusätzlich zu physischen Eigenschaften eine Handlung o. ä. verursachen. Für sämtliche Fälle von Entscheidungen, Wünschen oder anderen propositionalen Einstellungen, denen eine Handlung oder ein weiterer Gedanke folgt, würde jedoch systematisch eine Mehrfachbe-

stimmung vorliegen. Eine systematische Mehrfachbestimmung bzw. Überdeterminierung erscheint jedoch zunächst unplausibel.

### 3.4.3 Das Ausschlussargument in der Kritik

Das Ausschlussproblem wird in der philosophischen Literatur sehr vielfältig diskutiert. Es gibt Lösungsvorschläge, die versuchen, mentale Eigenschaften vor der Wirkungslosigkeit zu bewahren – bei gleichzeitiger Unterscheidung von ihren physischen Realisierern (vgl. z. B. Hawthorne 2001; Shoemaker 1980b). So wählt Yablo (1992) einen sprachlogischen Weg, in dem er in Anlehnung an W. E. Johnson (1964) zwischen „determinables" und „determinates" unterscheidet. Mentale Eigenschaften sind lediglich determinables von physischen Eigenschaften, wie z. B. die Farbe Rot von Purpur, weshalb sie nicht in einer kausalen Konkurrenz zueinander stehen (vgl. Kap. 4.6). Andere Ansätze versuchen, den kausalen Ausschluss über die Identität zwischen mentalen und physischen Eigenschaften zu vermeiden. Die Identität der Ereignisse bzw. Eigenschaften M und P wäre ein möglicher Weg aus dem kausalen Ausschluss heraus (vgl. Crane 1995, 215). In einem früheren Aufsatz versucht Kim (1992) zu zeigen, dass mentale Eigenschaften nicht von ihren physischen Realisierern im Sinne der Multirealisierbarkeit getrennt sind. Vielmehr gibt es viele Arten dieser mentalen Eigenschaften, weshalb eine „struktur-spezifische" Identität möglich ist (vgl. Jackson 1995).[78]

Warum aber sollte das Argument des kausalen Ausschlusses überhaupt gelten? Horgan merkt an, dass die Akzeptanz des Arguments eine Frage der intuitiven Überzeugung ist. Er weist darauf hin, dass im alltäglichen Gebrauch die Bedeutung kausaler Erklärungen innerhalb des jeweiligen Kontextes gesichert und passend hinsichtlich der beschriebenen Phänomene ist. In der philosophischen Diskussion dagegen wird die kausale Kraft mentaler und physischer Eigenschaften gleichzeitig hinterfragt. Es besteht somit kein fixierter, stabiler Kontext für die unterschiedlichen Begriffe bzw. Paradigmen. Äußerst intuitiv sei nun die Tendenz, der physischen Kausalität ausschlaggebendes Gewicht zu- und den mentalen Eigenschaften abzusprechen. (vgl. Horgan 1997, 180f.).

Seiner Struktur nach steht das Argument des kausalen Ausschlusses in der Tradition der Sparsamkeit und der Argumentationsökonomie. Ereignis $P_1$ ist kausal hinreichend für Ereignis $P_2$. Warum aber sollten wir akzeptieren, dass nur ein Ereignis ein anderes verursachen kann? Warum

[78] In der Konsequenz kann z. B. die mentale Eigenschaft ‚Schmerz' nicht mehr als eine natürliche Art aufgefasst werden. Um dies zu vermeiden, kann alternativ zwischen so genannten ‚Tropen' und ‚Typen' unterschieden werden: Entsprechend sind kausal relevante mentale Eigenschaften ‚Tropen', einzelne, objektspezifische Eigenschaften, und multirealisierbare mentale Eigenschaften ‚Typen', verschiedene Objekte vereinheitlichende Eigenschaften (vgl. z. B. Heil 1992; Heil/Robb 2003; Macdonald/Macdonald 1995).

sollten nicht zwei gleichzeitige Ereignisse $P_{1a}$ und $P_{1b}$ jeweils hinreichend Ereignis $P_2$ verursachen können? Der Einwand gegen eine mehrfachverursachte mentale Kausalität: Es erscheint höchst unwahrscheinlich, dass eine solche Mehrfachverursachung systematisch geschieht, da solche Fälle in der Natur äußerst selten zu beobachten sind. Die diskutierten Standardfälle für systematische Verursachung, wie der „Doppelmord", werden als besondere Ausnahmen betrachtet. Dennoch ist denkbar, dass verschiedene Neuronenensembles parallel ein weiteres neuronales Ereignis verursachen. Dies zeigt auch der Vorschlag einer simultanen Kausalität (vgl. Lowe 2003).

So ist die systematische Mehrfachverursachung für nicht-reduktive Physikalisten nur problematisch, wenn die zwei gleichzeitigen Ursachen unabhängig voneinander dieselbe Wirkung erzielen sollen. Dies wäre für mentale Kausalität eine unwahrscheinliche Koinzidenz. Sobald jedoch mentale und physische Eigenschaften in einer systematischen Abhängigkeit zueinander stehen, ist diese Problematik nicht mehr gegeben. Dies ist der Fall bei mentalen Eigenschaften, die zwar von physischen realisiert werden, aber doch nicht identisch mit ihnen sind: „[A]lthough that mental cause is realized by one or more physical events which are part of that sufficient physical cause, it is not identical with any part of that sufficient physical cause" (ebd., 146). Dieser zwar nicht unabhängige, aber doch eigenständige Charakter macht eine kausale Rolle für mentale Eigenschaften möglich. Dabei unterscheidet sich die Kausalität von den angesprochenen Standardfällen der Mehrfachverursachung, da hierbei eben nicht vollständig unabhängige Ursachen vorliegen. Lowe geht sogar noch einen Schritt weiter, indem er ein Modell der simultanen Kausalität für interaktionistische Dualisten einführt. Dieses fußt auf einer kausalen Abhängigkeit mentaler von physischen Ereignissen. Dabei kann ein mentales Ereignis M eine Ursache eines physischen Ereignisses E sein, da jenes auch die Wirkung eines oder mehrerer physischer Ereignisse $P_1$ und $P_2$ ist, welche ebenfalls Ursache für das physische Ereignis E sind. Entscheidend ist, dass Ereignis E über verschiedene Wege von verschiedenen Teilereignissen gemeinsam verursacht werden kann, wobei $P_1$ und $P_2$ nur dann E verursachen können, wenn zuvor M verursacht wurde, welches wiederum E mitverursacht (vgl. ebd., 148f.). Somit wird nochmals verdeutlicht, dass auch eine systematische Mehrfachverursachung zumindest nicht grundsätzlich auszuschließen ist.

Da das Argument des kausalen Ausschlusses auf den beiden physikalischen Grundannahmen beruht, alle Ereignisse seien physisch und die physische Welt sei ein geschlossenes System, ist es sehr schwierig zu kritisieren (vgl. Van Gulick 1999, 242). Das Ausschlussproblem ist stark verknüpft mit dem Prinzip der kausalen Geschlossenheit der Physik oder auch der Voraussetzung, dass kausale Beziehungen unter ein striktes Gesetz fallen müssen. Dagegen würde eine gemäßigte nicht-reduktive Position zwar die kausale Geschlossenheit der Physik anerkennen, aber zulassen,

dass Handlungen von mentalen und physischen Eigenschaften mehrfachverursacht wären (vgl. Loewer/LePore 1989).[79]

Es wiederholt sich hier die Kritik gegen die Dominanz des physikalischen Weltbildes. So muss sich der Verfechter des Arguments fragen lassen, ob er mit dem Ausschluss sämtlicher mentaler Eigenschaften auch die kausale Kraft der Eigenschaften der Spezialwissenschaften ausschließt. Die Ausschlussproblematik betrifft auch Eigenschaften, die nicht mental sind. So drohen alle Eigenschaften, die durch physische Eigenschaften multirealisierbar sind, kausal irrelevant zu sein. Durch das Ausschlussargument sind also auch die Eigenschaften der Einzelwissenschaften betroffen. Somit stellt sich die erweiterte Frage, ob wir akzeptieren müssen, dass auch biologische oder geologische Eigenschaften kausal wirkungslos sind und inwiefern mentale Eigenschaften in diesem Kontext eine Sonderstellung einnehmen (vgl. Fodor 1989; Baker 1993). Wenn neurologische, geologische und biochemische Eigenschaften epiphänomenal sind, ist es wiederum keine Besonderheit, dass mentale Eigenschaften epiphänomenal sind. „Why should psychological properties have to meet a higher standard for causal potency than biochemical, geological, or optical properties?" (Van Gulick 1999, 249)[80]

Um mentalen Eigenschaften über ihre Nicht-Reduzierbarkeit hinaus auch kausale Wirksamkeit zu ermöglichen, argumentiert auch Block gegen das Argument des kausalen Ausschlusses:

> First it is hard to believe that there is no mental causation, no physiological causation, no molecular causation, no atomic causation but only bottom level physical causation. Second, it is hard to believe that there is no causation at all if there is no bottom level of physics. (Block 2003, 138)

Sein erstes Argument hat eine starke intuitive Motivation. Es ist einfach schwer zu glauben, dass lediglich die kausale Erklärung der Physik eine echte Erklärung ist, und die kausalen Erklärungen der übrigen Einzelwissenschaften ausscheiden. Seiner Auffassung nach sollten die kausalen Gesetze der unterschiedlichen Ebenen nicht als kompetitiv, sondern als eigenständig akzeptiert werden (vgl. ebd., 135). Als Beispiel führt er in der Tradition des Arguments der Multirealisierbarkeit den Hurrikan Edna an, der zumindest ähnliche Schäden verursacht, auch wenn er von unterschiedlichen mikrophysischen Partikeln realisiert wird. Block folgert: „If you think there are lawlike relations at many levels, and you think laws have something to do with individuation of events and properties, you have some reason to endorse event individuation at levels other than the physical (ebd., 137; vgl. auch LePore/Loewer 1987). Horgan schlägt aus diesem Dilemma

[79] Nicht-Physikalisten könnten ihre Position mit der Betonung der nomischen Verbindungen zwischen mentalen und physischen Ursachen stützen. Eine starke Mehrfachdeterminierung von Handlung kann so über eine bloße Koinzidenz hinaus erklärbar werden.

[80] Weitere Argumente um die Diskussion der Generalisierbarkeit des Ausschlussarguments finden sich in: Noordhof (1999); Gillet (2001); Bontly (2002).

heraus eine weitere Position vor, den „robust causal compatibilism". Diese würde das Argument des kausalen Ausschlusses als falsch verwerfen und die Prämisse, dass nur physische Eigenschaften kausale Eigenschaften sind, nicht akzeptieren. Dies ist der Weg, den verschiedene Nicht-Reduktivisten gehen (vgl. z. B. Horgan 1997, 171).

Das zweite Argument zielt auf einen theoretischen Erklärungsregress, der durch die tatsächliche Forschungslage innerhalb der Physik gestützt wird. So ist es eine offene Frage, ob es überhaupt eine Grundebene elementarer Teile gibt (vgl. Dehmelt 1989; zitiert nach: Block 2003). In der zeitgenössischen Physik wird deshalb bezweifelt, dass das zurzeit gängige Standardmodell vollständig die Naturgesetze repräsentiert, und nicht die derzeit bekannten Entitäten in weitere kleinere Entitäten aufgeteilt werden müssen (vgl. Wilczek 1996; Walker 1996; alle zitiert nach: Block 2003). Block schließt, sofern es keine Grundebene gebe und jede Eigenschaft auf einer niedrigeren Eigenschaft superveniere, das kausale Ausschlussargument zeige, dass es – ohne Grundebene – überhaupt keine Kausalität geben könne. Da es seiner Meinung nach aber keine Frage ist, dass es Kausalität grundsätzlich geben müsse, scheint das kausale Ausschlussargument fehlerhaft (vgl. Block 2003., 138f.). [81]

***Quintessenz Argument des kausalen Ausschlusses***

- *Verschärfung der These der physikalischen Geschlossenheit*
  - *neben physischen weitere (mentale) Eigenschaften als Ursachen überflüssig*
  - *kausale Kraft nur für physische Eigenschaften*
  - *Ablehnung systematischer Mehrfachverursachung*
- *bei systematischer Abhängigkeit strukturspezifisch Möglichkeit einer gemeinsamen Verursachung*
- *Verursachung eines Ereignisses erfolgt über Teilereignisse*
- *kausale Kraft als funktionale bzw. kausale Rolle einer Eigenschaft*
- *Gefahr des Erklärungsregresses: keine letzten Basisebenen bestimmbar*

[81] Kim selbst reagiert auf diese Kritik, die sich vornehmlich gegen seine Version des kausalen Ausschlussarguments (vgl. Kim 1998) richtet. Er argumentiert gegen den Regresseinwand u. a. mit dem Hinweis, dass ähnlich zur metaphysischen Frage nach einem ersten Beweger bzw. einer ersten Ursache bei Aristoteles oder Thomas von Aquin auch für die Frage der Ebenen eine fundamentale Ebene definiert werden muss, um einen infiniten Regress zu vermeiden. Konsequenterweise kann innerhalb der Physik eine fundamentale Ebene nur durch eine weitere fundamentalere Ebene ersetzt werden, insofern diese nun als diejenige kausal geschlossene und relevante gilt (vgl. Kim 2003, 173f.).

### 3.5 Die Voraussetzung der strikten Gesetze

Eine weitere Voraussetzung für mentale Kausalität ist die Gesetzmäßigkeit von Kausalrelationen. Auch diese hängt eng mit den Voraussetzungen der kausalen Geschlossenheit der Physik sowie dem „kausalen Ausschlussargument" zusammen. Grundlegend ist Davidsons Beschreibung des Anomalen Monismus. Danach beruhen nur physische Beschreibungen auf strikten Gesetzen, während psychologische Beschreibungen keinen strikten Gesetzen entsprechen (vgl. Kap. 2.3.1; Kap. 2.3.2). Problematisch ist hier erneut die Frage, ob somit auch andere Einzelwissenschaften nicht zwingend kausale Erklärungen liefern, da sie keine strikten Gesetze beschreiben.

#### 3.5.1 Geltungsbereich von Rationalitätsprinzipien

Ereignisse, die in einem Ursache-Wirkung-Verhältnis miteinander verknüpft sein sollen, müssen ein Gesetz instanziieren, damit überhaupt sinnvoll bzw. wissenschaftlich von Kausalität gesprochen werden kann. Für einen Goldschmied ist es wichtig zu wissen, dass Gold immer bei 1064° C schmilzt, während der Schmelzpunkt von Silber bereits bei 960,8° C liegt. Es muss ein allgemeines Gesetz geben, das die Ereignisse der Art x (Temperatur von 1064° C) mit denen der Art y (Gold schmilzt) verbindet. Ein Goldschmied, der ein ihm unbestimmtes Metall erhitzt, wird davon ausgehen, dass es sich, wenn es bei 960,8° C schmilzt, um Silber handelt. Strikte wissenschaftliche Gesetze sollen insofern eine genau Prognose einer kausalen Relation oder umgekehrt eine hinreichend erklärende Beschreibung solcher Relationen liefern. Ein striktes Gesetz gilt nicht nur unter bestimmten Bedingungen, sondern ausnahmslos (vgl. LePore/Loewer 1994, 262f.).

Das Problem, das sich im Zusammenhang mit mentaler Verursachung stellt, ist, dass es scheinbar weder strikte mentale Gesetze noch strikte psychophysische Gesetze gibt oder geben kann. Davidson hat in seiner Theorie des Anomalen Monismus das Problem psychophysischer Gesetze untersucht (vgl. Davidson 1970). Die Ablehnung einer Existenz solcher Gesetze bezieht sich dabei auf intentionale mentale Ereignisse, wie z. B. Wünsche und Absichten. Nicht-intentionale mentale Ereignisse wie Schmerzen oder Juckreize sind dagegen durchaus gesetzesartig mit ihren physischen Korrelaten verknüpft. Ansonsten wäre die Vorstellung der „physischen Realisierung" der mentalen Ereignisse hinfällig; es wäre nicht plausibel, wieso Schmerzen durch das Feuern von C-Fasern entstehen sollen. Davidson beschreibt, dass der Bereich des Mentalen nicht von Gesetzen, sondern von Rationalitätsprinzipien gekennzeichnet ist. Das zentrale Argument gegen psychophysische Gesetze ist das von Davidson aufgestellte „Principle of Charity". Dessen für die intentionalistischen Diskurse konstitutiven Prinzipien sind nicht vereinbar mit den

Grundsätzen physikalischer Diskurse, denn mentale Ereignisse werden im Gegensatz zu physischen Ereignissen relational und normativ bestimmt (vgl. Kap. 2.3.2). Aussagen und Überzeugungen einer Person werden auf ihre Konsistenz und Kohärenz hinsichtlich weiterer Überzeugungen der Person überprüft. Damit unterscheiden sich der mentale und der physische Bereich grundlegend. Eine gesetzesartige Verknüpfung durch Brückengesetze ist nicht möglich (vgl. Kap. 3.3.1).

Hinter der Forderung, dass die Relation zweier Ereignisse nur als kausal beschrieben werden kann, wenn sie zusammen ein striktes Gesetz instanziieren, steckt auch das Prinzip eines hinreichenden Grundes. So qualifizieren sich Gesetze nur als strikte Gesetze, wenn sie zwei Bedingungen erfüllen: Sie müssen sowohl ausnahmslos gelten, als auch Teil einer in sich geschlossenen umfassenden Theorie sein. Es ist die Forderung nach Ausnahmslosigkeit, welche auf das Prinzip des hinreichenden Grundes verweist (vgl. Van Gulick 1999, 239). So liefert uns ein Gesetz T, das Ausnahmen zulässt, keine hinreichende Erklärung, warum dem Ereignis C das Ereignis D folgen soll. Ein Ausbleiben von D wäre genauso möglich, wie das Eintreffen von D, obwohl statt C das Ereignis B voranging. Solange die Antezedensbedingungen des Ereignisses D nicht als hinreichend bestimmt sind, ist eine vollständige kausale Erklärung der Relation zwischen C und D nicht gegeben. Da strikte Gesetze eine solche Erklärung liefern sollen, müssen sie ausnahmslos gelten. Dies ist aber für den Bereich des Mentalen nicht möglich. Die Rationalistätsprinzipien der Kohärenz und Konsistenz lassen stets verschiedene Vorbedingungen zu, denen wiederum verschiedene mentale Ereignisse folgen können. Eine Prognose im strengen wissenschaftlichen Sinne zu stellen, ist somit nicht möglich. Mein Gesprächspartner Peter äußert beispielsweise, dass er gerne mal wieder Essen ginge. Aus dieser Aussage ist weder ersichtlich, ob der Besuch einer Fast-Food-Kette oder ein Candlelight-Dinner beim teuersten Italiener der Stadt angestrebt wird, noch ob das „Essen gehen" sofort erfolgen soll oder lediglich in naher Zukunft. Wenn ich Peter nun näher kenne und weiß, dass er ein großer Bewunderer amerikanischer Fast-Food-Kultur und zudem ein sehr ungeduldiger Mensch ist, kann ich mir ausmahlen, dass der „Burger-Besuch" recht bald erfolgen wird. Nach wie vor bewege ich mich allerdings im Bereich der Spekulation, eine seriöse Prognose ist nicht möglich. Umgekehrt bleibt mir verborgen, warum Peter gerade jetzt auf die Idee kommt, „mal wieder Essen zu gehen". Die psychologischen Erklärungsmöglichkeiten scheinen vielfältig, womöglich ist er sich seiner Beweggründe selbst nicht bewusst.

Dass mentale Eigenschaften nicht unter strikte Gesetze fallen, hat in Verbindung mit der Unterscheidung zwischen kausal relevanten und irrelevanten Eigenschaften von Ereignissen dazu geführt, dass mentale Eigenschaften als kausal wirkungslos betrachtet werden. Unter den Bedingungen des anomalen Monismus erscheint es irrelevant, unter welche Beschreibung ein Ereignis gefasst wird – entscheidend ist, dass einer physischen Beschreibung auch ein striktes kausales

Gesetz zugrunde liegt (vgl. Kim 1984, 267; Sosa 1984, 277f.). Die physischen Eigenschaften eines Ereignisses sind aufgrund der strikten Gesetze hinreichend, jegliche kausale Verbindung zu gewährleisten.

Von der kausalen Wirkungslosigkeit sind aber nicht nur psychologische Eigenschaften, sondern letztlich sämtliche höherstufige Eigenschaften betroffen (vgl. Kap. 3.1.3). Jede Einzelwissenschaft beschreibt die Welt in einer eigenen Sprache, aus einer eigenen Perspektive – mit eigenen Gesetzen für die speziellen Entitäten, mit denen die Einzelwissenschaft arbeitet. Grundsätzlich gilt, dass Änderungen auf der Ebene der niedrigstufigen Eigenschaften Auswirkungen auf die Eigenschaften der höherstufigen Ebene haben. „Changes in the molecules of your brain, for instance, can have dramatic repercussions for your brain and for your state of mind as well." (Heil 1998, 119) Psychologische Gesetze beinhalten nicht die molekularbiologische Perspektive, während umgekehrt molekulare Ereignisse Auswirkungen auf psychologische Prozesse haben. Die Gesetze der Einzelwissenschaften betreffen jedoch aus der materialistischen Perspektive nur die Entitäten auf der Ebene, auf der die jeweilige Einzelwissenschaft operiert. Solange die Einzelwissenschaften als autonom und in ihren Gesetzmäßigkeiten als nicht-reduzierbar betrachtet werden, können deren Gesetze allerdings nicht ausnahmslos gelten – sie sind nicht strikt (vgl. ebd.).

Die Forderung nach strikten Gesetzen schließt mentale Eigenschaften als erklärungsrelevant für kausale Ereignisse also vollständig aus, da ausschließlich physische Objekte und Ereignisse hinreichende Bedingungen für kausale Verknüpfungen sein können. Die mentalen Eigenschaften bleiben somit de facto wirkungslos. Die Möglichkeit, dass ein Ereignis aufgrund seiner mentalen Eigenschaften (Überzeugungen, Wünsche, Absichten) in eine Kausalrelation eintritt, in die es ohne diese nicht eintreten würde, wird zurückgewiesen. Vielleicht sind mentale Ereignisse generell zu komplex, um sie gesetzmäßig mit ihren neuronalen Korrelaten zu verknüpfen. Um ihre kausale Kraft zu retten, müsste dagegen geprüft werden, ob sie nicht doch Gesetzen unterliegen, und ob nicht eventuell die Gesetzeskonzeption von Davidson oder gar das Konzept der nomologischen Kausalität zumindest für mentale Ereignisse zurückgewiesen werden muss.

### 3.5.2 Nicht-strikte Gesetze als Alternative

Verbunden mit dem idealisierten Wissenschaftsbegriff ist auch die Forderung nach strikten Gesetzen angreifbar. So scheinen die engen physikalischen Gesetze die vagen mentalen Zustände und Strukturen nicht wirklich beschreiben zu können (vgl. Hardcastle 1998). Psychologische Verallgemeinerungen scheinen grundsätzlich heterogen. Insofern können psychologische Kausalerklärungen Präzisionsansprüchen oder Universalisierbarkeitsforderungen von vornherein nicht

nachkommen. Sie lassen sich eher in Verallgemeinerungen erfassen, wie es soziologische Modelle versuchen (vgl. Audi 1999, 67ff.). Daneben birgt ein starker Gesetzesbegriff auch Probleme für die Supervenienztheorie oder funktionalistische Positionen. Die Behauptung, dass es keine psychophysischen Gesetze gibt, kann so interpretiert werden, dass nicht nur kausale Gesetze, sondern auch Korrelationsgesetze fehlen und somit das Verhältnis propositionaler Einstellungen und Hirnzustände betroffen ist. Dadurch wird auch die funktionalistische Relation zwischen Rollenspieler und kausaler Rolle in Frage gestellt (vgl. Kap. 2.2.3).

Man kann die Möglichkeit bzw. die Formulierung von Gesetzen in der Wissenschaft zwar auch als Basis für Typen-Identität bezeichnen. Danach gibt es für jede Ebene eine physische Beschreibung. Ob zwei Entitäten typen-verschieden sind oder nicht, hängt entscheidend von dem Gesetzescharakter ab, der für die entsprechende Ebene gilt. Wenn dies jedoch grundsätzlich gilt, scheinen die Bedingungen in der vorliegenden Form für die Typen-Identität psychologischer Zustände wiederum von der Referenz zu den psychologischen (und möglicherweise neurologischen) Gesetzen, welche für diese Zustände gelten, abzuhängen (vgl. Block/Fodor 1980, 248f.).

Die Annahme, wonach psychologische Zustände, die nach neurologischen Gesetzen verschieden sind, ipso facto nach psychologischen Gesetzen verschieden sind, trifft somit nicht zu. Unter Einbeziehung sämtlicher Wahrscheinlichkeiten können verschiedene neurologische Zustände funktional identisch sein. Typen-Verschiedenheit neurologischer Zustände garantiert nicht die Typen-Verschiedenheit psychologischer Zustände.

> [T]he fundamental problem with […] physicalism is that type identity is being determined relative to, at best, a subset of the laws which must be presumed to operate upon psychological states. The only justification for this restriction seems to lie in the reductionist biases of theses positions. (ebd.)

Sobald der Reduktionismus jedoch in Frage gestellt wird, zeigt sich, dass die nomologischen Anforderungen der Typen-Identität für psychologische Zustände äußerst komplex und vielfältig sind. Mit dem geringen Wissen, das über psychologische Gesetze zur Verfügung steht, scheint es nicht plausibel, dass die Grenzen psychologischer Zustände durch physische Gesetze bestimmt werden können bzw. den Grenzen aufgrund physischer Gesetze entsprechen (vgl. ebd., 249). Andererseits wird die Eigenständigkeit psychischer Gesetze auf diese Weise betont.

Davon abgesehen kann aber auch gefragt werden, warum ausschließlich strikte Gesetze eine notwendige Bedingung für (mentale) Kausalität sind. Eine nomologische Erklärung muss nicht die Basis für eine kausale sein; eine kausale Beziehung zwischen zwei Eigenschaften P und M kann auch auf einer anderen Relation beruhen, als auf einem strikten Gesetz (vgl. McLaughlin 1989; zitiert nach: Van Gulick 1999, 238). Als Alternative gelten die so genannten ceteris-paribus-

Gesetze. Bei diesen Gesetzen handelt es sich um vorläufige Generalisierungen, die unter Normalbedingungen zutreffen. Im Gegensatz zu den strikten Gesetzen lassen sie Ausnahmen zu und erfassen mit ihrer schlichten Form „Wenn F dann G" die Vorbedingungen für die heterogenen mentalen Zustände besser (vgl. Fodor 1989). Bei dem Fall „Wenn Fritz müde ist, dann geht er schlafen" muss die Vorbedingung „müde zu sein" erfüllt werden, damit als Konsequenz das Schlafengehen erfolgen kann. Eine absolute Generalisierung wie bei strikten Gesetzen ist jedoch nicht möglich, da sich das Antezedens durchaus ändern könnte. Ein Freund könnte Fritz anrufen und ihn spontan in eine Bar einladen, weshalb Fritz nicht schlafen geht. Ceteris-paribus-Gesetze lassen solche Ausnahmen zu. Damit sind sie allerdings nur bedingt voraussagefähig und erklären auch nur bedingt das Eintreffen von G nach F. So bleibt weiterhin offen, inwiefern die präzisen Kausalverbindungen physischer Ereignisse mit den vagen Verbindungen mentaler Ereignisse in eine kausale Relation geführt werden können.

Kutschera bemerkt, dass heute fast ausschließlich physikalische Naturgesetze angegeben werden können (vgl. Kutschera 2003, 32f.). Für einen Materialisten müssen sich aus den physikalischen Gesetzen psychophysische Gesetze entwickeln lassen – auf Basis einer physikalischen Analyse psychischer Zustände und Ereignisse. Allerdings können wir zumindest derzeit nicht auf solche Analyseverfahren zurückgreifen. Es gibt nur wenige psychophysische Gesetze, Kutschera nennt das Weber-Fechnersche Gesetz über den Zusammenhang von Reiz- und Empfindungsstärke. Die bekannten psychophysischen Gesetze sind ferner ungenau und beziehen sich nur selten auf komplexere mentale Zustände, wie propositionale Einstellungen. In der Neurowissenschaft kann immerhin bestimmt werden, welche Hirnregionen für welche geistigen Leistungen relevant sind. Allerdings sind derzeit nur wenige Korrelationen bekannt, wenn man die Vielzahl an mentalen Zuständen und Ereignissen bedenkt. Diese können nicht als wirkliche physikalische Analyse eines psychischen Problems gewertet werden, da vielfach Korrelationen zwischen geistigen Leistungen und Hirnregionen beschrieben werden. Die Aktivität innerhalb der Hirnregion sowie ihre Vernetzung zu anderen Hirnregionen ist noch zu erforschen.

Zudem zeigt aber das Vorhandensein zumindest einiger psychophysischer Gesetze, wie das Weber-Fechnersche Gesetz über die Abhängigkeit von Empfindungen von physischen Reizen, dass solche immerhin möglich sind. Die inzwischen bekannten psychophysischen Gesetze legen nahe, dass es auch psychophysische Kausalität gibt: „Auch wenn es nur sehr wenige psychophysische Gesetze gibt, ist das noch kein brauchbarer Einwand gegen eine psychophysische Kausalität." (Kutschera 2003, 77) Entgegen der These der Anomalie des Psychischen kann darauf hingewiesen werden, dass es auf Exaktheit im Sinne der Physik gar nicht ankommt. In dem Sinne sind solche Gesetze nicht exakt, als sie nur unter nicht genauer bestimmten Normalbedingungen gelten. Ein optischer Reiz ruft z. B. eine Farbempfindung bestimmter Qualität hervor – sofern die Versuchsperson wach ist, sich auf ihre Empfindungen konzentriert, nicht unter Einfluss von

Drogen steht, gesund ist etc. Diese letzte Unbestimmtheit der Konditionen – das et cetera – hat in exakten Gesetzen keinen Platz (vgl. ebd., 76, Fußnote 97).

Menzies (2003) weist darauf hin, dass unser kausales Denken von Abstraktion, genauer Generalisierung und Idealisierung geprägt ist. Um die Komplexität eines Systems zu verringern, wird es als ein bestimmtes System mit entsprechenden Gesetzen betrachtet. Implizit wird dabei angenommen, dass das, was für das einzelne System gilt, für alle Systeme derselben Art gilt. Das individuelle Armhochheben kann einerseits als ein rein neurophysiologisches System unter neurophysiologischen Gesetzen betrachtet werden und andererseits als eine rationale Handlung unter rational-intentionalen Gesetzen. In beiden Fällen wird die Komplexität des Einzelfalls reduziert, da ein bestimmter Aspekt generalisierend hervorgehoben wird (vgl. ebd., 205f.). Hinsichtlich der Idealisierung betont Menzies, dass die Gesetze, die (mentale) Kausalität betreffen, typischerweise unter ceteris-paribus-Bedingungen gelten. Die be-obachteten Systeme o. ä. entwickeln nur unter ganz bestimmten Bedingungen eine kausale Folge und verlaufen in einem kausalen Nexus, der nicht vom „Standardfall" abweichen darf. Dies ist insofern eine Idealisierung, als tatsächliche Systeme bzw. Ereignisse üblicherweise von weiteren Ereignissen gestört werden. Diese Idealisierung betrifft aber nicht nur psychologische, sondern auch physikalische Gesetze. Zeitgenössische Forschung hat gezeigt, dass die wissenschaftlichen Gesetze der Physik wie der Einzelwissenschaften unveränderliche ceteris-paribus-Gesetze sind (vgl. Cartwright 1983, 1999; Joseph 1980; Giere 1988; Suppe 1989; alle zitiert nach: Menzies 2003, 206).[82] Dahinter steht die Überlegung, dass der grundsätzliche Eingriff des Beobachters und die damit verbundene Kontingenz in jeglichem natürlichen System den Gedanken ausnahmsloser Gesetze unplausibel macht. Aufgrund der menschlichen Intervention können solche Gesetze lediglich unter ceteris-paribus-Bedingungen formuliert werden: „The laws of neurophysiology, as well as the laws of rational intentional psychology, hold only under the condition that there are no external interferences, including interventions by other agents." (ebd.; vgl. Kap. 3.3.3)

Hinsichtlich der Forderung nach strikten Gesetzen kann somit eingewendet werden, dass wir womöglich ebenso wenig psychologische, wie physische strikte Gesetze für mentale Kausalität vorliegen haben. Zwar wird gemeinhin vor dem Hintergrund einer theoretisch vollständigen und umfassenden Physik argumentiert (vgl. Quine 1948, 1951). Tatsächlich gibt es derzeit keine strikten physischen Gesetze, die im Sinne der Identitätstheorie oder auch des Funktionalismus Verhalten oder Gedanken und ihre kausale Wirksamkeit vorhersehbar und bestimmbar machen. Hier steht die Neurowissenschaft vor einem inversen Problem: Auch wenn aufgrund der Korrelation verschiedene mentale Prozesse zunehmend durch bildgebende Verfahren, wie z. B. funkti-

[82] In seiner Argumentation gegen das kausale Ausschlussargument hat Block zudem darauf hingewiesen, dass in der zeitgenössischen Physik bezweifelt wird, dass das zurzeit gängige Standardmodell vollständig die Naturgesetze repräsentiert, und nicht die derzeit bekannten Entitäten in weitere kleinere Entitäten aufgeteilt werden müssen (vgl. Block 2003; Kap. 3.4).

onelle Magnetresonanztomographie (fMRT; vgl. Kap. 5), bestimmbar sind, ist es nicht möglich, allein aufgrund solcher Daten eine genaue Prognose für Verhalten oder Gedanken zu wagen. So lässt sich auch argumentieren, dass es nicht nur eine theoretisch umfassende physikalische Theorie gibt, sondern dass auch eine ideale psychologische Theorie mentale Ereignisse vorhersehen und bestimmen kann. Schließlich gibt es auch auf diesem Feld klaren wissenschaftlichen Fortschritt. Das Problem der mental-physischen Kausalität wäre damit zwar nicht gelöst, aber die Nicht-Reduzierbarkeit wäre unterstrichen.

***Quintessenz Voraussetzung der strikten Gesetze***

- *Postulat, wonach wissenschaftliche Gesetze strikt sein sollen*
    - *Sicherstellung von Prognosefähigkeit und von hinreichender Erklärung*
- *weder strikte mentale noch psychophysische Gesetze für intentionale mentale Ereignisse vorhanden*
- *mentaler Bereich unterliegt Rationalitätsprinzipien*
- *ausnahmslose Gültigkeit für alle Einzelwissenschaften problematisch*
- *mit Fehlen strikter kausaler Gesetze auch keine strikten Korrelationsgesetze möglich*
- *ceteris-paribus-Gesetze als zweckmäßige Alternative*
    - *wenige (ungenaue) psychophysische Gesetze bestimmbar Hinweis auf psychophysische Kausalität*
    - *Gesetze für mentale Kausalität ohne Exaktheit physikalischer Gesetze: Aufstellung von Gesetzen unter Normalbedingungen für den Standardfall möglich*
- *strikte physische Gesetze aufgrund ihrer Generalisierung und Idealisierung streng genommen ebenfalls als ceteris-paribus-Gesetze*

### 3.6 Möglichkeiten für mentale Kausalität

Die bereits genannten Voraussetzungen, insbesondere das Argument der kausalen Geschlossenheit der Physik, das Argument der strikten Gesetze sowie das Argument des kausalen Ausschlusses lassen mentale Kausalität nahezu abwegig erscheinen: Dennoch, natürlich ist die Annahme, dass es eine Form der mentalen Kausalität gibt, immer noch Bestandteil der aktuellen Diskussion. In der Entwicklung spiegelt sich letztlich das Ringen wider, die kausale Kraft mentaler Ereignisse bzw. mentaler Eigenschaften zu gewährleisten. Dies war die ursprüngliche Motivation der Vertreter der Typen-Identität, die alltägliche Beobachtung, wonach wir Menschen aufgrund unserer bewussten Gedanken Entscheidungen fällen oder Handlungen ausführen, auf festen theoretischen Boden zu stellen (vgl. Kap. 2.1). In einer physischen Welt schien ihnen dies nur durch die (Typen-) Identität mentaler und physischer Ereignisse möglich zu sein. Insofern erweist sich das kausale Argument, welches z. B. von Papineau vertreten wird (vgl. Papineau 2004, 17f.), noch immer als guter Grund, der für die Identität mentaler und physischer Eigenschaften spricht. Wenn angenommen wird, dass Bewusstsein Handlung verursacht, und zugleich gilt, dass alles Physische allein und hinreichend durch Physisches (nicht mehrfach) verursacht wird, scheint die kausale Kraft von Bewusstsein bzw. mentalen Eigenschaften nur gegeben, wenn diese mit physischen Eigenschaften identisch sind.

#### 3.6.1 Die Gefahr des Epiphänomenalismus

Im Fortgang der Argumentation wurde es zunehmend schwieriger, die kausale Kraft für das Mentale zu verteidigen. Dies wird besonders deutlich in dem Argument des kausalen Ausschlusses, in dem sich auch die Voraussetzung der Geschlossenheit der Physik und die Forderung nach strikten physischen Gesetzen widerspiegeln. Zudem wird das Argument durch den Ausschluss systematischer Mehrfachrealisierung verstärkt. Es ist somit letztlich eine Verschärfung der These, dass jede physische Wirkung einen *hinreichenden* physischen Grund hat (vgl. Kap. 3.4, 3.4.2). Mit der Position des Funktionalismus liefen mentale Eigenschaften als „höherstufige" Eigenschaften oder Eigenschaften „zweiter Ordnung" stets Gefahr, lediglich epiphänomenale Eigenschaften, also kausal wirkungslos am Gängelband der physischen, der „niedrigstufigen" Eigenschaften „erster Ordnung", zu sein (vgl. Kap. 3.1). So stellt sich hinsichtlich des Ausschlussarguments die Frage, inwiefern mentale Eigenschaften als Eigenschaften der höheren Ebene eine kausale Wirkung haben können, wenn die eigentliche kausale Kraft bei ihren Realisierern liegt (vgl. Kap. 3.4.1). Die Position des Epiphänomenalismus ist dabei auch methodisch motiviert. Argumente der Sparsamkeit und der Kohärenz sprechen eher gegen ein Modell, das Mehrfachverursachung

zulässt: die Annahme einer autonomen kausalen Rolle für bewusste Ereignisse ist ein zusätzlicher metaphysischer Ballast, der zudem in Widerspruch zu dem vorherrschenden physikalischen Weltbild steht (vgl. Birnbacher 2006, 122).

Diese „Gefahr des Epiphänomenalismus" (vgl. Fodor 1989, 59) konnte auch die Position der Supervenienztheorie nicht beseitigen, die durch die starke asymmetrische Anbindung zunächst versuchte, mentalen Eigenschaften eine ontologisch solide Basis zu verschaffen und zumindest ihre „kausale Relevanz" zu garantieren. In der Ausprägung durch Kim scheint im Gegenteil die Supervenienztheorie sogar die Reduktion mentaler Eigenschaften zu verfolgen und geht damit letztlich sogar weiter als die Identitätstheorie, auch in Richtung Elimination mentaler Eigenschaften – zumindest tendenziell (vgl. z. B. Kim 2003). Und auch für Nicht-Reduktivisten ergibt sich das Dilemma, dass mentale Ereignisse entweder eigenständige Ursachen sind, was zur Verletzung des Prinzips der kausalen Geschlossenheit führt. Oder sie sind kausal wirkungslos, woraus schließlich Epiphänomenalismus folgt (vgl. Kap. 3.4.1). Die Position des anomalen Monismus bzw. der Tokenidentität scheint in der Konsequenz aus mentalen Eigenschaften bzw. Ereignissen Epiphänomene zu machen (vgl. McLaughlin 1999), auch wenn Davidson selbst mit seiner Position gerade die Eigenständigkeit mentaler Ereignisse zu bewahren sucht (vgl. Kap. 2.3). Schließlich müssen alle Formen einer Körper-Geist-Kausalität unter physischen Gesetzen subsumiert werden (vgl. Kap. 2.3.2; Kap. 3.5.1; Kap. 3.4.1).

Mentale Ereignisse werden zwar von physischen Ereignissen im Gehirn verursacht, haben selbst jedoch keine Wirkung auf physische Ereignisse, wie z. B. Bewegung oder Verhalten (vgl. Levine 2001, 22). Das Problem des Epiphänomenalismus entsteht unter der Annahme, dass einige, speziell die physischen Eigenschaften der Ereignisse kausal wirksam sind und andere, speziell die mentalen, nicht:

> The question is whether mentalistic properties are causally efficacious, or whether mentalistic descriptions of properties are relevant to understanding causation. Many philosophers have developed more confidence in the efficacy of the underlying neural properties than in that of the 'intentional aspects' of mental events. (Burge 1999, 98)

Der intentionale oder der phänomenale Aspekt mentaler Ereignisse scheint demnach für viele Philosophen irrelevant und epiphänomenal zu sein und von dem zugrunde liegenden echten kausalen Prozess auf einer physischen Ebene abzuhängen. Dabei betrifft der Vorwurf des Epiphänomenalismus psychologische und phänomenale Konzepte mentaler Ereignisse (vgl. Chalmers 1997, 11). Sie lassen sich insofern verknüpfen, als jemand eine Glaubensannahme über seinen phänomenalen Zustand hat; er weiß, dass er Schmerzen empfindet und nimmt deshalb die Hand von der Herdplatte. Unter der Annahme, dass Schmerzen keine physische Wirkung haben, können Schmerzen entweder diese Glaubensannahme nicht verursachen, oder die Glaubensannahme

ist epiphänomenal. Motiviert wird diese Position von der Überlegung, dass die zugrunde liegenden physischen Prozesse sich in jedem Fall ereignen, damit hinreichend sind, und die mentalen Ereignisse ihre kausale Wirksamkeit eben von den physischen Eigenschaften als echte kausale Größen erhalten.

Um das grundsätzliche Problem des Epiphänomenalismus zu beschreiben, betrachtet Block die Beschreibung einer Eigenschaft zweiter Ordnung (anstelle der funktionalen Eigenschaft), die aus Eigenschaften mit bestimmten kausalen Relationen zueinander besteht. Das Beispiel des roten Tuchs des Toreros wurde bereits erwähnt (vgl. Kap. 3.1.4). Ein weiteres Beispiel ist „Schläfrigkeit" (englisch: dormitivity) als eine Eigenschaft zweiter Ordnung, die im Besitz einer chemischen Eigenschaft (erster Ordnung) kausal relevant für Schlaf sein könnte. Somit wäre x ist schläfrig = x hat eine Eigenschaft, die kausal relevant für Schlaf ist. Allerdings ist in dieser Beziehung zunächst vor allem die Eigenschaft erster Ordnung entscheidend für die kausale Wirkung. Eine Schlaftablette, die ohne unser Wissen in unser Essen gemischt wird, ist kausal verantwortlich für unser Einschlafen aufgrund der chemischen Eigenschaft, nicht aufgrund der „Schläfrigkeit" an sich. Verschiedene Schlaftränke wirken aufgrund unterschiedlicher chemischer Substanzen, wie Valium oder Seconal. Aber solange man nicht von der Eigenschaft „Schläfrigkeit" der Tablette weiß, kann diese auch nicht relevant für das Einschlafen sein. Ausnahme ist hierfür der bekannte Placebo-Effekt. Unter funktionalistischen Gesichtspunkten ist das Gehirn „a syntactic engine driving a semantic engine" (Block 1990, 143), das heißt, die Syntax und Struktur der physischen Eigenschaften ist relevant für den kausalen Prozess, nicht ihre Bedeutung:

> Dormitivity has causal-explanatory relevance because it involves a causally efficious property. A causally inefficious property can nonetheless be causal-explanatory if it "brings in" a causally efficacious property. But the causally relevant property is the first-order one, not the second-order one […]. (ebd., 163)

Block geht es darum zu zeigen, dass in Standardfällen Eigenschaften zweiter Ordnung überhaupt nicht als kausal relevant in Frage kommen (vgl. ebd., 156). Ausnahmen scheinen vielmehr jene Fälle zu sein, in denen eine intelligente Person um sie weiß. Nach der nomologischen Konzeption von kausaler Relevanz sind Eigenschaften zweiter Ordnung nicht kausal relevant, da sie nicht für die Wirkung nomologisch hinreichend sind (vgl. Kap. 3.5.1; Kap. 4.1).[83] Die nomologische Theorie für kausale Relevanz wird von Block favorisiert. In der Konsequenz sieht er auch das Dilemma: „[W]e must either abandon functionalism or accept epiphenomenalism." (Block 1990, 161)

[83] Block weist darauf hin, dass „Schläfrigkeit" zumindest logisch hinreichend für Schlaf sein kann. Eine Tablette X verursacht „Schläfrigkeit", wenn sie die Eigenschaft hat, die kausal Schlaf garantiert, wenn die Tablette X eingenommen wird (Block 1990, 157f.).

Die Gefahr des Epiphänomenalismus für mentale Eigenschaften wird in der Diskussion der Voraussetzungen für mentale Kausalität verschärft deutlich. Dies gilt besonders in der Argumentation um das kausale Ausschlussargument (vgl. Kap. 3.4), aber auch die übrigen Voraussetzungen scheinen mentale Eigenschaften allenfalls als epiphänomenale, kausal wirkungslose Eigenschaften zuzulassen. Dies zeigt sich einmal im verwendeten ontologischen Grundmodell, wonach lediglich die physischen Eigenschaften die kausale Wirkung eines Ereignisses bestimmen. Das Schichten-Welt-Modell, wonach Eigenschaften durch die physischen Basiseigenschaften realisiert werden, liefert hierfür den geeigneten Hintergrund und bestätigt scheinbar die kausale Geschlossenheit der Physik. So ist es schließlich sogar fraglich, ob mentale Eigenschaften, da sie kausal wirkungslos sind, überhaupt als echte Eigenschaften betrachtet werden können (vgl. Kap. 3.1; Kap. 3.2; Kap. 3.3; Kap. 3.4). Auch in der Unterscheidung zwischen dem Bereich des Mentalen und dem Bereich des Physischen wurde aus physikalistischer Sicht die Wirkungslosigkeit des ersten Bereichs deutlich. Die Merkmale des Mentalen (vgl. Kap. 3.2.1), wie propositionale Einstellungen, Qualia oder subjektive Perspektive sind mit den Merkmalen des Physischen (vgl. Kap. 3.2.2), wie räumliche Ausdehnung oder Messbarkeit nicht vereinbar. In dieser empirischen Ausprägung scheinen naturgemäß physische Eigenschaften allein kausal relevant und hinreichend für die Wirksamkeit der mentalen Zustände zu sein – das Mentale ist nicht im Gehirn messbar (vgl. Kap. 3.2.3). Mentale und physische Erklärungen können allenfalls als alternative Beschreibungen gelten, echte Kausalerklärungen liefern aber nur physische Beschreibungen (vgl. Kap. 3.2). Die Erklärungslücke ist lediglich eine epistemische. Ontologisch sind aber nur die physischen Eigenschaften relevant, eine Tatsache, die einfach akzeptiert werden muss (vgl. Kap. 3.2.4). Auch aus der Voraussetzung der kausalen Geschlossenheit der Physik als Ergänzung zur These der Vollständigkeit der Physik folgt, dass mentale Eigenschaften allenfalls epiphänomenal sind: Demnach wird keine ein physisches Ereignis involvierende Kausalkette je die Grenze des Physischen zum Nicht-Physischen bzw. Mentalen überschreiten. Zentrales Argument ist die Annahme einer Erklärungsvollständigkeit der physikalischen Theorie. (vgl. 3.3). Die starke Physikalismusthese wird dabei als ein empirisches Argument aufgefasst. In einer induktiven Generalisierung wird formuliert, dass in jedem Fall, in dem die Naturwissenschaft eine Ursache für ein physisches Ereignis entdeckt hat, jene eine hinreichende physische Ursache für das Ereignis war (vgl. Kap. 3.3.1).

Aus Sicht des Physikalismus ist dies auch im Zusammenhang mit der mentalen Kausalität der Fall: Die physischen Eigenschaften liefern eine hinreichende Erklärung. Schließlich scheint auch die Forderung nach strikten Gesetzen dazu zu führen, dass mentale Eigenschaften lediglich epiphänomenal sein können. Da die Rationalitätsprinzipien nicht ausnahmslos gelten, gibt es weder strikte mentale Gesetze noch strikte psychophysische Gesetze. Die Ablehnung einer Existenz solcher Gesetze bezieht sich dabei auf intentionale mentale Ereignisse, wie z. B. Wünsche und Absichten. Dass mentale Eigenschaften nicht unter strikte Gesetze fallen, hat in Verbindung

mit der Unterscheidung zwischen kausal relevanten und irrelevanten Eigenschaften von Ereignissen dazu geführt, dass mentale Eigenschaften ebenso wie andere höherstufige Eigenschafen als kausal wirkungslos betrachtet werden. Schließlich kann nur unter Anwendung solcher Gesetze eine echte Kausalrelation beschrieben werden (vgl. 3.5.1.). An dieser Stelle kann nochmals auf das Beispiel der Sopranistin hingewiesen werden, die das hohe C singt, woraufhin Glas zerbricht. Hier würde man die Eigenschaft der hohen Frequenz und nicht die Bedeutungseigenschaft des Liedtextes als ursächlich für das Zerbrechen des Glases ansehen (vgl. Dretske 1988) – und somit weder das Ereignis als solches, noch seine mentale Eigenschaft, sondern lediglich seine physische Eigenschaft.

Halten wir also fest: Die mentalen Eigenschaften werden von den physischen erzeugt, sind von ihnen abhängig und haben nur über sie kausale Kraft. Sie sind nur aufgrund der neuronalen Aktivität, die komplexer ist als sie selbst, an Handlungen beteiligt. Andererseits: Die physischen Eigenschaften auf der niedrigstufigen Ebene können im Gegenzug auch als Spezifizierung der mentalen Eigenschaften auf der höherstufigen Ebene betrachtet werden (vgl. u. a. Hardcastle 1998, Fodor 1989, Jackson/Pettit 1990). Gegnern der epiphänomenalistischen Position erscheint es in einer starken Intuition absurd abzulehnen, dass Gedanken oder auch Empfindungen einen Unterschied in unserem Verhalten ausmachen. Die These des Epiphänomenalismus ist demnach nicht mit unserem Wissen über uns selbst vereinbar. Kurz gesagt widerspricht sich die Position insofern selbst, als unter der Hypothese, dass propositionale Einstellungen kausal wirkungslos sind, es letztlich keinen Grund gibt, die Position des Epiphänomenalismus zu vertreten oder nicht zu vertreten. Zudem hätten wir keinen Zugang zu einem Wissen über unsere eigenen Bewusstseinsereignisse (vgl. Birnbacher 2006, 123f.; Chalmers 1997, 192f.).

### 3.6.2 Die Voraussetzungen in der Kritik

Es muss betont werden, dass die Voraussetzungen für mentale Kausalität nach wie vor kritisch diskutiert werden und es gute Argumente gibt, die in Erwiderung auf die Voraussetzungen für eine kausale Kraft mentaler Eigenschaften plädieren.

Zunächst wurde deutlich, dass die Frage der mentalen Kausalität nur abhängig von den ontologischen Grundeinheiten – Substanzdualismus, Token- oder Typenidentität, funktionalen oder supervenienten Eigenschaften – beantwortet werden kann. Auch die Terminologie stellt ein Problem innerhalb der Diskussion dar. Beispielsweise ist nicht eindeutig festgelegt, was sich hinter der Bezeichnung „physische Eigenschaft“ verbirgt. In einem weiten Sinn können hiermit Relationen beschrieben werden, bei denen physische Eigenschaften ihre Umgebung umfassen, die in nicht-mentalistischen, nicht-intentionalen Begriffen beschrieben wird (vgl. Kap. 3.1). Der Er-

eignisbegriff ist sehr unterschiedlich – durch Davidson und Kim – geprägt. Während die erste Version aufgrund der Ereignisontologie die kausale Kraft der mentalen Ereignisse für gewährleistet hält, betont letztere Version aufgrund der Eigenschaftsinstanziierung deren kausale Wirkungslosigkeit. Ferner stehen beide konkurrierende Ontologien selbst als zirkulär oder als zu feinkörnig in der Kritik (vgl. Kap. 3.1.1). Ebenso ist der Eigenschaftsbegriff immer noch problematisch, da offen ist, ob Eigenschaften als allgemeine oder als partikuläre, an den konkreten Dingen vorkommende Entitäten existieren (vgl. Kap. 3.1.2). Im Rahmen des Schichten-Welt-Modells wird sehr unterschiedlich argumentiert. Nicht-Reduktivisten betonen die Eigenständigkeit der Eigenschaften auf höherstufigen Ebenen, um ihre kausale Kraft zu erhalten, bei gleichzeitigem Verweis auf ihre Realisierung durch Basiseigenschaften auf niedrigstufigen Ebenen, um sie nicht als kausale Konkurrenten zu etablieren und das Prinzip des Ausschlusses von Mehrfachkausalität nicht zu verletzen. Den einzelnen Ebenen bzw. den Einzelwissenschaften wird eine eigene Taxonomie mit echten eigenen Eigenschaften mit kausaler Wirkung zugesprochen (vgl. Kap. 3.1.3).

Daneben muss festgehalten werden, dass es definierte Merkmale des Mentalen gibt, die sich von den Merkmalen des Physischen unterscheiden lassen und – das erscheint dabei besonders wichtig – bislang nicht in einer physikalischen Terminologie oder auch Ontologie erfasst werden (vgl. Kap. 3.2.1; Kap. 3.2.2). Diesen Merkmalen des Mentalen, wie propositionalen Einstellungen, Qualia oder der subjektiven Perspektive, fällt im alltäglichen Verständnis eine kausale Rolle bzw. kausale Wirkung zu. Da nun mentale Eigenschaften nicht extensionsgleich mit physischen Eigenschaften sind, kann nicht zufrieden stellend dargestellt werden, wie physische Eigenschaften die kausalen Rollen der mentalen Eigenschaften übernehmen können. (vgl. Kap. 3.2.1; Kap. 3.2.2; Kap. 3.2.3; Kap. 3.2.4), zumal auch eine Gleichsetzung von physisch mit neuronal nicht zwingend gegeben scheint (vgl. Kap. 3.2.3). Hinsichtlich der Erklärungslücke zwischen dem mentalen und dem physischen Bereich wird aus nicht-reduktivistischer Perspektive prinzipiell an diesem Unterschied festgehalten. Das geringe Wissen über die Natur psychophysischer Korrelationen erschwert auch das Verständnis der explanatorischen Lücke. Zudem erscheinen psychophysische Korrelationen als sehr komplex, da allein die psychologischen Unterscheidungen oft anderen Kriterien unterliegen als die neurologischen Unterscheidungen (vgl. Kap. 3.2.4). Ein möglicher Ausweg scheint hier zu sein, mentale Phänomene als biologisch bzw. natürlich zu begreifen und sie somit nicht grundlegend von physischen Phänomenen zu unterscheiden. Auf diese Weise wäre zumindest eine kausale Wechselwirkung zwischen biologisch-mentalen und biologisch-physischen Eigenschaften möglich (vgl. ebd.).

Gegen die Voraussetzung der kausalen Geschlossenheit der Physik sprechen insbesondere die Argumente gegen psychophysische Reduktion und Brückengesetze. Der psychologische Theoriebereich scheint nicht nomologisch koextensional zum physikalischen zu sein. Zudem scheidet

eine Identität mentaler Eigenschaften aufgrund ihrer Multirealisierbarkeit aus (vgl. Kap. 3.3.2). Ein schwerwiegendes Argument ist, dass die Physik derzeit offensichtlich nicht als vollständig bezeichnet werden kann. Viele Wissenschaftler räumen ein, dass die Physik längst nicht alle Fragen beantwortet. Es ist ferner fraglich, inwiefern innerhalb der Physik zukünftig z. B. neue Kategorien für physische Ursachen ermitteln werden. Möglicherweise wird das bestehende physikalische Paradigma auch von einer zukünftigen Physik abgelöst oder erneuert. Die Aussagen der Wissenschaft verweisen nur auf eine mehr oder weniger geschlossene physische Welt. So geht die Quantenphysik davon aus, dass jede Be-obachtung eines mikrophysischen Geschehens einen Eingriff in dieses Geschehen darstellt. Es werden also gerade nicht geschlossene Systeme betrachtet, die Annahme einer kausalen Geschlossenheit der Physik kann somit nicht ein Prinzip dieser naturwissenschaftlichen Disziplin sein (vgl. 3.3.3).

Tatsächlich wird bezweifelt, ob das physikalische Weltbild mit all seinen Konsequenzen als maßgeblich zu akzeptieren ist oder nicht als ideales Modell gleichberechtigt neben anderen steht. Das Argument muss nicht beeindrucken, es ist zu nah am Physikalismus orientiert sowie keine philosophische These, sondern eine immerhin substanzielle empirische Vermutung. Es ist nicht absurd, zu behaupten, dass einige physische Wirkungen nicht-reduzierbare mentale Ursachen haben. Dies beinhaltet nicht zwangsläufig, dass Handlungen physikalische Gesetze brechen, sondern verweist auf die Möglichkeit psychophysischer Gesetze.

Dagegen ist es fraglich, ob ein physikalisches Kausalitätsmodell einfach auf psychologische Phänomene übertragen werden kann. Schließlich ist zweifelhaft, dass die rein physikalische Erklärung den psychologischen Prozessen gerecht wird. Auch die Schwierigkeiten, den Begriff „physisch“ einzugrenzen (vgl. Kap. 3.2.2), spielen hier herein. Dabei zeigt sich, dass die Gefahr eines Ebenen-Regresses besteht, da sämtliche Eigenschaften der höherstufigen Ebenen – und eben auch neurologische Eigenschaften – grundsätzlich auf Mikroeigenschaften reduzierbar sind. Allerdings kann keine Ebene als Basisebene eindeutig bestimmt werden, so dass theoretisch auch physische Eigenschaften weiter reduzierbar sind (vgl. Kap. 3.3.3). Intertheoretische Reduktion betrifft letztlich alle Wissenschaften, eine starke Physikalismusthese untergräbt somit die gesamte wissenschaftliche Weltsicht.

Das Argument des kausalen Ausschlusses kann zunächst wenig überzeugen, solange es sich darauf stützt, dass z. B. mentale Eigenschaften von physischen realisiert werden und deshalb keine eigene kausale Kraft besitzen. Hier bieten funktionalistische Positionen oder das Modell der supervenienten Kausalität mögliche Alternativen (vgl. Kap. 2.2; Kap. 2.4). Erst durch die zusätzliche Ablehnung der Möglichkeit einer systematischen Mehrfachverursachung kann das Ausschlussargument seine Aussagekraft entfalten. Aber auch hiergegen gibt es begründete Einwände. Zum einen wird argumentiert, dass der alltägliche Gebrauch kausaler Erklärungen innerhalb des

jeweiligen Kontextes hinsichtlich der beschriebenen Phänomene passend ist. Die Tendenz, der physikalischen Kausalität ausschlaggebendes Gewicht zu- und den mentalen Eigenschaften abzusprechen, ist aus Sicht eines robusten kausalen Kompatibilismus eine Frage der intuitiven Überzeugung (vgl. Kap. 3.4.2).

Zum anderen gibt es Lösungsvorschläge, die mentale Eigenschaften vor der Wirkungslosigkeit zu bewahren versuchen – bei gleichzeitiger Unterscheidung von ihren physischen Realisierern. Es reicht demnach nicht aus, unter Verweis auf das Prinzip der Sparsamkeit oder die Unwahrscheinlichkeit und Seltenheit in der Natur mehrfachverursachte mentale Kausalität abzulehnen. So können Physikalisten nicht grundsätzlich ablehnen, dass Mehrfachdeterminierung möglich ist; Beispiele sind der gleichzeitige Tod eines Ermordeten oder verschiedene gleichzeitige Ursachen eines Hausbrands. Zudem ist denkbar, dass verschiedene Neuronenensembles parallel ein weiteres neuronales Ereignis verursachen. So ist die systematische Mehrfachverursachung bei mentaler Kausalität für Nicht-Reduktivisten nicht problematisch, wenn die zwei gleichzeitigen Ursachen in einer systematischen Abhängigkeit zueinander stehen, was bei der Realisierungsrelation mentaler bzw. physischer Eigenschaften angenommen wird. Dabei unterscheidet sich die Kausalität von den angesprochenen Standardfällen der Mehrfachverursachung, da hierbei eben nicht vollständig unabhängige, aber auch nicht identische Ursachen vorliegen. Dies wird als Argument für mentale Kausalität durch mentale Eigenschaften aufgefasst. Diese Möglichkeit diskutiert auch der Vorschlag einer simultanen Kausalität (vgl. Kap. 3.4.2). Eine systematische Mehrfachverursachung ist demnach zumindest nicht grundsätzlich auszuschließen.

Schließlich wiederholt sich die Kritik gegen die Dominanz des physikalischen Weltbildes. So muss sich der Verfechter des Arguments fragen lassen, ob er mit dem Ausschluss sämtlicher mentaler Eigenschaften auch die kausale Kraft der Eigenschaften der Spezialwissenschaften ausschließt. Es ist einfach schwer zu glauben, dass lediglich die kausale Erklärung der Physik eine echte Erklärung ist, und die kausalen Erklärungen der übrigen Einzelwissenschaften ausscheiden. Den Taxonomien auf höherstufigen Ebenen werden deshalb eigene kausale Gesetze zugesprochen. In diesem Zusammenhang sieht sich auch das Argument des kausalen Ausschlusses mit einem Regressproblem konfrontiert. So ist es eine offene Frage, ob es überhaupt eine Grundebene elementarer Teile gibt. In der zeitgenössischen Physik wird deshalb bezweifelt, dass das zurzeit gängige Standardmodell vollständig die Naturgesetze repräsentiert, und nicht die derzeit bekannten Entitäten in weitere kleinere Entitäten aufgeteilt werden müssen (vgl. Kap. 3.3.2).

Schließlich gibt es auch gute und überzeugende Argumente, die gegen die Voraussetzung bzw. Forderung nach strikten Gesetzen aufgeführt werden. Zunächst kann bezweifelt werden, ob die Unterscheidung zwischen intentionalen mentalen Ereignissen ohne strikte Gesetze und phänomenalen mentalen Ereignissen mit strikten Gesetzen so zutreffend ist. Die Diskussion um die

Möglichkeit der Brückengesetze lässt hier zumindest einen anderen Schluss zu (vgl. Kap. 3.3.2; Kap. 3.5.1). Generell scheinen die engen physikalischen Gesetze die heterogenen vagen mentalen Zustände und Strukturen nicht wirklich zu beschreiben. Insofern können psychologische Kausalerklärungen Präzisionsansprüchen oder Universalisierbarkeitsforderungen von vornherein nicht genügen. Damit geraten auch die supervenienztheoretische Relation und die funktionalistische Determinierungsbeziehung in Schwierigkeiten und werden angreifbar (vgl. Kap. 3.5.2).

Außerdem wird betont, dass nicht allein strikte Gesetze, sondern auch ceteris-paribus-Gesetze als Basis für eine kausale Relation in Frage kommen. Unser kausales Denken ist von Abstraktion geprägt. Um die Komplexität eines bestimmten Systems zu verringern, wird es als ein bestimmtes System mit entsprechenden Gesetzen betrachtet. Diese Idealisierung betrifft aber nicht nur mentale Kausalität bzw. psychologische Gesetze, sondern auch physikalische Gesetze. Für jegliche beobachtete Systeme gilt, dass sie nur unter ganz bestimmten Bedingungen in einem kausalen Nexus verlaufen, der nicht vom „Standardfall" abweichen darf. Zeitgenössische Forschung legt den Schluss nahe, dass die wissenschaftlichen Gesetze der Physik wie der Einzelwissenschaften beide unveränderliche ceteris-paribus-Gesetze sind. Hinsichtlich der Forderung nach strikten Gesetzen kann somit eingewendet werden, dass wir womöglich ebenso wenig psychologische wie physische strikte Gesetze für mentale Kausalität vorliegen haben. Es wurde bereits angesprochen, dass es zumindest einige psychophysische Gesetze gibt. Diese bekannten psychophysischen Gesetze legen nahe, dass es auch psychophysische Kausalität gibt. Die Grenzen psychologischer Zustände allein durch physische Gesetze zu bestimmen ist jedoch nicht möglich (vgl. Kap. 3.5.2).

### 3.6.3 Homogenität als implizite Voraussetzung

Crane sieht den Konflikt zwischen der Geschlossenheit der Physik und der mentalen Kausalität vor allem aufgrund der Vorstellung, dass mentale und physische Kausalität einander entsprechen sollen. Gäbe es eine grundsätzliche ontologische Unterscheidung zwischen mentaler und physischer Kausalität, bestünde der Konflikt – wenn überhaupt – nicht in der vorliegenden Form. Der Physikalismus argumentiert, dass die Begriffe „mental" und „physisch", auf Kausalität angewendet, übertragbare Epitheta sind. Sie beschreiben lediglich die Relata der Kausalität und sind nicht die Kausalität selbst (vgl. Crane 1995, 219). So gesehen könnte das Prinzip der Homogenität als eine weitere Voraussetzung ergänzt werden (vgl. Menzies 2003, 200).

Entsprechend ist die Zurückweisung der Homogenität eine Möglichkeit, das Dilemma aufzulösen. Danach sind mentale Ursachen keine Ursachen auf dieselbe Weise wie physische Ursachen (vgl. insbesondere Crane 1995, 232ff.). Beispielsweise unterscheidet Dretske ausdrücklich

mentale bzw. intentionale Zustände als „structering causes" von Verhalten und physische Zustände als „triggering causes" (vgl. Dretske 1988). Jackson und Pettit unterscheiden wiederum zwischen der Erklärungsweise, die intentionale Phänomene als „program explanation" ermöglichen und üblicher (ordinary) kausaler Erklärung (vgl. Jackson/Pettit 1988). Auch die von Kim getroffene Unterscheidung zwischen einer Verursachung der supervenienten mentalen Eigenschaft und der subvenienten physischen Basis impliziert letztlich eine Zurückweisung der Voraussetzung einer Homogenität mentaler und physischer Kausalität (vgl. Kim 1998, 32ff.).[84] Ebenfalls impliziert Yablos Unterscheidung zwischen mentalen Eigenschaften als „determinables" und physischen Eigenschaften als „determinates" eine Zurückweisung der Homogenitätsvoraussetzung wie auch die Annahme, mentale Eigenschaften seien Eigenschaften zweiter Ordnung (vgl. Block 1990; Papineau 1993; Yablo 1992).

In der Zurückweisung der Voraussetzung der kausalen Homogenität des Mentalen und des Physischen laufen diese Versionen des Physikalismus jedoch noch immer Gefahr, mentalen Zuständen letztlich nur eine epiphänomenale Wirkung einzuräumen. Die eigentliche kausale Arbeit wird von den physischen Basen bzw. Realisierern geleistet.

Burge weist darauf hin, dass unser Verständnis mentaler Kausalität vor allem von unserem Verständnis für mentale Erklärung abhängt, unabhängig von jeglichem Wissen über die basalen Prozesse. Seiner Ansicht nach bestimmen jene Eigenschaften die kausale Wirkung eines Ereignisses, die letztlich als kausale Erklärung dienen. Physische Eigenschaften erklären zwar den kausalen Zusammenhang hinsichtlich des physischen Bereichs. Wenn physische Ereignisse jedoch mentale Eigenschaften haben, ist man nicht zu der Position verpflichtet, dass ausschließlich die physischen Eigenschaften die kausale Wirkung des physischen Ereignisses bestimmen. Es fehlt der Nachweis, dass die mentale Erklärung weder kausal ist noch die Muster der kausalen Eigenschaften beschreiben kann:

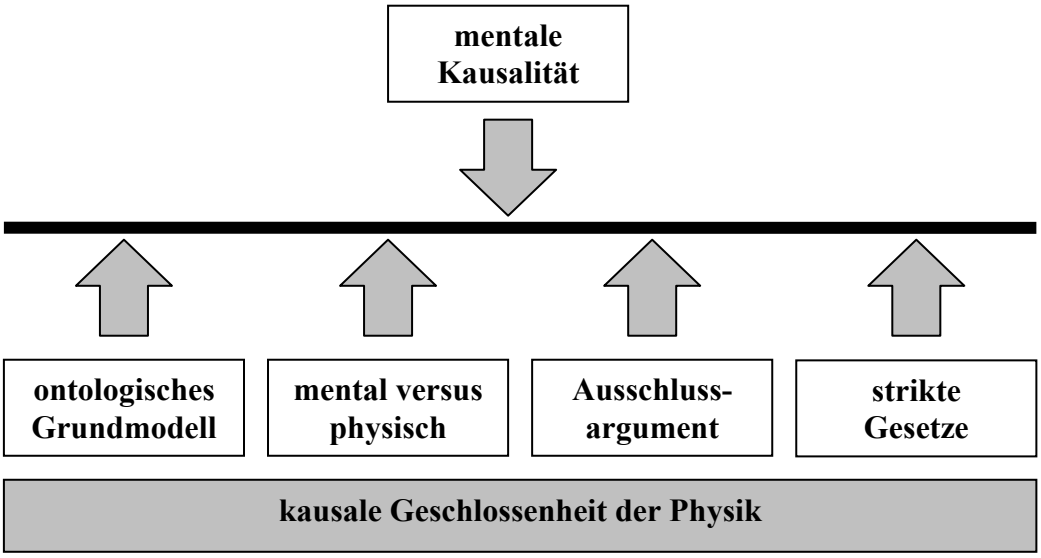
> For the causal powers of a physical event that is mental might include possible effects that are specified in mentalistic explanation. No one has shown that mentalistic explanation is either non-causal or non-descriptive. Nor is either view plausible. (Burge 1999, 101)

Intentionaler Inhalt ist demnach der erklärende und identifizierende Kern mentaler Ereignisse. Hierdurch wird systematisches Verstehen und Information über Ereignisse erst möglich – diese Erklärungsschemata sind die stärksten Indikatoren, die wir für mentale Kausalität haben. Unser Verständnis der mentalen Kausalität rührt nicht in erster Linie von einer Neubeschreibung durch

[84] Vgl. auch Audie 1999; Segal/Sober 1991 sowie Tye 2000, Kap. 8. Mittlerweile tendiert Kim eher dazu, die Voraussetzung, dass es mentale Kausalität gibt, zu verneinen und riskiert damit, eine Form des Epiphänomenalismus zu vertreten (vgl. Kim 1993b).

physische Begriffe her, sondern zunächst von unserem Verständnis für mentale Erklärungen (ebd., 111). Burge argumentiert hier von der epistemischen Ebene her, wenn er sich auf die Relevanz mentaler Erklärungen bezieht, und vermeidet mentale Kausalität als die metaphysische Frage, als die sie von den meisten Materialisten betrachtet wird (vgl. Kap. 3). Andererseits weist er auf eine strukturelle Schwäche dieser Position hin. Zwar ist die metaphysische Position der Materialisten in sich weitgehend konsistent, aber das eigentliche Problem kann nicht erklärt werden. Ironischerweise führt der Versuch, mentale Kausalität zu klären, für viele Materialisten zu dem Glauben, dass es eine ernste metaphysische Frage ist, ob mentale Charakterisierungen überhaupt eine kausale Relevanz haben (vgl. Burge 1999, 103).

Eine weitere Pointe in der Zurückweisung dieser Voraussetzung der Homogenität ist somit der Verlust der eigentlichen Motivation für diese physikalistische Position, da der Konflikt zwischen der mentalen Kausalität und der Geschlossenheit der Physik umgangen wird. Mentale Zustände sind kausal wirksam, in demselben Sinn, in dem physische Zustände kausal wirksam sind. Deshalb scheinen sie physische Zustände zu sein (vgl. Lewis 1983a, 101ff.) Sobald mentale Phänomene auf ihre eigene Weise verursachen, erscheint dies nur als terminologische Variante, die jedoch keine ontologische mentale Kausalität unabhängig von oder in Ergänzung zu der physischen beinhaltet. In der Konsequenz ist es auch fraglich, warum mentale Phänomene durch physische Phänomene realisiert werden (vgl. Crane, 1995, 235). Letztlich konnte die materialistische Position das Problem der mentalen Kausalität nicht lösen, obwohl das das eigentliche Ziel der Position war.




Grafik 7: Die zentrale Voraussetzung innerhalb der Diskussion um mentale Kausalität ist die Hypothese der kausalen Geschlossenheit der Physik. Sie ist im Grunde die Basis, die sich auch in den übrigen Voraussetzungen wiederfindet und diese in ihrer Ausrichtung bestimmt. Damit wird die Annahme einer wirkungsvollen mentalen Kausalität per definitionem ausgeschlossen.

Die Position, sich von der epistemischen Seite aus dem Problem der mentalen Kausalität zu nähern, erfährt von Papineau als ausgewiesenem Materialisten implizite Bestätigung, da dieser selbst einräumt, zumindest eine starke dualistische Intuition zu erfahren. Dass es sehr schwer fällt, die Identität phänomenaler und materieller Eigenschaften zu akzeptieren, bezeichnet Papineau als eine "intuition of distinctness", also eine intuitive Erfahrung unterschiedlicher Eigenschaften.[85] Auch er verweist also auf die alltägliche Erfahrung, wonach wir uns als Akteure unserer Handlung aufgrund mentaler bzw. bewusster Eigenschaften begreifen – eventuell sogar begreifen müssen. Diese starke Intuition ist schließlich auch der Grund dafür, dass Papineau auf der epistemischen Ebene den konzeptionellen Dualismus beibehält (vgl. Papineau 2004, 161). Den physischen Monismus mit dem Modell einer bewussten mentalen Kausalität zu versöhnen, scheint dennoch nicht möglich. So erscheint es schließlich als eine weitere Pointe der Diskussion um mentale Kausalität, dass nicht nur die Nicht-Reduktivisten vor einem Dilemma zwischen der Verletzung der Voraussetzungen und epiphänomenalen mentalen Eigenschaften stehen, sondern auch Vertreter identitätstheoretischer oder reduktivistischer Positionen die Frage der mentalen Kausalität nicht abschließend lösen können. Sie bewegen sich in einem argumentativen Zirkel.

Hinsichtlich der Voraussetzung der Geschlossenheit der Physik wird zwar die psychologische Fragestellung im Kontext der physischen Wissenschaft berücksichtigt. In Verbindung insbesondere mit den Voraussetzungen des kausalen Ausschlusses oder der strikten Gesetze ist jedoch kein Raum für die Kausalität mentaler Eigenschaften und – der Eindruck drängt sich auf – auch kaum Raum für mentale Phänomene. Sie stören vielmehr die gut aufeinander abgestimmten physikalischen Argumente, da sie nicht reduzierbar, eliminierbar oder identisch mit physischen Eigenschaften sind. In der Konsequenz kann somit der Bereich der mentalen Phänomene nicht genau erklärt werden und wird lediglich (alternativlos) ausgeklammert, ohne etwas Positives an seine Stelle zu rücken. Implizit wird dabei zudem ein heimlicher Dualismus aufrechterhalten, indem das erklärbare Physische von dem Mentalen, das ab einem gewissen Punkt nicht weiter erklärbar ist und deshalb als wissenschaftliche Erklärung ausscheidet, abgegrenzt wird (vgl. Searle 2001, 59ff.). Andererseits gibt es gute Gründe anzunehmen, dass es in der Welt mentale Phänomene gibt (vgl. Kap. 3.2.1), die bislang aus dem geschlossenen physikalistischen Erklärungsmodell herauszufallen scheinen. Es ist gut möglich, dass diese Phänomene in einer zukünftigen Physik eingebunden und erklärt werden können. „If there is good evidence that the world contains fundamental mental properties, then physicists would or at least should try to figure out

[85] Diese Intuition führt Papineau zurück auf den Fehlschluss, dass materielle Konzepte bzw. Be-griffe die Bewusstseinseigenschaften nicht adäquat beschreiben. Die Erklärung für den Fehlschluss, den Papineau als „antipathetic fallacy" bezeichnet, ist, dass die phänomenalen Konzepte zum Teil dieselben neuronalen Zustände aktivieren, wie die originalen Bewusstseinszustände; sie sind lediglich deren schwache Kopie. Den materiellen Konzepten scheint somit „das Gefühl" für diese Zustände zu fehlen. Papineau ist aber der Ansicht, dass es sich hierbei eben um einen Fehlschluss handelt, da sich die materiellen Konzepte (auf der epistemischen Ebene) ebenso adäquat wie die phänomenalen Konzepte auf die physischen Eigenschaften (auf der ontologischen Ebene) beziehen (vgl. Papineau 2004, 161ff.).

how these properties work." (Montero 2003, 178) Die Frage der Evidenz wird allerdings, wie wir gesehen haben, von den verschiedenen physikalistischen Positionen sehr unterschiedlich beantwortet.

Diese Diskussion verläuft paradoxerweise mitunter stark deduktiv, obwohl die Physikalisten grundsätzlich in einer empirischen Tradition mit einer induktiven Vorgehensweise stehen. Ziel muss deshalb die Vermeidung der dogmatischen Falle sein, in die die Diskussion hineinzugeraten droht. Deshalb wird als Ziel dieser Arbeit ein eigenständiges Modell der mentalen Kausalität entwickelt, das nicht homogen zur physischen Kausalität ist (vgl. Kap. 7). Das bedeutet nicht, dass sämtliche metaphysische Voraussetzungen abgelehnt werden. Aber das bedeutet, dass mentale Kausalität als eigenständiges Phänomen unter eigenen, separaten Gesichtspunkten betrachtet werden muss. Auf dem Weg zu diesem eigenständigen Modell ist es wichtig, theoretische und empirische Ansprüche zu berücksichtigen. Deshalb werden im nächsten Schritt theoretische Kriterien für mentale Kausalität herausgearbeitet, um diese später in Bezug zu Beobachtungssätzen auf ihre empirische Validität hin zu prüfen.

***Quintessenz Möglichkeiten für mentale Kausalität***

- *Gefahr des Epiphänomenalismus: Ringen um kausale Kraft mentaler Ereignisse bzw. mentaler Eigenschaften scheinbar erfolglos*
    - *lediglich punktuelle Kritik an Voraussetzungen möglich*
- *Teufelskreis der physikalistischen Position:*
    - *keine Erklärung für mentale Kausalität, obwohl Position durch Phänomen motiviert*
    - *dogmatische Falle: mentale Kausalität wird alternativlos ausgeklammert*
- *irrige Annahme der Homogenität von physischer Kausalität und mentaler Kausalität*
    - *andere Art der Erklärung für Intentionalität als identifizierenden Kern notwendig*
- *epistemische Differenzierung zwischen mentaler und physischer Kausalität als notwendiger Schritt für Konzeption eines Modells einer mentalen Kausalität*

**Kapitel 4**

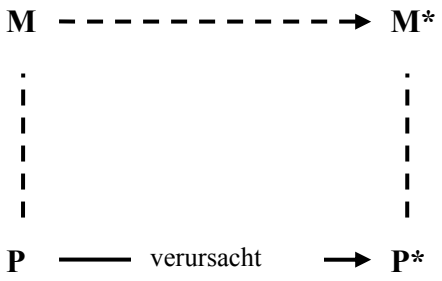
**Kriterien für mentale Kausalität**

Anknüpfend an den Aspekt der Homogenität, lässt sich feststellen, dass mentale Kausalität als eine physische Kausalität in der Tradition einer causa efficiens diskutiert wird. Sie wird als eine Wirkursache im mechanistischen Sinn verstanden. Ein solches Verständnis der mentalen Kausalität ist auf das physikalische Weltbild sowie die metaphysische Variante der hierarchischen Schichten-Welt zurückzuführen (vgl. Kap. 3.1.3). So beschreibt die Atomtheorie Systeme als Ansammlungen von Teilchen, deren raumzeitliche Begrenzung durch Kausalbeziehungen bestimmt sind (z. B. Regentropfen, Gletscher, Elefanten). Viele Eigenschaften der großen können durch das Verhalten der kleinen kausal erklärt werden. Searle unterscheidet (a) eine links/rechts Makro/Makro-Erklärung mit einem früheren Ereignis zur Erklärung eines späteren Ereignisses von (b) einer unten/oben Mikro/Makro-Erklärung, bei der so genannte Oberflächen-Phänomene durch Mikro-Phänomene erklärt werden (vgl. Searle 1993, 2001; Kap. 4.5.2).

Das implizite Kausalitätsmodell lässt sich anhand der Idee, die hinter dem Supervenienzkonzept steht, exemplifizieren. Die grundlegende Idee geht dabei nach Kim auf „Edward's Diktum" zurück. Danach gibt es eine Spannung bzw. einen Widerspruch zwischen vertikaler Determination und horizontaler Kausalität. Tatsächlich ist horizontale Kausalität aufgrund vertikaler Determination ausgeschlossen (Kim 2003, 153). Unter vertikaler Determination ist zu verstehen, dass beispielsweise die Eigenschaften einer Statue, wie Farbe, Form, Textur, Dichte, Härte oder elektrische Leitfähigkeit, „vertikal" durch synchrone Mikroeigenschaften bestimmt sind. Im Gegensatz dazu werden diachrone kausale Relationen üblicherweise horizontal von links (Vergangenheit) nach rechts (Gegenwart) dargestellt. Die Vergangenheit bestimmt die Zukunft, und die Zukunft hängt von der Vergangenheit ab (vgl. ebd. 153f.). Körper-Geist-Supervenienz entspricht dabei der vertikalen Determination, während mentale Kausalität bzw. die Kausalität auf der höheren Ebene der horizontalen Kausalität entspricht.

Eigenschaften als solche sind nicht Bestandteil kausaler Relationen, aber Eigenschaftsinstanziierungen. Wenn es heißt, M verursacht M*, ist dies die kurze Form der Formulierung, die Instanziierung von M verursacht die Instanziierung von M*. Von der Supervenienzbeziehung wissen wir, dass M als Supervenienzbasis P und M* als Supervenienzbasis P* hat. Auch wenn man, wie Kim, davon ausgeht, dass M nicht M* horizontal verursacht, so scheint doch auf der unteren Ebene P kausale Ursache von P* zu sein und zwar in dem Sinne, dass P* von P „angestoßen" wird. So kann in einem Konditional formuliert werden, dass ohne P auch kein P* bzw. P* nur durch eine andere physische Eigenschaft verursacht wird, selbst wenn es in einer möglichen Welt W für das System S die mentalen Zustände M und M* gibt (vgl. Kim 2003, 157, 160f.).

Das Kausalitätsmodell, welches innerhalb der funktionalistischen Position mehr oder weniger explizit verwendet wird, ähnelt dem Modell der Supervenienztheorie. Allerdings ist es notwendigerweise um die input-/output-Relation ergänzt, da diese Relation für den Funktionalismus essentiell ist. Zudem wird eindeutig auf ein inneres Netzwerk von mentalen bzw. funktionalen Zuständen hingewiesen, um das die Relation eines jeden einzelnen funktionalen Zustandes zu erweitern ist. Auch hierbei handelt es sich um eine wirkmechanistische Konzeption, innerhalb derer mutmaßlich die Eigenschaften auf der niedrigeren Ebene die eigentliche kausale Relation darstellen (vgl. Block 1980b, 1990). Jackson geht sogar so weit, dass er Menschen mit Pflanzen vergleicht, die auf Objektreize determiniert reagieren, wie z. B. Pflanzen, die sich in Richtung Sonne bewegen. In ihnen würde man einen Zustand finden, der dieses „Verhalten" erklärt. Es wäre ein (philosophischer) Fehler zu behaupten, diese internen Zustände würden dieses Verhalten nicht erklären. Selbiges soll für Tiere gelten und schließlich auch für Menschen. Zwischen dem Input und dem Output besteht im Unterschied zu den Pflanzen beim Menschen lediglich eine sehr komplexe neuronale Verarbeitung (vgl. Jackson 1998, 95f.).[86]




Grafik 8: In der theoretischen Diskussion über mentale Kausalität werden die mentalen und physischen Eigenschaften als kausale Konkurrenten beschrieben. Unter den gegebenen Voraussetzungen stellt sich die Frage, welche Möglichkeit für die mentalen Eigenschaften bestehen bleibt, kausal wirksam zu sein. Schließlich kann allein auf der physischen Ebene eine kausale Ursache-Wirkungs-Relation beschrieben werden. Um mentale Kausalität tatsächlich zu fassen, muss jedoch die Homogenität einer mentalen und einer physischen Kausalität bestritten werden. Deshalb müssen eigenständige Kriterien für mentale Kausalität herausgearbeitet werden.

Searle kritisiert, dass in der zeitgenössischen Diskussion mit dem „Standardmodell der Verursachung" (Searle 2001, 75) kein adäquates Modell für mentale Kausalität entwickelt wurde. Er verweist auf Piaget, der in seiner Arbeit zur frühkindlichen Entwicklung gezeigt hat, dass der primitivste Verursachungsbegriff des Kindes ein „Stoß-Zug"-Begriff ist (vgl. Jean Piaget 1927). Die Erfahrung des Kindes, dass die Welt durch mechanistische Prinzipien funktioniert, wird durch ein wissenschaftliches Verständnis zu einer reicheren Konzeption kausaler Beziehungen erweitert. Nicht nur Gegenstände, sondern auch Krankheiten oder soziale Zustände können ver-

[86] Allerdings sollte bedacht werden, dass 1.) Pflanzen keine Gehirne haben, hier liegen andere interne Prozesse zu Grunde, weshalb auch ein größerer und ggf. vollständig anderer Spielraum entsteht; 2.) Pflanzen (und auch einige Tiere) kein reflexives Bewusstsein haben und somit nicht die Möglichkeit, sich von der Sonne weg zu drehen, eine Handlungsfreiheit, die wir den Menschen zuschreiben.

ursacht werden (vgl. Searle 2001, 76).[87] Hinter der Kritik des Epiphänomenalismus, dass Bewusstsein keine physikalische Wirkung, wie z. B. Armheben, verursachen kann, verbirgt sich nach Searles Ansicht deshalb die für die Beschreibung von Bewusstseinsverursachung ungeeignete Kombination aus einem „Rest-Dualismus mit einer extrem naiven Konzeption von Kausalbeziehungen" (ebd.).

Die gängigen Beispiele für mentale Kausalität übertragen einfache wirkmechanistische Modelle auf Handlungen, Empfindungen oder letztlich auf Denkprozesse. Es ist beispielsweise der Gedanke an ein Glas Milch, der den Anstoß gibt, dass der Körper nächtens aus dem Bett aufsteht, um in der Küche jenes zu trinken. Diesem Gedanken an die Milch liegt möglicherweise ein Durstgefühl zugrunde. Das entsprechende Hirnareal registriert, dass der Körper Flüssigkeit benötigt und bewirkt, dass andere Areale dafür sorgen, dass der Körper wach wird, aufsteht und Flüssigkeit zu sich führt. Andererseits signalisiert der Schmerz der Hand auf der heißen Herdplatte den motorischen Arealen, schnellstens die Hand wegzuziehen – idealerweise instinktiv und nahezu unbewusst. Der Gedanke an „Mozart" verursacht den Gedanken an „Don Giovanni", was wiederum verschiedene Erinnerungen oder Planungen und Handlungen, wie z. B. den Kauf einer Konzertkarte, nach sich zieht.

Es handelt sich hierbei zwar um verschiedene kausale Richtungen, die beschrieben werden. Gemeinsam ist den Beispielen aber allen, dass gemeinhin ein kausaler Wirkmechanismus impliziert wird, wonach ein Ereignis ein anderes „anstößt", oder eine Eigenschaft eine andere „anstößt", oder ein Teilchen ein anderes „anstößt" und somit verursacht. Aufgrund des reduktionistischen Ansatzes in der Philosophie des Geistes könnte eine Definition von mentaler Kausalität entsprechend lauten:

$$\text{mentale Kausalität}_{\text{Def.}} = \quad \text{mentale Eigenschaften + causa efficiens unter physischen Voraussetzungen.}$$

Beispiel ist hier unter anderem das Kausalitätsmodell von Papineau, wonach mentale Eigenschaften nur aufgrund ihrer Identität mit physischen Eigenschaften kausal wirksam sein können – unter den gegebenen metaphysischen Voraussetzungen (vgl. Kap. 2.1.2). Für die Kausalität mentaler Eigenschaften gilt für ihn das Prinzip der Transitivität: Wenn die belebten Körperwirkungen der Bewusstseinsursachen wiederum unbelebte Wirkungen haben, folgt daraus direkt, dass die Bewusstseinsursachen diese unbelebten Wirkungen haben:

[87] Für Searle geht der Ursachebegriff jedoch weiter: „Kurz, Verursachung ist nicht nur eine Sache von Stoßen und Ziehen; bei Verursachung geht es darum, dass etwas dafür verantwortlich ist, dass etwas anderes geschieht." (Searle 2001, 76) Deshalb ist sein Ursachebegriff im Kern mit Intentionalität verbunden (vgl. Kap. 4.5.2).

> „Thus, if some conscious desire causes my arm to move, and this movement in turn has such inanimate effects as a stone flying through the air, […] then my conscious desire itself will cause these inanimate effects." (Papineau 2004, 44)

Bewusste Systeme können also kausal wirksam sein, sofern ihre bewussten Eigenschaft – aufgrund ihrer Identität mit physischen Eigenschaften – unbelebte oder nicht-mentale Wirkung haben. Papineau macht sich ausdrücklich dafür stark, dass bewusste mentale Eigenschaften physische Wirkung haben.

Es ist problematisch, verbindliche Kriterien für mentale Kausalität anzugeben, weil mit dem Begriff der „mentalen Kausalität" nicht nur verschiedene Ansprüche verbunden werden, sondern auch verschiedene kausale Ebenen benannt werden. So hat die Beschreibung einer mentalen Kausalität zwischen Ereignissen oder Eigenschaften verschiedene Konsequenzen. Es kann angenommen werden, dass mentale Ereignisse sehr wohl kausal wirksam sind, da sie – wie Davidson es formuliert – als Token-Ereignisse beschreibungsunabhängig und identisch mit physischen Ereignissen sind. Oder mentale Ereignisse sind aufgrund ihrer physischen Eigenschaften kausal wirksam, wie das Supervenienzargument in seiner ursprünglichen Form nahe legt (vgl. Kim 1982, 1984). Oder aber es wird vertreten, dass mentale Ereignisse aufgrund ihrer mentalen Eigenschaften wirksam sind (vgl. Menzies 2003). Man kann die Frage der mentalen Kausalität aber auch direkt und unabhängig von mentalen Ereignissen auf der Ebene der mentalen Eigenschaften diskutieren, wobei sich die grundsätzliche Frage stellt, ob mentale Eigenschaften kausal wirksam sind oder nicht, und wie sie wirksam sein können.

Deshalb ist es problematisch, wenn die Ereignis-Situation auf Eigenschaften übertragen wird. Wenn zwei Ereignisse in einer Kausalbeziehung zueinander stehen, kann untersucht werden, welche Eigenschaften des ersten Ereignisses für die Eigenschaften des zweiten Ereignisses verantwortlich sind. Ereignisse sind die Instanziierungen von Eigenschaften durch Dinge zu einer bestimmten Zeit (vgl. Kim 1976; vgl. Kap. 3.1.1). Beispielsweise durchschlägt ein Stein ein Fenster, so dass es zerbricht. Zu einem bestimmten Zeitpunkt hat der Stein unter anderem die Eigenschaften Farbe, Masse und Geschwindigkeit. Die letzteren beiden Eigenschaften sind für das Zerbrechen des Fensters – neben der Eigenschaft des Fensters, zerbrechlich zu sein – mitverantwortlich, die Eigenschaft Farbe jedoch nicht. Wer mentale Eigenschaften für nicht identisch und nicht-reduzierbar mit bzw. auf physische(n) Eigenschaften hält, scheint eine Bringschuld leisten zu müssen, da die physikalischen Voraussetzungen anscheinend keinen Spielraum für mentale Eigenschaften als kausale Kräfte zulassen.

Dabei spielt auch eine Rolle, ob mentale Kausalität auf einer rein psychischen, auf einer rein neuronalen Ebene oder auf einer mikrobiologischen Ebene vorkommt.[88] Oder aber sollte der Begriff der mentalen Kausalität möglichst alle verschiedenen Ebenen umfassen?

Wenn wir die Frage der mentalen Kausalität diskutieren, müssen wir genauer die Kausalrelationen unterscheiden, in denen mentale Ereignisse bzw. mentale Eigenschaften vorkommen (vgl. Kim 1998, 139):

1. mental-physische Kausalität,
2. physisch-mentale Kausalität
3. mental-mentale Kausalität
4. physisch-physische Kausalität

Sowohl das Modell einer Schichten-Welt, als auch das Prinzip der Multirealisierbarkeit bestimmen mentale Kausalität vorwiegend im Sinne von 2. Reduktionisten und Eliminativisten würden mentale Kausalität dagegen vorwiegend im Sinne von 4. verstehen, die fehlerhafte psychologische Sprache verschwände ganz. Da mentale Eigenschaften physisch realisiert sind, ist 3. schwer akzeptierbar, und im Alltagsverständnis vermutlich als eine Kette von Glaubensannahmen zu erklären. Letztendlich hängt die Möglichkeit von 3. sehr stark von 1. ab. Bei der Frage nach mentaler Kausalität ist demnach vor allem die Frage spannend, inwiefern die Relationen 1. und 2. möglich sind, also eine Kausalität von mentalen Eigenschaften auf physische.

Bei dem Problem der mentalen Verursachung muss also auch zwischen der Verursachungsrichtung unterschieden werden. Traditionell bestand für Descartes das Problem, wie der ausdehnungslose Geist ohne Masse und Energie kausalen Einfluss auf den ausgedehnten Körper haben kann, bzw. wie umgekehrt bestimmte körperliche Zustände eine Wirkung auf das geistige Erleben haben können. In der zeitgenössischen Diskussion wird dieses Problem zwar nicht mehr als ein Problem zwischen zwei Substanzen betrachtet. Aber je nach Position ergibt sich die Unterscheidung zwischen einer Aufwärtsverursachung und einer Abwärtsverursachung zwischen der Mikro-, bzw. Makro-Ebene der physischen bzw. mentalen Eigenschaften (vgl. Kim 1999b, 2003). Interessanterweise scheinen die Schwierigkeiten einer Verursachung vom Körperlichen zum Geistigen bzw. von physischen zu mentalen Eigenschaften weniger zu bestehen, als in der umgekehrten Richtung. Es scheint in der Debatte weithin akzeptiert, dass physische Eigenschaften als Realisierer, subveniente Basen oder Eigenschaften erster Ordnung gegebenenfalls die mentalen,

[88] Die neuronale bzw. psychische Ebene ist allein deshalb von Kausalität zwischen mentalen oder physischen Eigenschaften zu unterscheiden, da es grundsätzlich auf der neuronalen Ebene mentale Eigenschaften und auf der psychischen Ebene physische Eigenschaften geben kann. Diese Möglichkeit impliziert zumindest die Position des biologischen Naturalismus (vgl. z. B. Searle 1992, 1998; Hornsby 1986, 1997).

supervenierenden Eigenschaften zweiter Ordnung verursachen. Da der Kausalitätsbegriff eng mit der Vorstellung einer Energieübertragung zusammenhängt, fällt es intuitiv schwerer, eine Verursachung vom Geistigen zum Körperlichen zu vertreten, da Geist an sich keine Energie besitzt, die übertragen werden kann. Diese Intuition täuscht jedoch, da aufgrund der vorausgesetzten Verschiedenheit mentaler und physischer Eigenschaften eine Verursachung in beide Richtungen problematisch ist (vgl. Eccles 1994; Lowe 2003). Sowohl die Bewegung des Körpers aufgrund einer Idee, als auch eine Veränderung des Wohlbefindens aufgrund einer physischen Veränderung sind per se erklärungsbedürftig.[89] Diese Problemlage bleibt eben auch für die zeitgenössische Diskussion bestehen, wenn nicht mehr von unterschiedlichen Substanzen, sondern von unterschiedlichen Eigenschaften ausgegangen wird.

Mögliche Kriterien für mentale Kausalität sollten sich vor diesem Hintergrund an den Voraussetzungen orientieren, die common sense für die gängigen Positionen und in der Diskussion sind, und an den Kriterien für mentale und physische Ereignisse bzw. Eigenschaften selbst. So werden als Kriterien für mentale Eigenschaften angegeben, dass sie bewusst oder unbewusst, nicht räumlich, intentional, privat, extrinsisch, propositional, qualitativ, anomal und multirealisierbar sind (vgl. Kap. 3.2.1). Kriterien für physische Eigenschaften sind ihre räumlich Ausdehnung, Messbarkeit sowie empirische Überprüfbarkeit (vgl. Kap. 3.2.2). Mentale Kausalität wird somit zum Diener zweier Herren. Die Kriterien für mentale Kausalität sollten nicht als eine Addition der Kriterien mentaler und physischer Eigenschaften verstanden werden, allerdings sollten sie den physikalischen Anforderungen, sowie dem spezifisch „Mentalen" der kausalen Relation gerecht werden. Da sich der Eigenschaftsbegriff in der Diskussion durchgesetzt hat (vgl. Kim 1976, 1999a; Davidson 1999; Kap. 3.1.3), soll im Folgenden zunächst eine stärkere Anbindung der Kriterien für mentale Kausalität an diesem erfolgen, ohne jedoch den Ereignisbegriff gänzlich zu vernachlässigen. Als Kriterien für mentale Kausalität können genannt werden:

1. Regularität
2. Multikausalität
3. Inhalt
4. Relationalität
5. Intentionalität
6. Proportionalität

[89] Im Rahmen eines Substanzdualismus oder eines Epiphänomenalismus ist Verursachung deshalb problematisch, weil aufgrund des Energieerhaltungsgesetzes bei jeglicher Form der Verursachung Energie übertragen und umgewandelt werden muss. Bei Veränderungen des Geistes ist dies jedoch nicht der Fall, sie werden ohne Energieübertragung bewirkt. Vor diesem Hintergrund führt der Epiphänomenalismus die Verursachung mentaler Zustände durch andere auf die gesetzmäßige Folge physischer Zustände zurück, wobei diese physischen Zustände bestimmte mentale verursachen. Die mentalen Zustände selbst sind jedoch wirkungslos (vgl. Kap. 3.6.1).

Diese Kriterien werden im Folgenden hinsichtlich ihrer Anwendbarkeit auf mentale Kausalität überprüft und diskutiert. Problematisch ist allerdings eine genaue Zuordnung, wie sich zeigen wird, da in der Diskussion um Eigenschaften nach wie vor strittig ist, ob Eigenschaften intrinsisch oder extrinsisch sind, ob der Inhalt von Eigenschaften intern oder extern ist. Yablo weist darauf hin, dass es grundsätzlich problematisch ist, für Kausalität notwendige Bedingungen zu bestimmen, abgesehen davon, dass es Ursache und Wirkung gibt und beide voneinander zu unterscheiden sind (vgl. Yablo 1992, 277). Es ist auch deshalb schwierig, für mentale Kausalität Kriterien zu bestimmen, da allein für das Mentale keine allgemeingültigen Kriterien feststehen (vgl. z. B. Searle 1984, 17f.). Einige Kriterien des Mentalen können auch auf mentale Kausalität übertragen werden. In der Konsequenz ist die Entwicklung von Kriterien für mentale Kausalität als Hypothesenbildung zu verstehen, die eine weitere Ausdifferenzierung ermöglicht. Einen Absolutheitsanspruch können die genannten Kriterien dagegen nicht erfüllen. Es ist vermutlich möglich, weitere Kriterien zu benennen, die eine Relation zu erfüllen hat, um als Variante der mentalen Kausalität zu gelten.

### 4.1 Das Kriterium der Regularität

Das Kriterium der Regularität ergibt sich bereits aus der Voraussetzung der strikten Gesetze, aber auch in Verbindung mit der kausalen Geschlossenheit[90] und dem ontologischen Grundmodell. Wenn wir von Kausalität sprechen, erwarten wir eine Relation zwischen einer Ursache A und einer Wirkung B, die nicht zufällig, sondern regelmäßig und unter Anwendung bestimmter Gesetze stattfindet. Dieser Anspruch gilt auch für mentale Kausalität. Wenn eine hinreichende Ursache auftritt, soll unter gegebenen Umständen auch die Wirkung eintreten. Die Ursache-Wirkung-Beziehung wird insofern als notwendige Folge verstanden (vgl. Kap. 7). Allerdings hat bereits David Hume auf den Umstand hingewiesen, dass Kausalurteile über die menschliche Erfahrung hinausgehen und induktiv geschlossen werden (vgl. Kienzle 1997, 118ff.; Beauchamp 1981, 33ff.). Durch Beobachtung lassen sich keine Notwendigkeiten feststellen, sondern lediglich das zeitliche Aufeinanderfolgen zweier Ereignisse in ihrer Regelmäßigkeit. Aus dieser Regelmäßigkeit wird dann die Kausalrelation abgeleitet. In zeitgenössischen Varianten der Regularitätstheorie wird wieder die notwendige Folge als naturgesetzliche Folge betont, wie in Grundsatz G* zum Ausdruck kommt:

[90] Hinsichtlich der kausalen Geschlossenheit soll an dieser Stelle lediglich angemerkt werden, dass nach dieser Voraussetzung keine ein physisches Ereignis involvierende Kausalkette die Grenze zum Nicht-Physischen überschreiten darf (vgl. Kap. 2.5.3). Bei einer mentalen Kausalität scheint diese Grenze jedoch per definitionem überschritten zu werden. Ein Ausweg ist konsequenterweise Searles Vorschlag, die Definition, was Mental ist, soweit zu ändern, dass mentale Zustände bzw. Bewusstseinszustände als biologisches Merkmal verstanden werden und somit in gewisser Weise auch physisch sind (vgl. Searle 1993, 109; Hornsby 1999).

Grundsatz (G*):

> Man sagt, ein Ereignis E sei Ursache eines Ereignisses E', wenn es Naturgesetze gibt und Umstände, die beim Eintreten von E bestehen, so dass E' logisch aus den Gesetzen, den Umständen und E folgt (aber auch nur aus ihnen allen zusammen), und wenn das Ereignis E' später eintritt als E. (Kutschera 2003, 72)

Problematisch ist, dass zum einen nicht zwischen einer Ursache und einer früheren Wirkung einer gemeinsamen Ursache unterschieden werden kann, und dass zum anderen die überwiegenden kausalen Erklärungen im Alltag wie in der Wissenschaft nicht die strenge theoretische Form haben. Wir akzeptieren als kausale Erklärung, dass Herr P. einen Autounfall hatte, weil eine schwarze Katze über die Straße lief, ohne genaue Gesetze und Antezedenzbedingungen zu kennen (vgl. ebd.). Diese Diskrepanz zwischen theoretischem Anspruch und praktischer Umsetzung spielt auch im Rahmen der Diskussion über mentale Kausalität eine wichtige Rolle, beispielsweise hinsichtlich der Voraussetzung strikter Gesetze. An dieser Stelle soll nun untersucht werden, inwiefern vertikale bzw. horizontale Regularitäten diese Anforderungen erfüllen können. Es wird dabei bewusst der Begriff der „Regularität" verwendet, da nicht abschließend beurteilt werden kann, ob es sich bei den Relationen um strikte, oder nicht-strikte Gesetze oder sogar um Realisierungsrelationen handelt. Der Begriff der Regularität verweist insofern auf die Regelmäßigkeit, mit der mentale und physische Ereignisse bzw. Eigenschaften auftreten und sich möglicherweise verursachen.

### 4.1.1 Vertikale Regularität

Übliche Beispiele für mentale Kausalität sind die Initiierung der Handlung, ein Glas Milch zu trinken, oder der Plan, am Abend ins Kino zu gehen. Solche propositionalen Einstellungen sind gegebenenfalls mit qualitativen Empfindungen verbunden, wonach Person P sehr gut weiß, wie es ist, ein Glas Milch zu trinken, bzw. die Atmosphäre eines Filmpalastes zu genießen versteht. Aus der Perspektive der vertikalen Regularität soll nun eine gesetzmäßige Abhängigkeit der mentalen Eigenschaften von den physischen Eigenschaften festgestellt werden, falls eine solche gesetzmäßige Abhängigkeit besteht. Aufgrund der Anwendung des oben aufgestellten Grundsatzes (G*) muss nun prinzipiell gezeigt werden, dass es Naturgesetze und Umstände gibt, aus denen folgt, dass physische Eigenschaften die Ursachen der mentalen Eigenschaften sind.

Allerdings besteht darin das Problem: Davidson hat analysiert, dass mentale Ereignisse bzw. Eigenschaften zumindest keinen strikten Gesetzen unterliegen, und es in der Folge auch keine psychophysischen Gesetze geben kann (vgl. Kap. 2.3; 3.5.1). Die Ablehnung der Existenz psychophysischer Gesetze bezieht sich auf intentionale Ereignisse. Die physische Realisierung

mentaler Ereignisse, wie z. B. Schmerzen durch das Feuern von C-Fasern, soll gesetzesartig gewährleistet bleiben. Hier stellt sich die Frage, ob diese Trennung bei mentalen Ereignissen, die psychologische und phänomenale Begriffe einschließen, so durchzuhalten ist (vgl. Chalmers 1997, 25ff.; Papineau 2004, 97ff.; Kap. 1.1). Unter welche Regelung würden solche „hybride Eigenschaften" fallen? Vermutlich dürfte es auch für sie keine strikten Gesetze geben, da die Anomalität der psychologischen Eigenschaften die Gesetzmäßigkeit der phänomenalen Eigenschaften aufheben dürfte. Außerdem sollte bedacht werden, dass für den mentalen Bereich bzw. für psychologische Eigenschaften zwar keine strikten Gesetze, aber zumindest Rationalitätsprinzipien gelten (vgl. Kap. 2.3). Dennoch schreibt Davidson mentalen Ereignissen eindeutig kausale Wirkung zu, nämlich als singulären Ereignissen, die mental und physisch beschreibbar sind. Die Ereignisse unterliegen als Bestandteil der kausalen Welt wiederum nomologischen Anforderungen; schließlich müssen Ereignisse ein Gesetz instanziieren, damit wissenschaftlich sinnvoll von einer Ursache-Wirkung-Relation gesprochen werden kann (vgl. Kap. 3.5.). Diese Lösung vermeidet allerdings eine genauere Bestimmung der Relation zwischen physischen und mentalen Eigenschaften und wird deshalb kritisiert (vgl. Kap. 2.3.3).

Als eine Alternative werden immer wieder die so genannten ceteris-paribus-Gesetze aufgeführt. Schließlich handelt es sich hierbei nicht um strikte Gesetze, sondern um Generalisierungen, die unter Normalbedingungen zutreffen. Sie scheinen zumindest der Vagheit der mentalen Eigenschaften eher gerecht zu werden, da sich bei ceteris-paribus-Gesetzen das Antezedens durchaus verändern kann (vgl. Kap. 3.5.2). Aber auch sie betonen die Regelmäßigkeit, mit der ein (mentales) Ereignis bzw. eine mentale Eigenschaft ein(e) weitere(s) verursachen kann. Damit bieten sie zumindest eine Grundlage, auf der unter idealen Bedingungen auch eine gesetzmäßige vertikale Relation zwischen physischen und mentalen Eigenschaften beschrieben werden kann. Problematisch bleibt für ceteris-paribus-Gesetze, wie auch für strikte Gesetze, darzustellen, inwiefern eine physische Eigenschaft hinreichend für die Verursachung einer mentalen Eigenschaft wäre.

Offensichtlich gibt es hier große Unterschiede in der Natur, denn nicht alle physischen Eigenschaften scheinen qualifiziert, um bei dem zugehörigen Objekt oder auch Lebewesen die entscheidenden psychologischen Vorgänge zu initiieren. Ein Stein, eine Amöbe oder ein Fisch haben eben keine oder eine andere Mentalität als Kühe, Seehunde oder Menschen. Es sind also ganz bestimmte physische Eigenschaften gefordert, damit mentale Eigenschaften verursacht werden können. Möglicherweise spielt hier einmal die Beschaffenheit, der Stoff, eine Rolle – Gehirne eignen sich besser als Erzklumpen. Daneben scheint auch die systematische Anordnung der Eigenschaften bei der Verursachung von Mentalität relevant zu sein (vgl. Kap. 3.2.3). Offen sind auch die Umstände, unter denen physische Ereignisse mentale verursachen könnten. Man kann zumindest schließen, dass sie aus einer biologischen Perspektive konstant sein sollten, beispiels-

weise muss das Gehirn eines Menschen ausreichend mit Blut bzw. Sauerstoff versorgt sein, der Energiebedarf des Körpers muss befriedigt werden.

Alternativ zu einer nomologischen Relation lässt sich die vertikale Regularität auch als Realisierungsrelation zwischen mentalen und physischen Eigenschaften beschreiben. Dies leisten funktionalistische und auch supervenienztheoretische Positionen. Die starke Bindung zwischen mentalen und physischen Eigenschaften wird durch die gleichzeitige Instanziierung beider Eigenschaften gewährleistet, wobei erstere Realisierungen der letzteren sind (vgl. Kap. 2.2). Für die Supervenienzbeziehung gilt in der allgemeinen Form nur die Anforderung der Korrelation, dennoch soll diese Beziehung möglichst stark und eng beschrieben werden. Schließlich sollen sich die supervenierenden Eigenschaften die kausale Kraft von den subvenienten Eigenschaften leihen – dies ist zumindest das Prinzip der „supervenienten Kausalität" (vgl. Kim 2003, 153f.). Wer, wie Kim inzwischen, diese ablehnt und die starke Supervenienzbeziehung als Form der Reduktion betrachtet, muss ebenfalls eine starke Beziehung annehmen, die über die bloße Korrelation hinausgeht (ebd.). Auch der Funktionalismus operiert im Grunde unter der Annahme einer engen Relation. Anders wären die Rollen und ihre Realisierer nicht aneinander zu binden. So sollen auf der vertikalen Ebene die höherstufigen Eigenschaften immer die Rolle innehaben, die sie durch die niedrigstufigen Eigenschaften umsetzen können (vgl. Kap. 2.2.1). Innerhalb einer nomologischen Relation bleibt eine gewisse ontologische Unabhängigkeit zwischen Ursache und Wirkung bestehen, da das Auftreten der Ursache nicht notwendig zum Eintreten der Wirkung führt (vgl. Levine 2001, 14). Es ist fraglich, ob auch das mereologische Weltmodell mit seiner hierarchischen Struktur die nomologische Beziehung zwischen den einzelnen Ebenen benötigt, um die kausale Geschlossenheit der physikalischen Welt zu sichern, oder ob hierfür die Realisierungsrelation zwischen mentalen und physischen Eigenschaften nicht ausreicht (vgl. Kap. 3.1.3; Kap. 3.2). Deshalb wird in der Forschung inzwischen die Realisierungsrelation favorisiert. Immerhin wird für jede niedrigstufige wie höherstufige Ebene eine eigene Taxonomie mit eigenständigen Gesetzen veranschlagt, was von reduktionistischer Seite jedoch bestritten wird. Kausalgesetze zwischen den Ebenen scheinen somit nicht mehr notwendig zu sein. Gesichert scheint dagegen vielmehr die Abhängigkeit der höherstufigen Ebenen von den kausalen Gesetzmäßigkeiten auf der niedrigstufigen Ebene bzw. der höherstufigen Eigenschaften von den niedrigstufigen. Problematisch ist auch hier, dass aufgrund der Multirealisierung mentaler Eigenschaften keine eindeutige Relation beschrieben werden kann. Als physische Eigenschaften kommen in allen möglichen Welten menschliche Gehirne ebenso wie Tintenfischgehirne oder Marsmenschengehirne in Frage, um mentale Eigenschaften zu realisieren. Deshalb sind höherstufige Eigenschaften nicht reduzierbar (vgl. Kap. 4.2.1).

Es scheinen also in der Essenz keine psychophysischen Gesetze bzw. keine kausale Abhängigkeit notwendig zu sein, um die Relation zwischen mentalen und physischen Eigenschaften zu

erfassen. Alternativ lässt sich die Relation auch sehr gut als Realisierung beschreiben. Damit wird auch das Problem singulärer psychophysischer Gesetze, die zumindest eine Token-Relation garantieren, vermieden.[91]

### 4.1.2 Horizontale Regularität und mentale Kausalität

Die Relation zwischen Ursache und Wirkung auf der horizontalen Ebene soll nomologisch gesichert sein. Die Instanziierung von Eigenschaften kann im Rahmen des Schichten-Welt-Modells einmal explanatorisch, darüber hinaus aber auch als nomologisch notwendig aufgefasst werden (LePore/Loewer 1989, 179; vgl. Kap. 3.1.3). Auf diese gesetzmäßige Regelmäßigkeit weisen sowohl Davidsons als auch Kims Ereignisbegriffe hin. Hierin liegt der entscheidende Unterschied zwischen der vertikalen und der horizontalen Regularität, denn Eigenschaften als solche sind nicht Bestandteile von Kausalrelationen, sondern lediglich Eigenschaftsinstanziierungen (vgl. Kap. 3.1.1; Kap. 3.1.2). Gerade Kim betont, dass Ereignisse als Eigenschaftsinstanziierungen Elemente von Kausalbeziehungen sind, weshalb sein feinkörniger Ereignisbegriff die Veränderung innerhalb einer Zeitspanne betont. Dabei handelt es sich um konkrete, partikulare Instanziierungen, die bei geringer Veränderung eine andere Wirkung hervorrufen. Allerdings kann man sich dann schon fragen, warum nur den physischen Eigenschaften kausale Kraft zu- und den mentalen Eigenschaften kausale Kraft abgesprochen wird, wenn letztlich doch nur die Ereignisse als Eigenschaftsinstanziierungen Bestandteil einer Kausalrelation sind.

Die Anforderung des Grundsatzes (G*) betrifft auch die horizontale Regularität. Gerade für zeitlich aufeinander folgende mentale Ereignisse muss prinzipiell gezeigt werden, dass es Naturgesetze und Umstände gibt, die sie zu einer kausalen Relation machen. Es ist eben immer der Fall, dass – auf der höherstufigen bzw. supervenienten Ebene – der mentale Zustand $M_1$ den mentalen Zustand $M_2$ bzw. die Handlung $H_2$ verursacht. Und es ist immer der Fall, dass – auf der niedrigstufigen bzw. subvenienten Ebene – der physische Zustand $P_1$ den physischen Zustand $P_2$ bzw. die motorische Bewegung $B_2$ verursacht. [92] Problematisch ist, dass zunächst a) nur auf der physischen Ebene die nomologische Sicherung möglich erscheint, b) in der Folge die Gesetze auf der physischen Ebene die mentale Ebene nicht wirklich beschreiben sowie c) eine Cross-Level-

[91] Wie können solche „Einzelfallgesetze" aussehen? Sind darunter ceteris-paribus-Gesetze zu verstehen? Dies wohl eher nicht, da für den Fall des Eintretens der bestimmten Relation auch gesetzmäßig beide Eigenschaften der Relation eintreten sollen und nicht lediglich ceteris paribus.

[92] In der Literatur lassen sich weitere Beispiele finden, die für eine gesetzmäßige Relation der mentalen Kausalität sprechen. So bevorzugt Block (1990) aus guten Gründen Nomologizität gegenüber kontrafaktischen Konditionalen für die Beschreibung mentaler Kausalität. Auch das Kausalargument der Identitätstheorie kann als Indiz dafür betrachtet werden, dass mentale Kausalität eine nomologische Relation ist (vgl. Kap. 2.1.2).

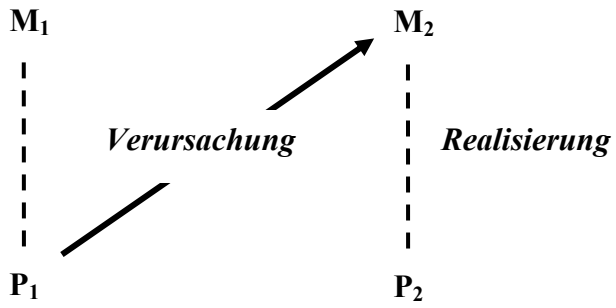
Kausalität, also eine vertikale Kausalität zwischen den Ebenen ausgeschlossen scheint (vgl. Kap. 3.5).

Dass a) eine nomologische Relation nur auf der physischen Ebene möglich erscheint, kann jedoch als epistemisches Problem aufgefasst werden. Schließlich ist es uns derzeit nicht möglich, strikte gesetzmäßige Voraussagen für psychologische Verhaltensweisen bzw. mentale Eigenschaften zu treffen, was jedoch nicht bedeutet, dass es hier keine nomologische Relation gibt. Immerhin können auf Basis von ceteris-paribus-Gesetzen für einige psychologische Dispositionen inzwischen wissenschaftliche Aussagen und Vorhersagen getroffen werden; Beispiel ist hier das Weber-Fechnersche Gesetz über den Zusammenhang von Reiz- und Empfindungsstärke (vgl. Kutschera 2003, 32f.; Kap. 3.5.2). Im Grunde wiederholen sich einige der Überlegungen zur vertikalen Regularität zu strikten bzw. nicht-strikten Gesetzen. Ein Fortschritt in der Forschung lässt auch hierfür einen Wissenszuwachs erwarten. Schließlich ist es derzeit auch nicht möglich, allein aufgrund neurologischer Daten genaue Prognosen für weitere neurologische Aktivität oder Handlungen zu bestimmen. Auf der physischen Ebene wird schließlich auch unter der Annahme einer vollständigen Physik gearbeitet (vgl. Quine 1951, 1995), die als substanzielle empirische Vermutung gelten kann. Dennoch kann ihr auch eine ideale vollständige Psychologie gegenübergestellt werden, in der gesetzmäßige Regularitäten zwischen mentalen Ereignissen beobachtet werden können.

Dass b) die Gesetze auf der physischen Ebene die mentale Ebene nicht wirklich beschreiben, ist ebenfalls vorrangig ein epistemisches Problem, welches jedoch nicht direkt die Nomologozität betrifft. Sollte es, wie für a) beschrieben, für psychologische Dispositionen gesetzmäßige Beschreibungen geben, ist es unter der Annahme der Multirealisierbarkeit nicht verwunderlich, dass allein auf Basis physischer Gesetze keine Aussage über psychologische Gesetze getroffen werden kann. Wenn mentale Eigenschaften aufgrund ihrer Multirealisierbarkeit nicht typenidentisch mit physischen Eigenschaften sind, ist auch keine Reduktion möglich.[93] Eine zusätzliche Schwierigkeit ergibt sich dabei aus der Überlegung, dass für unendliche Bereiche die Angabe der kausalen Rolle schwer möglich ist, da die Summe aller Ursachen und Wirkungen berücksichtigt werden muss. Diese setzt sich jedoch aus einer unabsehbar großen Zahl anderer Zustände und Ereignisse zusammen, für die wiederum eine entsprechende Summe angegeben werden müsste (vgl. Kap. 3.1.2; Kap. 2.2). Hinsichtlich der mentalen Kausalität gilt dieses Problem theoretisch sowohl für den mentalen, als auch für den physischen Bereich. So scheint es insgesamt nicht umsetzbar, strikte Gesetze zu formulieren, die auf neurologischer Basis die nahezu unendlichen Möglichkeiten der Verknüpfung innerhalb eines Gehirns und darüber hinaus auf alle existierenden Gehirne und jene in allen möglichen Welten darstellen. Dies kann als ebenso problematisch betrachtet werden, wie die Vagheit des mentalen Bereichs für den, nach Davidson, lediglich Rati-

[93] Diese Frage wird ausführlich unter dem Kriterium der Multikausalität diskutiert (vgl. Kap. 4.2).

onalitätsprinzipien gelten (vgl. Kap. 2.3.2). Somit erscheint es sinnvoll, dass mentale Kausalität nicht zwingend strikten Gesetzen unterliegen muss. Die ausnahmslose Gültigkeit, die solche Gesetze einfordern, erscheint für den mentalen Bereich, aber auch für den physischen Bereich eher ideal und nur schwer bzw. nicht erfüllbar zu sein (vgl. Kap. 3.5.2). So wurde auch darauf hingewiesen, dass die kausale Geschlossenheit der Physik eine substanzielle empirische Vermutung ist. Natürlich müssen auch Gesetze, die mentale und physische Eigenschaften miteinander verbinden, mit denen der Physik konsistent sein. Andererseits scheinen rein physikalische Erklärungen den psychologischen Prozessen nicht gerecht zu werden. In der Konsequenz müssen gesetzmäßige physische und gesetzmäßige mentale Beschreibungen ergänzend herangezogen werden, um mentale Kausalität sinnvoll zu beschreiben. Eine Schnittmenge beider Bereiche ist zwar denkbar, liefert jedoch keine vollständige Beschreibung.[94] Wir können nicht sicher sein, dass die mentale Beschreibung „Ich plane, heute Abend ins Kino zu gehen" mit der physischen Beschreibung bestimmter neuronaler Korrelate extensionsgleich ist.




Grafik 9: Eine mögliche Cross-Level-Kausalität als Aufwärtsverursachung zwischen der physischen Eigenschaft $P_1$ und der mentalen Eigenschaft $M_2$ gerät in Konflikt mit der dargestellten Realisierungsrelation $P_2/M_2$.

Auch die Möglichkeit der c) Cross-Level-Kausalität bzw. der Kausalität zwischen den Ebenen wurde bereits ansatzweise angesprochen. Im Rahmen der horizontalen Regularität muss dabei die mögliche kausale Verbindung zwischen der mentalen Eigenschaft $M_1$ und der physischen Eigenschaft $P_1$ eines Ereignisses $E_1$ sowie des zeitlich folgenden Ereignisses $E_2$ mit der mentalen Eigenschaft $E_2$ und der physischen Eigenschaft $P_2$ untersucht werden. Grundsätzlich wird eine Abwärtsverursachung z. B. von $M_1$ auf $P_2$ auch unter Verweis auf die Voraussetzungen für mentale Kausalität ausgeschlossen (vgl. Kim 1999b; Kap. 2.4.3; Kap. 3.4). Auf der anderen Seite ist nicht plausibel, warum es eine Aufwärtsverursachung in der umgekehrten Richtung von $P_1$ auf $M_2$ geben kann. Diese wäre zumindest unter ceteris-paribus-Gesetzen möglich. Allerdings wäre $P_1$ in doppelter Verantwortung, einmal als Realisierer von $M_1$ und dann als kausale Ursache von

[94] Dies klingt, als ob der Eigenschaftsdualismus auf epistemischer und ontologischer Ebene propagiert würde. Dagegen kann jedoch gesagt werden, dass es sich ontologisch nicht um verschiedene Eigenschaftsarten, sondern lediglich um verschiedene Einzeleigenschaften handelt, wenn man neurologische und psychologische Eigenschaften voneinander unterscheidet. Beide Klassen können als physische Eigenschaften gehandelt werden, sind jedoch voneinander als nicht-identisch zu unterscheiden.

$M_2$; dies widerspräche sowohl dem Sparsamkeitsgebot, als auch der Ablehnung von $P_1$ als vertikale Ursache von $M_1$ – wieso sollte man die vertikale Verursachung ablehnen, aber eine Cross-Level-Verursachung zulassen? Zudem würde man in diesem Fall implizit mentale Eigenschaften als ontologisch unabhängig von physischen Eigenschaften einordnen, da die nomologische Relation beinhaltet, dass das Auftreten der Ursache ein Eintreten der Wirkung zur Folge haben kann, aber nicht muss. Insofern stünde diese Variante der Cross-Level-Kausalität einem mentalen Emergentismus nahe.[95] Außerdem stünde nun auch die mentale Eigenschaft $M_2$ in einer doppelten Relation, einmal als Wirkung von $P_1$ und zusätzlich als Realisierung von $P_2$. Es erscheint nicht plausibel, dass eine mentale Eigenschaft in einer Variante der Mehrfachverursachung sowohl verursacht, als auch realisiert werden muss, $P_1$ und $P_2$ stünden in einer unsinnigen Konkurrenz zueinander (vgl. Kap. 3.4.2).

***Quintessenz Kriterium Regularität***

- *Mentale Kausalität als regelmäßige Relation*
- *theoretischer Anspruch genauer Antezendenzbedingungen – praktisch nicht erfüllbar*

*vertikale Regularität:*

- *keine strikten psychophysischen Gesetze; ceteris-paribus-Gesetze unter Idealbedingungen*
- *alternativ: Realisierungsrelation bei gleichzeitiger Instanziierung*
- *konkrete physische Eigenschaften für mentale Eigenschaften notwendig (Stoff und systematische Anordnung)*

*horizontale Regularität:*

- *gesetzmäßige Relation zwischen Ereignissen als konkrete, partikulare Instanziierungen*
- *nur einige Eigenschaften kausal verantwortlich für nachfolgende Wirkungen*
- *epistemisches Problem: Erkennen bzw. Wissen von strikten Gesetzen beschränkt; kein Analogieschluss von Gesetzen auf physischer Ebene auf Vorgänge auf mentaler Ebene möglich*
- *strikte Gesetze für mentale Kausalität wenig sinnvoll wegen a) mentaler Vagheit und b) neurologisch unendlichen Möglichkeiten*
- *Notwendigkeit horizontaler ceteris-paribus-Gesetze für mentale Kausalität*

Dass dennoch zumindest eine nomologische Relation zwischen Ereignis $E_1$ und Ereignis $E_2$ bestehen muss, ist ein minimaler Konsens. Üblicherweise wird so argumentiert, dass es die physi-

[95] Der Emergentismus kann als Vorläufer der Supervenienztheorie betrachtet werden. Der Emergenzbegriff wurde maßgeblich von Charles D. Broad (1925) entwickelt (vgl. Broad 1925; Alexander 1927).

schen Eigenschaften $P_1$ des ersten Ereignisses sind, die das Eintreten des zweiten Ereignisses $E_2$ über seine physischen Eigenschaften $P_2$ bewirken. Also muss es neben der Gesetzmäßigkeit zwischen $E_1$ und $E_2$ auch eine Gesetzmäßigkeit zwischen $P_1$ und $P_2$ geben. Die mentalen Eigenschaften bleiben in dieser physikalistischen Version der mentalen Kausalität wirkungslos (vgl. Kim 2003, 159f.).

### 4.2 Das Kriterium der Multikausalität

Dass Ursachen verschiedene Wirkungen bzw. Wirkungen verschiedene Ursachen haben können, ist zunächst eine triviale Aussage. Dennoch lohnt hinsichtlich der mentalen Kausalität ein zweiter intensiver Blick auf dieses Merkmal von Kausalrelationen. Dabei zeigt sich, dass die Multikausalität als Kriterium modal bedingt ist, beispielsweise aufgrund der Multirealisierbarkeit. Sie spiegelt sich zudem aber auch empirisch in der Plastizität des Gehirns.

Es wurde bereits gezeigt, dass Multirealisierbarkeit ein wichtiges Kriterium für mentale Eigenschaften ist. So hat insbesondere die These der Multirealisierbarkeit mentaler Eigenschaften die Position der Typen-Identität in Frage gestellt und die funktionalistischen Positionen befördert (vgl. Fodor 1974; Putnam 1967a; Kap. 2.1; Kap. 2.2). Mentale Prozesse werden im Sinne des Rollenfunktionalismus vorrangig über die Funktionen, die sie erfüllen, definiert und nicht über ihre Realisierer. Jede mentale Eigenschaft ist eine höherstufige Eigenschaft, die eine niedrigstufige physische Eigenschaft besitzt, welche die erforderliche funktionale Rolle spielt. In Verbindung der Multirealisierbarkeit mit der horizontalen Ebene wird die hohe Flexibilität, die mentale Ereignisse auszeichnet, für die mentale Kausalität zusätzlich betont.

#### 4.2.1 Das Prinzip der Multirealisierbarkeit

Multirealisierbarkeit ist eine Kernfrage in der Debatte um die Relation mentaler und physischer Eigenschaften. Zum einen wird dabei die Realisierbarkeit mentaler Eigenschaften durch physische Eigenschaften diskutiert. Als metaphysische Relation[96] handelt es sich um einen Spezialfall

[96] Neben dem metaphysischen Begriff der Realisierung können der mathematische und der logisch-semantische unterschieden werden. Im mathematischen Sinn wird gefordert, dass ein Zustand einen anderen realisiert, da Elemente des letzteren den Elementen des ersteren entsprechen. Block hat allerdings gezeigt, dass es problematisch ist, mentale Zustände auf nicht-bewusste Gruppen zu übertragen (vgl. Block 1978). Chalmers fordert eine eins zu eins Entsprechung zwischen den Eigenschaften der unterschiedlichen Bereiche sowie eine kausale Struktur der realisierenden Eigenschaften, um den notwendigen Übergang zum computationalen Zustand zu gewährleisten (vgl. Chalmers 1994). Im logisch-semantischen Sinn wird die Realisierung in interpretierbare symbolische Objekte übersetzt. Dennett entwickelt ein Modell der Handlungsinterpretation, das auf rationale Symbole

der Determinierung, wobei grundsätzlich durch die Realisierung eine sehr enge Beziehung beschrieben wird. Instanziierungen der realisierenden und realisierten Eigenschaften sollen zur selben Zeit auftreten, wobei die realisierte Eigenschaft eindeutig durch die realisierte bestimmt wird (vgl. Tye 1995). Darüber hinaus wird angenommen, dass die kausale Kraft der realisierten Eigenschaft von der realisierenden abgeleitet wird – mentale Eigenschaften hängen somit von der kausalen Kraft ihrer Realisierer ab (vgl. Shoemaker 2001).

Die These der Multirealisierbarkeit besagt nun, dass die Realisierer von mentalen Eigenschaften unterschiedlich sein können – es besteht eine „Eigenschaftsvariabilität". Verschiedene physische Eigenschaften können dieselbe funktionale Rolle in verschiedenen Individuen spielen. So kann bei einem Computer dieselbe computationale Funktion von verschiedenen technischen Systemen ausgefüllt werden. Die Instanziierung einer physischen Eigenschaft kann zwar hinreichend sein für die computationale Eigenschaft, sie ist aber nicht notwendig, da auch andere Eigenschaftsinstanziierungen jene computationale Eigenschaft ausfüllen können.

Hinsichtlich der Multirealisierbarkeit mentaler Eigenschaften ist entscheidend, dass diese keine nomologisch koextensiven physischen Eigenschaften haben und deshalb nicht auf ihre entsprechenden Basen reduziert werden können. Es gibt keinen speziellen neuronalen Zustand N, der, sämtliche physische Organismen umfassend, Schmerz realisiert. Vielmehr wird von einer Vielzahl neurophysischer Arten $N_m$, $N_r$, $N_a$, usw. ausgegangen, wobei $N_m$ Schmerzen in Menschen realisiert, $N_r$ Schmerzen in Reptilien und $N_a$ Schmerzen in Marsianern. Dahinter steht die Annahme, dass die biologischen Systeme zu weitreichend sind, als dass eine einzelne physisch-biologische Basis zur Reduktion ausreicht. Sogar innerhalb eines einzelnen Systems kann es immer wieder zu Änderungen der neuronalen Basis für Schmerzen kommen (vgl. Kim 1993a, 313). Menschen können also genauso denken, dass 2 + 2 = 4 ist, wie Marsianer mit Gehirnen aus grünem Schleim oder Computer. Damit wird von einer großen offenen Menge an physischen Systemen ausgegangen, die denken können, dass 2 + 2 = 4 ist. Der These zufolge sind sämtliche mentalen Eigenschaften durch verschiedene physische zu realisieren. Also ist keine Reduktion von mentalen auf physische Eigenschaften möglich (vgl. Stich 1996, 28). Das Argument der Multirealisierbarkeit ist insofern ein modales Argument, als die geforderte Identität von Eigenschaften nicht für alle möglichen Welten gilt. Es ist nicht möglich, einen einzigen physischen Zustand zu bestimmen, der z. B. der Realisierer von Schmerz in allen aktuellen und möglichen Welten ist. Nicht nur die neurologischen Zustände der Menschen unterscheiden sich sehr subtil in ihren physischen Konditionen, sondern auch die der Menschen von anderen Lebewesen, wie Hunden, Katzen oder Tintenfischen, ganz zu schweigen von Marsmenschen. Trotz dieser unterschiedli-

---

des Systems abzielt. So ist es innerhalb der Interpretation rational für ein System, bestimmte Symbole und Operationen zu besitzen (vgl. Dennett 1978).

chen physischen Konditionen würde man aber immer noch von Schmerzen sprechen (vgl. Kap. 2.2).

Es gibt verschiedene Einwände gegen Multirealisierbarkeit.[97] So verweist Kim darauf, dass Eigenschaftsvariabilität nicht bedeutet, dass überhaupt keine physische Eigenschaft realisiert wird und auch keine mental-physische Identität besteht, wenn eine mentale Eigenschaft realisiert wird. Er vertritt die Möglichkeit so genannter Interlevel-Identität mit einer minimalen Eigenschaftsvariabilität.[98] Der Grundannahme, dass mentale Eigenschaften von verschiedenen Zuständen, wie chemischen, biologischen oder elektronischen realisiert werden können, stimmt Kim zwar zu, jedoch ist er der Meinung, dass die Folgerungen, die aus dieser Überlegung gewonnen werden, nicht weit genug gehen. Im Gegensatz zu der Annahme, dass Multirealisierbarkeit ein Argument für die Autonomie der Spezialwissenschaft Psychologie ist, erscheint ihm das Prinzip, konsequent zu Ende gedacht, eher als Argument gegen deren Autonomie (vgl. Kim 1993a, 311). Allein die Terminologie erweist sich schon als verräterisch. Es ist nicht metaphysisch neutral, davon zu sprechen, mentale Eigenschaften würden realisiert. Das impliziert ein ontologisches Bild, wonach von vornherein die Abhängigkeit der mentalen von den physischen Eigenschaften bestimmt wird. Konsequenterweise werden mentale Eigenschaften auch als Eigenschaften zweiter Ordnung bezeichnet, die nicht in einer ebenso zwingenden Verbindung stehen wie ihre physischen Implementierungen (vgl. Block 1990, 155f.).

Zusammengefasst besagt die These physischer Realisierung, dass mentale Zustände in einem System genau dann erscheinen, wenn die geeigneten physischen Bedingungen in dem System gegeben sind, und die nomologische Relation jener durch die kausal-nomologischen Verbindungen letzterer erklärt werden können. Nach der These der Multirealisierbarkeit kann die mentale Eigenschaft Schmerz von verschiedenen physischen bzw. neuronalen Eigenschaften $N_m$, $N_r$ und $N_a$ realisiert werden. Kim kommt zu der Überlegung, dass diese Eigenschaften sich nicht nur voneinander unterscheiden, sondern dass sie darüber hinaus die disjunkte Menge $N_m \vee N_r \vee N_a$ der Eigenschaft N bilden.[99] Wenn nun angenommen wird, dass Schmerz nomologisch äquiva-

[97] Neben der hier vorgestellten Skepsis von Kim kann auf Patricia Churchland verwiesen werden, die davon ausgeht, dass Bereiche der Psychologie und der Neurowissenschaft sich in paralleler Entwicklung so stark annähern, wodurch auf neurowissenschaftlicher Seite psycho-funktionale Kriterien bestimmt werden können, um Gehirnstrukturen zu identifizieren und dadurch mental-physische Korrelationen zu etablieren (vgl. Churchland 1986). Zudem werden von der üblichen wissenschaftlichen Reduktion abweichende Variationen vorgeschlagen, um Eigenschaftsidentität jenseits gesetzmäßiger Koextension zu ermöglichen (vgl. z. B. Bickle 1998).

[98] Beispielsweise ist Temperatur mit kinetischer Energie in Gasen identisch, wobei zwei Aggregatzustände der Moleküle mit derselben Temperatur sich physisch aufgrund ihrer konstituierenden Moleküle mit unterschiedlichen Positionen und Ausrichtungen unterscheiden (Kim 1972). Nicht-Reduktionisten bestreiten jedoch, dass der Fall der mentalen Eigenschaften mit der Eigenschaftsidentität kinetischer Energie und der Temperatur idealer Gase vergleichbar ist.

[99] Üblicherweise wird angenommen, dass Disjunktionen im Gegensatz zu Konjunktionen für die Fälle, auf die sie zutreffen, keine Ähnlichkeit garantieren, obwohl die Idee einer Eigenschaft gerade mit der Idee der Ähnlichkeit verbunden wird. Kim ist jedoch der Ansicht, dass Eigenschaften nicht schon an sich disjunkt oder konjunkt sind. Eigenschaften können wiederum Disjunktionen oder Konjunktionen anderer Eigenschaften bilden. „The point about disjunctive properties is best put as a closure condition on properties: the class of properties is not

lent zu N ist, einer völlig disjunkten und nicht-nomologischen Eigenschaft, kann auch der Schmerz selbst als heterogen und nicht-nomologisch eingeordnet werden (vgl. Kim 1993a, 323). Multirealisierbarkeit bedeutet, dass mehr als eine physische Eigenschaft bzw. eine offene Gruppe physischer Eigenschaften eine mentale Eigenschaft realisiert. Eine Eigenschaft P hat demnach die Eigenschaften $P_1$, $P_2$, ..., $P_n$, was aber identisch ist mit der Disjunktion $P_1 \vee P_2 \vee ... \vee P_n$. Die Konsequenz, die sich demnach aus der Multirealisierbarkeit ergibt, ist über die Tatsache hinaus, dass Schmerz nicht-nomologisch ist, dass die Eigenschaft, Schmerz zu haben, nicht in strikten Gesetzen formuliert werden kann. Eine wissenschaftliche Theorie, die Kausalgesetze und -erklärungen für psychologische Ereignisse liefern soll, lehnt Kim ab. Zwar kann es einzelne Theorien über Schmerzen von Menschen, Reptilien oder Marsianern geben, aber keine umfassende Theorie mit einem vereinheitlichenden Anspruch, und damit ist schließlich das Ziel, die Psychologie als nichtreduzierbare autonome Wissenschaft zu bewahren, gescheitert (vgl. ebd., 325). Außerdem argumentiert Kim, dass zwar nicht Schmerz an sich, aber Schmerz für Menschen eine gesetzliche koextensive neurophysiologische Eigenschaft ist (ders. 1972; 1992; 1993a). Physische Eigenschaften werden dabei als disjunkte Eigenschaften vorgestellt. Die Eigenschaft, Schmerz zu haben, ist gesetzlich koextensiv mit der disjunkten Eigenschaft, ein bestimmtes menschliches Neuronenensembel oder ein bestimmtes marsianisches Neuronenensemble zu haben. In diesem Fall umfasst die disjunkte Eigenschaft jede möglich Realisierung von Schmerz (vgl. ders. 1998). Letztlich sind es für Kim das Prinzip der kausalen Erbschaft und das Prinzip der Individuation von Arten, welche in gegenseitiger Verbindung die kausale Kraft mentaler Eigenschaften und damit auch diese selbst in Schwierigkeiten bringen. Realisiert von oder supervenierend auf derselben physischen Basis sind die Instantiierungen von M Angehörige derselben Art. Haben sie jedoch unterschiedliche physische Basen mit unterschiedlicher kausaler Kraft, fallen sie unter verschiedene Arten. Letzterer Fall soll nach dem Prinzip der Multirealisierbarkeit zutreffen, woraus sich ergibt, dass mentale Arten aufgrund ihrer Nicht-Nomologizität keine kausalen Arten sein können und sich somit als wissenschaftlich strukturierbare Arten disqualifizieren (vgl. ders. 1993, 327).

Doch auch wenn Kim vertritt, dass die These ein Argument für die Reduktion auf lokale disjunkte Eigenschaften ist (vgl. ders. 2003), so stimmt er der Grundannahme zu, dass mentale Eigenschaften von verschiedenen Zuständen, wie chemischen, biologischen oder elektronischen realisiert werden können. Für Kim folgt aus dieser Argumentation nicht, dass Psychologie von vornherein als unwissenschaftlich abzulehnen ist, schließlich hat sie ein physisches Fundament. Eine psychologische Theorie kann demnach lokal reduziert werden auf die physikalische Theorie der entsprechenden Spezies. Diese Reduktion kann über lokale Brückengesetze erfolgen, die er

---

closed under disjunction [...]. Thus, there may well be properties P and Q such that P or Q is also a property, but its being so doesn't follow from the mere fact that P and Q are properties." (Kim 1993a, 321f.)

verschiedener empirischer Einwände zum Trotz zumindest metaphysisch für möglich hält. Schließlich folgt aus der These physischer Realisation, dass mentale Zustände eines Systems auf physischen Zuständen beruhen müssen:

> [T]his reduction would tell us how human psychology is physically implemented, how the causal connections between our psychological events and processes work at the physical-biological level, what biological subsystems subserve our cognitive capacities and functions, and so forth. (Kim 1993b, 274)

Es ist zweifelhaft, ob disjunkte physische Eigenschaften überhaupt existieren, da sie keine inhaltlichen Aussagen über Ähnlichkeiten von Objekten oder plausible Aussagen über deren kausale Kräfte zulassen (vgl. Teller 1983). Block argumentiert zudem, dass disjunkte Prädikate keine natürlichen Arten ausdrücken; bestimmbare Prädikate natürlicher Arten werden jedoch für wissenschaftliche Vorhersagen und Erklärungen benötigt (vgl. Block 1997).

Nicht-Reduktionisten argumentieren, dass eine grundsätzliche Variabilität aufgrund der funktionalen Natur mentaler Phänomene sowie der aktuellen Auffassung über die physisch variable Zusammensetzung des Gehirns äußerst plausibel bleibe. Nach wie vor gilt beispielsweise für den Fall der Computation, dass Wissenschaftler keine einzelne notwendige und hinreichende physische Bedingung für eine computationale Funktion bestimmen können. Sie weisen darauf hin, dass neurowissenschaftliche Forschung lediglich mental-physische Korrelation, jedoch nicht die für die Eigenschaftsidentität benötigte Koextension nachweisen. Selbst wenn künstliche Systeme ausgeschlossen wären und die psychologische Theorie sich ausschließlich auf biologische Systeme bzw. Säugetiere bezöge, und selbst wenn die Neurowissenschaft psychofunktionale Kriterien zur Identifizierung der Gehirnstruktur von Säugetieren entwickeln könnte, müssten solche Identifizierungen mit der kompensatorischen und erfahrungsabhängigen Plastizität sowie evolutionsabhängigen physischen Variationen kompatibel sein (vgl. Rosenberg 2001; Horgan 2001). Die neuronale Plastizität des Gehirns ist inzwischen gut belegt. Es können kognitive Funktionen von Bereichen, die aufgrund von Verletzungen oder Krankheit beschädigt sind oder ausfallen, von anderen Bereichen übernommen werden (vgl. Rauschecker 1995; Johnson 1993). Hinsichtlich der erfahrungsabhängigen Plastizität zeigt sich, dass die Basisvernetzung des Gehirns durch die individuelle sensorische Erfahrung verfeinert wird und individuelle Unterschiede bezüglich der Realisierung mentaler Funktionen schafft (vgl. King 1999).[100] Diese Beobachtungen und Anforderungen machen eine Identität bestimmter funktional-mentaler Typen mit koextensiven funktional-physischen Typen höchst unwahrscheinlich.

[100] Dennett führt die Eigenschaftsvariabilität auf evolutionäre Kräfte zurück, wonach das Gehirn aufgrund der Realisierung kognitiver Funktionen eine Variabilität ausgebildet hat, um die Fähigkeit des Organismus zu erhöhen, sich an eine veränderte Umwelt anzupassen (vgl. Dennett 1991).

### 4.2.2 Multikausale Relationen und mentale Kausalität

Die hohe Variabilität des Gehirns in Verbindung mit der Multirealisierbarkeit mentaler Eigenschaften lässt sich auch auf die kausalen Relationen übertragen, die diese Eigenschaften eingehen. Multikausalität scheint nun insofern auch ein Kriterium für mentale Kausalität zu sein, als die Eigenschaften der Ursachen für die verschiedenen Wirkungen vielfältig sind. Dies lässt sich innerhalb der vertikalen Relation zwischen höherstufigen und niedrigstufigen Eigenschaften bzw. supervenienten und subvenienten Eigenschaften beschreiben. Aber in der Konsequenz gilt dies auch für die horizontale Ebene, da mentale Eigenschaften eben durch verschiedene mentale Eigenschaften verursacht werden, die wiederum multirealisierbar sind. Damit ist letztlich auch das Ereignis als Eigenschaftsinstanziierung multikausal. Beispielsweise sieht ein Schachspieler in einer Partie durch die Dame des Gegners seinen Läufer in Gefahr und überlegt deshalb, den Läufer durch einen Springer zu schützen. Der Gedanke $G_1$ „Läufer durch Dame in Gefahr" ist hierbei als ursächlich für den Gedanken $G_2$ „Läufer durch Springer schützen" zu verstehen. Nun gibt es verschiedene Fälle, in denen mentale Kausalität „multikausal" ist:

> 1) Auf der psychologischen Ebene kann der Gedanke $G_2$ auch durch andere Gedanken bei dem Spieler S verursacht werden, wie den Gedanken $G_3$ „Es ist immer besser, den Läufer proaktiv zu schützen" oder den Gedanken $G_4$ „Eventuell wird der Gegner versuchen, meinen Läufer mit dem Springer zu attackieren".
>
> 2) Unter Supervenienzbedingungen und der Annahme, dass die Gedanken $G_1$ und $G_2$ auf physischen Eigenschaften $P_1$ und $P_2$ supervenieren, wird $G_2$ entweder a) direkt von $P_1$ verursacht, wobei $P_1$ verschiedene physische Eigenschaften sein können, da $G_1$ multirealisierbar ist, oder b) $G_2$ wird vertikal von $P_2$ klassisch multirealisiert oder c) $G_2$ wird indirekt von $P_1$ über $P_2$ verursacht bzw. $P_1$ verursacht $P_2$, wobei auch hier Multirealisierbarkeit gilt, da $G_1$ multirealisierbar ist und somit auch $P_2$ durch verschiedene physische Eigenschaften verursacht werden kann.

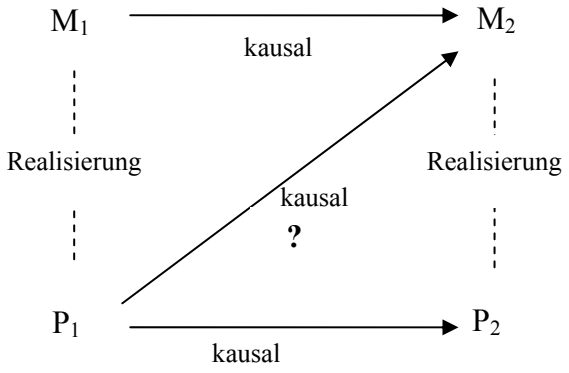
Im Unterschied zu rein physischen kausalen Relationen, in denen einer Wirkung genau eine Ursache vorausgeht, kann für mentale Kausalität angenommen werden, dass sie multirealisierbar ist. In der Verallgemeinerung kann als Prinzip für die Multirealisierbarkeit mentaler Kausalität formuliert werden:

> Für mentale Kausalität gilt, dass Eigenschaft F in einem System A als Ursache Eigenschaft G oder Eigenschaft H oder Eigenschaft I usw. haben kann, wobei die Eigenschaften G, H, I usw. für F hinreichend sein müssen, auch wenn sie selbst durch eine bestimmte Zahl variabler Eigenschaften realisiert sind.

Wie bereits im vorhergehenden Kapitel angesprochen, birgt die Variante der Cross-Level-Kausalität am meisten Brisanz, da hierbei traditionell die Überschreitung des physischen in den

mentalen Bereich, vice versa, thematisiert wird. Grundsätzlich kann hierfür zunächst Multikausalität in beide Richtungen angenommen werden: Verschiedene physische Eigenschaften verursachen eine mentale Eigenschaft, aber auch verschiedene mentale Eigenschaften verursachen eine physische Eigenschaft. Dass verschiedene mentale Eigenschaften physische Eigenschaften verursachen, könnte sich aus einer empirischen Perspektive darin niederschlagen, dass verschiedene mentale Eigenschaften bzw. propositionale Einstellungen, Gedanken etc. die Aktivierung eines bestimmten Neurons als physische Eigenschaft verursachen können. Tatsächlich wird dieser Fall unter Verweis auf die Voraussetzungen für mentale Kausalität, insbesondere das Ausschlussargument respektive die Geschlossenheit der Physik, abgelehnt.

Deshalb wird die Cross-Level-Kausalität der physischen Eigenschaften auf mentale Eigenschaften favorisiert. Danach können nun verschiedene physische Eigenschaften eine mentale Eigenschaft verursachen, wobei es sich nicht um eine zeitgleiche Variante der Multirealisierung handelt, sondern um eine zeitliche, kausale Folge. Ein Beispiel wäre also nicht das mentale Ereignis „Schmerzen haben", sondern gegebenenfalls der daraus resultierende Wunsch, den Schmerz zu beenden, z. B. durch kaltes Wasser nach einer Verbrennung. Dabei wäre die physische Eigenschaft $P_1$ in doppelter Verantwortung: Zum einen als Realisierer der mentalen Eigenschaft $M_1$ = Schmerzen haben, zum anderen als Ursache der nachfolgenden mentalen Eigenschaft $M_2$ = Wunsch, Schmerz zu beenden. Als multikausale Ursachen für dieses Ereignis können nun verschiedene Schmerzzustände bzw. deren physische Realisierungen gelten. Allerdings gibt es hierbei zumindest zwei Einwände: Zum einen scheint nicht nur $P_1$ in doppelter Verantwortung zu sein, auch $M_2$ scheint eine mehrfache Relation einzugehen, zu $P_1$ als Wirkung und zudem zu der physischen Eigenschaft $P_2$ als dessen Realisierung.




Grafik 10: Die physische Eigenschaft $P_1$ ist in mehrfacher Verantwortung, da sie einmal Realisierer der mentalen Eigenschaft $M_1$ ist und zugleich Ursache der mentalen Eigenschaft $M_2$ bzw. der physischen Eigenschaft $P_2$. Zusätzlich ist die mentale Eigenschaft $M_2$ mehrfach bestimmt, einmal als Wirkung der physischen Eigenschaft $P_1$, einmal als Realisierung der Eigenschaft $P_2$. Unter der Idee einer starken Realisierungsrelation mit einer gleichzeitigen Eigenschaftsinstanziierung lässt sich dieses Dilemma auflösen, indem $M_1$ und $P_1$ proportional ihren mentalen bzw. physischen Nachfolger verursachen (vgl. Kap. 4.6). Zusätzlich muss berücksichtigt werden, dass innerhalb der vorliegenden Untersuchung $M_2$ bzw. $P_2$ nicht ausschließlich von $P_1$ verursacht werden kann, da hauptsächlich externe Stimuli als kausale Ursache entscheidend sind.

Es lässt sich bezweifeln, ob allein aus Gründen der ontologischen Sparsamkeit eine Konkurrenz zwischen physischer Ursache und physischem Realisierer sinnvoll ist.[101] Zum anderen kann man nachfragen, was genau die nachfolgende mentale oder die nachfolgende physische Eigenschaft verursacht. Hier scheint es unter den gegebenen Überlegungen plausibler zu sein, dass nicht nur jedes physische Ereignis nur eine physische Ursache hat, sondern auch jede physische Ursache nur eine physische Wirkung. $P_1$ verursacht also nur $P_2$ und keine mentale Eigenschaft $M_2$. Damit würde eine Cross-Level-Kausalität zumindest zurückgewiesen, aber nicht zwingend ausgeschlossen. Problematisch ist allerdings, dass die Kontinuität einer Abfolge mentaler Eigenschaften ausschließlich von der kausalen Kontinuität ihrer multiplen Realisierer abhinge. Nicht nur die Eigenschaften $M_1$ und $M_2$ würden von verschiedenen physischen Eigenschaften $P_1$, $P_2$, …, $P_n$ realisiert. Es ist ferner möglich, dass einige Realisierer von $M_1$ keine Realisierung von $M_2$ verursachen. Dies ist zumindest ein möglicher Grund dafür, dass einzelwissenschaftliche Gesetze im Allgemeinen nicht ausnahmslos gelten. Dies ist aber auch eine problematische Konsequenz, da somit psychologische Erklärungen (und auch diejenigen der Einzelwissenschaften) überhaupt keine Relevanz hätten, schließlich würden sie von den nahezu willkürlich anmutenden multikausalen Prozessen auf der physischen Ebene der Realisierer abhängen. Und wenn mentale Eigenschaften multirealisierbar sind, wäre für alle möglichen Welten nicht auszumachen, welche physischen Multirealisierer Ursache oder Wirkung welcher physischen Multirealisierer wären.

Diese Konsequenz mündet aber im Absurden, da schließlich alle alternativen physischen Realisierer „blind" weitere alternative physische Realisierer verursachten. Dies scheint aber auch der Voraussetzung der strikten Gesetzen für den physischen Bereich zuwiderzulaufen (vgl. Kap. 3.5.1) Außerdem scheint es sehr wohl einige psychische bzw. psychophysische Gesetze zu geben sowie auch logische Gedankenfolgen oder mathematische Gesetzmäßigkeiten – diese sind jedoch nicht zwingend von physischen Gesetzmäßigkeiten abhängig. Die „psychologische Taxonomie" (vgl. Kap. 3.1.4) verläuft nicht rein assoziativ oder gar willkürlich. Einen möglichen Ausweg bietet hier die Idee der zeitgleichen, gleichberechtigten Instanziierung: als Ereignisse werden mentale und physische Eigenschaften gleichzeitig und kausal gleichberechtigt instanziiert. Beide sind Wirkung des vorhergehenden Ereignisses A, welches einfach das folgende Ereignis B verursacht. Da aber das Ereignis A spezifische mentale Eigenschaften $M_a$ und physische Eigenschaften $P_a$ aufweist, ist das Ereignis B nachfolgend zumindest „eingeschränkt" determiniert.

Trotz dieser Widersprüche scheint sich für mentale Kausalität aus dem Prinzip der Multirealisierbarkeit sowie der Plastizität des Gehirns zu ergeben, dass sie eine multikausale Relation ist. Dabei gilt es weiterhin zu prüfen, welche Relevanz der psychologischen Ebene und mentalen

[101] Andererseits ist dieser Fall der Verursachung bei gleichzeitiger Realisierung insofern für das Gehirn denkbar, als a) ein Neuronenensemble $P_1$ ein weiteres verursacht und dabei die Realisierungsrelation von $M_2$ und $P_2$ zeitgleich erscheint oder b) eine extramentale physische Eigenschaft $P_1$ eines Umweltereignisses die mentale Eigenschaft $M_2$, die von der physischen Eigenschaft $P_2$ realisiert wird, verursacht.

Eigenschaften zukommen kann, und wie stark mentale Kausalität durch multikausale Prozesse auf der physischen Ebene bestimmt wird.

***Quintessenz Kriterium Multikausalität***

- *modale Multirealisierbarkeit in möglichen Welten*
- *empirische Plastizität des Gehirns*
- *Multikausalität auf horizontaler Ebene für Ereignisse*
- *Ereignisinstanziierung sichert kausale Kraft der Realisierer bei Eigenschaftsvariabilität (nicht koextensiv)*
- *kein Chauvinismus: Variabilität für alle Arten („extern") und für jedes Individuum („intern") gültig*
- *Multikausalität fußt auf Multirealisierbarkeit: verschiedene multirealisierte Ereignisse als Ursache für eine „mentale" Wirkung*
- *Ohne mentale Eigenschaften sind physische Eigenschaften „blind", ohne physische Eigenschaften sind mentale Eigenschaften „leer" – das entsprechende Ereignis wird in beiden Fällen nicht realisiert*

- *Prinzip der Multirealisierbarkeit mentaler Kausalität:*

  *Für mentale Kausalität gilt, dass Eigenschaft F in einem System A als Ursache Eigenschaft G oder Eigenschaft H oder Eigenschaft I usw. haben kann, wobei die Eigenschaften G, H, I usw. für F hinreichend sein müssen, auch wenn sie selbst durch eine bestimmte Zahl variabler Eigenschaften realisiert sind*

### 4.3 Das Kriterium des Inhalts

Auf die Relevanz von Inhalt hinsichtlich mentaler Kausalität haben Block oder auch Jackson hingewiesen.[102] Sie ergibt sich zumindest implizit aus den diskutierten theoretischen Positionen (vgl. Kap. 2). Die Frage lautet zunächst, wie neuronale Prozesse, die per se nicht empfänglich für Inhalte sind, eine Wirkung bzw. Änderung des gesamten Systems haben können. Und wie kann Inhalt eine Wirkung auf die neuronalen Prozesse eines Systems haben? „Rather, the point is that when mental events have effects, they typically have those effects (rather than different effects) because the mental events have the contents that they have, rather than some other contents." (Block 1990, 140) Wenn Glaubensannahmen oder Gedanken Inhalte haben, die unterschiedlich

[102] So hat nach Jackson das Problem der mentalen Kausalität in der Diskussion nun zwei Formen: 1.) die Voraussetzung eines Funktionalismus mit weiten oder engen mentalen Eigenschaften und 2.) die Voraussetzung eines begründbaren starken Externalismus über intentionalen Inhalt (vgl. Jackson 1996, 96).

verwendet werden, sollten sie auch unterschiedliche Wirkungen erzielen. Würde jemand annehmen, überall, außer in den USA, sei es lebensgefährlich, würde er dieses Land nicht verlassen.

Der Inhalt unserer Gedanken wird häufig über propositionale Ausdrücke und den alltäglichen Gebrauch der Begriffe erklärt. Die Generalisierungen der Alltagspsychologie erscheinen erfolgreich in der Vorhersage und Erklärung von Verhalten und darüber hinaus in der Beschreibung von Glaubensannahmen, Wünschen, Zweifeln, Ängsten, Wahrnehmungen und anderen mentalen Zuständen oder Ereignissen. Sie besitzen, so wird angenommen, Inhalt. Die meisten Inhalte werden mit propositionalen Ausdrücken in Verbindung gebracht, was bedeutet, dass Verben wie glauben, wünschen, hoffen oder denken in der Regel ein Finalsatz folgt, der den Inhalt wiedergibt. Beispielsweise wünsche ich, es würde regnen. Oder: Peter befürchtet, es gebe Krieg.

Glaubensannahmen und andere propositionale Eigenschaften umfassen Beziehungen von Gedanken, als deren Konstituenten Konzepte bzw. Begriffe aufgefasst werden können. So muss jemand, der glaubt, dass der Bundestag sich versammelt, ein Konzept von Bundestag und sich versammeln haben. Konzepte werden deshalb auch als Worte der mentalen Sprache vorgestellt, die sowohl bei reiner Denktätigkeit, als auch für propositionale Eigenschaften verwendet werden (vgl. Fodor 1990b).[103] Von konzeptionellen bzw. begrifflichen Inhalten werden nicht-begriffliche Inhalte unterschieden. Letztere werden auf phänomenale Eigenschaften bzw. Qualia zurückgeführt und vor allem im Zusammenhang mit Sinneseindrücken aufgeführt. [104] Einige Vertreter dieser Position gehen in ihren Überlegungen sogar soweit zu behaupten, dass begriffliche Repräsentationen nicht ohne phänomenale Eigenschaften auskommen können (vgl. z. B. Chalmers 1997; Searle 1993). In diesem Falle müsste auch untersucht werden, welche Rolle phänomenale Eigenschaften bei der Bestimmung von begrifflichen Repräsentationen spielen. Allerdings gibt es in der zeitgenössischen Diskussion starke Zweifel, ob es überhaupt phänomenale Eigenschaften und damit verbundene nicht-begriffliche Repräsentation von Inhalt gibt, bzw. welche Rolle sie überhaupt spielen könnten (vgl. z. B. Dennett 1988; McDowell 1994). Deshalb wird sich die Bestimmung des Kriteriums „Inhalt" an dieser Stelle zunächst auf den begrifflichen Inhalt und seine Definition als „enger" oder „weiter" Inhalt beschränken.

---

[103] Nach Evans ist es möglich, dass einige Konzepte mentaler Zustände nicht als mentale Sprache artikuliert werden. Visuelle Wahrnehmung scheint solche Prozesse und Zustände zu umfassen, da solche Wahrnehmungen reicher sind, als in einem Gedanken ausgedrückt werden kann. So unterscheidet Evans auch zwischen konzeptionellem und nicht-konzepionellem Inhalt, wobei gerade letzterer eine entscheidende Rolle hinsichtlich Wahrnehmung und subdoxastischen kognitiven Prozessen spielt (vgl. Evans 1982).

[104] Die Vertreter der Qualia-Position können in Repräsentationalisten bzw. Intentionalisten (z. B. Dretske 1995; Lycan 1996; Tye 2000) und Phänomenalisten (z. B. Block 2003; Chalmers 1997) unterteilt werden. Während erstere die phänomenalen Eigenschaften der mentalen Ereignisse auf intentionalen Inhalt reduzieren wollen, propagieren die Phänomenalisten die Eigenständigkeit und Nicht-Reduzierbarkeit phänomenaler Eigenschaften. Eine vermittelnde Position nimmt ggf. die Position „phänomenaler Konzepte" ein, die begriffliche und phänomenale Eigenschaften sozusagen als hybride Eigenschaften zusammenführt (vgl. z. B. McGinn 1991).

Aufgrund ihres Inhalts haben mentale Zustände einige bemerkenswerte Eigenschaften: Sie beziehen sich auf die Realität außerhalb des mentalen Zustands, sie sind wesentlich normativ, sie unterliegen Wahrheitsbedingungen, sie verursachen aufgrund ihres Inhalts weitere mentale Zustände, und schließlich hat eine Person privilegierten Zugang zu den Inhalten ihrer mentalen Zustände. Da Inhaltseigenschaften prima facie so verschieden sind von physischen oder auch biologischen Eigenschaften, ist es immer noch eine offene Frage, ob Inhaltseigenschaften naturalisiert werden können oder nicht. Sollten sie nicht naturalisiert werden können, könnten sie prinzipiell nicht mehr Teil wissenschaftlicher Annahmen sein – falls ihnen dann überhaupt noch Existenz zugesprochen werden kann (vgl. Stich 1983). Einen möglichen Weg, Inhalt zu naturalisieren, bietet die „Representational Theory of Mind" (RTM) an. Mentale Repräsentation wird beispielsweise als mentale Bilder, Karten oder sprachliche Ausdrücke vorgestellt. Als eine Hauptströmung innerhalb der Diskussion umfasst die Theorie weitere Positionen, wie z. B. die „Language of Thought Hypothese" (LOT-Hypothese). Diese wird im Folgenden in einer kurzen Übersicht dargestellt.

### 4.3.1 Die Language of Thought Hypothese

Die Kernfrage hinsichtlich des Inhalts mentaler Zustände lautet prinzipiell: Wie gelangen mentale Zustände bzw. mentale Repräsentationen zu ihrem Inhalt? Die Antwort auf diese Frage soll im Rahmen des physikalischen Weltbilds geleistet werden, weshalb die Relation zwischen abstrakten, mentalen Inhalten und den Objekten, die sie repräsentieren, in einem naturalistischen Ansatz erfolgen soll – der mentale Inhalt soll naturalisiert werden. Die Positionen der „Language of Thought Hypothesis" sowie verschiedene teleologische Theorien versuchen dies zu leisten.

Nach Fodors Hypothese einer „Sprache des Denkens" ähneln sich Struktur und Eigenschaften von Sprache und Gedanken. Die Gedanken werden als ein System mit mentalen Symbolen bzw. Repräsentationen aufgefasst, die physisch im Gehirn des Denkenden verankert sind und somit die neuronale Basis für Gedanken schaffen. Dieses besteht aus einer kombinatorischen Syntax, weshalb die Repräsentationen nur mit syntaktischen Eigenschaften kausal interagieren können. Es ist die Kausalrelation, die den Inhalt der Repräsentation bestimmt. Damit ist die LOT-Hypothese eine kausale Theorie der Repräsentation von Gedanken. Ein Gedanke ist letztlich ein Zeichen bzw. ein Symbol einer Repräsentation, die über eine konstituierende syntaktische Struktur mit einer geeigneten Semantik verfügt. Denken ereignet sich der Struktur nach also in Form einer mentalen Sprache, dem so genannten „Mentalese". Es findet als syntaktische Operation von Repräsentationen statt. Dem propositionalen Ausdruck „Der General liebt den Krieg"

entspricht im Gehirn des Sprechers ein vorhandener Gedanke mit einer ähnlichen syntaktischen Struktur.

Für die Repräsentation von Gedanken wird angenommen, dass es für jede propositionale Einstellung A genau eine bestimmte psychologische bzw. computationale/funktionale Relation R gibt, und für alle Propositionen P und Subjekte S, dass S und A nur dann P sind, wenn es eine mentale Repräsentation #P# gibt, wobei S R mit #P# verknüpft und #P# bedeutet, dass P. Denken als mentaler Prozess wiederum besteht aus kausalen Zeichensequenzen solcher mentaler Repräsentationen (vgl. Fodor 1987, 17). Jede propositionale Einstellung wird somit als eine funktionale Rolle bestimmt, die von einem mentalen Satz, dem unmittelbaren Objekt der Einstellung, eingenommen wird. Was einen mentalen Satz zu einer konkreten Glaubensannahme macht, sind ganz bestimmte computationale bzw. funktionale Operationen, die eben nur für Glaubensannahmen geeignet sind, ebenso wie nur ganz bestimmte Operationen für Wünsche etc. geeignet sind (vgl. Schiffer 1981).

Den Kern der LOT-Hypothese macht die Annahme aus, dass mentale Repräsentationen wie eine Sprache syntaktisch und semantisch strukturiert sind. Mentale Repräsentationen sind als unmittelbarer Gegenstand von propositionalen Einstellungen komplex strukturierte, syntaktische und semantische Symbole. Deshalb wird die LOT-Hypothese auch als mentale Sprache „Mentalese" bezeichnet. Damit beschreibt die Hypothese im Schwerpunkt die Natur von Denkprozessen mit propositionalem Inhalt und weniger Phänomene wie Erfahrungen, Qualia oder mentale Bilder – auch wenn dies nicht grundsätzlich ausgeschlossen ist. Mentale Repräsentationen, die die direkten Objekte der propositionalen Einstellungen bilden, sind Teil eines repräsentationalen bzw. symbolischen Systems (vgl. Fodor/Pylyshyn 1988, 12f.). Dieses umfasst zunächst eine kombinatorische Syntax und Semantik. Komplexe Repräsentationen werden systematisch aus einfachen Zeichen gebildet: Der semantische Inhalt komplexer Repräsentationen ist die Funktion des semantischen Inhalts seiner einfachen Konstituenten in Verbindung mit einer bestimmten Syntax. Die Operation der Repräsentationen ist kausal sensibel für die Syntax der Repräsentationen, die schließlich über eine kombinatorische Syntax bestimmt werden.[105] Auf einer bestimmten Ebene sind mentale Repräsentationen somit als funktionale Einheiten bestimmbar, die von den physischen, gegebenenfalls neurophysiologischen Eigenschaften derjenigen Person realisiert werden, die jene propositionale Einstellung besitzt. Damit unterscheidet die LOT-Hypothese eine dreifache Relation zwischen der Person, dem Mentalese-Satz und der Proposition. Auf diese Weise soll letztlich das Problem gelöst werden, wie der Inhalt von einzelnen Begriffen festgelegt wird: die Mentalese-Sätze verfügen als physische, syntaktische Struktur über die relevante kausale Kraft, abstrakte semantische Inhalte zu erfassen und zu übertragen bzw. zu encodieren. Denk-

[105] Hierin liegt der Grund, dass die LOT-Hypothese häufig mit der Computational Theory of Mind (CTM) identifiziert wird. Denken wird als computationaler Prozess verstanden.

prozesse werden also kausal von ihrer syntaktischen Struktur bestimmt, die dabei jedoch auf semantische Inhalte ausgerichtet ist.

Der wesentliche Vorteil der LOT-Hypothese scheint zu sein, dass sie die Naturalisierung der propositionalen Inhalte ermöglicht und somit für die empirische Forschungsarbeit der Kognitionspsychologie einen theoretischen Rahmen bietet, in dem Denkprozesse als Teil der physischen Welt ohne nicht-reduzierbare psychische Eigenschaften o. ä. beschrieben werden können. Denken ist nach der RTM wie auch der LOTH ein kausales Fortschreiten von Zustand zu Zustand, das einen semantischen Sinn ergibt, weshalb die Übertragung solcher Zustände in z. B. computationale Zeichen die semantischen Eigenschaften erhalten muss, um als echtes Denken zu gelten. Gedanken sind also selbst nicht nur kausal, sondern auch semantisch, normativ bzw. rational miteinander verbunden, damit sie einen Sinn ergeben – es wird semantische Kohärenz eingefordert. Die LOTH bietet hierfür eine kausale Lösung, indem sie Denken mit computationalen Prozessen vergleicht: Denken besteht aus im Gehirn realisierten physisch prozessierten Repräsentationen; diese bilden ein formales System bzw. eine Art Sprache mit einer geeigneten Syntax und genau definierten Ableitungsregeln, die aus den syntaktischen Grundzügen der Repräsentationen bestehen. Die physisch kodierten, formalen Eigenschaften der computationalen Symbolverarbeitung sollen somit die semantischen Eigenschaften bzw. die semantische Kohärenz der symbolischen Verarbeitung sichern. Das Gehirn wird pointiert als ein syntaktischer Motor beschrieben, der einen semantischen Motor betreibt (vgl. Block 1990, 143).

Als weitere Vorteile bietet die LOT-Hypothese Erklärungen für die Produktivität und die Systematik von Sprache. Produktivität ist die Fähigkeit, komplexe Gedanken durch die Verknüpfung einfacher Gedanken herzustellen, während Systematik die Fähigkeit umfasst, Gedanken zu denken, die in einer systematischen Beziehung zueinander stehen, wie „liebt der General den Krieg“ und „ein Krieg braucht Generäle“. Prinzipiell kann eine Person eine unbegrenzte Anzahl an Gedanken „produzieren“, was sich auch in der weit reichenden Fähigkeit, neue Begriffe und Sachverhalte zu erfassen, ausdrückt. Diese Anforderung wird innerhalb der LOTH durch den Verweis auf die kombinatorische Syntax erklärt. Ferner scheint die Fähigkeit, manche Gedanken zu denken, intrinsisch bzw. systematisch mit der Fähigkeit, andere Gedanken zu denken, verbunden zu sein (vgl. Kap. 4.4.1). Für Mentalese bedeutet das übertragen, dass die Zeichen einer Repräsentation mit den Zeichen weiterer Repräsentationen systematisch verbunden sind. Da propositionale Aussagen aber nicht nur syntaktisch, sondern auch semantisch sinnvoll sind und systematisch miteinander verknüpft sein müssen, wird in Ergänzung zu einer kombinatorischen Syntax eine kombinatorische Semantik postuliert (vgl. Fodor/Pylyshyn 1988).

### 4.3.2 Internalismus versus Externalismus

Gemeinhin wird angenommen, dass psychologische Erklärungen sowohl kausal als auch inhaltlich basiert sind. Der Wunsch, ein Glas Wasser zu trinken, erklärt in Verbindung mit dem Wissen, dass sich im Kühlschrank Wasser befindet, dass eine Person in die Küche an den Kühlschrank geht und sich das Glas Wasser einschenkt. Solange die Inhalte solcher mentaler Repräsentationen extrinsisch/extern sind, scheint es schwierig, ihnen kausale Kraft zuzusprechen, da argumentiert wird, dass nur intrinsische/interne Eigenschaften kausal wirksam sind (vgl. Stich 1983; Fodor 1982, 1987, 1994). Doch zunächst stellt sich die Frage, ob propositionale Eigenschaften tatsächlich extern oder intern sind, bzw. ob mentale Repräsentationen von intrinsischen Faktoren abhängenden „engen Inhalt" oder von extrinsischen Faktoren abhängenden „weiten Inhalt" haben.

Internalisten vertreten die Position, dass der epistemische normative Status von Glaubensannahmen[106] ausschließlich von relevanten internalen Faktoren der Perspektive des Glaubenden abhängen. Ob eine Person eine Eigenschaft hat oder nicht, hängt ausschließlich von ihren intrinsischen Eigenschaften ab, auf denen jene ggf. superveniert. Die äußere Umgebung einer Person spielt keine Rolle bei der Konstituierung von mentalen Inhalten. Wohlgemerkt schließt Internalismus nicht zwangsläufig aus, dass die Umwelt nicht eine solche Eigenschaft verursachen kann, aber die Eigenschaft selbst ist ein innerer physischer Zustand. Ein Schlag in die Magengrube kann Bauchschmerzen verursachen, aber die Bauchschmerzen selbst sind ein rein innerer physischer Zustand. Nach einer verwandten Definition ist eine mentale Eigenschaft F internal für eine Person S, wenn F Teile des mentalen Erlebens von S begründet. Demnach würde eine Wahrnehmung, eine Glaubensannahme oder eine Repräsentation, die Person S hat, internal sein, solange diese Zustände Bestandteil des mentalen Erlebens sind. Für diese Perspektive ist entscheidend, ob die Zustände subjektiv passen.

Externalisten vertreten die Ansicht, dass der Inhalt verschiedener mentaler, insbesondere intentionaler Zustände von der Beziehung einer denkenden Person zu ihrer äußeren Umwelt abhängt. In verschiedenen Gedankenexperimenten über Zwillingswelten wurden weit reichende Konsequenzen der Einbettung von Individuen in sprachliche oder soziale Kontexte hinsichtlich der Natur mentaler Zustände beschrieben. Ausgehend von einem semantischen Externalismus wurde diese Position auf mentale Inhalte hin ausgeweitet und gefestigt (vgl. Burge 1979; McGinn 1977; Putnam 1975c).[107] Externalisten vertreten die Position, dass der epistemische Status von

[106] Es gibt verschiedene Arten des epistemisch normativen Status. So können wir sagen, dass eine Glaubensannahme gerechtfertigt, rational, verständlich oder verantwortbar ist. Damit ist es möglich, hinsichtlich einiger epistemischer Status Internalist und hinsichtlich anderer Externalist zu sein.

[107] Während Putnams semantischer Externalismus im Grunde von einem objektiven Betrachter ausgeht, der die Unterschiede zwischen Wasser/$H_2O$ und Zwillingswasser/XYZ kennt und die falsche Referenz entlarven kann,

Glaubensannahmen nicht ausschließlich von jenen internalen Faktoren bestimmt wird. Mentaler Inhalt wird zumindest teilweise durch die vom Denkenden externalen Bedingungen individuiert (vgl. Putnam 1975c; Burge 1979). Die externalistische Perspektive betont, dass eine Glaubensannahme nur als Wissen zählt, wenn sie wahr ist – und somit von externalen Wahrheitsbedingungen abhängt. Externale Faktoren umfassen die Genauigkeit, die Zuverlässigkeit und die passenden kausalen Beziehungen zur Umgebung der Person. Sie lassen sich objektiv aus der Dritte-Person-Perspektive ermitteln. Es wird auf die Ätiologie der einzelnen Glaubensannahmen verwiesen, wonach es entscheidend ist, wie eine Person zu einer bestimmten Überzeugung gelangt. Grundsätzlich ist die Position des Externalismus mit verschiedenen Leib-Seele-Theorien vereinbar, auch wenn hier unterschiedlich argumentiert wird. So nimmt Davidson für die Tokenidentitätstheorie in Anspruch, dass mentale und physische Token miteinander identisch sein können, auch wenn die äußeren Ursachen für einen bestimmten Zustand wechseln können. Sein Beispiel ist der Sonnenbrand, der mit einem bestimmten Körperzustand identisch ist, auch wenn dieser Körperzustand von etwas anderem als Sonnenlicht verursacht werden könnte (vgl. Davidson 1989; Kap. 2.3). Für die funktionalistische Position ergibt sich allerdings die Schwierigkeit, dass die funktionale Rolle mentaler Zustände von der input-output-Relation innerhalb des Gehirns abhängt. Deshalb erscheint es sinnvoll zu sein, diese Relation auf die äußere Umwelt hin auszuweiten, damit sie mit Externalität kompatibel wird. Eine andere Möglichkeit ist, den Unterschied zwischen dem internen mentalen Zustand und seinem externen Inhalt stark zu machen. So wird ein innerer Glaubenszustand auf seine interne funktionale Rolle zurückgeführt, während sein Inhalt jedoch von der Beziehung zur äußeren Umwelt abhängt (vgl. Fodor 1987, Kap. 3).

Auf die Herausforderung des Externalismus haben einige Philosophen reagiert, indem sie zwar zugestanden haben, dass mentale Zustände external individuierte Inhalte haben, jedoch zusätzlich „engen Inhalt" besitzen, der nicht external individuiert ist. Die moderate internalistische Position erlaubt, dass es neben dem weiten Inhalt des Alltagsgebrauchs zusätzlich engen Inhalt gibt (vgl. Fodor 1987; Loar 1987). So genannter „enger Inhalt" hängt nicht von der äußeren Umgebung einer Person ab. Damit ist auch die Überlegung verbunden, dass die kausale Kraft intrinsischer Zustände, die zwei Personen miteinander teilen, dieselbe sein muss. Der enge Inhalt einer bestimmten Glaubensannahme ist derjenige Inhalt, der vollständig von den individuellen intrinsischen Eigenschaften abhängt. Der enge Inhalt eines mentalen Zustands bleibt unabhängig von Veränderungen der äußeren Umgebung derselbe, er ändert sich nur aufgrund einer intrinsischen Veränderung innerhalb der Person. Insofern wird angenommen, dass zwei in ihren physischen Eigenschaften identische Doppelgänger auch dieselben mentalen Inhalte haben. Der enge

---

geraten bei McGinn wie bei Burge die individuellen epistemischen Limitationen sowie deren Abhängigkeit von unterschiedlichen äußeren Wahrheitsbedingungen in den Fokus (vgl. McGinn 1977; Burge 1979). Dabei wird gezeigt, dass trotz identischer intrinsischer Bedingungen die Glaubensannahmen unterschiedlich sind, was die Position des Externalismus bestätigt.

Inhalt einer Repräsentation wird durch ihre intrinsischen Eigenschaften bzw. die ihrer syntaktischen Struktur bestimmt.

Es gibt verschiedene Gründe für den Ansatz von „engem Inhalt", z. B. die Erste Person-Autorität. Aus internalistischer Sicht wird der Person ein privilegierter Zugang hinsichtlich ihrer Glaubensannahmen zugesprochen. Demnach kann ausschließlich die Person selbst wissen, ob sie einen bestimmten mentalen Zustand hat. „Enger Inhalt" ist damit im Grunde nicht beschreibbar (vgl. Fodor 1987; Block 1986). Ein weiterer wichtiger Grund ist, dass enger Inhalt für kausale Erklärungen benötigt wird. Wenn die kausalen Kräfte mentaler Zustände den Inhalt der Zustände umfassen, muss dieser Inhalt möglicherweise „eng" sein. Mentale Zustände sollen aufgrund ihres Inhalts das menschliche Verhalten erklären. Der Inhalt unserer Glaubensannahmen oder Wünsche wird als zentral für unsere Handlungen betrachtet. Allerdings wird nur intrinsischen Eigenschaften die kausale Kraft, Wirkung zu erzielen, zugesprochen. Kausalität ist lokal und kontextabhängig, weshalb die Eigenschaften der äußeren Umwelt nur über die individuellen intrinsischen Eigenschaften einer Person Einfluss auf deren Handlung haben können. Dagegen wird argumentiert, dass externale Eigenschaften sehr wohl mit lokaler Kausalität und Individuation vereinbar sind; kausale Kraft wird vielmehr relativ zur äußeren Umgebung bestimmt (vgl. Burge 1989). Allerdings ist es nach wie vor schwierig, eine zufrieden stellende Beschreibung von „engem Inhalt" zu formulieren – nicht zuletzt, da Sätze der natürlichen Sprache nur externale Inhalte umfassen. Obwohl die Identität intrinsischer neuronaler Zustände hinreichend für die Identität von „engem Inhalt" ist, ist sie keine plausible notwendige Bedingung. So wäre es beispielsweise unmöglich, dass zwei verschiedene Individuen überhaupt denselben mentalen Zustand mit „engem Inhalt" haben, oder dass ein Zustand seinen Inhalt durch Verstehen o. ä. erhält. Es wird kontrovers diskutiert, ob es überhaupt „engen Inhalt" gibt. So ist es nach wie vor schwierig, engen Inhalt genau zu bestimmen. Es gibt verschiedene, konkurrierende Definitionen. So wird vorgeschlagen, „engen Inhalt" als detaillierte Beschreibung von Glaubensannahmen als begriffliche bzw. konzeptionelle Rolle (englisch: „conceptual role") oder als „mapping conception"[108] (vgl. z. B. Block 1986; Fodor 1987) zu bestimmen. Doch für diese verschiedenen Ansätze für „engen Inhalt" ergeben sich u. a. die Probleme, „engen" deutlich von „weitem Inhalt" abzugrenzen, oder dem normativen Charakter und den damit verbundenen Wahrheitsbedingungen von Inhalt gerecht zu werden. Die Ansätze für „engen Inhalt" scheinen mitunter zu syntaktisch zu sein. Es ist schwer zu präzisieren, was genau „enger Inhalt" letztlich ist.

[108] Eine detaillierte Beschreibung der Glaubensannahme „Wasser löscht Durst" wäre z. B. „etwas Klares, Farbloses, Geschmacksloses löscht Durst". Allerdings sind die detaillierten Beschreibungen selbst unter Umständen „weiter Inhalt". Eine konzeptionelle Rolle eines mentalen Zustands beschreibt seine kausalen Verbindungen zu weiteren Zuständen, bzw. wie der Ausdruck sich hinsichtlich Sinneseindrücken und Verhaltensweisen mit weiteren Ausdrücken verbindet und interagiert. Die „mapping conception" betont die Passung zwischen mentalem Inhalt und äußerer Umwelt. „Enger Inhalt" wird als die Eigenschaft verstanden, die „weiten Inhalt" speziell für die aktuelle Umwelt bestimmt.

### 4.3.3 Inhalt und mentale Kausalität

Eine zentrale Kritik an der Language of Thought-Hypothese ist, dass sie in einen Erklärungsregress mündet. Die Hypothese will erklären, wie natürliche Sprachen bzw. deren propositionale Inhalte entstehen, gelernt bzw. bedeutungsvoll werden. Dies soll der LOT-Hypothese zufolge durch die mentale Sprache Mentalese, die als syntaktisch-physisches Medium zwischengeschaltet ist, erreicht werden. Natürliche Sprachen haben dadurch Bedeutung, dass sie in ihrer Struktur mit der Gedankensprache korrespondieren. Dagegen kann eingewendet werden, dass jede erfolgreiche Erklärung für die Entstehung und das Funktionieren der LOT prinzipiell auch für die natürliche Sprache gelten müsste. Anders herum ist nicht einsichtig, warum die Schwierigkeiten der natürlichen Sprache in der Erklärung ihrer semantischen Inhalte nicht auch für die Gedankensprache gelten sollten.[109] Eine mögliche Antwort wäre, dass die Gedankensprache ihre Bedeutungen über einen anderen Weg erhält als die natürliche Sprache, nämlich über kausale Relationen, die die Mentalese-Sätze aufgrund einer bestimmten computationalen Struktur repräsentieren (vgl. Fodor 1975). Der Verweis auf die kausale Relation ist zentral für die Entstehung der Gedankensprache, wenn auch diese Begründung nicht unproblematisch ist. Schließlich scheint sie nicht flexibel genug, um der tatsächlichen biologischen Kapazität und den Möglichkeiten des menschlichen Gehirns, auch in seinen „Fehlleistungen", gerecht zu werden. Deshalb scheint sich die LOT-Hypothese als kausale Theorie naheliegend zunächst auf die Repräsentation von Wahrnehmungen und Sinneseindrücken zu beziehen, während bezweifelt werden kann, dass sie hinreicht, nicht natürliche Arten bzw. komplexe, z. B. ideologische präpositionale Einstellungen kausal zu erklären – auch wenn dieses in Anspruch genommen werden kann.[110] Möglicherweise werden manche komplexe oder basale Gedanken und Begriffe überhaupt nicht durch Kausalrelationen, sondern definitorisch bestimmt. Tatsächlich wird wenig darüber gesagt, wie die Verursachung einer propositionalen Einstellung als Ereignis selbst möglich ist, wobei dieser Vorgang als vor allem empirische Angelegenheit aufgefasst wird. Es sind mutmaßlich bestimmte Neuronenensembles, die bestimmte Sinnesdaten erhalten und diese als Repräsentation speichern.

Problematisch ist dabei die Frage, wie z. B. der Begriff „Katze" gespeichert wird. Warum bezieht sich der Begriff auf genau ein konkretes Tier bzw. eine bestimmte Tierart und nicht auch

[109] Weitere Kritik äußert Dennett, indem er einwendet, dass es sowohl propositionale Einstellungen ohne bewusste Repräsentationen geben kann als auch bewusste Repräsentationen ohne propositionale Einstellungen (vgl. z. B. Dennett 1981). Damit fordert er ein, dass die LOTH auch eine Antwort auf die Herausforderung dispositioneller propositionaler Eigenschaften und umgekehrt sublimer neuronaler Prozesse ohne direkter Zuordnung zu mentalen Inhalten finden muss.

[110] Der Unterschied zwischen der direkten Wahrnehmungsrepräsentation z. B. eines Baumes und der indirekten Repräsentation einer Ideologie kann dadurch verdeutlicht werden, dass ein Kommunist sich mit einiger Literatur u. a. von Marx auseinandersetzen muss. Wird diese Tätigkeit kausal interpretiert, setzt sie immer noch voraus, dass bereits ein gehöriger Bestand an repräsentationalen Strukturen vorhanden ist, auf dem aufgebaut werden kann. Es ist schwer zu beurteilen, ob dabei insbesondere abstrakte Begriffe tatsächlich kausal repräsentiert werden können oder nicht.

auf weitere Eigenschaften, die damit verbunden sind? Gegebenenfalls wird viel mehr repräsentiert als nur ein Tier. Ferner muss auch untersucht werden, was genau den Begriff „Katze" auslöst, welches sind die notwendigen oder hinreichenden Ursachen dafür, dass in meinem Gehirn der Begriff „Katze" repräsentiert wird – muss hierfür eine Katze oder eine bestimmte Katzeneigenschaft vorhanden sein? Welche direkten oder indirekten Ursachen müssen zusammenkommen, damit ein Begriff verursacht wird und nicht ein anderer? Besonders deutlich werden solche Schwierigkeiten der kausalen Theorie hinsichtlich der Begriffsbestimmung bei dem Problem der Fehlrepräsentation. So sind nach der LOT-Hypothese eigentlich keine Fehlrepräsentationen möglich, da Propositionen etc. durch die Kausalrelation prinzipiell als gültige Repräsentationen festgelegt sind. Tatsächlich kann aber mein aktueller Begriff „Katze" auch von etwas anderem als einer Katze bzw. einer relevanten Katzeneigenschaft, z. B. dem Schatten einer Katze oder sogar – bei schlechten Sichtverhältnissen – einem Hund verursacht werden (vgl. für die weitere Diskussion dieses Problem Kap. 4.5)[111]

Das kausale Problem für Inhalt besteht im Schwerpunkt für externalistische Positionen. Einerseits sollen unsere Überzeugungen oder Wünsche unser Verhalten kausal erklären (vgl. Davidson 1999; Kap. 2.3.1), andererseits sollen nur lokale, innere Eigenschaften unseres Körpers über die notwendige kausale Kraft verfügen (vgl. Stich 1983). Wenn Inhalt zumindest teilweise eine Sache der Dinge außerhalb eines Subjekts ist, wie z. B. Geschichte, eine Sprachgemeinschaft oder etwaige Vorfahren, dann lautet die Frage, wie Inhalt kausale Erklärung liefern kann. Alle Annahmen, wonach Inhaltseigenschaften nicht intrinsisch, syntaktisch oder neurologische Eigenschaften sind, sind zunächst inkonsistent mit der Annahme, dass Inhalt kausal wirksam ist. Eine übliche Einschätzung lautet folglich, dass Inhalt selbst nicht kausal wirksam ist, aber eine Eigenschaft, die eng mit ihm verbunden ist.

> Something of the shape, syntactical or neurological category, encodes the relevant semantic property, and the encoding property does the causing – that is how content gets to be causally relevant without acutally causing. (Jackson 1998, 99)

In der Vorstellung gibt es neben der "weiten" extrinsischen Inhaltseigenschaft eine intrinsische Eigenschaft, die den extrinsischen Inhalt entschlüsselt und somit als „engen Inhalt" für kausale

[111] In der Variante der Asymmetric Dependence Theory (ADT) versucht Fodor dieses Problem zu lösen, indem er der inhaltlich wertvollen Kausalbeziehung zwischen Objekt und Repräsentation weitere Kausalbeziehungen zur Seite stellt, die jedoch nicht für die tatsächliche Begriffsbildung, sondern allenfalls für mögliche Fehlrepräsentationen in Frage kommen (vgl. Fodor 1987, 1990a). Diese zweiten Kausalbeziehungen sind von der ersteren asymmetrisch abhängig, da sie nur aufgrund dieser möglich sind, die relevante Kausalbeziehung aber sehr wohl ohne die zusätzlichen besteht. Nach dem gewählten Beispiel wären die Objekte „Schatten" oder „Hund" als Fehlrepräsentationen für den Begriff „Katze" asymmetrisch von der tatsächlichen Kausalrelation zwischen dem Objekt „Katze" und dem Begriff „Katze" abhängig. Die Kausalbeziehung, die die meisten Merkmale des Begriffs erfüllt, verursacht den entsprechenden Inhalt. Einerseits macht ADT als atomistischer Ansatz die tatsächliche Kausalrelation metaphysisch sehr stark. Andererseits kann eingewendet werden, dass aufgrund optimaler Konditionen der naturalistische Ansatz konterkariert wird.

Prozesse verfügbar macht. Allerdings bleiben die Standardeinwände auch unter Verwendung des Entschlüsselungsmodells bestehen. So wird Inhalt mitunter als per se epiphänomenal verstanden. Tatsächlich sollte der Zusammenhang zwischen syntaktischen Zuständen und Inhalt theoretisch stichhaltig begründet werden:

> Of course, to make good this defence of the coding response, its proponents need a theory of content, a theory about what kinds of relations between syntactical states and their surroundings necessitate that subjects in those syntactical states have mental states with such and such contents. (ebd., 100)

In jeder moderaten internalistischen bzw. externalistischen Version ist der Inhalt von Wünschen und Glauben breit und schließt äußere kausale Geschichte und Umgebung von Subjekten mit ein, während psychologische Doppelgänger dagegen ausgeschlossen werden. Damit scheint eine intrinsische physische Kausalität für mentale Zustände aufgrund ihres Inhaltes nicht möglich zu sein. In den meisten Fällen versuchen wir, über breite bzw. externale Begriffe kausale Erklärungen zu finden, z. B. beim Spurenlesen eines wilden Tieres durch „perzeptiven Kontakt", oder bei der Lust auf ein Eis im Angesicht eines Eisladens. Breite Begriffe, scheint es, liefern Informationen über kausale Ursprünge. Im Subjekt werden diese Eigenschaften möglicherweise von informations-sensitiven Zuständen enkodiert (ebd., 105). Die Verbindung von weitem und engem Inhalt ist eine mögliche Lösung. Wenn jemand sich in einem intentionalen Zustand mit weitem Inhalt befindet, macht erst die Kombination mit einem engen intentionalen Zustand diese Verbindungen nach Außen bzw. zu dem kausalen Ursprung wahr.

> If we can identify this narrow local fact with a relatively intrinsic feature, presumably a neurological property of some kind, or something of a more structural kind that underlies broad content will be causally efficacious; if we cannot, then the narrow content will be causally relevant without being efficacious, but at least it will be local. (ebd., 106)

Sollte die Kombination zwischen „engem" und „weitem Inhalt" nicht möglich sein, bleibt die Unterscheidung zwischen kausaler Wirksamkeit und kausaler Relevanz, wobei Inhalt lediglich relevant ist. Manche Eigenschaften werden in der neurowissenschaftlichen Terminologie und andere in der Psychologie beschrieben. Letztere sind nicht kausal von ersteren verursacht, sondern werden von ersteren implementiert. Die Herausforderung wäre dann, kausale Relationen zwischen neuronalen Eigenschaften zu finden, die induzieren, dass es passende Relationen zwischen psychologischen Eigenschaften gibt.[112]

[112] Das Argument für „engen Inhalt" wurde im Grunde für wissenschaftliche Erklärungen der Psychologie und der kognitiven Wissenschaft benötigt. Inzwischen stimmt Fodor allerdings zu, dass die Zwillingswelt-Gedankenexperimente gezeigt haben, dass viele mentale Zustände weiten Inhalt haben (vgl. Fodor 1987, 1994).

***Quintessenz Kriterium Inhalt***

- *Konzepte und Begriffe als Konstituenten von Gedanken bzw. Beziehungen von Gedanken*
- *Unterscheidung begriffliche und nicht-begriffliche Inhalte*
- *mentale Inhaltszustände: normativ, unterliegen Wahrheitsbedingungen, verursachen weitere mentale Zustände, gewähren privilegierten Zugang*
- *Denken als kausales Fortschreiten von Zustand zu Zustand mit semantischem Sinn*
- *physische und formale Eigenschaften sichern semantische Kohärenz*
- *psychologische Erklärungen sollen kausale und inhaltliche Basis haben*
- *Ätiologie der Inhalte: intrinsische Eigenschaften kausal wirksam, aber von externalen Wahrheitsbedingung abhängig*
- *externale Eigenschaften mit lokaler Kausalität und Individuation vereinbar*
- *kausale Kraft relativ zu äußerer Umgebung*

*Language of Thought-Hypothese:*

- *Kausalrelationen bestimmen Inhalte, funktionale Einheiten*
- *kombinatorische Syntax als Basis*
- *Gedanke = Symbol einer Repräsentation mit syntaktischer Struktur für semantischen Inhalt*
- *Zahlreiche Kritikpunkte: Fehlrepräsentationen u. a.*
- *Angreifbares Modell: Encodierung äußerer Inhalt zu engem Inhalt*
- *Alternativ: Unterscheidung äußerer Inhalt (äußerer Stimulus als Ereignis) und innerer Inhalt (innerer Gedanke als Ereignis)*

Der Versuch, Inhalt durch entschlüsselnde Eigenschaften intrinsische kausale Kraft zukommen zu lassen, zeigt – unabhängig davon, ob dieser Lösungsansatz trägt – die Schärfe des Problems. Denn sollte jeglicher Inhalt, da er weit und extrinsisch ist, epiphänomenal sein, würde dies fundamental am menschlichen Selbstbild, aber auch Weltbild rütteln. Schließlich sind es durchweg Inhalte (oder Informationen?), die wir verarbeiten, mit denen wir die Welt beschreiben – und die wir als Ursache unserer Handlung verstehen. (Und letztlich ist auch die Aussage „Inhalt ist epiphänomenal" ein Inhalt, der uns bei seinem Zutreffen streng genommen auch nicht weiter stören müsste.) Hinsichtlich der mentalen Kausalität scheint relevant zu sein, dass „äußerer Inhalt" Verhalten bewirken kann, wie z. B. nach einem Telefonat als äußerem Ereignis Opernkarten zu kaufen, oder nach dem Lesen von Reiseprospekten einen bestimmten Urlaub zu buchen.[113]

[113] An dieser Stelle kann gefragt werden, ob nicht auch die Verbrennung bzw. ein Schmerz ein Inhalt sein kann, der bewirkt, dass die Hand von der Herdplatte genommen wird. Hierbei würde es sich allerdings um einen sehr weiten Inhaltsbegriff handeln, für Qualia würde man wohl eher von einer physischen Information sprechen, wohingegen Inhalt aufgrund seiner Abstraktheit schwer zu naturalisieren ist.

Aber auch „innerer Inhalt“ kann, wie z. B. ein plötzlicher Wunsch, eine Idee oder ein Geistesblitz als inneres Ereignis Körperbewegung und Verhalten verursachen. Der innere Inhalt kann grundsätzlich auf länger zurückliegende äußere Inhalte zurückgeführt werden. Aber beispielsweise im Falle des Nachdenkens kann eine Kombination zweier Gedanken und somit ein neuer Gedanke, eine neue propositionale Einstellung entstehen.

### 4.4 Das Kriterium der Relationalität

Das Kriterium der extrinsischen Relationalität stößt in den Kern mentaler Kausalität vor. Mentale Eigenschaften werden als extrinsisch aufgefasst. Damit wird beschrieben, dass sie abhängig von äußeren Vorgängen sind. Im Gegensatz dazu hängen intrinsische Eigenschaften nur von dem Ding, Zustand oder Ereignis ab, durch welches sie instanziiert werden, sie sind somit unabhängig von äußeren Vorgängen (vgl. Lewis 1983b, 197; Yablo 1998, 479ff.). Da jedoch nur intrinsische Eigenschaften – wie Masse, Form oder Gewicht – kausal wirksam sind, sind mentale Eigenschaften in der Schlussfolgerung kausal wirkungslos (vgl. z. B. McGinn 1989, 133). In Abgrenzung zu der Differenzierung zwischen internalen und externalen Eigenschaften beziehen sich intrinsische und extrinsische Eigenschaften nicht nur auf den propositionalen Inhalt eines mentalen Zustands. Dagegen kann Inhalt gegebenenfalls als extrinsische Eigenschaft neuronaler Aktivität eingestuft werden, da sowohl Internalisten als auch Externalisten nicht ablehnen, dass Inhalt grundsätzlich einen Bezug zur Umwelt hat; sie gewichten den Einfluss der Umwelt lediglich unterschiedlich.

Die Darstellung mentaler Eigenschaften als extrinsisch begegnet uns innerhalb der Supervenienztheorie. Die Supervenienzbeziehung selbst ist zwar eine intrinsische Relation, hinsichtlich der mentalen Eigenschaften wird jedoch ein Bezug zu äußeren Bedingungen hergestellt. Schließlich wird mentalen bzw. supervenienten Eigenschaften kausale Kraft abgesprochen, da sie extrinsisch sind (vgl. Kim 1982; vgl. Kap. 2.4.3). Nach Davidson können psychische Ereignisse nicht unter strikte Gesetze gefasst werden, da sie Rationalitätsprinzipien unterliegen und gemäß des so genannten „principle of charity“ Teil eines weiten Referenzrahmens sind (vgl. Kap. 2.3; Kap. 2.3.1). Insofern müssen auch sie als relational und extrinsisch verstanden werden. Innerhalb der funktionalistischen Position wird über die kausale Rolle argumentiert, die höherstufigen Eigenschaften zufällt. Diese kausale Rolle kann insofern als extrinsisch aufgefasst werden, als hierfür ein übergeordnetes höherstufiges System den Referenzrahmen liefert und somit auch hier eine äußere Abhängigkeit, die über das direkte Objekt hinausgeht (vgl. Kap. 2.2; Kap. 2.2.3).

Für mentale Kausalität ist das Kriterium der Extrinsität bedeutsam, weil hiermit letztlich verbunden ist, ob mentale Eigenschaften oder lediglich physische Eigenschaften kausal wirksam

sein können. Dabei wird innerhalb der Diskussion deutlich, dass die Begriffe „intrinsisch“ und „extrinsisch“ weniger deutlich definiert und voneinander abgegrenzt werden können, als zunächst angenommen.

### 4.4.1 Instrinsische versus extrinsische Eigenschaften

Unter einer intrinsischen Eigenschaft wird traditionell verstanden, dass sie zu einem bestimmten Zeitpunkt von einem Objekt instanziiert wird und nichts weiter als dieses Objekt und seine Teile umfasst. Beispiele für intrinsische Eigenschaften sind Form, Masse, Ladung oder innere Struktur. Ein Ding, Zustand oder Ereignis hat eine intrinsische Eigenschaft aufgrund seiner selbst. Aus der Humeschen Perspektive ist eine intrinsische Beschreibung des Zustands der Welt jederzeit vollständig und frei von Implikationen des Zustands zu einer anderen Zeit leistbar: „Festigkeit, Ausdehnung, Bewegung, diese Eigenschaften sind alle in sich abgeschlossen und weisen nie auf ein anderes Ereignis hin, das aus ihnen hervorgehen könnte.“ (Hume 1984, 78). Eine extrinsische Eigenschaft umfasst dagegen etwas, das noch über das Objekt hinausgeht. Die Eigenschaften eines Balles, „rot und rund zu sein“, sind intrinsisch, während die Eigenschaften, weniger als zwei Meter von einem Baum zu liegen und einem kleinen Jungen zu gehören, als extrinsische Eigenschaften gelten. Weitere Beispiele für extrinsische Eigenschaften sind „ein Bruder sein“, „verschuldet sein“ oder „an Berlin denken“ (vgl. Lewis 1983b, 197). Damit können intrinsische und extrinsische Eigenschaften scheinbar klar von einander abgegrenzt werden:

> „The intrinsic properties of something depend only on that thing; whereas the extrinsic properties of something may depend, wholly or partly, on something else.“ (ebd.)

Dennoch ist eine Unterscheidung allein deshalb problematisch, da derzeit verschiedene Bedeutungen von „intrinsisch” und „extrinsisch” diskutiert werden. [114] So können ein relationaler, ein innerer oder ein qualitativer Begriff für Intrinsität unterschieden werden. Paradoxerweise scheint es insbesondere schwierig, eine präzise Charakterisierung für intrinsische Eigenschaften zu bestimmen, obwohl gerade mit ihnen als basale Eigenschaften kausale Kraft bzw. echte Veränderung verbunden sein soll (vgl. Geach 1969). Der relationale Begriff besagt, dass ein Objekt extrinsische Eigenschaft aufgrund seiner Beziehung zu anderen Objekten hat, während intrinsische Eigenschaften keine Relation zu etwas, das nicht Teil ihres Bezugsobjektes ist, haben dürfen und deshalb nicht-relational sind (vgl. Franscecotti 1999). Dies ist allerdings eine sehr strikte und

[114] Eine extreme Position hat Leibniz vertreten, der aufgrund seiner metaphysischen Position ablehnte, dass es überhaupt einen extrinsischen Nennwert gibt (vgl. Leibniz 2007). Popper dagegen glaubte, dass physische Eigenschaften wesentlich dispositional und somit extrinsisch sind (vgl. Popper 1959, 424f.).

schwer zu erfüllende Anforderung. Die Kraft, Türschlösser einer bestimmten Art zu öffnen, beinhaltet Beziehungen zu äußeren Dingen anderer Art. Intrinsisch zu sein und relational zu sein, scheint sich nicht zwingend auszuschließen, und viele Eigenschaften umfassen beides. Beispielsweise steht auch die Bewegung eines Objekts in Relation zum Zustand der Welt zu einem anderen als dem gegebenen Zeitpunkt. So ist es problematisch, einen Zustand, ein Ereignis oder auch eine Eigenschaft völlig abgetrennt von anderen zu betrachten. In vielen Fällen stellen sich scheinbar intrinsische Eigenschaften als relational heraus.[115] Begriffe wie alt, groß oder unvollständig scheinen weder relativ, noch external dominiert zu sein, und doch beinhalten sie eine stillschweigende Relation (vgl. Locke 2006, Bd. 1, Kap. 25, 28). Intrinsische Eigenschaften würden somit eine sehr feine und lediglich begriffliche Unterscheidung machen.

Der Begriff der inneren oder äußeren Eigenschaften unterscheidet eine metaphysische und eine epistemische Bedeutung für intrinsische oder extrinsische Eigenschaften. Danach ist eine intrinsische Eigenschaft metaphysisch eine Eigenschaft eines Objekts, wenn dieses die Eigenschaft aufgrund seiner selbst und unabhängig von anderen Dingen hat. Eine intrinsische Eigenschaft ist epistemisch eine Eigenschaft eines Objekts, wenn sie durch Beobachtung des Objekts selbst bestimmt werden kann (vgl. Dunn 1990). Ausschlaggebend ist hierbei die metaphysische Bestimmung, danach sind z. B. „identisch mit X sein" oder auch „X als einen Teil haben" intrinsische Eigenschaften im Gegensatz zu „ein Doppelgänger von X sein".[116]

Der qualitative Begriff der Intrinsität bzw. Extrinsität schließlich bezieht sich auf die Feststellung, dass intrinsische Eigenschaften Doppelgänger gemeinsam haben und diese sich bei ihnen niemals unterscheiden. Danach ist ein Objekt x intrinsisch P, wenn alle seine Doppelgänger P sind, wenn also alle Objekte dieselben intrinsischen Eigenschaften wie das Objekt haben, P sind (vgl. Humberstone 1996, 227). Daraus ergibt sich scheinbar auch, dass koextensive Eigenschaften gleichsam intrinsisch oder extrinsisch sind, sowie die Transitivität der Doppelgängerrelation – der Doppelgänger des Doppelgängers des Originals ist ebenfalls Doppelgänger des

[115] Idealisten gehen sogar soweit in ihrer Argumentation, dass die physischen Eigenschaften die eigentlich extrinsischen Eigenschaften sind. So ist jede physische Entität, wie Protonen, Neutronen oder Elektronen ausschließlich über ihre mathematischen Eigenschaften definiert – und damit abhängig von einer mathematischen Theorie. Folglich ist ein Proton extrinsisch, da es über extrinsische Eigenschaften wie Masse, Form oder Drehung definiert wird und nicht über intrinsische Qualitäten. Es ist nach jener Perspektive lediglich ein Konstrukt, das vollständig durch die Kombination extrinsischer mathematischer Eigenschaften bestimmt wird (vgl. Lloyd 2006, 122).

[116] Wie schwierig es ist, mentale Eigenschaften als intrinsisch oder extrinsisch zu bestimmen, wird auch dadurch deutlich, dass Internalisten über mentalen Inhalt vertreten, dass eben der Inhalt mentaler Zustände stark auf intrinsischen Eigenschaften des Subjektes superveniert. Externalisten halten dagegen, dass zwei Subjekte innerhalb einer möglichen Welt intrinsische Doppelgänger und zugleich unterschiedlich in ihrem mentalen Inhalt sein können (vgl. Putnam 1975c). Insofern müssen Externalisten Supervenienz ablehnen. Dabei bietet sich die Supervenienzrelation an, die Unterscheidung zwischen internalen und externalen Relationen zu präzisieren. Internale Relationen (wie z. B. kleiner als) supervenieren stark auf intrinsischen Eigenschaften ihrer Relata, während externale Relationen (wie z. B. drei Kilometer entfernt sein von) nicht einmal schwach auf den intrinsischen Eigenschaften ihrer Relata supervenieren (vgl. Lewis 1986).

Originals. Problematisch an dieser „globalen Definition" intrinsischer Eigenschaften ist, dass einige Eigenschaften nicht erfasst werden können. Angenommen P ist die Eigenschaft, quadratisch zu sein und Q ist die Eigenschaft quadratisch zu sein oder in einem goldenen Berg zu sein. Für einige Objekte Q beruht ihr „Q sein" allein auf intrinsischen Eigenschaften, für andere dagegen auf extrinsischen Eigenschaften außerhalb des Objekts selbst. Es sind vor allem solche disjunkte Eigenschaften, die einer qualitativen Definition der Intrinsität Probleme bereiten. Schließlich gilt der Grundsatz für extrinsische Eigenschaften, dass der Besitz solcher Eigenschaften die Existenz anderer Objekte voraussetzt und umfasst. Alternativ kann anstelle einer globalen auch eine „lokale Definition" für Intrinsität getroffen werden, wonach x eine Eigenschaft Q intrinsisch besitzt, wenn x auch eine Eigenschaft P intrinsisch besitzt und „Eigenschaft P haben" auch „Eigenschaft Q haben" umfasst. Es wird dabei deutlich, dass der intrinsische Gehalt eines Objekts davon abhängt, welche Eigenschaften es intrinsisch besitzt, nicht welche intrinsischen Eigenschaften es besitzt.

Intrinsische Eigenschaften zeichnen sich dadurch aus, dass sie „einsam" sind und für sich alleine stehen, während extrinsische Eigenschaften mit dem Begriff der „Begleitung" kompatibel sind (vgl. Lewis 1983b, 198; Kim 1982; zitiert nach: Lewis 1983b; Yablo 1998). Um diesen Punkt deutlicher herauszustellen und auch dem Problem der disjunkten Eigenschaften zu begegnen, haben Langton und Lewis betont, dass intrinsische Eigenschaften natürliche Eigenschaften sein müssen, wozu disjunkte Eigenschaften, wie „einsam und rund sein oder begleitet und kubisch", nicht gehören (vgl. Langton/Lewis 1998; Kap. 3.1.2). Es sind hiernach lediglich vollständig natürliche Eigenschaften intrinsisch und es sind vor allem intrinsische Eigenschaften, die sich niemals zwischen Doppelgängern unterscheiden:

> Then we can say that two things are duplicates if (1) they have exactly the same perfectly natural properties, and (2) their parts can be put into correspondence in such a way that corresponding parts have exactly the same perfectly natural properties, and stand in the same perfectly natural relations […]. (Lewis 1986, 61f.)

Diese Einschätzung qualitativer Intrinsität fußt nicht nur auf der Aussage, dass alle vollständig natürlichen Eigenschaften intrinsisch sind, sondern auch implizit auf der Aussage, dass vollständig natürliche Eigenschaften für eine vollständige Beschreibung der Welt hinreichend sind. Diese Ergänzung ist notwendig, damit keine Objekte, die dieselben natürlichen Eigenschaften teilen, sich hinsichtlich ihrer intrinsischen Eigenschaften unterscheiden. Diese Prämisse verweist auch auf die Voraussetzung der Geschlossenheit der Physik, wonach physische Eigenschaften hinreichend für eine (kausale) Beschreibung der Welt sind (vgl. Kap. 3.3). Damit schließt sich der Kreis. Extrinsische Eigenschaften, die nicht natürliche Eigenschaften bzw. zumindest nicht im

Kern natürliche Eigenschaften sind, scheinen für eine Beschreibung der Welt nicht notwendig zu sein. Damit sind mentale Eigenschaften per definitionem ausgeschlossen.

### 4.4.2 Relationalität und mentale Kausalität

Im Unterschied zu gängigen Meinungen argumentiert Searle, dass auch mentale Eigenschaften bzw. intentionale Zustände intrinsische Tatsachen sind. Ihnen kommt eine intrinsische Intentionalität zu (vgl. Searle 2001, 111; Searle 1993, 11). Er möchte auf diese Weise Bewusstseinszustände naturalisieren. Denn ebenso wie Bewusstsein versteht Searle auch Intentionalität zunächst als ein beobachter-unabhängiges Merkmal des Geistes in dem Sinne, dass die Existenz dieses Merkmals nur von dem denkenden Subjekt und nicht einem zusätzlichen Beobachter abhängt. Konsequenterweise vergleicht Searle intrinsische Intentionalität mit den physischen Merkmalen Gravitation und Elektromagnetismus oder den biologischen Merkmalen Mitose, Meiose und Photosynthese (vgl. Searle 2001, 114f.).

Es muss gezeigt werden, inwiefern die intrinsische Intentionalität Teil der natürlichen Welt sein kann und mentale Phänomene auf natürlichen Eigenschaften beruhen. Searle beginnt mit den biologisch primitivsten Formen von Intentionalität, nämlich Wünschen, die körperliche Bedürfnisse, wie Hunger und Durst repräsentieren. Er stellt fest, dass es eine biologische Tatsache ist, dass neurobiologische Prozesse Bewusstseinszustände wie Hunger oder Durst verursachen. Einwände, hier würde ein Kluft zwischen Körper und Geist bzw. eine Erklärungslücke zwischen Physischem und Geistigem bestehen, lässt Searle nicht gelten (vgl. Kap. 3.2.4). Biologische Eigenschaften werden dagegen als intrinsische und damit prinzipiell natürliche Eigenschaften interpretiert, die Körper und Geist vereinen. Searle verfolgt die Strategie zu zeigen, dass die Form der natürlichen biologischen Intentionalität nicht nur für primitive, sondern auch komplexe biologische Formen wie Sinnesmodalitäten gilt (vgl. Searle 2001, 117). So verursachen Hirnvorgänge zwar visuelle Erlebnisse, aber das visuelle Erlebnis selbst ist die intrinsische Intentionalität als biologische Form:

> Das heißt, sobald wir das visuelle Erlebnis samt all seinen Merkmalen haben, und sobald wir für diese Merkmale neurobiologische und psychologische Erklärungen gegeben haben, kann es keine weitere interessante Frage danach geben, wie das Erlebnis ein Fall von Jemandem-so-vorkommen-dass-er-etwas-sieht sein kann; […] Das Erlebnis ist einfach ein Erlebnis des Jemandem-so-vorkommen-dass-er-etwas-sieht. (ebd., 119)

Intrinsische Intentionalität ist ein Phänomen, das zur biologischen Beschaffenheit von Menschen und einigen anderen Lebewesen gehört. Wie sie verwendet werden oder was sie über sich selbst

denken, oder auf welche Weisen der Selbstbeschreibung sie verfallen – daran liegt es nicht, ob sie Intentionalität haben. Es ist eine Tatsache, dass Tiere Hunger oder Durst bekommen, dass sie Dinge sehen, sich fürchten o. ä. (vgl. Searle 1993, 98). Der durchaus übliche gegensätzliche Sprachgebrauch von „intrinsisch" und „relational" kann nicht aufrechterhalten werden. So hat der Mond intrinsischerweise eine Masse, er ist aber nicht intrinsischerweise ein Satellit. Ein Satellit ist er nur relativ zur Erde. Bedürfnisse, Wünsche, visuelle Erlebnisse sind zwar relativ zu anderen Objekten, aber auch intrinsische intentionale Zustände „[…] und das bedeutet einfach, dass sie echte Intentionalität haben und nicht bloß etwas, das echter Intentionalität mehr oder weniger gleicht (als ob), bzw. nicht bloß etwas, das daraus resultiert, dass jemand anders etwas in gewisser Weise verwendet oder zu etwas bestimmte Einstellungen hat (abgeleitet)." (Searle 2001, 99.)

Man kann zwar in einer ersten Intuition unterscheiden, dass extrinsische Eigenschaften relationale und intrinsische Eigenschaften nicht relationale Eigenschaften sind. Jedoch scheinen viele intuitiv intrinsische Eigenschaften in bestimmter Weise auch relational zu sein. So umfasst die Quadratheit eines Objektes auch die Relation zwischen seinen Seiten. Andererseits scheint es ein gutes und klares Kriterium für intrinsische Eigenschaften zu sein, dass sie keine anderen Beziehungen haben als zu dem Objekt, auf das sie sich beziehen. Man kann also festhalten, dass

a) einige Eigenschaften, die vornehmlich intrinsisch sind, auch eine extrinsische Komponente haben und
b) einige Eigenschaften, die vornehmlich extrinsisch sind, auch eine intrinsische Komponente haben.

Grundsätzlich können intrinsische Eigenschaften als extrinsische aufgefasst werden, vice versa – zumindest, wenn man einen relationalen bzw. nicht-relationalen Begriff von intrinsisch und extrinsisch verwendet. Einige intrinsische Eigenschaften sind zumindest teilweise extrinsisch oder relational. „Properties may be more or less extrinsic; being a brother has more of an admixture of intrinsic structure than being a sibling does, yet both are extrinsic." (Lewis 1983b, 197) Sider hat in Ergänzung hierzu beobachtet, dass viele Eigenschaften in der Alltagssprache bzw. -psychologie in gewisser Weise „maximal" und nicht intrinsisch sind (vgl. Sider 2003). Maximale Eigenschaften eines Objektes umfassen auch minimale Eigenschaften eines Objektes, sie werden also nicht von allen Objekten desselben Typs in der maximalen Version geteilt. Beispielsweise wird die maximale Version eines Hauses von einer Version mit der Eigenschaft „Haus ohne Fenster" unterschieden. Erstere schließt letztere mit ein. Ein Fels zu sein, ist ein weiteres Beispiel für solch eine Eigenschaft – ein Fels kann wiederum Teil eines größeren Felsens sein. Deshalb ist „ein Haus sein" oder „ein Fels sein" extrinsisch, denn Doppelgänger des großen Felsens können ein Fels sein, wenn sie vom übrigen Teil des Felsens getrennt werden können (vgl. ebd. 2001,

357, 360). Maximale Eigenschaften können insofern als Spezialfall von „Grenz-Sensitivität" betrachtet werden. Eine Eigenschaft ist „grenz-sensitiv", wenn ihre Instanziierung von Vorgängen im äußeren Grenzbereich des zugehörigen Objekts bzw. Ereignisses abhängt (ebd., 358). Hinsichtlich der mentalen Kausalität stellt sich hierbei die Frage, inwiefern mentale oder physische Eigenschaften oder auch neuronale Aktivität als maximal eingeordnet werden kann. Welche maximale neuronale Aktivität kann mit einer minimalen neuronalen Aktivität für die Prozessierung eines Gedankens kontrastiert werden? In der Korrelation können physischen und mentalen Eigenschaften unterschiedliche Geltungsbereiche zugeschrieben werden, wie bereits in der Diskussion um die Ko-Extension bzw. die Identität mentaler und physischer Eigenschaften angedeutet wurde (vgl. Kap. 2.1.3).

Die Feststellung, dass intrinsische und extrinsische Eigenschaften „gemischt" sein können, müsste sich grundsätzlich auch auf den Bereich der mentalen Kausalität anwenden lassen. Für einige physische intrinsische Eigenschaften würde gelten, dass sie als solche extrinsisch, für einige mentale extrinsische Eigenschaften, dass sie als solche intrinsisch sein können. Es ist allerdings schwer einzugrenzen, wann ein mentaler Zustand intrinsisch ist und wann nicht. Darf ein mentaler Zustand nur rein neuronale „Informationen" umfassen, was ist dann mit Informationen über z. B. innere Körperfunktionen (wie Herzschlag, Durst, Schmerz). Ist der Körper etwa ein Marker für Intrinsität? Oder lassen sich die Grenzen Körper und Außenwelt „fließend" vorstellen? Schließlich ist neuronale Aktivität an sich auch in Relation zur biologischen Aktivität des Körpers zu setzen. Insofern mentalen Eigenschaften propositionaler Gehalt zufällt, sind sie als propositionale Eigenschaften extrinsisch, somit durch äußere Bedingungen bestimmt und von diesen abhängig. Sie sind damit abhängig von der Funktion oder Rolle, die ihnen ein Beobachter verleiht. Der Begriff „Badewanne" ist insofern extrinsisch, als über die bloße Form und das Material dieses Gegen-stands eine gesellschaftliche Konvention besteht, dass Badewannen nicht der Lagerung von Büchern dienen, sondern der Körperreinigung (vgl. Searle 1993, 10f.). Oder sie sind als funktionale Eigenschaften extrinsisch, da sie auf der höherstufigen Ebene die kausale Rolle erfüllen, die ihnen aufgrund der externen Bedingungen abverlangt wird.

Wenn man die „globale Definition" intrinsischer Eigenschaften auf mentale Kausalität überträgt, ergeben sich gewichtige Unterschiede abhängig davon, ob man mentale Eigenschaften, physische Eigenschaften oder mentale Ereignisse mit der Definition konfrontiert. Ist es nicht auch zulässig zu sagen, dass

> ein mentales Ereignis x über die mentale Eigenschaft M intrinsisch verfügt, wenn alle seine Doppelgänger über M verfügen und alle Ereignisse, die die gleiche Eigenschaft intrinsisch haben, auch über die Eigenschaft M verfügen?

Damit ist die mentale Eigenschaft M zwar selbst noch keine intrinsische Eigenschaft, aber sie kann dennoch unter lokalen Aspekten intrinsisch zu einem mentalen Ereignis gehören. Diese Unterscheidung erscheint insofern bedeutend, als bei mentaler Kausalität genau zu fragen ist, ob in der kausalen Relation die Eigenschaftsebene von der Ereignisebene losgelöst betrachtet werden kann, und zusätzlich die mentalen und physischen Ereignisse vollständig getrennt analysiert werden können. In den vorhergehenden Kapiteln wurde bereits darauf hingewiesen, dass die Realisierungsrelation sehr eng ist. Dieses scheint ein entscheidender Punkt zu sein. So kann auch formuliert werden:

> Ein mentales Ereignis x verfügt intrinsisch über die Eigenschaft R, wobei R die Realisierungsrelation zwischen der mentalen Eigenschaft M und der physischen Eigenschaft P beschreibt, weshalb R eine intrinsisch-extrinsische Eigenschaft ist, wenn alle Doppelgänger von x über R verfügen und alle Ereignisse, die die gleichen Eigenschaften intrinsisch haben, auch über die Eigenschaft R verfügen.

Problematisch ist diese Beschreibung allerdings unter der Bedingung der Multirealisierbarkeit mentaler Eigenschaften. Schließlich sind Doppelgänger der mentalen Eigenschaft M denkbar, die sich aber in der physischen Eigenschaft P unterscheiden, und umgekehrt Doppelgänger in der physischen Eigenschaft P, die sich in der mentalen Eigenschaft M unterscheiden. Vielleicht erscheint es allein deshalb geboten, eher eine „lokale Definition" für Intrinsität bzw. Extrinsität der mentalen Kausalität zu verwenden, wonach jedes mentale Ereignis X, das intrinsisch über eine intrinsisch-extrinsische Eigenschaft R verfügt, auch über die intrinsische Eigenschaft P und über die extrinsische Eigenschaft M verfügt, wenn R sowohl P als auch M umfasst. So wäre beispielsweise ein empirischer Bezug denkbar, dass Neuronen intrinsisch über mentale und physische Eigenschaften verfügen und deshalb entscheidend für mentale Kausalität sind, bzw. mentale Ereignisse prozessieren. Dabei stellt sich andererseits die Frage, ob nicht die physischen Eigenschaften eines Neurons auch seine mentalen Eigenschaften umfassen und somit ebenfalls per definitionem extrinsisch sind: Denn jede Eigenschaft, deren Existenz andere zugehörige Objekte voraussetzt, gilt streng genommen als extrinsisch (vgl. Yablo 1998, 487).

Man kann einwenden, dass es sich hierbei letztlich um ein definitorisches Problem handelt, welche Eigenschaften vornehmlich wie zu bezeichnen sind. Jenseits der definitorischen Frage ist aber zu klären, ob mentale – und vornehmlich extrinsische – Eigenschaften nicht eben doch kausal wirksam sein können, bzw. ob der Hinweis, diese Eigenschaften seien extrinsisch, sie zwangsläufig als kausal wirkungslos qualifiziert. Hierauf verweist Geachs Analyse, wonach zwischen einer echten Veränderung und einer eher begleitenden Veränderung unterschieden werden kann (vgl. Geach 1969). In dem bekannten Beispiel wird Xanthippe zum Zeitpunkt von Sokrates Tod zur Witwe. Das Ereignis, eine Witwe zu werden, geschieht nicht aus sich selbst heraus, es ist

abhängig vom Tod des Sokrates und wird somit als extrinsisch charakterisiert. Es sind hiernach nur die intrinsischen Eigenschaften, denen die Veränderung eines Ereignisses zugeschrieben werden können. Die Frage ist nun, ob diese Unterteilung auch für mentale Kausalität relevant ist, dass eben nur die physischen Eigenschaften eines mentalen Ereignisses kausal relevant sind und nicht die mentalen Eigenschaften. Vor diesem Hintergrund ist die Unterscheidung zwischen intrinsisch und extrinsisch interessant, die Ellis trifft. Er weist darauf hin, dass zwischen Eigenschaften, die unabhängig von äußeren, einwirkenden Kräften sind, und Eigenschaften, die abhängig von äußeren, einwirkenden Kräften sind, unterschieden werden kann (vgl. Ellis 2001). Dabei wird deutlich, dass für viele Objekte bzw. Ereignisse gilt, dass sie von äußeren Kräften, die auf sie einwirken, abhängig sind.

Als Beispiel kann ein auseinander gezogenes Gummiband aufgeführt werden, dass von den äußeren Kräften, die es dehnen und ihm somit seine Form verleihen, abhängt. Dieses Beispiel macht deutlich, dass es nichts Ungewöhnliches sein muss, dass ein Objekt aufgrund äußerer Kräfte seine Eigenschaften erhält. Möglicherweise kann diese Analyse auch auf mentale Ereignisse und ihre mentalen Eigenschaften übertragen werden. Schließlich ist es die Übertragung bzw. Repräsentation äußerer Ereignisse, die mentale Ereignisse auszeichnet. Mentale Eigenschaften werden als extrinsisch aufgefasst, da sie sich auf etwas Äußeres beziehen, das über messbare intrinsische Qualitäten hinausgeht. Sie sind relational. Hierin liegt aber bereits das Problem, da die extrinsische Qualität per definitionem nicht in einen physischen Kausalprozess integrierbar scheint; der relationale Charakter ist nicht integrierbar. Dies wäre jedoch eine Anforderung, die eine spezifisch mentale Kausalität erfüllen müsste. Da mentale Ereignisse in Relation zu Umwelt bzw. äußeren Ereignissen stehen, geht jede Kausalität, die dieses Moment ausklammert, an dem wesentlichen Punkt einer mentalen Kausalität vorbei.[117] Allein auf der Basis neurologischer Prozes-se ist es aufgrund des inversen Problems bislang nicht möglich, die Relation zwischen mentalen und physischen Eigenschaften, dem untersuchten mentalen Ereignis und dem entsprechenden Umweltereignis zu beschreiben. Wer sich allein auf eine neurologische Beschreibung als Bezug auf physische Eigenschaften beschränkt, wird ontologisch der mentalen Relation nicht gerecht. Deshalb schreibt Dretske dem menschlichen Gehirn auch extrinsische und intrinsische Eigenschaften zu. So wie ein Foto ein gewöhnliches physisches Objekt mit intrinsischen Eigenschaften wie Masse, Farbe, Größe etc. aufgrund seiner extrinsischen Eigenschaften etwas anderes repräsentiert, so ist ein Gedanke ein gewöhnlicher neuronaler Gehirnzustand mit üblichen intrin-

[117] Wenn Eigenschaften tatsächlich intrinsisch sind, wie können sie dann etwas repräsentieren, das außerhalb ist? Eine konsequent intrinsische Position dürfte nicht von „repräsentieren" sprechen, sondern müsste eine Art Solipsismus vertreten. Die mentalen Zustände würden nicht die (Außen-) Welt repräsentieren, sondern sie wären die einzige (intrinsische) Welt des Denkenden. In diesem Punkt streift die Frage, ob eine mentale Eigenschaft intrinsisch oder extrinsisch ist, auch die Frage nach der Externalität von Gedanken. Das Kriterium der extrinsischen Relation steht dabei in engem Zusammenhang mit den Kriterien „Intentionalität" und „Inhalt" (vgl. Kap. 4.3; Kap. 4.5).

sischen Eigenschaften, aufgrund dessen extrinsischer Eigenschaften etwas weiteres in der Welt repräsentiert wird (vgl. Dretske 1994, 133; Kap. 6.4).

***Quintessenz Relationalität***

- *extrinsische Eigenschaften abhängig von äußeren Vorgängen: kausale Rolle über das Objekt hinaus*
- *intrinsische Eigenschaften kausal wirksam: nur Teile des instanziernden Objektes betreffend*
- *Besitz extrinsischer Eigenschaften setzt andere Objekte voraus („Begleitung")*
- *Ablehnung des gegensätzlichen Sprachgebrauchs „intrinsisch versus relational": intrinsische Eigenschaften mit extrinsischer Komponente, extrinsische Eigenschaften mit intrinsischer Komponente*
- *maximale Eigenschaften umfassen minimale Eigenschaften („grenzsensitiv")*
- *Einstufung mentaler Eigenschaften als biologisch-intrinsische Eigenschaften*
- *extrinsische und intrinsische Eigenschaften für das Gehirn*
- *Integration des relationalen Charakters als zentrale Anforderung einer spezifisch mentalen Kausalität*
- *Fragestellung: Umfassen physische Eigenschaften eines Neurons auch mentale Eigenschaften?*

- *Lokale Definition:*

  *Jedes mentale Ereignis X, das intrinsisch über eine intrinsisch-extrinsische Eigenschaft* R *verfügt, verfügt auch über die intrinsische Eigenschaft P und verfügt über die extrinsische Eigenschaft M, wenn* R *sowohl P als auch M umfasst.*

### 4.5 Das Kriterium der Intentionalität

Eng verbunden mit dem Kriterium der Externalität ist das Kriterium der Intentionalität für mentale Kausalität. Insbesondere Searle (1993; 2001) hat auf diesen Aspekt im Zusammenhang mit der kausalen Funktion des Bewusstseins hingewiesen. Intentionalität beschreibt das mentale Vermögen, zu repräsentieren bzw. über Dinge, Eigenschaften oder Zustände zu sein. Nach einer traditionellen Definition umschreibt der Begriff der Intentionalität die Spannung, mit der der menschliche Geist sich auf Objekte bezieht. Aufgrund der Intentionalität kann ein psychologischer bzw. mentaler Zustand auf ein intentionales Objekt gerichtet sein – mentale Zustände beziehen sich auf *etwas*. (vgl. Brentano 1874).[118] In einem weiten Begriff können alle möglichen

[118] Brentano ging sogar so weit, Intentionalität als das entscheidende Kriterium des Mentalen zu bestimmen. Danach ist Intentionalität eine Eigenschaft aller mentaler Zustände und Ereignisse, wie Glaubenszustände, Urteile, Wünsche oder Wahrnehmungen. Nicht-Intentionalisten argumentieren dagegen, dass der phänomenale Charakter der Qualia nicht durch Intentionalität erfasst wird (vgl. Searle 1992). In einer moderaten Position wird

Dinge intentionale Objekte sein: konkrete physische Objekte in Raum und Zeit genauso wie abstrakte Objekte, wie z. B. Zahlen, oder auch fiktionale Objekte und sogar irrationale Objekte, wie z. B. Einhörner oder quadratische Kreise. Deshalb wird auch von einer intentionalen Nichtexistenz gesprochen, da intentionale Objekte nicht in Raum und Zeit existieren müssen (vgl. Chisholm 1957). Entscheidend ist für alle Objekte dabei, dass sie im Geist von einer Person vorgestellt werden. Franz Brentano hat Intentionalität maßgeblich als zentrales Merkmal für mentale Phänomene definiert. Mentale Zustände beziehen sich wesentlich auf Dinge außerhalb ihrer selbst. Sie beinhalten somit stets auch das Objekt, auf das sie sich beziehen. Diese Objekte haben die Eigenschaft der intentionalen Nichtexistenz, wodurch Intentionalität möglich wird. Auch die intentionale Nichtexistenz ist ein Merkmal ausschließlich mentaler Phänomene. Hieraus ergeben sich allerdings Zweifel, ob intentionale Objekte tatsächlich existieren.

Es gibt eine wichtige Asymmetrie zwischen intentionalen und nicht-intentionalen Beziehungen. Nicht-intentionale Beziehungen gelten bzw. existieren als konkrete Ereignisse in Raum und Zeit. Wenn es heißt, „Kleopatra küsste Cäsar", wird beiden, Kleopatra wie Cäsar, eine Existenz zugesprochen. Heißt es dagegen, „Kleopatra liebte Cäsar", ist der Fall nicht mehr ganz so eindeutig. Zwar wissen wir, dass es Cäsar gab und können für diesen Fall auch eine konkrete Beziehung bestimmen, doch ist es auch möglich, dass jemand eine fiktionale Person, wie Anna Karenina liebt, und in diesem Fall ist die Beziehung nicht mehr in Raum und Zeit zu bestimmen. Für die intentionale Relation reicht es aus, wenn die denkende Person existiert, während das gedachte Objekt nicht existieren muss.[119]

Anknüpfend an Brentanos Annahme einer „intentionalen Nichtexistenz" wird bezweifelt, ob es überhaupt intentionale Objekte gibt. Insbesondere Quine (1963) hat die Existenz intentionaler, abstrakter Objekte bezweifelt – und damit auch, dass Intentionalität das entscheidende Kriterium des Mentalen ist. So scheint es nicht möglich, intentionale Phänomene mit nicht-intentionalen Begriffen zu bestimmen bzw. auf physische Begriffe zu reduzieren. Ursprung dieser Feststellung ist die Überlegung, ob Intensionalität ein Kriterium für Intentionalität ist (vgl. Chisholm 1957; Quine 1960, 220). Quine präsentierte ein Dilemma, wonach entweder intentionale Ausdrücke unentbehrlich sind, über eine eigenständige Wissenschaft verfügen und deshalb eine physikalische Ontologie zurückgewiesen wird, oder aber die Position des Physikalismus akzeptiert und in der Folge auf die leeren intentionalen Begriffe und eine damit verbundene Wissen-

---

angenommen, dass zumindest viele mentale Zustände sowohl intentionale, als auch phänomenale Eigenschaften haben (vgl. Block 1997; Chalmers 1997).

[119] Ein weiteres Beispiel für die intentionale Beziehung zwischen Gedanken und Dingen, die nicht existieren, sind zutreffende Verneinungen. Wenn jemand zutreffend die Existenz eines Objektes oder einer Person verneint, ist dies ein Beleg dafür, dass intentionale Zustände sich (auch) auf nicht-existierende Objekte beziehen können. So kann jemand verneinen, dass ein Pegasus existiert.

schaft verzichtet wird (vgl. ebd., 221).[120] Ein möglicher Ausweg aus diesem Dilemma ist, das Phänomen der Intentionalität mit der physikalistischen Position in Einklang zu bringen. Wenn intentionale bzw. semantische Eigenschaften zwar keine fundamentalen Eigenschaften der Welt sein können, soll zumindest gezeigt werden, dass physische Systeme dennoch intentionale Zustände enthalten können (vgl. Fodor 1987; Dretske 1981). Hieraus resultieren verschiedene Bemühungen, Intentionalität zu naturalisieren, dass heißt zu zeigen, dass auch nicht-mentale Dinge intentional sein können, wie z. B. Dretskes informationstheoretischer Ansatz oder Millikans teleosemantischer Ansatz (vgl. z. B. Dretske 1988, 1994; Millikan 1984, 2000).

### 4.5.1 Intentionalität unter teleologischen Gesichtspunkten

Allgemein gilt für teleologische Theorien, dass der Inhalt durch die Funktion mentaler Repräsentationen bzw. der Mechanismen, die diese Repräsentationen verarbeiten, festgelegt werden soll. Dretskes informationstheoretischer Ansatz ist eine mögliche Variante, Intentionalität unter teleologischen Gesichtspunkten zu naturalisieren. In Abgrenzung zu repräsentationalen Theorien beschreibt er eine Indikatorrelation. Die Indikatorbeziehung ist eindeutig und hat im Unterschied zur Repräsentation eine hundertprozentige Wahrscheinlichkeit, weshalb sie keine Fehlrepräsentation zulässt. Hierin liegt ein offensichtlicher Vorteil der Indikatorrelation. Falsche Wahrnehmungen, beispielsweise, dass jemand in der Dämmerung einen Hund für eine Katze hält, sind nun insofern möglich, als die entsprechende Wahrnehmungsrepräsentation ihre Funktion nicht hinreichend erfüllt hat. Die der Wahrnehmungsrepräsentation untergeordnete Indikator-Funktion ist jedoch fehlerfrei.

> The fundamental idea is that a system, S, represents a property, F, if and only if S has the function of indicating (providing information about) the F of a certain domain of objects. The way S performs its function (when it performs it) is by occupying different states $s_1$, $s_2$, … $s_n$ corresponding to the different determinate values $f_1$, $f_2$ … $f_n$, of F. (Dretske 1995, 2)

Im Kern besagt der Ansatz, dass ein Signal S genau dann Informationen über die Instanziierung der Eigenschaft F enthält, wenn das G-sein von S mit der Instanziierung von F nomologisch korreliert. Wenn S nicht G ist, da die Eigenschaft F nicht instanziiert wurde, dann birgt die Mög-

[120] Quine verdeutlicht dies in einem Gedankenexperiment am Beispiel eines „radikalen Übersetzers", der die Sprache und Überzeugungen eines fremden Stammes erforscht. Quine nimmt an, dass verschiedene, miteinander unvereinbare Übersetzungssysteme möglich sind, die insofern alle richtig sind, als sie mit dem Verhalten der Eingeborenen übereinstimmen. Aus den verschiedenen Übersetzungen folgen verschiedene Inhalte und Überzeugungen. Dieser Umstand verstößt jedoch gegen Quines Existenzkriterium, wonach die Individuation von Dingen nur eindeutige Ergebnisse umfassen darf (vgl. Quine 1960, Kap. 7). In der Konsequenz lehnt Quine die Existenz intentionaler Zustände ab.

lichkeit von S, auch G zu sein, Informationen über F, bzw. S ist ein Indikator für F. Anders formuliert indiziert ein Zustandstyp S einen weiteren Zustandstypen F dann, und nur dann, wenn ein G vorhanden ist, auch ein F vorhanden ist. Üblicherweise wird diese Indikatorenrelation auf eine Kausalrelation zurückgeführt, wonach G die Wirkung von F ist. Dies muss aber nicht sein, es können G und F auch eine gemeinsame Ursache haben. Ferner muss die Relation nicht nomologisch sein. Dretskes Beispiel hierfür ist die Türklingel: solange immer jemand hinter der Tür steht, wenn die Türklingel läutet, ist dieses ein Indikator dafür, dass jemand hinter der Tür steht. Sobald jedoch aufgrund eines zufälligen Ereignisses die Türklingel läutet, ohne dass jemand hinter der Tür steht, ist die Indikatorrelation aufgehoben (vgl. Dretske 1988). Es soll ein bestimmtes Abhängigkeitsverhältnis zwischen dem Sachverhalt und dem Indikator gewährleistet sein, wonach S ein Indikator für den Sachverhalt F ist, wenn S nur dann existiert, wenn F existiert. F kann auch ohne S existieren, aber nicht umgekehrt. Insofern ist eine Indikationsbeziehung eine asymmetrische, aber sehr strikte Relation, da S immer nur ein Indikator für den Sachverhalt F sein soll.

Ein Fingerabdruck trägt Informationen über einen bestimmten Menschen, die Tankanzeige in einem Auto Informationen über die verbleibende Menge an Benzin im Tank, eine Kompassnadel trägt Informationen über die Position des magnetischen Nordpols. In all diesen Beispielen korreliert eine Eigenschaft eines physischen Gegenstands mit einer physischen Eigenschaft, die in ihrer Umgebung instanziiert wurde. Im Unterschied zu der Korrelation zwischen der Tankanzeige und der verbleibenden Menge an Benzin gibt es bei einer Korrelation zwischen Eisbären, die am Nordpol leben, und einer Kompassnadel keine gesetzliche Relation: Die Kompassnadel zeigt nach Norden, da dort der entsprechende Pol ist, nicht weil dort Eisbären leben. Die beiden Informationen verfügen nur eingeschränkt über eine intensionale Schnittmenge, die Relation ist nicht nomologisch. Diese sehr strikte Bestimmung der Indikation ist allerdings insofern problematisch, als eine Indikationsrelation zu hundert Prozent zutreffen soll. Dies schließt auch geringere Wahrscheinlichkeiten aus und erscheint somit in ihrer Entwicklung unrealistisch. Es ist fraglich, ob die Entwicklungssituation selbst, innerhalb der eine solche Relation entstehen soll, den geforderten Ansprüchen der Indikation genügen würde. Um dem Vorwurf zu entgehen, dass sein Ansatz aufgrund seiner ätiologischen Bezüge lediglich für vorhandene Funktionen, nicht aber für zukünftige, neue Anforderungen gelten kann, hat Dretske seinen Ansatz um die Funktion von Lernprozessen erweitert. Lebewesen sind u. a. durch Konditionierung in der Lage, unterschiedliche neue epistemische Wege zu entwickeln. Damit versucht er, insbesondere auf das „functional indeterminacy problem“ zu reagieren, wonach entfernte bzw. komplexe Inhalte aufgrund disjunkter Eigenschaftspfade für Indikatorrelationen problematisch, da nicht eindeutig bestimmbar sind (s. u.).

Ausgangspunkt für Millikans teleosemantische Theorie ist Brentanos Überlegung der Nichtexistenz intentionaler Objekte. So wird ihrer Ansicht nach auch in der Biologie ein Bezug zu Nichtexistentem hergestellt. Intentionalität wird als Verknüpfung einer biologischen Norm mit der Funktionsweise eines Organs aufgefasst. Diese Norm erklärt sich allein aus der evolutionären Entwicklung eines Organs, die zu der entsprechenden Funktion geführt hat. Es ist die biologische Funktion des Herzens, Blut zu pumpen, auch wenn es dazu temporär nicht in der Lage ist. Eine solche Funktion ist notwendigerweise in einem evolutionären Prozess entstanden, weshalb sie intentional auf die Umwelt außerhalb des Organs bzw. Objekts selbst verweist (vgl. Millikan 1984).[121] Aufgrund dieser biologischen Funktionsbeschreibung gilt Intentionalität nicht mehr allein für den geistigen Bereich, sondern auch naturalisiert für im weiteren Sinne physische Bereiche – es können auch nicht-mentale Dinge intentional sein.

Millikan verfolgt einen konsumenten-orientierten Ansatz hinsichtlich mentaler Repräsentationen, was bedeutet, dass Systeme Repräsentationen zur Umsetzung ihrer geeigneten Funktionen nutzen. Ein grundsätzliches Argument für diesen konsumenten-orientierten Ansatz ist, dass teleologische Funktionen durch evolutionäre Selektion erreichte Wirkungen sind, weshalb ein repräsentationaler Ansatz sich eben auch auf die Wirkung bzw. die Funktion von Repräsentationen konzentrieren sollte: entscheidend ist die Verarbeitung, nicht die Produktion von Repräsentationen (vgl. Millikan 1989b, 85). Außerdem ist nach Millikan für eine Repräsentation entscheidend, dass sie als solche genutzt wird (vgl. Millikan 1989a, 84-90). Eine Repräsentation wird sogar durch ihre Nutzung maßgeblich bestimmt. Der Inhalt der Repräsentationen ist zwingend mit der Ausführung der entsprechenden Funktion bzw. Körperbewegung des Konsumenten verbunden. Ein mögliches Beispiel hierfür ist der Bienentanz, der innerhalb einer Bienenpopulation aufgrund eines bestimmten Mechanismus Angaben über Richtung und Entfernung einer möglichen Nektarquelle liefert und die „interpretierenden Bienen" zum Hinflug veranlasst. Im Unterschied zum Bienentanz, bei dem zwischen Inhalt produzierenden und konsumierenden Bienen unterschieden werden kann, ist eine innere Repräsentation bei kognitiven Systemen vollständig enthalten. Aber auch hier wird die entschlüsselte Repräsentation in eine geeignete Funktion, ihre Nutzung, übertragen.

Um den Inhalt einer Repräsentation näher zu bestimmen, muss die Funktion der Konsumenten betrachtet werden. Diese haben ihre Funktion, da in der Vergangenheit Token dieses Typs die Erhaltung und das Überleben dieses Systems gesichert haben. Um den Inhalt zu bestimmen, müssen die Ereignisse betrachtet werden, die zur Entwicklung der entsprechenden Funktion bei den Konsumenten unter dem Gesichtspunkt der Systemerhaltung führte. Welche

[121] Eine Fehlrepräsentation eines intentionalen Zustands wäre mit einer Fehlfunktion gleichzusetzen. Da Millikan allerdings hauptsächlich über evolutionär erfolgreiche Funktionen argumentiert, kann es solche falschen Repräsentationen innerhalb ihres Ansatzes nicht geben.

Übereinstimmung zwischen der Repräsentation und der Umwelt war hierfür vonnöten? Der Bienentanz ist eine Re-präsentation, die einer Bienenpopulation auch über größere Distanzen die Versorgung mit Nektar ermöglicht. Das grundlegende Ereignis für die Entwicklung einer solchen Repräsentation nennt Millikan die „normale Kondition", die mit dem Inhalt der Repräsentation gleichzusetzen ist.

Allerdings scheint es problematisch, diese „normale" Kondition für eine Re-präsentation zu isolieren und zu bestimmen. Sie scheint in der Kritik zu spezifisch und damit als Inhalt nicht plausibel. So muss eine giftige, ungenießbare Fliege für den Frosch von einer Fliege als „Froschfutter" unterschieden werden. Außerdem wird kritisiert, dass allgegenwärtige förderliche Hintergrundbedingungen nicht ausreichend gewürdigt bzw. aus der Theorie ausgeschlossen werden. Als Konsequenz müssten sie für jeden Beitrag der Anpassung bzw. für jede Entwicklung einer Repräsentation miteinbezogen werden. Eine Repräsentation bedeutet nicht nur „Froschfutter", sondern „Froschfutter unter der Bedingung von Sauerstoff, Schwerkraft etc.". Ein weiterer Einwand betrifft die Verlagerung auf den Interpreten einer Repräsentation. Es stellt sich die Frage, ob die dargestellte Ebene der Repräsentation auch bei vielen Konsumenten funktioniert bzw. zu einem eindeutigen Inhalt führt oder im Gegenteil zu mehr Ambiguität führt als gewünscht. Ein grundsätzlicher Einwand ist schließlich, dass Millikan evolutionäre und intentionale Erklärungen unzutreffend miteinander verknüpft, obwohl evolutionäre Entwicklungen nicht zwangsläufig zu den entsprechenden intentionalen Zuständen führen. Es ist problematisch, dass Repräsentationen bzw. intentionale Zustände und deren Inhalte im Schwerpunkt über ein „Belohnungssystem" erzeugt werden. Jemand kann Sex wollen, ohne an Bindung oder Babys interessiert zu sein – wobei letztere den Erhalt der Menschheit fördern.

Generell lässt sich gegen teleologische Theorien einwenden, dass die Funktionen zu unbestimmt sind, um Inhalt und Intentionalität zu bestimmen. Die funktionale Unbestimmtheit ergibt sich Fodor zufolge aus der Tatsache, dass natürliche Selektion extensional ist: wenn es für einen Organismus M möglich ist, die Handlung O während einer bestimmten Umweltsituation mit dem Merkmal F zu machen, und F verhält sich koextensiv zu dem Merkmal G, dann kann die Handlung O genauso gut von dem Organismus M in Anwesenheit von G durchgeführt werden. Teleologische Theorien können jedoch nicht zwischen koextensiven Inhaltskandidaten innerhalb der Umwelt unterscheiden, in der ein Lebewesen sich entwickelt. So schnappt ein Frosch sowohl nach Fliegen als auch nach kleinen, dunklen, sich bewegenden Teilchen – der Inhalt und die intentionale Relation sind allein aufgrund der Funktion nicht eindeutig zu bestimmen (vgl. Fodor 1990b).

Ein weiteres Problem ist, dass organische Systeme komplexe kausale Selektionsprozesse durchlaufen. Beispielsweise wurde ein typisches Merkmal einer Antilope selektiert, weil a) sich die

Form eines Hämoglobins verändert hat, was b) zu einer schnelleren Sauerstoffversorgung geführt hat, was c) der Antilope ermöglicht, höhere Ebenen zu ersteigen, wo sie d) im Sommer besseres Weideland erreicht und somit e) bessere Überlebenschancen hat (vgl. Neander 1995). Nach einer ätiologischen Theorie kann die Frage, welche vergangenen Schritte zu jener Eigenschaft führten, mit allen Schritten von a) bis e) beantwortet werden. Alle Optionen sind zutreffende Funktionsbeschreibungen, eine eindeutige intentionale Relation ist epistemisch nicht bestimmbar. In Verbindung mit den alternativen Funktionsentwicklungen des Frosches können nun auch zwei legitime disjunkte Entwicklungspfade für repräsentationale Funktionen genannt werden. Alle funktionalen biologischen Eigenschaften haben das Ziel der Überlebenssicherung, aber es gibt spezifischere Dinge, aufgrund derer sie selektiert wurden. Einer Kaskade von Eigenschaften entspricht eine Vielzahl spezifischer Funktionsbeschreibungen.

Der schwerwiegendste Einwand gegen teleologische Theorien ist schließlich, dass nicht einsichtig ist, wie sie intellektuell anspruchsvolle Begriffe und Vorstellungen erklären können. Keine naturalistische Theorie des Inhalts kann derzeit darstellen, wie wir z. B. über Demokratie, Ethik, Quarks oder zukünftige Lebenswelten nachdenken können. Teleologische Ansätze versagen hier (vgl. Peacocke 1992). Sie beziehen sich im Schwerpunkt auf Inhalte, die die evolutionäre Tauglichkeit und Überlebenschancen von Systemen betreffen. Inhalte, die nicht direkt mit evolutionären Ansprüchen in Verbindung gebracht werden können, werden somit systembedingt ausgeschlossen. Zukünftige Inhalte sind zu vage, Quarks möglicherweise zu unspezifisch für direkte evolutionäre Ansprüche. Dennoch sind solche Gedanken ständig in unserer Umwelt gegenwärtig. Es ist fraglich, ob sich diese vielfältigen, verschiedenen intellektuell anspruchsvollen Begriffe auf einfache Inhalte zurückführen lassen.

### 4.5.2 Intentionalität und mentale Kausalität

Auch Searle kritisiert die teleologische Position, wonach Intentionalität lediglich als Spezialfall der evolutionsbiologischen Selektion von Funktionen betrachtet wird. Er sieht einen grundsätzlichen Unterschied zwischen der Funktion, Blut zu pumpen und dem Wunsch, ein Glas Wasser zu trinken. Eine evolutionäre Funktion ist möglicherweise eine notwendige, aber keine hinreichende Bedingung für semantische Gerichtetheit bzw. Intentionalität. Searle hält dagegen, dass die den intentionalen Begriffen innewohnende Normativität ein Symptom dafür ist, dass Intentionalität nicht als teleologische Funktion zu naturalisieren ist. Intentionale Begriffe setzen vielmehr Standards der Wahrheit, Rationalität, Widerspruchsfreiheit etc. Solche Standards können unmöglich einem System zu Eigen sein, das aus nackten, blinden, nichtintentionalen Kausalbeziehungen

besteht – bei der „Billardkugel-Verursachung“ gibt es keinen normativen Bestandteil (vgl. Searle 1993, 69f.).[122]

Die kausale Frage für Intentionalität stellt sich im Zusammenhang mit der Frage, ob externale Zustände kausal wirksam, oder nur lokale Prozesse kausale Prozesse sein können (vgl. Kap. 4.3.2). Warum lediglich die inneren Eigenschaften einer Ursache in einem Prozess kausal wirksam sind und nicht auch ihre äußeren Eigenschaften, kann am Wert einer Euromünze dargestellt werden. Neben ihren inneren, physischen Eigenschaften, wie der Form und der Materialbeschaffenheit, wird der Wert der Euromünze durch ihre zeitliche, äußere Beziehung u. a. zur Europäischen Zentralbank bestimmt. Wirft man jedoch eine Euromünze in einen Kaffeeautomaten, sind es nur die physischen Eigenschaften der Münze, die bewirken, dass tatsächlich eine Tasse Kaffee aus dem Automaten ausgeworfen wird. Die äußeren Eigenschaften scheinen der Argumentation zufolge keine relevante kausale Wirkung in diesem Prozess zu erzielen, sie sind lediglich epiphänomenal. Analog wird geschlossen, dass auch die äußeren bzw. intentionalen Eigenschaften mentaler Zustände kausal wirkungslos, epiphänomenal sind. Eine epiphänomenalistische Perspektive auf das Phänomen der Intentionalität vertritt Stich (1983). Er betont unter Verweis auf die Language of Thought-Hypothese, dass ausschließlich die syntaktischen Eigenschaften, nicht die semantischen Eigenschaften der mentalen Symbole kausal wirksam sind. Eine übliche Erwiderung gegen diese Schlussfolgerung ist, darauf hinzuweisen, dass in den Einzelwissenschaften dennoch viele erfolgreiche Erklärungsmodelle entwickelt werden. Deshalb muss die Annahme der Epiphänomenalität äußerer Eigenschaften nicht akzeptiert werden (vgl. Baker 1993; Burge 1986, 1993).

In einer direkten Erwiderung auf das Münzbeispiel kann darauf verwiesen werden, dass auch der Kaffeeautomat äußere, zeitliche Eigenschaften hat, die den Wert von Euromünzen betreffen und unter anderem in Beziehung zur Europäischen Zentralbank stehen. Schließlich wurde das Münzsystem des Automaten zu einem bestimmten Zeitpunkt, der sich aus den äußeren Bedingungen und einem aktuellen Wert der Münzen ergibt, eingerichtet. Erst aufgrund einer äußeren Konvention werden die physischen Eigenschaften des Automaten ausgerichtet, und funktioniert das physische Zusammenspiel von Münze und Automat. In eine ähnliche Richtung argumentiert auch Dretske mit seiner Unterscheidung zwischen Körperbewegung und Verhalten. Propositionale Einstellungen liefern ihm zufolge zwar keine direkte Erklärung für eine einzelne körperliche Bewegung, spielen aber als zeitlich vorhergehende Eigenschaften, die solche Bewegungsabläufe mitbestimmen, eine Rolle für die kausale Erklärung (vgl. Dretske 1988). Danach wird Verhalten nicht mit Körperbewegung identifiziert, sondern umfasst letztere. Verhalten

[122] Darwins bedeutendste Leistung besteht für Searle gerade darin, dass er Zweck und Teleologie aus der Evolution entfernt und an ihre Stelle natürliche Formen der Auslese eingesetzt hat. Selbst Begriffe wie „biologische Funktion“ sind immer relativ zu einem Beobachter, der den Kausalvorgängen einen normativen Wert zuordnet (vgl. Searle 1993, 69f.).

ist der Prozess, innerhalb dessen die Körperbewegung einer Person durch einen inneren, gegebenenfalls intentionalen Zustand verursacht wird. Dretske argumentiert, dass intentionale Wünsche oder Glaubensannahmen das Ergebnis ontogenetischer, nicht-intentionaler selektiver Lernprozesse sind. Solche mentalen Zustände ermöglichen Anpassung, da sie über wertvolle Informationen über die Umwelt des Individuums verfügen. Insofern können die propositionalen Einstellungen einer Person aufgrund ihrer Intentionalität als Ursachen für ihre Körperbewegung betrachtet werden (vgl. Dretske 1995). Intentionalität ist insofern kausal relevant, als sie bestimmt, warum verschiedene Bewegungstypen regelmäßig von verschiedenen inneren Zuständen verursacht werden.[123]

Intentionale Zustände unterliegen grundsätzlich normativen Kontexten und somit Wahrheitsbedingungen. Sie beziehen sich, zumindest im Standardfall, auf Dinge oder Personen, die tatsächlich existieren und können deshalb überprüft werden. Wenn beispielsweise Peter glaubt, dass in Deutschland eine Bundeskanzlerin regiert, die Angela Merkel heißt und zugleich Vorsitzende der CDU ist, sind dies verschiedene einzelne Glaubensannahmen, die auch untereinander konsistent und wahr sein müssen. Peters Glaubensannahmen stehen in einer konkreten Relation zu anderen Personen in Raum und Zeit. Andererseits sind seine Glaubensannahmen auch einfach innere Zustände – wie ist es aber möglich, dass etwas Extramentales wesentlich für die inneren Zustände einer Person ist?

Die Kausaleigenschaften intrinsischer Intentionalität verbinden für Searle das Kausale mit dem Normativen. Intentionale Phänomene wie das Regelfolgen und das überzeugungs- und wunschgeleitete Handeln sind echte Kausalphänomene; als intentionale Phänomene stehen sie aber in wesentlicher Beziehung zu normativen Phänomenen wie Wahrheit und Falschheit, Erfolg und Misslingen, Widerspruchsfreiheit und Widerspruch, Rationalität, Täuschung und Erfüllungsbedingungen im allgemeinen. Kurz, die wirklichen Tatsachen der Intentionalität enthalten normative Bestandteile; bei funktionalen Erklärungen gibt es hingegen keine Tatsachen, außer den nackten physischen Fakten; Normen gibt es nur in uns und nur von unserem Standpunkt aus (vgl. Searle 1993, 262).

Intentionale Verursachung wird benötigt, um die Erklärung menschlichen Verhaltens zu verstehen. Rationales menschliches Verhalten beruht auf Gründen, aber diese Gründe erklären

[123] Wenn man diese Position mit der Ansicht Dretskes zusammenführt, dass auch der Münzautomat eine Geschichte hat bzw. von einem Programm abhängt, lässt sich feststellen, dass die Funktionsweise physischer Eigenschaften offenbar von ihrem jeweiligen Algorithmus abhängt – eine Position, die dem Computerfunktionalismus nahe kommt (vgl. Kap. 2.2). Es stellt sich allerdings die Frage, warum die Syntax in dem Antagonismus „mental versus physisch" in der Regel unproblematisch mit dem physischen Bereich verbunden, sogar als teilhaftig dargestellt wird. Müssten nicht die syntaktischen Eigenschaften als abstrakte Eigenschaften stärker von physischen Eigenschaften abgegrenzt werden? Sind sie im Zusammenhang mit Intentionalität nicht sogar eher Teil der Intentionalität, als intentionales Programm? Schließlich scheint die Syntax eine abstrakte Zielvorgabe für die physischen Eigenschaften vorzugeben.

das Verhalten nur, wenn die Beziehung zwischen dem Grund und dem Verhalten eine logische und eine kausale ist. „In Erklärungen rationalen menschlichen Verhaltens wird somit wesentlich vom Werkzeug der intentionalen Verursachung Gebrauch gemacht." (Searle 2001, 129) Intentionale Verhaltenserklärungen implizieren ihrer Form nach jedoch nicht, dass eine Handlung eintreten muss. Intentionale Ursachen sind nicht hinreichend, um eine Handlung zu determinieren, so dass zwischen den Ursachen einer Entscheidung und der Entscheidung selbst sowie zwischen der Entscheidung und dem Vollzug einer Handlung jeweils eine Lücke besteht. Somit enthalten Absichten keine kausal hinreichenden Gründe für die aus ihnen folgende Handlung.

Typischerweise funktioniert ein Bewusstseinszustand dadurch, dass er ein Ereignis der Art, das er verursacht, repräsentiert. Die Handlung bzw. Wirkung, Wasser zu trinken, wird von dem Wunsch bzw. der Ursache, Wasser zu trinken, repräsentiert. Diese Fähigkeit, Gegenstände und Sachverhalte in der Welt zu repräsentieren und aufgrund dieser Repräsentation zu handeln, bezeichnet Searle als eines der wichtigsten Merkmale von Bewusstsein:

> Intentionalität ist das Merkmal des Geistes, durch das Geisteszustände auf Sachverhalte in der Welt gerichtet sind, von ihnen handeln, sich auf sie beziehen oder auf sie abzielen. (ebd., 82)

Das Ideal einer Theorie des intentionalen Gehalts wäre es, wenn sie einzig und allein mit Rückgriff auf zweierlei Kausalbeziehungen formuliert wäre: einerseits Kausalbeziehungen zwischen Menschen untereinander, andererseits Kausalbeziehungen zwischen Menschen und Gegenständen (bzw. Sachverhalten) in der Welt.

Die Struktur intentionaler Zustände unterteilt Searle in die Unterscheidung zwischen dem Typ und dem Gehalt, in die Ausrichtung sowie in die Erfüllungsbedingungen intentionaler Zustände. Alle drei Strukturmerkmale sind miteinander verknüpft. Für jeden intentionalen Zustand muss zwischen dem Gehalt des Zustands und unterschiedlichen intentionalen Modi unterschieden werden. Die Unterscheidung zwischen Gehalt und Modus gilt für alle intentionalen Zustände, wie Überzeugungen, Wünsche, visuelle Wahrnehmungen oder Absichten. Man kann hoffen, glauben oder befürchten, dass es regnen wird. Die unterschiedlichen Zustandstypen haben denselben Gehalt. Bei dem Gehalt handelt es sich um eine Proposition – z. B. „es wird regnen" – mit Wahrheits- bzw. Erfüllungsbedingungen. Searle bevorzugt den Terminus „Erfüllungsbedingungen", da aufgrund der unterschiedlichen Zustandstypen intentionale Zustände nicht nur wahr oder falsch, sondern auch (nicht-)erfüllt oder (nicht-)realisiert sein können. Wünsche können in einem strengen Sinn nicht wahr werden (vgl. ebd., 121). Der Begriff der Erfüllungsbedingung ist zentral für das Verständnis von Intentionalität, da er letztlich Auskunft über den Zustandstyp und den Umstand des intentionalen Zustands gibt: „Ein intentionaler Zustand ist erfüllt, wenn die Welt so ist, wie der intentionale Zustand sie repräsentiert." (ebd., 125)

Searle unterscheidet weiter zwischen unterschiedlichen Arten, zu denen intentionale Gehalte durch unterschiedliche Zustandstypen in Beziehung stehen. Es kann zwischen zwei Ausrichtungsarten unterschieden werden: Überzeugungen, Wahrnehmungen und Erinnerungen haben die Geist-auf-Welt-Ausrichtung; ihr Ziel ist es, die Dinge zu repräsentieren, wie sie sind. Wünsche und Absichten haben die Welt-auf-Geist-Ausrichtung; ihr Ziel ist es, die Dinge zu repräsentieren, wie sie gewünscht und beabsichtigt werden. Überzeugungen und Wahrnehmungen heißen in Abhängigkeit von der wirklichen Welt wahr oder falsch, während Wünsche und Absichten auch unabhängig von der Welt, also unabhängig davon, ob der intentionale Zustand umgesetzt werden konnte, bestehen (vgl. ebd., 122):

> Wenn wir Gegenstände sehen, verursachen die Gegenstände, die wir sehen, unsere visuellen Erlebnisse von ihnen. Wenn wir uns an Ereignisse erinnern, die wir erlebt haben, dann verursachen diese vergangenen Erlebnisse unsere gegenwärtigen Erinnerungen. Wenn wir die Absicht haben, unsern Körper zu bewegen, dann verursachen diese Absichten die Körperbewegungen. (ebd., 127)

Die beschriebenen Fälle vereinen jeweils eine kausale und eine intentionale Komponente. Die Form dieser systematischen Verknüpfung von Repräsentationsvermögen des Geistes und Kausalbeziehung zur Welt ist die intentionale Verursachung. Searle ist der Ansicht, dass sie evolutionäre Qualität besitzt und das menschliche Überleben sichert.

Die intentionale Verursachung unterscheidet sich grundsätzlich von der mechanistischen Wirkkausalität. Ursache und Wirkung funktionieren in wechselseitiger Ergänzung. Die Ursache ist die Repräsentation der Wirkung, oder aber die Wirkung ist die Repräsentation der Ursache. Die Erfüllungsbedingung des Wunsches, Wasser zu trinken, wird von diesem bewirkt und repräsentiert. Wenn ich meinen Durst durch Wassertrinken stille, verursacht mein Wunsch, Wasser zu trinken als Geisteszustand, dass es der Fall ist, dass ich Wasser trinke.

Die Absicht, den Arm zu heben, ist dagegen erst erfüllt, wenn die Absicht selbst in dem Moment das Armheben bewirkt, dass heißt, sie selbst die Erfüllungsbedingungen bewirkt. Deshalb nennt Searle Absichten kausal selbstbezüglich. Kausale Selbstbezüglichkeit liegt bei „volitiven" Zuständen, wie Absichten, vor, sowie bei „kognitiven" Zuständen, wie Wahrnehmung und Erinnerung. Der volitive Zustand einer Absicht ist nur dann erfüllt, wenn unter der Absicht, den Arm zu heben, der Arm gehoben wird, und die Absicht, den Arm zu heben, die Ursache des Armhebens ist. Der kognitive Zustand der Wahrnehmung ist nur dann erfüllt, wenn z. B. der Umstand, dass in Sichtweite ein Baum steht, das visuelle Erlebnis selbst verursacht, das die Erfüllungsbedingungen hat, dass in Sichtweite ein Baum steht. Bei kognitiven Zuständen mit kausaler Selbstbezüglichkeit ist zwar die Ausrichtung vom Typ Geist-auf-Welt, aber die Kausalität verläuft von der Welt auf den Geist. Bei volitiven Zuständen mit kausaler Selbstbezüglichkeit dagegen ist

die Ausrichtung vom Typ Welt-auf-Geist, aber die Kausalität verläuft von dem Geist auf die Welt. Eine Absicht passt zur Welt nur dann, wenn die Absicht selbst das Ereignis des Armhebens verursacht (vgl. ebd., 128).

Intentionale Zustände wirken dabei nicht isoliert, sondern in Verbindung mit anderen intentionalen Zuständen. Verschiedene Überzeugungen sind Bestandteil einer Überzeugung, wie die Grundlage des Wunsches, ein Eis zu essen, die Überzeugung ist, dass es Eis gibt, das Eis gegessen werden kann etc. . Neben der Menge an intentionalen Zuständen sind weitere Fähigkeiten und Voraussetzungen notwendig, um mit der Welt zurechtzukommen. „Diese Menge von Fähigkeiten, Fertigkeiten, Neigungen, Gewohnheiten, Dispositionen als selbstverständlich unterstellten Voraussetzungen und allgemein „Know-how" ist es" (ebd., 130f.), was Searle als „Hintergrund" bezeichnet. Nur vor diesem Hintergrund an Fähigkeiten legen intentionale Zustände ihre Erfüllungsbedingungen fest und erfüllen ihre Rolle. Jede Absicht setzt einen enormen metaphysischen Apparat voraus. Unterhalb der bewussten Gedanken an der Oberfläche gibt es einen riesigen Apparat, der grundlegend für die Wünsche und Überzeugungen ist (vgl. ebd., 131f.).

***Quintessenz Kriterium Intentionalität***

- *mentales Vermögen zu repräsentieren bzw. über etwas zu sein – auf ein intentionales Objekt gerichtet*
- *im Geist einer Person, nicht in Raum und Zeit*
- *echtes Kausalphänomen: z. B. wertvolle Informationen über Umwelt des Individuums*
- *Welt wegen Bewusstsein repräsentiert (Gehalt, Modi, Erfüllungsbedingungen)*
- *Richtungen: Geist-auf-Welt/Welt-auf-Geist*
- *Keine Isolierung: in Verbindung mit anderen intentionalen Zuständen/Hintergrund an Fähigkeiten*

- *Dretskes Indikatorrelation: mit 100-prozentiger Wahrscheinlichkeit auf Kausalreduktion zurückführbar; Problem: Ätiologie; Funktion von Lernprozessen*
- *Millikans teleosemantischer Ansatz: Begründung über biologische Normen, Funktionsweise der Organe und Evolutionsprozesse; Fokus auf Nutzung der Repräsentation; Problem: „normale Kondition" lässt Hintergrundbedingungen außen vor; evolutionäre-intentionale Verknüpfung unzutreffend; intellektuell anspruchsvolle Begriffe und Vorstellungen nicht erklärbar*
- *„Billiardkugelverursachung" deckt normativen Bereich nicht ab: Verhalten und Körperbewegung als ein (auf Gründen beruhender) Prozess; Verhalten umfasst Körperbewegung*

### 4.6 Das Kriterium der Proportionalität

Es ist wichtig für unser Selbstverständnis, in die kausale Wirksamkeit unserer Glaubensannahmen und Wünsche, Emotionen und unseres Bewusstsein zu vertrauen. Wir verstehen uns als Handelnde in der Welt, ohne Mentalität auf Muster neuronaler bzw. elektrochemischer Aktivität zu reduzieren. Kim kritisiert jedoch die Nicht-Reduktivisten, dass dies lediglich eine Wunschliste sei und auch erst der Startpunkt der Debatte um mentale Kausalität (vgl. Kim 2003, 172).

Das Beharren, dass mentale Ereignisse oder Eigenschaften kausal wirksam sein müssen, wird tatsächlich stark von dieser Intention motiviert. Dahinter steckt allerdings auch der Anspruch an mentale Kausalität, dass die Relation zwischen Ursache und Wirkung adäquat beschrieben und erfasst wird. Burge bemerkt, dass es nicht entschieden ist, dass ein physikalisches Modell ein passendes Modell für mentale Kausalität ist. Man kann verschiedene Wege spezifizieren, wonach mentale Ursachen eine Veränderung bewirken, ohne in Konflikt mit physikalischen Erklärungen zu geraten. Dieser Unterschied, den psychologische Ursachen ausmachen, benötigt ihm zufolge weder Lücken in der physischen Kausalkette, noch konkurrieren mentale und physikalische Erklärungen miteinander. Es wäre unsinnig, einen Erklärungsausschluss anzunehmen, da beide kausalen Erklärungen dieselbe physische Wirkung als Ergebnis zweier unterschiedlicher Erklärungsmuster beschreiben.

> Neither type of explanations makes essential, specific assumptions about the other. So the relation between the entities appealed to in the different explanations cannot be read off the causal implications of either or both types of explanation. (Burge 1999, 116)

Psychologische und physikalische Erklärungen konkurrieren demnach nicht miteinander und schließen sich auch nicht wechselseitig aus. Sie sind vielmehr gleichberechtigt.[124] Burge schließt von der epistemischen auf die ontologische Ebene und sieht die Zulässigkeit psychologischer Erklärungen letztlich auf beiden Ebenen gegeben. Er ist sogar optimistisch, dass, solange mentalistische Erklärungen Wissen und Verständnis liefern und diese Erklärungen kausal sind, auch mentale Kausalität als Teil der Welt betrachtet werden kann (vgl. ebd., 116f.).

Implizit wird die Forderung nach einer adäquaten bzw. proportionalen Kausalität in dem Generalisierungsargument formuliert, wonach nicht nur die Psychologie, sondern sämtliche Einzelwissenschaften durch das Ausschlussargument von kausaler bzw. explanatorischer Irrelevanz betroffen wären (vgl. Kap. 3.4). Auch Davidsons beschreibungsunabhängige Ereignisontologie kann als Ansatz verstanden werden, der zumindest für eine proportionale Beschreibung der kau-

[124] Eine ähnliche Einschätzung vertritt auch Davidson, für den psychologische Erklärungen durchaus zulässig sind. Allerdings kann für ihn aufgrund des Anomalismus des Mentalen kein Einklang von physikalischer und psychologischer Erklärung bestehen, Widersprüche sind denkbar. Schließlich ist auf der ontologischen Ebene die „echte" Kausalität nur eine Ereigniskausalität (vgl. Kap. 2.3).

salen Relation argumentiert (vgl. Kap. 2.3). Auf ihn beziehen sich auch spätere Autoren, die gegen das kausale Ausschlussargument betonen, dass epistemische Mehrfachdetermination unproblematisch sei, und es auf der epistemischen Ebene sehr wohl mehrere gleichberechtigte Kausalmodelle geben könne, die nicht in direkter Konkurrenz zueinander stehen (vgl. z. B. Menzies 2003; Raymont 2003). So sind die Eigenschaften selbst nicht Teil des kausalen Prozesses, sondern lediglich beschreibend auf der Erklärungsebene relevant. Es sind nur die Träger der Eigenschaften, also die Ereignisse, die z. B. nach Davidsons Ereignisontologie kausal relevant sind. Solange aber nur Ereignisse kausal relevant und mentale wie physische Ereignisse (token-) identisch sind, kann es keine kausale Mehrfachdeterminierung geben (vgl. Raymont 2003, 234f.; vgl. auch Kap. 2.3; Kap. 3.1.1).

Nach Menzies ist es ein Missverständnis, dass das Konzept der Kausalität als eine kategorische, absolute Relation aufgefasst wird. Er argumentiert dagegen, dass Kausalrelationen als Entitäten aufgefasst werden, die über eine bestimmte funktionale Rolle in Bezug auf abstrakte Muster definiert werden sollten. Damit ist es möglich, mehrere kausale Ebenen zu unterscheiden, zum einen eine Ebene auf der mentale Zustände über distinkte psychologische Wege Verhalten verursachen, und zum anderen eine Ebene, auf der physische Gehirnzustände über distinkte neuronale Wege Verhalten verursachen. Menzies ist der Ansicht, dass diese verschiedenen kausalen Ebenen nicht in Konkurrenz zueinander stehen (vgl. Menzies 2003, 196). Problematisch ist bei solchen Ansätzen jedoch, dass sie sich vornehmlich auf die epistemische Ebene beziehen, während das kausale Ausschlussargument auf die ontologische Ebene abzielt. Um für die kausale Kraft mentaler Eigenschaften zu argumentieren, müsste deshalb gezeigt werden, wie mentale Kausalität ontologisch proportional zu bestimmen ist.

### 4.6.1 Das Prinzip der Proportionalität

Eine ontologisch stärkere Position bezüglich adäquater bzw. proportionaler mentaler Kausalität diskutieren Yablo (1992, 1997, 2003) und Schröder (2003, 2007). Das Prinzip der Proportionalität richtet sich zunächst gegen das Argument der kausalen Geschlossenheit der Physik, sowie gegen das Argument des kausalen Ausschlusses. Da nach dem Prinzip der kausalen Geschlossenheit der Physik jede physische Wirkung kausal von ihrer physischen Ursache erzielt wird, bleibt scheinbar keine kausale Arbeit mehr für mentale Ursachen übrig. Die supervenierenden mentalen Eigenschaften werden dabei durch ihre subvenienten physischen Basen gewissermaßen ausgeschlossen (vgl. Yablo 1997, 253). Nach Schröder wird dabei im Kern bezweifelt, dass mentale Eigenschaften als intentionale bzw. semantische Eigenschaften kausal wirksam sein können (vgl. Schröder 2003, 174).

Yablo verbindet in Abgrenzung zur Supervenienzrelation, zur Identitätsrelation und anderen reduktiven Ansätzen für seinen Ansatz der mentalen Kausalität das Prinzip der Eigenschaftsproportionalität mit der determinable-determinate-Relation.[125] In der logischen Tradition ist beispielsweise die Farbe rot nicht identisch mit einer ihrer spezifischen Farbschattierungen scharlach-, karmesin- oder purpurrot, die rot erfordern. Umgekehrt hängt rot von purpurrot ab, wobei purpurrot sein bereits rot sein beinhaltet. Purpur ist eine Variante von rot bzw. ein determinate von rot. Ebenso wie scharlach-, karmesin- oder purpurrot determinates von Rotsein sind, sollen vor dem Hintergrund der Multirealisierbarkeit auch die verschiedenen Gehirnzustände C $\{C_1, C_2, \ldots C_n\}$ determinates für Schmerz sein. Damit wird für Yablo eine Relation beschrieben, wonach Gehirnzustände plausibel mentale Zustände tragen, ohne dass zugleich eine kausale Konkurrenz entsteht. Beispielsweise picken Möwen auf rote Formen. Niemand würde das Rotsein eines Dreiecks als irrelevant für das Picken bezeichnen vor dem Hintergrund, dass die Wirkung tatsächlich durch eine spezielle Rotvariante erzielt wurde. Es kann festgehalten werden, dass determinates nicht ihren determinables zuvorkommen; das kausale Ausschlussargument greift nicht für diese Relation (vgl. Yablo 1997, 256f.). „Since this is a relation in which physical and mental states plausibly stand, my pain can […] be relevant to effects for which my brain state suffices." (ebd., 257)

Wenn nun mentale Eigenschaften determinables der physischen Eigenschaften sind, können nur sie proportional zu einer abstrakten Verhaltenseigenschaft sein, wohingegen die physischen Eigenschaften eines Gehirnzustands, die determinate sind, nicht proportional sind. Damit wäre die kausale Wirksamkeit mentaler Eigenschaften gesichert (vgl. Yablo 1992, 277f.). Dabei lehnt Yablo das Ausschlussargument auch deshalb ab, da mentale und physische Eigenschaften aufgrund der determinable-determinate-Relation nicht miteinander um kausalen Einfluss konkurrieren. „[R]ather than competing for causal honors, determinables and their determinates seem likelier to share in one another's success." (ebd., 272) Insofern mentale Eigenschaften physische Eigenschaften umfassen, besitzen sie beide gemeinsam die kausale Kraft, die sie teilen – so wie eine Enzyklopädie den Regalraum mit ihren Einzelbänden teilt.

Yablo räumt ein, dass aufgrund des Prinzips der Proportionalität Gehirnzustände zwar einen quantitativen Vorteil haben und durchaus hinreichend für verschiedene Wirkungen sind, allerdings sind sie hinsichtlich zusätzlicher Wirkungen eines Ereignisses o. ä. in einer schlechteren Ausgangslage. Seiner Auffassung nach kann ein Zustand X, selbst wenn er relevant und hinreichend für einen zweiten Zustand Y ist, diesen erst dann kausal verursachen, wenn er sich auch proportional zu ihm verhält. So kann X relevant für Y sein, auch wenn einige wichtige Faktoren

[125] Mögliche Übersetzung des Begriffs „determinate" wären „Bestimmter" oder „Bestimmender" und eine mögliche Übersetzung für „determinable" wäre „Bestimmbarer". In der fortlaufenden Arbeit wird allerdings an der englischsprachigen Terminologie festgehalten, da diese als Fachterminologie die spezifische Relation zwischen determinates und determinables sinnvoll beschreibt.

für das Erscheinen von Y fehlen, und so kann X hinreichend für Y sein, auch wenn X verschiedene unbedeutende Details beinhaltet (vgl. Yablo 1997, 257, 266). Es wird argumentiert, dass der Gehirnzustand, der überflüssige mikrophysische Details beinhaltet, im proportionalen Vergleich gegenüber dem mentalen Zustand sogar schlechter abschneidet. So würde ich auch mit einer anderen mikrophysischen Beschreibung klagen, wenn ich Schmerzen hätte, während die Frage, wie ich mich verhalten würde, wenn lediglich der Gehirnzustand ohne den Schmerzzustand erschiene, nicht erörtert werden kann, da der Gehirnzustand einen Schmerzzustand einschließt (vgl. ebd. 257f.) Unter dem Gesichtspunkt der Proportionalität scheinen physische Zustände mit unnötigen mikrostrukturellen Details überladen zu sein. Schließlich müssen streng genommen verschiedene mikrostrukturelle Alternativen berücksichtigt werden, die Schmerzen verursachen. Mentale Zustände dagegen vereinfachen bzw. abstrahieren die Relation, wodurch ihnen kausale Relevanz zukommt (vgl. ebd., 254). Das Prinzip der Proportionalität betrifft mentale Eigenschaften. Es sichert grundsätzlich, dass eine ursächliche Eigenschaft und eine Wirkung, die eine bestimmte Beschreibung erfüllt, in ihrer Kausalrelation zusammenpassen. Das Prinzip kann folgendermaßen formuliert werden:

> For any instance of causation the property of the cause which is to be causally relevant for the instantiation of a given property of the effect must be proportional to this property of the effect. (vgl. Schröder 2003, 178)

Eine Eigenschaft verhält sich dann proportional zu einer anderen, wenn sie denselben Grad an Abstraktion erfüllt. Wenn die Eigenschaft B einen hohen Abstraktionsgrad hat, muss Eigenschaft A ebenfalls einen hohen Abstraktionsgrad haben, um sich proportional zu verhalten. Ist der Abstraktionsgrad von B dagegen gering, sollte er auch für A gering sein. Das Zerbrechen einer Fensterscheibe kann ohne Details, wie z. B. das Einschlagen eines konkreten Gegenstands, erklärt werden. Der Schwung bzw. Impuls, mit dem ein Gegenstand einschlägt, wäre die einzige Eigenschaft, die zur kausalen Erklärung herangezogen würde. Sollte dagegen erklärt werden, auf welche bestimmte Art und Weise das Fenster zerbrach, ist der Impuls alleine nicht hinreichend, diese weitere Eigenschaftsinstanziierung zu verursachen, wohingegen Form und Oberfläche des einschlagenden Gegenstands erklären würden, dass die Fensterscheibe auf eine bestimmte Art und Weise zerbrach. Grundsätzlich gilt also, je detaillierter die Eigenschaft ist, die erklärt werden soll, desto detaillierter müssen auch die erklärenden Eigenschaften sein (vgl. ebd.).

Das Prinzip der Eigenschaftsproportionalität wirkt dabei in zwei Richtungen. Einerseits zwingt es Erklärungen, die notwendigen Details zu liefern, die benötigt werden, um die Details einer Wirkung angemessen zu erklären. In dieser Richtung wird das Prinzip weitgehend akzeptiert. Andererseits verlangt das Prinzip auch nach der höchsten abstrakten Eigenschaft, mit der

die Eigenschaft der Wirkung angemessen erklärt werden kann.[126] Dabei besteht einerseits die Gefahr der Übergeneralisierung und andererseits der Überspezifizierung. Nach dem Prinzip der Proportionalität soll eine Ursache ausreichend, aber auch nicht zu viel für eine Wirkung sein. Anscheinend fallen auch Gesetze unter das Prinzip der Proportionalität. Die Aussage „Materie leitet ceteris paribus Elektrizität" ist eine Übergeneralisierung, während die Aussage „Centstücke leiten ceteris paribus Elektrizität" eine Überspezifizierung ist. Es ist das Kupfer, das die Elektrizität leitet (vgl. Yablo 1997, 259).

Schröder betont vor diesem Hintergrund, dass der Abstraktionsgrad einer Eigenschaft für ihre proportionale kausale Wirkung entscheidend ist. Wenn z. B. die meisten Fensterscheiben eines bestimmten Typs durch den Aufprall eines Gegen-stands mit einem Impuls von mehr als 5 kg/sek. zerbrechen, dann ist die Eigenschaft „ein Impuls von mehr als 5 kg/sek." proportional zur Eigenschaft des Fensterzerbrechens. Gegenstände mit einem Impuls von 6, 7 oder 8 kg/sek. führen alle zum Zerbrechen der Fensterscheibe, da sie alle die Eigenschaft mit dem höheren Abstraktionsgrad „ein Impuls von mehr als 5 kg/sek." haben. Wenn also dieselbe Eigenschaft durch die Wirkung von unterschiedlichen Ereignissen instanziiert wird, ist es die „sparsamste" Erklärung, dass dieselbe Eigenschaft der Ereignisse kausal verantwortlich ist (vgl. Schröder 2003, 178f.).[127] Das Prinzip der Sparsamkeit kann somit als ein Argument für das Prinzip der Eigenschaftsproportionalität gewertet werden. Die Erklärungspraxis wird tatsächlich von dem zweiten Teil des Grundsatzes bestimmt, nach dem die höchste abstrakte Eigenschaft entweder hinreichend für die Instanziierung einer anderen Eigenschaft, oder Teil mehrerer hinreichender Eigenschaften ist.

Das Prinzip der Proportionalität steht nach Ansicht Schröders nicht in Widerspruch zum Prinzip der kausalen Geschlossenheit der Physik. Denn in seiner üblichen Formulierung besagt das Prinzip zwar, dass es für jede physische Wirkung eine hinreichende physische Ursache gibt. Aber es besagt nicht, dass physische Wirkungen in einer bestimmten Art und Weise beschrieben werden müssen, bzw. dass es für jede physische Beschreibung einer Wirkung eine hinreichende physische Ursache gibt. Damit sind einerseits mentale Beschreibungen für physische Wirkungen generell möglich, sowie andererseits mentale Ursachen unter physischen Beschreibungen. Das Prinzip der kausalen Geschlossenheit kann somit zumindest in einer schwachen Variante akzeptiert werden. Danach gibt es für jede physische Wirkung eine hinreichende physische Ursache unter einigen physischen Beschreibungen. Diese Formulierung ist kompatibel mit Fällen, wonach es einige

[126] So betont auch Yablo, dass in der determinable-determinate Relation das determinate eine Spezifizierung des determinable ist: „*P* determines *Q* iff: for a thing to be *P* is for it to be *Q*, not *simpliciter*, but in a specific way." (Yablo 1992, 252)

[127] Ein weiteres Beispiel, wonach zwei Wirkungen aufgrund ihrer Korrelation eine Ursache haben können, ist der Beginn von schlechtem Wetter und die Veränderung des Barometers – beide Ereignisse werden von etwas drittem, dem Fall des atmosphärischen Drucks – verursacht.

physische Beschreibungen von Wirkungen gibt, die so abstrakt sind, dass es keine abstrakte physische Beschreibung ihrer proportionalen Ursachen gibt. Und auch auf solche Fälle sollte das Prinzip der kausalen Geschlossenheit anwendbar sein. Das heißt, Schröder fordert ein, dass das Prinzip grundsätzlich auch der epistemischen Ebene, den Beschreibungen von Kausalrelationen, Rechnung trägt und darauf anwendbar ist (vgl. ebd. 183). Aus der schwachen Formulierung des Prinzips der kausalen Geschlossenheit ergibt sich, dass jedes konkrete Ereignis bzw. jede Instanziierung einer Eigenschaft zumindest eine Eigenschaft besitzt, für die es eine proportionale physische Ursache gibt. In diesem Sinne gibt es keine Unvollständigkeit der Physik und wird das Prinzip der kausalen Geschlossenheit nicht verletzt.

Auch das Argument des kausalen Ausschlusses scheint durch das Prinzip der Proportionalität entkräftet. Wenn überhaupt, kämen nur abstrakte physische Eigenschaften als Konkurrenten für den Inhalt einer Repräsentation hinsichtlich proportionaler kausaler Verursachung in Frage. Sie würden diese als Verursacher ausschließen. Eine solche abstrakte physische Eigenschaft sollte trivialerweise nicht einfach eine Disjunktion verschiedener einfacher physischer Eigenschaften sein. Eine solche Disjunktion würde im Grunde auch dem Prinzip der Proportionalität widersprechen, da in einem modalen Sinn sämtliche alternativen intrinsischen Eigenschaften berücksichtigt werden müssten (vgl. Schröder 2003; Yablo 1997). In einem Gedankenexperiment führt Schröder die Überlegung aus, dass abstrakte physische Eigenschaften letztlich zu unspezifisch sind, um als proportionale Eigenschaften für Verhaltenseigenschaften in Frage zu kommen. Angenommen, zwei Repräsentationen haben dieselben intrinsischen Eigenschaften, aber unterschiedliche Bedeutung und unterschiedliche Verbindungen zum motorischen Kortex. Physische Eigenschaften könnten dann nicht erklären, worin der Unterschied dieser Verhaltensweisen zu den zwei Repräsentationen bestünde. Sie wären zu unspezifisch, wären als intrinsische Eigenschaften nicht kausal relevant für die Instanziierung und würden deshalb die proportionale Bedingung nicht erfüllen.[128] Umgekehrt zeigt das Prinzip der Multirealisierbarkeit, dass psychologische Prozesse in ihrem Ablauf von den mentalen Zuständen bis hin zur entsprechenden Handlung verschieden realisierbar sind. Insofern kann die empirische Information über einen neurophysiologischen Prozess allenfalls erklären, wie ein bestimmter psychologischer Prozess in einem bestimmten Individuum realisiert ist – bei anderen Individuen wird derselbe Prozess jedoch ganz anders realisiert (vgl. Menzies 2003, 217f).

[128] Auch abstrakte physische Eigenschaften, die gewissermaßen den Durchschnitt entsprechender physischer Eigenschaften bilden, scheiden als proportionale Kandidaten für die Erklärung von Verhaltenseigenschaften aus, da sie künstlich konstruiert und somit letztlich trivial sind. Schließlich wird für ihre Konstruierung eine äquivalente Klasse von Verhaltensweisen benötigt, von der sie abhängen. Im Gegensatz dazu ist eine Erklärung der Körperbewegung durch physische Eigenschaften vollständig und unabhängig von Verhaltensbeschreibungen zu leisten. Dieser Umstand zeigt, dass physische Eigenschaften eben proportional auf dieser Detailebene Erklärungen leisten, aber – auch als Durchschnittseigenschaften – keine proportionale Erklärung für Verhaltenseigenschaften liefern können (vgl. Schröder 2003).

### 4.6.2 Proportionalität und mentale Kausalität

Bevor das Prinzip der Proportionalität als Bedingung für adäquate Kausalität entwickelt war, reichte es für die hinreichende Erklärung einer bestimmten Bewegung aus, dass unbestimmte physische Eigenschaften aufgeführt werden konnten. Mit dem Prinzip hat sich die Situation nun insofern verändert, als nun gezeigt werden muss, dass es bestimmte abstrakte physische Eigenschaften gibt, die semantische Eigenschaften als proportionale Eigenschaften von mentaler Verursachung ausschließen (vgl. Schröder 2003, 185).

Allerdings reicht es nicht aus, dass proportionale Eigenschaften notwendig für die Erklärung der kausalen Relation sind. Es soll darüber hinaus gezeigt werden, dass sie auch notwendig für die Instanziierung der zu erklärenden Eigenschaften selbst sind. Marras argumentiert zutreffend, dass Ereignisse nur deshalb kausale Kraft haben, weil sie ein bestimmter Ereignistyp sind und ganz bestimmte Eigenschaften besitzen. Ereignisse sind Kraft ihrer Eigenschaften kausal wirksam. Es reicht nicht aus zu zeigen, dass mentale Eigenschaften eine wichtige Rolle in der kausalen Erklärung von Handlungen spielen, oder dass sie nomologischen oder kontrafaktischen Bedingungen genügen. Damit mentale Eigenschaften eine echte kausale Rolle spielen, muss gezeigt werden, dass sie nicht nur einen explanatorischen, sondern auch einen kausalen Unterschied machen (vgl. Marras 2003, 249-254). Deshalb versucht er vor dem Hintergrund einer Implementierungs- und Realisierungsrelation, funktionalen bzw. mentalen Eigenschaften auf einer metaphysischen Ebene zu kausaler Kraft zu verhelfen (ebd. 259ff.). Im Ergebnis leihen sich die mentalen Eigenschaften allerdings nur ihre kausale Kraft von ihren subvenienten Basen, auf die sie zumindest nicht reduzierbar sind.

Yablo versucht als Alternative zur Supervenienzrelation die determinates-determinables Relation zu etablieren (vgl. Yablo 1992). Man kann hier jedoch einwenden, dass sich diese Relation in erster Linie auf die Eigenschaften der Ursache in ihrer logischen Dimension bezieht. Problematischerweise kann nicht gezeigt werden, wie mentale Eigenschaften determinables der physischen Hirneigenschaften sein können. Wie beispielsweise die Eigenschaft, ein rotes Haus zu repräsentieren, determinable der Eigenschaft sein kann, ein bestimmtes neuronales Ereignis zu sein, bleibt allerdings offen. Die logische Verbindung eignet sich somit nur bedingt für mentale Kausalität.[129]

Schröder versucht, die Relation als eine Relation zwischen semantischen Eigenschaften und Aktivierungsvektoren zu spezifizieren. Diese Spezifizierung soll neutral hinsichtlich der Diskussion zwischen klassischen repräsentationalen Theorien sowie konnektionistischen Ansätzen sein (vgl. Schröder 2003, 181; Kap. 4.3). Insofern können die Aktivierungsvektoren auch als Imple-

[129] Während bei den Ausdrücken „krank" und „schwer krank" der erste Ausdruck determinable des zweiten Ausdrucks ist, kann eine solche Schlussfolgerung nicht bei Ausdrücken gemacht werden, die sich auf semantische und physische Eigenschaften mentaler Repräsentationen beziehen.

mentierungen symbolischer Repräsentationen gedacht werden. In dieser Funktion repräsentieren sie Intentionen bzw. die Schlussfolgerungen, die sich aus Glaubensannahmen und Wünschen für Handlungen ergeben. Im Grunde ist diese Position aber recht speziell, und man kann das Prinzip der Proportionalität durchaus bereits für eine allgemeine Relation geltend machen: die Supervenienzrelation. „Usually, not much attention is given to this, but it is assumed that there is a base property *P* on which another property *M* supervenes." (Schröder 2003, 180) Damit kann den vorangegangenen Überlegungen Rechnung getragen werden, wonach das Prinzip der Proportionalität in Bezug zur mentalen Kausalität für verschiedene metaphysische oder auch nomologische Positionen anwendbar ist. In der Begründung, warum dem Prinzip der Proportionalität nicht nur epistemische, sondern auch ontologische Relevanz zukommt, bleibt Schröder allerdings vage. Wenn die speziellen oder auch die abstrakten Eigenschaften betrachtet werden, soll bestimmt werden, wovon eine bestimmte Wirkung bzw. eine bestimmte Eigenschaft tatsächlich abhängt. Es soll bestimmt werden, ob die scheinbare Abhängigkeit einer Eigenschaft immer noch bestätigt werden kann, wenn der Detailumfang der Eigenschaft verändert wird. Diese Suche nach einer Eigenschaft, von der die Instanziierung einer weiteren Eigenschaft tatsächlich abhängt, scheint für das Prinzip der Eigenschaftsproportionalität auch auf ontologischer Ebene zu sprechen (vgl. ebd.). Die Proportionalität von Eigenschaften ist demnach ein Grundsatz für die Kausalität und nicht nur für die kausale Erklärung. Es ist allerdings fraglich, ob diese Argumentation ausreicht, um einen Physikalisten vom ontologischen Gehalt des Prinzips der Proportionalität zu überzeugen.

Vielversprechender erscheint dagegen ein weiterer Argumentationsgang, der auf einen epistemischen Mehrwert semantischer bzw. mentaler Eigenschaften hinweist. In seiner Argumentation gegen den kausalen Ausschluss mentaler Eigenschaften kritisiert Schröder die Aussage, wonach die physischen Eigenschaften die kausale Arbeit anstelle der mentalen Eigenschaften verrichten, als eine irreführende Metapher. Genauer gesagt bezieht er sich auf den Umstand, wonach mentale Eigenschaften auf physischen Eigenschaften supervenieren und diese keine kausale Arbeit zu verrichten haben, da die kausale Arbeit bereits von den physischen Eigenschaften erledigt wird. Schröder geht es darum zu zeigen, dass die physischen Eigenschaften keine angemessene Entsprechung bzw. kein geeigneter Ersatz für mentale Eigenschaften sind. Seinen Standpunkt verdeutlicht er am Beispiel des Tankens: wenn ein Auto zur Tankstelle gefahren werden muss, um zu tanken, und Jones dies tut, muss Smith diese Arbeit nicht mehr erledigen. Bei der Arbeit handelt es sich um ein bestimmtes Ereignis, welches neben der Eigenschaft „zur Tankstelle fahren und tanken" möglicherweise eine Vielzahl von weiteren Eigenschaften instanziiert. Jones muss als Handelnder sicherstellen, dass die möglichen in Frage kommenden physischen Ereignisse dazu beitragen, die erforderliche Beschreibung zu erfüllen. Eine mögliche

physische Entsprechung für Jones wäre nach Schröder jedoch nicht in der Lage, diese Anforderung zu erfüllen:

> What the physical property P1 makes sure is that another event has a certain property P2. It does not and it cannot assure that a different property P3 equally satisfies a certain description B, that is, a behavioural description like 'grasping a cup of tea'. (ebd., 177)

Der entscheidende Unterschied zwischen Jones und seiner physischen Entsprechung, bzw. mentalen Eigenschaften und ihrer physischen Entsprechung ist der mögliche Zugriff auf eine Vielzahl verschiedener physischer Ereignisse. Während eine Beschreibung B von verschiedenen physischen Eigenschaften $P_2$ und $P_3$ erfüllt wird, kann die physische Eigenschaft $P_1$ nur gewährleisten, dass ein weiteres Ereignis die Eigenschaft $P_2$ hat, aber nicht, dass es eine weitere Eigenschaft $P_3$ gibt. Dagegen verfügen mentale Repräsentationen über Kriterien, die Beschreibung eines physischen Ereignisses zu erfüllen. Damit, so Schröder, ist es letztlich der Inhalt der mentalen Repräsentationen, der dazu beiträgt, dass Jones das Auto tankt. Mentale bzw. intentionale Eigenschaften ermöglichen es, dass ein bestimmtes Ziel unter Instanziierung einer bestimmten physischen Eigenschaft $P_i$ unter einer Vielzahl möglicher physischer Eigenschaften P bzw. äquivalenter Verhaltensweisen erreicht wird. Dabei ist es wichtig, dass zum Erreichen des Ziels auch die äquivalente Eigenschaft $P_j$ instanziiert werden kann, sollte die Eigenschaft $P_i$ nicht möglich sein. Diese Option haben intrinsische physische im Gegensatz zu mentalen Eigenschaften nicht.

Dass mentale Eigenschaften eine Handlung sicherstellen, bedeutet, dass verschiedene neuronale motorische Pfade mit einer bestimmten Zielrepräsentation verknüpft sind. Die Entwicklung solcher alternativer Verknüpfungspfade ist darauf zurückzuführen, dass die Zielrepräsentation einen bestimmten repräsentationalen Inhalt hat (vgl. Dretske 1988; zitiert nach: Schröder 2003, 177). Intentionaler Inhalt ist letztlich verantwortlich für die Ausbildung der „gesamten Struktur", wonach verschiedene alternative Verbindungen mit einer bestimmten Zielrepräsentation verknüpft werden und auch tatsächlich ein alternativer motorischer Pfad aktiviert wird, falls ein anderer motorischer Pfad nicht aktiviert werden kann. Auch Marras argumentiert vorsichtig in eine solche Richtung. So wird die kausale Kraft mentaler Eigenschaften zwar durch die kausale Kraft physischer Eigenschaften bestimmt, aber sie ist weder identisch noch äquivalent, da mentale Eigenschaften Bestandteil von Generalisierungen sein können, die *nicht* allein in den Begriffen physischer Eigenschaften gefasst werden können. Deshalb gibt es keinen Grund für die Annahme, dass intentionale und physikalische Erklärungen miteinander in Konflikt geraten (vgl. Marras 2003, 261). An dieser Stelle muss letztlich überlegt werden, welche Rückschlüsse epistemische Beschreibungen auf die ontologische Ebene zulassen. Zumindest lässt sich vermuten, dass Mehrinformationen, die wir aufgrund mentaler Beschreibungen erhalten, nicht überflüssige, sondern relevante Informationen proportional zum ontologischen Ereignis sein können.

Es konnte gezeigt werden, dass das Prinzip der Proportionalität ein wichtiges Kriterium für mentale Kausalität darstellt, wenn nicht sogar für kausale oder nomologische Prozesse überhaupt. Dem Prinzip kommt insofern eine besondere Bedeutung zu, als hierbei verschiedene Voraussetzungen sowie Kriterien für mentale Kausalität miteinander verknüpft werden können. So argumentiert das Prinzip entscheidend für die kausale Relevanz intentionaler, semantischer Eigenschaften. Deren proportionale kausale Wirkung muss auch unter dem Gesichtspunkt der Multirealisierbarkeit betrachtet werden, die nicht nur innerhalb z. B. der Supervenienzrelation, sondern auch hinsichtlich der kausalen Beschreibung der Eigenschaften der Wirkung erfolgen muss. Dabei wird deutlich, dass das Prinzip der kausalen Geschlossenheit der Physik auch in seiner Zuspitzung, als Argument des kausalen Ausschlusses, angreifbar ist.

Das Prinzip der Proportionalität soll nicht nur die kausale Erklärung, sondern die kausale Relation selbst beschreiben. Sollte dies zutreffen, wären semantische Eigenschaften sogar besser geeignet bzw. relevant als intrinsische physische Eigenschaften vis-a-vis bestimmte Verhaltenseigenschaften kausal zu verursachen und zu erklären (vgl. Schröder 2003, 185). Denn übertragen auf das Phänomen der mentalen Kausalität ergibt sich als Konsequenz der Eigenschaftsproportionalität eine Änderung des Ausschlussarguments (vgl. Kap. 2.5.4). Anstatt superveniente mentale Eigenschaften von der kausalen Wirksamkeit auszuschließen, scheinen die detaillierten subvenienten Eigenschaften hinsichtlich der zu erklärenden Verhaltenseigenschaften irrelevant zu sein.[130] Wenn beispielsweise die Eigenschaft „trinken" erklärt werden soll, ist eine detaillierte Beschreibung des physischen Zustands, auf welchem die Intention zu trinken superveniert, irrelevant für die zu instanziierende Eigenschaft. Wenn andererseits die genauen Verbindungen, die das Trinken als Ereignis realisieren, erklärt werden sollen, dann ist die Intention zu trinken als Erklärung nicht detailliert genug, und die Angabe eines physischen Zustands z. B. im präfrontalen Kortex wäre proportional zur bewirkten Eigenschaft.[131] Wenn semantische Eigenschaften physischer Strukturen proportional zu den Verhaltenseigenschaften hinsichtlich einer Erklärung sind, und wenn keine Eigenschaften, die nur physisch beschreibbar sind, proportional zu diesen sind, dann müssen die semantischen Eigenschaften kausal relevant für die Instanziierung der Verhaltenseigenschaften sein, die physischen jedoch nicht. Schröder zeigt, dass, wenn die intrinsischen physischen Eigenschaften proportional zu den neuronalen Verbindungen sind, sie nicht

[130] Ähnlich argumentiert auch Yablo, wenn er sagt, dass üblicherweise von zwei Ursachekandidaten derjenige präferiert wird, der sich stärker proportional zur Wirkung verhält. Beispielsweise hat die Entscheidung, die Türklingel zu benutzen anstatt zu klopfen, eine physische Determination P. Yablo würde aber dahingehend argumentieren, dass das Türklingeln auch mit einer anderen physischen Realisierung Q erfolgt wäre. Somit würde P nicht benötigt werden und wäre nicht proportional zum Türklingeln als Wirkung (vgl. Yablo 1992, 277f.).

[131] Schließlich stützt das Prinzip der Eigenschaftsproportionalität auch die Position der Einzelwissenschaften, da diese mit proportionalen Eigenschaften echte kausale Wirksamkeit haben können. Die kausale Wirksamkeit von Eigenschaften, die durch die Einzelwissenschaften bestimmt werden, wird nach Schröder sogar garantiert, da diese letztlich aufgrund ihrer Proportionalität zu ihrer Wirkung innerhalb der jeweiligen Wissenschaft akzeptiert sind.

proportional zu den abstrakten Verhaltenseigenschaften sein können. Aber die mentalen bzw. semantischen Eigenschaften können dies leisten und sind von daher auch kausal relevant für die Verhaltenseigenschaften. Solange semantische Eigenschaften von physischen Strukturen proportional für die Erklärung von Verhaltenseigenschaften sind und keine ausschließlich in physischen Begriffen beschreibbaren proportionalen Eigenschaften gefunden werden, sind die semantischen Eigenschaften kausal relevant für die Instanziierung der Verhaltenseigenschaften – im Gegensatz zu den physischen Eigenschaften.

***Quintessenz Kriterium der Proportionalität***

- *adäquate Relation von Ursache und Wirkung: keine Konkurrenz; keine Lücken in physischer Kausalkette; Gleichberechtigung auf epistemischer Ebene*
- *derselbe Grad an Abstraktion: weder Übergeneralisierung noch Überspezifizierung*
- *determinable-determinate-Relation: Zurückweisung des kausalen Ausschlussarguments*
- *Proportionalität abstrakter Verhaltenseigenschaften: Vermeidung unnötiger (alternativer) mikrostruktureller Details; Vereinfachung mentaler Zustände entscheidend für mentale Kausalität*
- *Wirksamkeit abstrakter proportionaler Eigenschaften entgegen Argument der kausalen Geschlossenheit oder kausalem Ausschlussargument möglich*
- *Wirkung abhängig von Instanziierung bestimmter Eigenschaften (Detailumfang)*
- *epistemischer Mehrwert: mentale Eigenschaften als Zielvorgabe für Instanziierung physischer Eigenschaften; Sicherstellung von Handlung sichert Ausbildung neuronaler Struktur*

**Kapitel 5**

**Neuronale Prozessierung und Repräsentation**

Für die empirische Diskussion der mentalen Kausalität wurden zwei Studien hinzugezogen, die im Kern das Phänomen der emotionalen Aufmerksamkeit, sowie die neuronale Prozessierung verschiedener emotionaler Dimensionen untersuchen. Diese beiden neurowissenschaftlichen Paradigmen „emotionale Aufmerksamkeit" und „emotionale Beurteilung" wurden dahingehend ausgewählt, dass sich die verschiedene Kriterien, die mit der Frage der mentalen Kausalität verbunden sind, in einem weiteren Schritt auf diese beziehen lassen und einen spezifischen Vergleich ermöglichen (vgl. Kap. 6).

Die erstere der beiden neurowissenschaftlichen Studien konzentrierte sich mit der Untersuchung des „Aufmerksamkeitseffekts" insbesondere auf die Auswirkungen der Aufmerksamkeits-, wie der Bildstimuli unter einer emotionalen Kondition auf der neuronalen Ebene bzw. veränderte Aktivitätsmuster im Gehirn. Deshalb wurde als zweite Studie die Beurteilungsstudie hinzugezogen, die zusätzlich die Korrelation zwischen neuronaler Aktivität und subjektiver Bewertung bestimmte. In beiden Studien wird die neuronale Prozessierung (emotionaler) Aufmerksamkeit untersucht. Die Erwartung bzw. Aufmerksamkeit auf zukünftige(r) Ereignisse erlaubt den Menschen die Optimierung der Geschwindigkeit und Akkuratheit des Informationsprozesses, sowie die Vorbereitung, bzw. Bereitschaft zu früher rechtzeitiger motorischer und autonomer Reaktion. Insofern kann Aufmerksamkeit als möglicher Bestandteil kausaler Handlungsfolgen betrachtet werden. Sie ist kein fester Bestandteil, da unbewusste und bewusste menschliche Handlungen, einschließlich der Wahrnehmung, sowohl mit, als auch ohne vorhergehende Aufmerksamkeit ausgeführt werden. In einer philosophischen Definition wird Aufmerksamkeit als das durch die Eigenart eines Gegenstandes oder willentlich herbeigeführte Verharren des Bewusstseins zum Zweck genauer Auffassung verstanden. Es handelt sich dabei um einen Zustand erhöhter geistiger Aktivität, da die Aufmerksamkeit verschiedene Intensitätsgrade bis zur stärksten Konzentration annehmen kann (vgl. Eisler 1904, 567ff.). Es ist grundsätzlich auch eine unterschwellige Aufmerksamkeitssteigerung durch unbewusst wahrgenommene Warnsignale denkbar; so werden Straßenschilder von routinierten Autofahrern weniger bewusst wahrgenommen als von Fahranfängern. Trotzdem stellen sie sich zumindest in der Regel auf die durch die Signale angekündigten Situationen ein.

Die Aufmerksamkeitsstudie „Dissociable Networks for the Expectancy and Perception of Emotional Stimuli in the Human Brain" untersuchte die Wirkung von Aufmerksamkeits- und Wahrnehmungsstimuli auf den Faktor „Emotion", die durch entsprechend ausgewählte Bilder jeweils induziert wurden (vgl. Bermpohl et al. 2006). Ziel der Studie war es einerseits zu bestim-

men, welche Gehirnareale bei emotionaler Aufmerksamkeit involviert sind, und andererseits diese Regionen von denen zu unterscheiden, die auf die Wahrnehmung emotionaler Stimuli selbst reagieren. Dabei war die Kondition der emotionalen Aufmerksamkeit insofern unsicher, als durch das entsprechende Signal für emotionale Aufmerksamkeit keine genaue Information gegeben wurde, ob ein positives oder negatives Bild erscheinen würde. Die Ergebnisse der Studie lassen vermuten, dass unbestimmte emotionale Aufmerksamkeit ein neuronales Netzwerk einbezieht, das von Hirnregionen, die sich auf die emotionale Bildwahrnehmung beziehen, abgegrenzt werden kann. Diese Abgrenzung verweist auf die mögliche Trennung zwischen antizipatorischen und perzeptiven Komponenten der emotionalen Stimulus-Prozessierung (vgl. ebd.).

Die Beurteilungsstudie „Segregated neural representation of distinct emotion dimensions in the prefrontal cortex – an fMRI study" untersuchte die genaue neuronale Repräsentation der emotionalen Dimensionen Valenz, Intensität und Wiedererkennung (vgl. Grimm et al. 2006). Ein übliches Problem der Bestimmung präfrontaler cortikaler Repräsentationen ist, dass diese Regionen auch bei kognitiven Funktionen, die mit emotionaler Prozessierung verbunden sind, aktiv sind. Deshalb wurde eine fMRT-Studie entwickelt, die, wie auch die Aufmerksamkeitsstudie, emotionale Bilder mit einbezieht. Hierdurch konnten die kognitiven Funktionen bzw. die Prozessierung von Beurteilung und vorhergehender Aufmerksamkeit kontrolliert und abgegrenzt werden. In einem weiteren Schritt wurde die gemessene neuronale Aktivität mit zuvor gemessenen subjektiven Post-Scanning-Bewertungen der emotionalen Dimensionen durch die Probanden korreliert. Im Ergebnis konnten in verschiedenen präfrontalen cortikalen Regionen getrennte neuronale Repräsentationen für die verschiedenen emotionalen Dimensionen bestimmt werden (vgl. ebd.).

Beide Studien wurden mit einem funktionellen Kernspintomographen bzw. einem funktionellen Magnetresonanztomographen (fMRT) durchgeführt. Vorteil der funktionellen Bildgebung ist eine hohe räumliche und zeitliche Auflösung sowie die nicht-invasive Methodik, weshalb den Probanden zum Beispiel kein Kontrastmittel o. ä. injiziert werden muss. Allerdings kann die neuronale Aktivität mit dieser Methode nicht direkt gemessen werden. Die funktionelle MRT bedient sich des so genannten Blood Oxygenation Level Dependent-Effekts (BOLD-Effekts). Gemessen wird die paramagnetische Eigenschaft des Desoxyhämoglobins, bei dem es sich um eine Form des Hämoglobins handelt, die keinen Sauerstoff transportiert. Neuronale Aktivität führt zu einer Erhöhung des Sauerstoffverbrauchs und somit zu einem Anstieg des Desoxyhämoglobins. Dadurch steigt wiederum der regionale Blutfluss an, was zu einer Zunahme der Oxygenierung bei gleichzeitiger Abnahme des Desoxyhämoglobingehalts führt. Nach ungefähr sechs Sekunden nach einer Stimuluspräsentation ist das Maximum des Desoxyhämoglobinabfalls erreicht, nach weiteren sechs Sekunden befindet sich der Gehalt wieder auf dem Niveau vor der Stimulusprä-

sentation (vgl. Fox et al. 1988; Rosen et al. 1993; Di Salle et al. 1999). Bei der Messung wird insofern eine Korrelation zwischen der neuronalen Aktivität und dem Desoxyhämoglobingehalt im Blut angenommen. So können durch die funktionelle MRT Regionen als aktiviert dargestellt werden, die in einer physiologischen Messung keinerlei Aktionspotenziale aufweisen und damit zu falschen positiven Ergebnissen führen.[132] Dagegen wurde eine Korrelation zwischen der Aktivität lokaler Feldpotenziale der Neuronen und der hämodynamischen Reaktion bzw. dem BOLD-Effekt gemessen. Da diese als Input neuronaler Verarbeitung betrachtet werden, scheint die funktionelle MRT vorwiegend hinsichtlich des Inputs neuronaler Verarbeitung aussagekräftig zu sein. Weil die Aktivität lokaler Feldpotenziale meistens mit der Aktivität der Aktionspotenziale einhergeht, kann man den BOLD-Effekt auch auf letztere beziehen (vgl. Logothetis et al. 2001).

Weitere Schwierigkeiten des fMRT sind mit der räumlichen sowie der zeitlichen Auflösung der durch den BOLD-Effekt induzierten Signale verbunden. Die räumliche Auflösung ist von der Relation der Blutgefäße zu den Nervenzellen abhängig. Eine eindeutige Auflösung einer Nervenzelle ist nur möglich, wenn diese von einem eigenen Blutgefäß versorgt wird – dies trifft jedoch nur sehr selten zu. Bei der zeitlichen Auflösung unterscheiden sich BOLD-Signal und neuronale Aktivität deutlich: erstere weist einen Zeitverlauf von 12 Sekunden auf, letztere lediglich von wenigen Millisekunden. Immerhin scheint die zeitliche Relation zwischen neuronaler Reaktion und BOLD-Effekt konstant zu sein. Der neuronale Zeitverlauf kann folglich aus dem BOLD-Effekt abgeleitet werden (vgl. Heinzel 2002, 159f.).

Ziel neurowissenschaftlicher Studien mit dem fMRT ist die Darstellung neuronaler Aktivität unter einer bestimmten Kondition. Andere gleichzeitige neuronale Aktivität, die nicht in direktem Zusammenhang mit der eigentlichen Aufgabe steht, soll ausgeblendet werden. Ein übliches Verfahren hierfür ist die kognitive Subtraktion. Hierfür muss das experimentelle Design mindestens zwei verschiedene Aufgaben enthalten: die Aktivierungsaufgabe und die Baselineaufgabe. Die Aktivierungsaufgabe bezieht sich auf den zu untersuchenden Prozess, während die Baselineaufgabe als Kontrollkondition der Aktivierungsaufgabe in jeder Hinsicht gleicht, außer dass der zu untersuchende Prozess nicht enthalten ist. Dieser Prozess kann nun durch einfache Subtraktion der Baselineaufgabe von der Aktivierungsaufgabe bestimmt werden. Problematisch ist dabei, dass während der Messung trotz konkreter Anweisung bei den Probanden unter beiden Konditionen der zu untersuchende Prozess aktiviert wird. So ist es in der vorliegenden Beurteilungsstudie möglich, dass die Probanden auch während der Phase der bloßen Bildwahrnehmung unbewusst eine Bildbeurteilung vornehmen. Außerdem können auch stets weitere neuronale Prozesse zusammen mit dem zu untersuchenden Prozess herausgerechnet werden, so dass auch neu-

[132] Es wurde gezeigt, dass es keine quantitative Beziehung zwischen der Frequenz der Aktionspotenziale der Neuronen und der hämodynamischen Reaktion bzw. dem BOLD-Effekt gibt. Solange die Frequenz der Aktionspotenziale als Output neuronaler Verarbeitung aufgefasst werden, kann über die fMRT-Messung keine direkte Aussage getroffen werden (vgl. Logothetis et al. 2001).

ronale Aktivität über den eigentlichen Prozess hinaus zu diesem hinzugerechnet wird. Eine Alternative zur kognitiven Subtraktion ist deshalb die kognitive Konjunktion. Hierbei wird die gemeinsame neuronale Aktivität aus mindestens zwei Konditionen ermittelt. Die Anforderung an die Baseline ist nun insofern geringer, als sie sich lediglich von dem zu untersuchenden Prozess der Aktivierungsaufgabe unterscheidet. In einer Konjunktionsanalyse wird beispielsweise die Subtraktion (Aktivierungsaufgabe 1 > Baselineaufgabe 1) mit der Subtraktion (Aktivierungsaufgabe 2 > Baselineaufgabe 2) addiert. Durch die kognitive Konjunktion zeigt sich dabei der Prozess, der in beiden Subtraktionen abläuft.

Grundsätzliches Problem beider neurowissenschaftlicher Studien ist die Interpretation der Ergebnisse im Abgleich mit anderen, bereits durchgeführten Studien. Diese können teilweise die neuen Ergebnisse unterstützen, andererseits aber auch abweichende oder sogar konträre Positionen enthalten. Der Vergleich ist problematisch, da die Konditionen ähnlicher Studien eben nur ähnlich sind, aber im Detail voneinander abweichen. Insofern ist die Interpretation neurowissenschaftlicher Daten, wie z. B. bei bildgebenden Verfahren, ein sehr schwieriges, mitunter etwas vages Geschäft, zumal die Studien häufig auf eine relativ geringe Teilnehmerzahl zurückgreifen müssen und nur selten wiederholt werden.

Eine fundamentale Kritik an der üblichen Praxis der Entwicklung von fMRT-Studien üben Hardcastle und Stewart (vgl. Hardcastle/Stewart 2002, 2003, 2005). Sie kritisieren, dass durch die theoretische Dominanz der kognitiven Psychologie und deren Verknüpfung mit fMRT-Studien im Aufbau das volle Potenzial der Neurowissenschaft als eigenständiges theoretisches Paradigma nicht ausgeschöpft werden kann und fMRT-Studien deshalb lediglich als Lokalisationsinstrument genutzt werden. Sie betonen, dass das Gehirn mehr sei, als eine statische Maschine, die unseren kognitiven Fähigkeiten zugrunde liegt, und fordern eine veränderte Perspektive (vgl. ebd. 2005, 37). Methodisch kritisieren sie ferner, dass das Gehirn sich in „Schnappschüssen" nicht in seiner Plastizität vollständig offenbart und viele Verbindungen – allenfalls die dominanten – nicht ermittelt werden. Ein Gehirnareal kann verschiedene Funktionen innehaben, abhängig davon, was sich im übrigen Gehirn ereignet. Im Gehirn wird innerhalb desselben Areals, abhängig von den Umständen, ein Prozess stärker vorangetrieben als ein anderer. Es ist nicht sinnvoll, zu fragen, was ein Areal macht, da die Antwort stets situationsabhängig ist (vgl. ebd.).

Methodische Schwierigkeiten gibt es somit auch für beide vorliegenden neurowissenschaftlichen Studien. Emotionale Stimulusprozessierung ist ein multidimensionales Konstrukt, welches auch andere Dimensionen als Valenz, Intensität oder Wiedererkennung umfassen kann. Durch gezeigte emotionale Stimuli werden autobiographische assoziiert. Es wurde gezeigt, dass das Wiedererlangen autobiographischer Erinnerungen das posteriore Cingulum und seine benachbarten Regionen wie das Retrosplenium und den Praecuneus umfasst (vgl. Fink 1996; Maddock

1999; Maddock et al. 2003; Piefke et al. 2003; Gilboa 2004; Lou et al. 2004; alle zitiert nach: Grimm et al. 2006, 337). Auch die Kontrollkonditionen für beide neurowissenschaftlichen Studien müssen vorsichtig betrachtet werden. Es kann argumentiert werden, dass die Phase der Bildbetrachtung immer auch eine Art der bewussten oder unbewussten Beurteilung beinhaltet. Auch wenn die Phasen in der Analyse erfolgreich voneinander getrennt wurden, kann zumindest eine implizite unbewusste Beurteilung während der Phase der Bildbetrachtung nicht vollständig ausgeschlossen werden. (vgl. ebd., 338).

Methodische Grenzen scheinen der funktionellen Magnetresonanztomographie vor allem im Mikrobereich gesetzt. Auf der höheren Organisationsebene bzw. der Ebene der Zellverbände oder Neuronenensemble liegen dagegen die Stärken dieser Methode (vgl. Heinzel 2002, 167). Gerade für diese Ebene lässt die funktionelle MRT verschiedene konzeptionelle Studiendesigns zu, wie auch anhand der nachfolgenden neurowissenschaftlichen Studien deutlich wird.

## 5.1 Das Paradigma der emotionalen Aufmerksamkeit[133]

Schnelle Identifizierung motivational relevanter Information und ihre Übersetzung in prompte Handlung ist überlebensnotwendig (Darwin 1872). Die Aufmerksamkeit und Erwartung (Antizipation) zukünftiger Ereignisse erlaubt den Menschen die Optimierung von Geschwindigkeit und Genauigkeit ihrer Reaktionen (Ingvar 1985). Aufmerksamkeit kann als vorhergehende Erwartung eines aufkommenden Stimulus, der durch situative und kontextuelle Zeichen angekündigt wird, verstanden werden. Bereits erworbenes Wissen zusammen mit aktuellen Informationen über die Umwelt schaffen die Basis, um Aufmerksamkeit zu generieren (vgl. Pavlov/Anrep 1927). Aufmerksamkeit kann in verschiedenen Bereichen, wie z. B. Sehen, Schmerz, Belohnung und Emotion beobachtet werden. Emotionale Aufmerksamkeit betrifft die Erwartung emotional auffallender Ereignisse. Sie ist vorbereitend für fokussierte affektive und kognitive Informationsprozessierung sowie für frühe motorische und autonome Reaktion. Obwohl funktionale Bildgebungsverfahren für die Untersuchung verschiedenster neuronaler Korrelate emotionaler Prozessierung eingesetzt wurden (vgl. Phan et al. 2002), steht die Untersuchung der Hirnregionen, die mit der Aufmerksamkeit emotionaler Bild-Stimuli in Zusammenhang gebracht werden, noch an einem Anfang.

Verschiedenen Studien ist dabei gemein, dass sie der Frage nach der Beziehung zwischen Aufmerksamkeits- und Wahrnehmungsaktivität im menschlichen Cortex nachgehen. Beispiels-

[133] Das folgende Kapitel ist eine zusammenfassende Übersetzung der Studie „Dissociable Networks for the Expectancy and Perception of Emotionale Stimuli in the Human Brain", die bereits in der Fachzeitschrift NeuroImage veröffentlicht wurde (vgl. Bermpohl et al. 2006).

weise wurde die Aufmerksamkeit emotionaler Bild-Stimuli untersucht (vgl. Ueda et al. 2003). In dem Versuchsparadigma wurde eindeutig angekündigt, ob ein positiver oder negativer Bild-Stimulus folgen würde. Die Ergebnisse zeigten eine mögliche Überlappung zwischen neuronalen Netzwerken, die hinsichtlich der Aufmerksamkeit sowie der Wahrnehmung emotionaler Stimuli involviert waren. Es gibt zwei mögliche Beschreibungen dieser Relation. William James (1892) entwickelte eine Theorie, wonach Aufmerksamkeit von der Aktivität desselben Netzwerkes abhängt, das aktuelle Wahrnehmungsstimuli prozessiert. Alternativ lässt sich annehmen, dass Aufmerksamkeit und Wahrnehmung verschiedene getrennte Hirnregionen einbeziehen. Es scheint, dass das Ausmaß der Überschneidung zwischen den Netzwerken Aufmerksamkeit und Wahrnehmung stark von dem untersuchten Bereich und dem entwickelten Paradigma abhängt (vgl. Bermpohl et al. 2006, 589).

Ein Vergleich zwischen den Aufmerksamkeits- und Wahrnehmungsnetzwerken innerhalb einer Studie wurde allerdings noch nicht ausgearbeitet. Dieser Vergleich wurde in der vorliegenden Studie durchgeführt: es wurde die Wirkung auf den Faktor ‚Emotion' durch Aufmerksamkeits- und Wahrnehmungsstimuli untersucht, die durch entsprechend ausgewählte Bilder jeweils induziert wurden. In der vorliegenden Studie wurde vor allem valenzunspezifisch die emotionale Aufmerksamkeit ohne Unterscheidung der positiven oder negativen Aufmerksamkeit untersucht. Dabei wurde die Aufmerksamkeit emotionaler Bilder im Vergleich zur Wahrnehmung dieser Stimuli untersucht. Für diesen Zweck wurde auf Unterschieden und Gemeinsamkeiten zwischen den involvierten neuronalen Netzwerken geachtet. Als Kontrollkondition dienten die neutrale Aufmerksamkeit und die neutrale Bildwahrnehmung, um die antizipatorischen und perzeptiven Komponenten spezifisch für die emotionale Prozessierung zu identifizieren. Als Signale für emotionale Aufmerksamkeit wurden in dem Paradigma unbestimmte Zeichen eingesetzt, die keine Informationen enthielten, ob ein positives oder negatives emotionales Bild folgen würde. Insofern kann die Kondition der emotionalen Aufmerksamkeit als eine Aufmerksamkeit eines Bild-Stimulus mit unbestimmtem emotionalem Inhalt betrachtet werden.

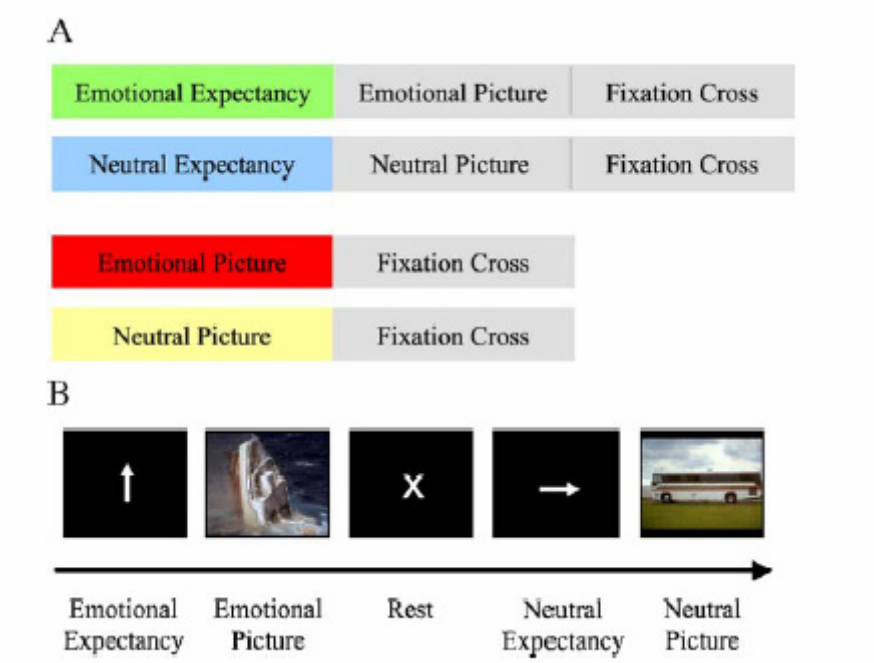
Die eingesetzte Aufmerksamkeits-Kondition unterscheidet sich von Pavolovscher Konditionierung in zweifacher Hinsicht: Die Probanden wurden über die Assoziation zwischen den Aufmerksamkeitszeichen und den folgenden Bild-Stimuli instruiert. In einem Testlauf wurde diese Folge eingeübt. Außerdem folgten dem Zeichen für emotionale Aufmerksamkeit in gleicher Proportion positive und negative Bild-Stimuli. Bei konventioneller Konditionierung werden dagegen die Zeichen entweder mit aversiven oder appetitiven Stimuli verknüpft. Die vorliegende Kondition für emotionale Aufmerksamkeit kann also eindeutig von Belohnungserwartung unterschieden werden. Während in Belohnungsparadigmen die Teilnehmer üblicherweise annehmen, das Erscheinen der Probe hängt von ihrer Erfüllung der Aufgabe ab, waren sich die Teilnehmer der vorliegenden Studie darüber bewusst, dass sie keinen Einfluss auf die gezeigten Stimuli hat-

ten. Deshalb wurde außerdem die Aufmerksamkeit und Wahrnehmung in Relation zu standardisierten und validierten emotionalen Stimuli gesetzt und verglichen, die keine Belohnungskomponenten o. ä. enthielten. Hierfür wurde angenommen, dass eine weitere Isolierung der emotionalen Komponente durch den Vergleich zwischen den aktiven Netzwerken bei Aufmerksamkeit bzw. Wahrnehmung möglich sei. Das Kernergebnis der vorliegenden Studie zeigt, dass unbestimmte emotionale Aufmerksamkeit speziell neuronale Aktivität im supracallosalen anterioren cingulären Cortex (SAC), im cingulären motorischen Areal (CMA) sowie im parieto-occipitalen Sulcus (POS) hervorruft. Dieses Netzwerk ist abtrennbar und unterscheidbar von dem Netzwerk, welches während der emotionalen Bildwahrnehmung aktiv ist (vgl. Bermpohl et al. 2006, 589f.).

### 5.1.1 Untersuchungsmethode und Datensammlung

Im Rahme der Studie wurden siebzehn gesunde Freiwillige unterschiedlichen Alters (21-37) und unterschiedlichen Geschlechts (9 weiblich, 8 männlich) als Probanden untersucht. Bei ihnen waren zum Zeitpunkt der Studie keine neurologischen oder psychiatrischen Erkrankungen bekannt. Als Bild-Stimuli wurden standardisierte Fotografien des International Affective Picture System (IAPS, Lang et al. 1999) verwendet. Diese Bilder sind abgestimmt hinsichtlich ihrer Valenz von Farbe, Helligkeit und Komplexität sowie ihrem semantischen Gehalt. Das Paradigma unterschied zwischen emotionalen Bildern (positiv und negativ) und neutralen Bildern (als Kontroll-Kondition). Jedem gezeigten Bild folgte eine Erholungsphase von 8,5 Sekunden Dauer, die den Probanden eine emotionale Erholung von dem Bild erlaubte. Bei der Hälfte der Bilder wurden zuvor aufmerksamkeitsgerichtete Zeichen präsentiert: Ein aufwärtszeigender Pfeil signalisierte, dass ein emotionales Bild folgen würde („emotionale Aufmerksamkeit"). Ein horizontaler Pfeil signalisierte, dass ein neutrales Bild folgen würde („neutrale Aufmerksamkeit"). Die Anweisung für aufwärtszeigende Pfeile lautete, sich auf das nachfolgende emotionale Bild einzustellen. Die Anweisung für die horizontalen Pfeile lautete, sich auf die nachfolgenden neutralen Bilder einzustellen. Die andere Hälfte der Fotografien wurde ohne vorhergehenden Pfeil gezeigt. Diese Proben ohne vorhergehende Aufmerksamkeit bildeten die Kondition „emotionale Bildwahrnehmung" und „neutrale Bildwahrnehmung". Insofern unterschied das Paradigma zwischen Aufmerksamkeit und Wahrnehmung sowie zwischen emotionalen und neutralen Konditionen. Neutrale Aufmerksamkeit und neutrale Bildwahrnehmung dienten als Kontrollkonditionen. Jede dieser Konditionen (emotionale Aufmerksamkeit, neutrale Aufmerksamkeit, emotionale Bildwahrnehmung, neutrale Bildwahrnehmung) umfasste 64 Proben, die innerhalb von acht Durchläufen gezeigt wurden. Die Konditionen wurden innerhalb und über die Durchläufe pseudorandomisiert und ausgeglichen. Der aufwärtszeigende und der horizontale Pfeil hatten als

nicht-bildliche Stimuli die gleiche Form, Größe, Farbe und Helligkeit und wurden vor einem schwarzen Hintergrund zentriert. Die Dauer der Aufmerksamkeitsphase und der Bildpräsentation betrug jeweils fünf Sekunden. Die relativ lange Dauer der Bildpräsentation wurde gewählt, um die Dauer der Aufmerksamkeitskondition und der Wahrnehmungskondition anzupassen. Außerdem wurde während der vorhergehenden behaviouralen Pilottestung ermittelt, dass einige komplexere Bilder eine längere Prozessierungszeit erforderten, um vollständig und bewusst verstanden zu werden, was gerade für das emotionale Verstehen der Bilder von Bedeutung ist. Vor dem eigentlichen Experiment wurden die Probanden mit dem Paradigma vertraut gemacht, indem sie einen Testlauf mit 20 Proben absolvierten. Sie wurden angewiesen, umgehend den Knopf zu drücken, sobald sie eine Fotografie sahen. Mit diesem Knopfdruck konnte die Aufmerksamkeit der Probanden überprüft werden. Der Knopfdruck beinhaltete kein spezielles Urteil, da sich eine solche kognitive Anforderung mit der emotionalen Stimulusprozessierung überschnitten hätte (vgl. Taylor et al. 2003).




Grafik 11: (A) In dem fMRT-Paradigma wurden emotionale und neutrale Aufmerksamkeit sowie emotionale und neutrale Wahrnehmung als Konditionen voneinander unterschieden. Emotionale und neutrale Aufmerksamkeitsintervalle (aufrechter, horizontaler Pfeil) gingen der emotionalen und neutralen Bildwahrnehmung voran. Der aufrechte Pfeil legte nicht fest, ob ein positives oder negatives emotionales Bild folgte. (B) In dem Beispiel einer Stimulus- bzw. Bildfolge werden die Konditionen „emotionale Aufmerksamkeit" mit dem darauffolgenden emotionalen Bild und „neutrale Aufmerksamkeit mit dem darauffolgenden neutralen Bild dargestellt (vgl. Bermpohl et al. 2006, 589).

Einen Tag nach der fMRI-Sitzung wurde das Paradigma den Probanden ein weiteres Mal gezeigt. Dieses Mal folgte auf jedes Bild eine Aufgabenphase, bestehend aus emotionaler Valenz- und Intensitäts-Einschätzung sowie einem zufälligen Erinnerungstest. Die Valenz- und Intensitätseinschätzung wurde auf einer visuellen 9-Punkte-Analogskala bewertet, innerhalb der zwischen (1)

als ‚sehr negativ' oder ‚wenig intensiv', über (5) als ‚neutral' oder ‚durchschnittlich intensiv' bis hin zu (9) als ‚sehr positiv' oder ‚sehr intensiv' unterschieden wurde.

Obwohl diese Post-hoc-Einschätzungen nicht die gleichzeitige Ausführung und Leistung während der Messung (Scanning) widerspiegeln, wird nachvollziehbar angenommen, dass die Probanden während der Messung und der Post-hoc-Befragung ähnliche Erfahrungen erleben. Die durchschnittlichen Ergebnisse der Valenz-Einschätzung für die eingesetzten negativen, neutralen und positiven Bilder betrugen 1,81 (± 0.54, Standardabweichung (S.D.)), 5,14 (± 0.30), und 7,26 (± 0,73). Die Post-hoc-Intensitätsergebnisse zeigten einen durchschnittlichen Wert von 5,99 (± 0,96) und 3,08 (± 1,05) für emotionale und neutrale Bilder. Die Erinnerungsaufgabe prüfte die Wiedererkennung der Bilder, die während der fMRI-Sitzung gezeigt wurden. Die durchschnittlichen Ergebnisse lagen bei 0,74 (± 0,00) und 0,63 (± 0,00) mit einer Fehlerquote von 0,08 (± 0,02) und 0,06 (±0,01) für emotionale und neutrale Bilder. Diese relativ hohen Wiedererkennungswerte legen nahe, dass die Probanden während der fMRI-Sitzung aufmerksam waren.

Die Bilder wurden mit einem Standard-Ganzkörperscanner (3 Tesla GE VH/1) aufgenommen; die Bildentwicklung und statistische Analyse erfolgte mit dem Programm SPM99. Neben den vier Hauptkonditionen des Experiments wurde die emotionale und neutrale Bildphase nach den Aufmerksamkeitssymbolen moduliert. Obwohl sie nicht den Kernkontrast der Studie beinhalten, wurden die beiden Konditionen moduliert, um eine mögliche Verwechslung der Aufmerksamkeit durch nachfolgende bildbezogene BOLD-Reaktion zu reduzieren (Bermpohl et al. 2006, 590f.). Da das Paradigma nicht zwischen positiver und negativer Valenz während der emotionalen Aufmerksamkeit unterschied, wurde in der Datenanalyse nicht zwischen positiven und negativen Bildphasen unterschieden.

Die parametrischen statistischen Karten wurden zunächst für die beiden Kontraste „Aufmerksamkeit > Wahrnehmung" sowie „Wahrnehmung > Aufmerksamkeit" ermittelt. Im zweiten Schritt wurden die während der emotionalen Kondition aktiven Hirnregionen (Aufmerksamkeit versus Wahrnehmung) identifiziert. Die Gruppen-Effekte „(emotionale Aufmerksamkeit > neutrale Aufmerksamkeit) > (emotionale Wahrnehmung > neutrale Wahrnehmung)" wurden in der Subtraktionsanalyse berechnet, um die speziell während der emotionalen Aufmerksamkeit aktiven Hirnregion zu identifizieren. Um die speziell für die emotionale Bildwahrnehmung aktiven Hirnregionen zu bestimmen, wurde der Kontrast „(emotionale Wahrnehmung > neutrale Wahrnehmung) > (emotionale Aufmerksamkeit > neutrale Aufmerksamkeit)" berechnet. In einem dritten Analyseschritt wurde schließlich bestimmt, welche Hirnregionen bei der emotionalen Aufmerksamkeit und bei der emotionalen Bildwahrnehmung aktiv sind. Hierfür wurde eine randomisierte Konjunktionsanalyse zwischen den Kontrasten „emotionale Aufmerksamkeit > neutrale Auf-

merksamkeit“ und „emotionale Wahrnehmung > neutrale Wahrnehmung“ durchgeführt.[134] [135] Als Kernregion wurden das anteriore Cingulum (BA 24, BA 32), der Dorsolaterale Präfrontale Cortex (DLPFC; BA 46, BA 9), die Amygdala und der Laterale Occipitale Komplex (LOC) bestimmt. Die Regionen anteriores Cingulum und Dorsolateraler Präfrontaler Cortex wurden als Kernregionen definiert, da sie bereits in verschiedenen anderen Aufmerksamkeitsparadigmen Aktivität erkennen ließen (vgl. Ploghaus et al. 1999; Kirsch et al. 2003; Porro et al. 2003; Ueda et al. 2003). Die Region Amygdala wurde miteinbezogen, da sie in aversive und appetitive Konditionierung involviert ist (vgl. Buchel et al. 1998; LaBar et al. 1998; Buchel et al. 1999; Parkinson et al. 2000; Gottfried et al. 2002). Der DLPFC wurde ausgewählt, da seine Implikation bei Aufmerksamkeit und Emotion bereits dokumentiert ist (vgl. z. B. Davidson and Irwin 1999; Carlsson et al. 2000; Ueda et al. 2003). Der LOC repräsentiert schließlich eine perzeptive Hirnregion auf einer höheren Ebene, die als Kontrollregion dient. Die Region wurde gewählt, da ihre Rolle bei der Objektwidererkennung z. B. bei Bilderkennung dokumentiert ist (vgl. Amedi et al. 2002).

### 5.1.2 Abgrenzung emotionaler Aufmerksamkeit

Hinsichtlich der behaviouralen Leistung wurden die Probanden aufgefordert, sofort einen Knopf zu drücken, wenn sie eine Fotografie sehen. Die Reaktionszeiten zeigten einen signifikanten Effekt der Aufmerksamkeit in der ANOVA ($F(1; 13) = 41{,}6$; $P < 0{,}0005$) mit schnelleren Reaktionen in der Aufmerksamkeitskondition. Der durchschnittliche Unterschied zwischen den Konditionen mit und ohne Aufmerksamkeit betrug für alle Probanden 104 Millisekunden bei emotionalen Bildern und 116 Millisekunden bei neutralen Bildern. Ein signifikanter Effekt hinsichtlich des Faktors Emotion oder hinsichtlich der Interaktion zwischen Aufmerksamkeit und Emotion wurde dagegen nicht beobachtet (vgl. Bermpohl et al. 2006, 591f.).

Um die Aufmerksamkeit unabhängig von der Emotion mit der Wahrnehmung zu vergleichen, wurde die neuronale Aktivität während der Aufmerksamkeit mit der neuronalen Aktivität während der Wahrnehmung kontrastiert. Dieser Kontrast zeigte signifikante Aktivität in einem weiten neuronalen Netzwerk, bestehend aus der Midline des Gehirns, einschließlich dem Dorsomedialen Präfrontalen Cortex (BA 9, 10), dem perigenualen und supracallosalen anterioren

[134] Es muss darauf hingewiesen werden, dass die Konditionen der Bildwahrnehmung (emotionale Wahrnehmung, neutrale Wahrnehmung) nur Bilder ohne vorhergehende Aufmerksamkeitsphase umfasste. Bildphasen nach vorhergehender Aufmerksamkeit wurden von den Vergleichen ausgeschlossen, um einen zusätzlichen Effekt einer aufmerksamkeitsbezogenen Aktivität zu vermeiden (vgl. Bermpohl et al. 2006, 591).

[135] Um eine Analyse der Kernregion auf Basis eines „unbiased Kontrast“ zu bestimmen, wurde der Kontrast „alle Konditionen/stimulusinduziert versus Baseline“ bestimmt. Für diesen Kontrast wurde zusätzlich die Baselinekondition als ein Co-Variat moduliert. Um Übernahmeeffekte von der vorhergehenden Bildphyse zu vermeiden, wurde die Baselinekondition über die finalen fünf Sekunden des Fixierungskreuzes nach neutralen Bildern bestimmt (vgl. Bermpohl et al. 2006, 591).

Cingulum (BA 24, 32), dem cingulären motorischen Areal (BA 24), dem supplementärmotorischen Areal (BA 6), dem posterioren Cingulum (BA 23, 29, 30, 31), dem Praecuneus, dem parieto-occipitalen Sulcus (BA 7, 19, 31) sowie dem Thalamus. Außerdem zeigte der Kontrast unterschiedliche Aktivität in lateralen parietalen Regionen (Gyrus angularis und supramarginalis, Lobulus parietalis inferior; BA 39, 40), Gyrus superior und medialis temporalis (BA 21, 22, 38), Gyrus praecentralis (BA 6) sowie dem DLPFC (BA 9, 10). Eine äußerst ähnliche Aktivierung wurde durch den Kontrast „emotionale Aufmerksamkeit > emotionale Wahrnehmung" bestimmt (vgl. Bermpohl et al. 2006, 592). Der umgekehrte Kontrast zeigte unterschiedliche Aktivität im occipitalen und lingualen Gyrus (BA 17, 18, 19), im Gyrus fusiformis (BA 20, 37), im Gyrus temporalis inferior und im Gyrus temporalis medialis (BA 20, 37, der Insula, dem Brocaschen Areal (BA 44, 45), dem Ventrolateralen Präfrontalen Cortex (BA 11, 47), dem Cerebellum, der Amygdala, dem parahippocampalen Gyrus, dem Uncus (BA 28, 34) sowie dem Mittelhirn. Auch hierbei wurde eine äußerst ähnliche Aktivierung durch den Kontrast unter der emotionalen Kondition beobachtet.

Um die Hirnregionen zu bestimmen, die insbesondere während der Aufmerksamkeit unter der emotionalen Kondition aktiviert waren, wurde eine serielle Subtraktion „(emotionale Aufmerksamkeit > neutrale Aufmerksamkeit) > (emotionale Wahrnehmung > neutrale Wahrnehmung)" durchgeführt. Diese Analyse zeigte unterschiedliche Aktivität im supracallosalen anterioren cingulären Cortex (SAC), im cingulären motorirschen Areal (CMA), im parieto-occipitalen Sulcus (POS) und im Gyrus temporalis superior und medialis (vgl. Tabelle 1). Der serielle Subtraktionsterm zeigt möglicherweise Aktivierung, die sich hauptsächlich auf den Kontrast „neutrale Wahrnehmung > emotionale Wahrnehmung" neben der Aktivierung des Kontrasts „emotionale Aufmerksamkeit > neutrale Aufmerksamkeit" bezieht. Um den jeweiligen Beitrag dieser beiden Konstituenten auf den seriellen Subtraktionsterm zu bestimmen, wurde der Term in einem weiteren Schritt weiter zerlegt.

Table 1
(Emotional > neutral expectancy) > (emotional > neutral perception)

| Region (Brodmann area) | Coordinates (MNI) | | | Peak *t* value | # of voxels |
|---|---|---|---|---|---|
| | *x* | *y* | *z* | | |
| Supracallosal anterior cingulate (BA 24, BA 32) | −4 | 12 | 36 | 4.36 | 22 |
| Cingulate motor area (BA 24) | −4 | −8 | 44 | 4.85 | 13 |
| Parieto-occipital sulcus (BA 7, 19, 31) | 0 | −76 | 36 | 4.98 | 17 |
| Superior/middle temporal Gyrus (BA 21, 22) | 52 | −32 | 8 | 4.92 | 29 |
| | −64 | −12 | −4 | 5.00 | 19 |

Peak voxel coordinates, $P < 0.05$ FDR-corrected.

Tabelle 1: Serielle Subtraktion zur Bestimmung der Hirnregionen mit spezifischer Aktivität während der Aufmerksamkeit unter der emotionalen Kondition in den Regionen SAC, CMA, POS und Gyrus temporalis superior und medialis (vgl. Bermpohl et al. 2006, 592).

Der Kontrast „emotionale Aufmerksamkeit > neutrale Aufmerksamkeit“ zeigte Aktivierung im SAC, im CMA (einschließlich dem supplementären und prä-supplementären motorischen Areal) und im POS. Der Kontrast „neutrale Wahrnehmung > emotionale Wahrnehmung“ zeigte Aktivität im linken und rechten Gyrus temporalis superior, was nahe legt, dass die in dieser Region beobachteten Interaktionseffekte sich eher auf Wahrnehmungs- als auf Aufmerksamkeitskonditionen beziehen.[136]

Um die spezielle Aktivität der Hirnregionen während der Bildwahrnehmung unter der emotionalen Kondition zu bestimmen, wurde die umgekehrte serielle Subtraktion vorgenommen: „(emotionale Wahrnehmung > neutrale Wahrnehmung) > (emotionale Aufmerksamkeit < neutrale Aufmerksamkeit)“. Diese Analyse zeigte unterschiedliche Aktivität in der Amygdala, der Insula, dem medialen und lateralen präfrontalen Cortex, dem Putamen, dem Mittelhirn, dem Cerebellum und occipitotemporalen visuellen Regionen (vgl. Tabelle 2).[137]

Table 2
(Emotional > neutral perception) > (emotional > neutral expectancy)

| Region (Brodmann area) | Coordinates (MNI) | | | Peak | # of |
|---|---|---|---|---|---|
| | *x* | *y* | *z* | *t* value | voxels |
| Right amygdala | 24 | −4 | −24 | 4.95 | 9 |
| Left amygdala | −32 | −4 | −24 | 4.45 | 28 |
| Right insula | 40 | 4 | 16 | 5.2 | 9 |
| Medial prefrontal cortex (BA 10) | −16 | 52 | 0 | 3.78 | 9 |
| Left premotor cortex (BA 6) | −52 | −4 | 28 | 4.16 | 22 |
| Right Broca's area (BA 44, 45) | 52 | 28 | 8 | 5.04 | 43 |
| Left Broca's area (BA 44, 45) | −44 | 32 | 8 | 3.93 | 14 |
| Right inferior/middle temporal and fusiform gyri (BA 20, 37) | 52 | −64 | 0 | 7.78 | 214 |
| Left inferior/middle temporal and fusiform gyri (BA 20, 37) | −48 | −60 | −20 | 6.92 | 243 |
| Right occipital and lingual gyri (BA 17, 18, 19) | 8 | −88 | −8 | 7.61 | 527 |
| Left occipital and lingual gyri (BA 17, 18, 19) | −12 | −88 | −8 | 7.47 | 385 |
| Right cerebellum | 24 | −68 | −20 | 6.37 | 203 |
| Left cerebellum | −44 | −48 | −28 | 5.66 | 77 |
| Left putamen | −20 | −4 | 4 | 3.61 | 7 |
| Right midbrain | 20 | −24 | −4 | 4.17 | 8 |

Peak voxel coordinates, $P < 0.05$ FDR-corrected.

Tabelle 2: Umgekehrte serielle Subtraktion zur Bestimmung der Hirnregionen mit spezifischer Aktivität während der Bildwahrnehmung unter der emotionalen Kondition in den Regionen Amygdala, Insula, medialer und lateraler präfrontaler Cortex, Putamen, Mittelhirn, Cerebellum und occipitotemporale visuelle Regionen (vgl. Bermpohl et al. 2006, 593).

[136] Schließlich wurde der serielle Subtraktionsterm exklusiv mit dem Kontrast „neutrale Wahrnehmung > emotionale Wahrnehmung“ kontrastiert, um speziell die Aktivierung hinsichtlich der emotionalen Aufmerksamkeit zu isolieren. Diese Vorgehensweise ist eher konservativ im Vergleich zur oben geschilderten Zerlegung, da es alle Voxel entfernt, die einen Signifikanzwert von P < 0,05 im Maskierungskontrast (Abdecken gemeinsamer Aktivierung) erreichen. Dieser Kontrast zeigte einzig Aktivierung im SAC (vgl. Bermpohl et al. 2006, 592).

[137] Im Maskierungskontrast des seriellen Subtraktionsterms mit „neutraler Aufmerksamkeit > emotionaler Aufmerksamkeit“ wurde eine ähnliche Aktivität beobachtet, wenn auch mit kleineren Clustern in der Amygdala und fehlenden Effekten im medialen präfrontalen Cortex und im Mittelhirn (vgl. Bermpohl et al. 2006, 593).

Während die bisherige Analyse vor allem die Unterschiede zwischen dem Aufmerksamkeits- und Wahrnehmungsnetzwerk untersuchte, wurde im nächsten Schritt bestimmt, ob eine mögliche Überlappung zwischen den bei der Aufmerksamkeit und bei der Wahrnehmung emotionaler Stimuli beteiligten neuronalen Netzwerken bestimmt werden konnte. Hierzu wurde eine Konjunktions-Analyse zwischen dem Aufmerksamkeits-Kontrast (emotionale Aufmerksamkeit < neutrale Aufmerksamkeit) und dem Wahrnehmungskontrast (emotionale Wahrnehmung > neutrale Wahrnehmung) durchgeführt. Aufgrund der Analyse kann von einer Abgrenzung der Netzwerke gesprochen werden, da sie keine überlappenden Voxel bei $P < 0{,}05$ FDR-corrected zeigte. Wurde allerdings die Schwelle auf $P < 0{,}001$ uncorrected verringert, zeigte die Analyse zumindest gemeinsame Aktivität innerhalb des rechten prä-supplementären motorischen Areals sowie im prämotorischen Cortex.

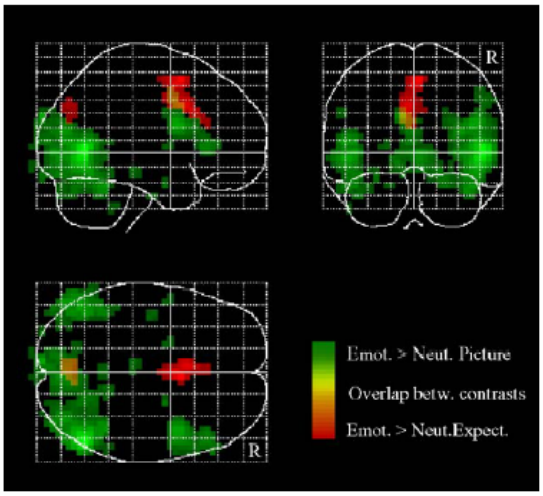
Die bisherige Analyse zeigt, dass insbesondere der SAC während der Aufmerksamkeit unter der emotionalen Kondition aktiviert wurde. Um den zeitlichen Verlauf der Aktivierung der SAC Peakvoxel zu bestimmen, wurden die Zeiteinheiten des BOLD-Signals in 2-Sekunden-Einheiten unterteilt. Der Kontrast wurde gewählt, da er eine Subtraktion des grundsätzlichen Aufmerksamkeitseffekts und die Isolierung des spezifisch emotionalen Aufmerksamkeitseffekts erlaubte. Das Histogramm des zeitlichen Verlaufs des Aufmerksamkeits-Kontrasts zeigte, dass die beobachtete SAC-Aktivierung mit 4-5 Sekunden lange vor dem Erscheinen des Bild-Stimulus erfolgte. Zur weiteren Untersuchung der Aktivitätsmuster des SAC wurden die Kontraste der Parameter im Durchschnitt zu jeder einzelnen Kondition im Vergleich zur Baseline bestimmt. In Übereinstimmung mit den oben genannten Ergebnissen war der Signalanstieg während der emotionalen Aufmerksamkeit größer als bei der neutralen Aufmerksamkeit. Im Vergleich konnte jedoch kein deutlicher Unterschied der Signalintensität zwischen der emotionalen und neutralen Bildwahrnehmung gemessen werden. Ein ausbleibender „emotionaler Effekt" betrifft auch die Wahrnehmung der erwarteten Bilder (vgl. Bermpohl et al. 2006, 594).[138]

[138] In einem nächsten Schritt wurde untersucht, ob die Aktivitätsmuster auch vor dem Hintergrund eines unbiased Kontrast vorhanden war (s. o.). Hierfür wurden die Peakvoxel des anterioren Cingulum in Relation zu dem Kontrast ‚alle Konditionen versus Baseline' gesetzt. Zum Vergleich wurden die Peakvoxel in den drei weiteren Kernregionen DLPFC, Amygdala und Lateraler Occipitaler Komplex (LOC) in denselben unbiased Kontrast gesetzt. Für jeden dieser unbiased Peakvoxel wurde der Kontrast der Parameter im Durchschnitt durch den Vergleich jeder einzelnen Kondition zur Baseline bestimmt. Obwohl der unbiased SAC Peakvoxel etwas mehr seitlich-dorsal bestimmt wurde als im vorher ermittelten biased SAC Peakvoxel, wurde ähnliche Musteraktivität ermittelt. Auch im unbiased SAC Peakvoxel wurde stärkere Aktivität während der emotionalen Aufmerksamkeit als während der neutralen Aufmerksamkeit gemessen. Zugleich wurde kein entscheidender Unterschied zwischen der emotionalen und der neutralen Bild-Kondition ermittelt. Ebenfalls betraf der ausbleibende emotionale Effekt sowohl die Wahrnehmung der erwarteten, als auch der unerwarteten Bilder. Eine unterschiedliche Musteraktivität wurde im DLPFC, in der Amygdala und im LOC beobachtet: In diesen Regionen wurde keine entscheidende Aktivität während der emotionalen, im Vergleich zur neutralen Aufmerksamkeit gemessen. Konsistent zu den vorherigen Ergebnissen zeigten diese Regionen stärkere Aktivität während der emotionalen im Vergleich zur neutralen Bildwahrnehmung. Zusätzlich zeigte die Amygdala stärkere Aktivität während der erwarteten als während der nicht erwarteten emotionalen Bildwahrnehmung (vgl. Bermpohl et al. 2006, 594f.).

### 5.1.3 Diskussion der Ergebnisse

Die vorliegende fMRT-Studie untersuchte die neuronalen Korrelate der Aufmerksamkeit emotionaler Bildstimuli im Vergleich zur Wahrnehmung dieser emotionalen Stimuli. Neutrale Aufmerksamkeit und neutrale Bildwahrnehmung dienten als Kontrollkondition, um diejenigen Hirnregionen zu identifizieren, die im Unterschied zur Wahrnehmung spezifisch während der Aufmerksamkeit unter der emotionalen Kondition aktiv sind. Die Analyse zeigte, dass spezifisch der supracallosale anteriore cinculäre Cortex (SAC), das cinguläre motorische Areal (CMA) und der parieto-occipitale Sulcus während der Aufmerksamkeit unter der emotionalen Kondition als ein „emotionales Aufmerksamkeitsnetzwerk" aktiv sind. Ein weiteres spezifisches neuronales Netzwerk konnte unter der Kondition der emotionalen Bildwahrnehmung identifiziert werden. Dieses „emotionale Bildwahrnehmungsnetzwerk" umfasst eine Reihe Hirnregionen, die bereits in bildgebenden Studien zur emotionalen Wahrnehmung erfasst wurden (vgl. Phan et al. 2002), einschließlich der Amygdala, der Insula, dem medialen und lateralen präfrontalen Cortex, dem Cerebellum und occipitotemporalen Arealen. Unter Anwendung der Konjunktionsanalyse konnte keine Überschneidung dieser beiden Netzwerke beobachtet werden. Deshalb lassen die Ergebnisse zwei getrennte Netzwerke für die Aufmerksamkeit sowie die Wahrnehmung emotionaler Bildstimuli vermuten (vgl. Bermpohl et al. 2006, 595).

Dieses Ergebnis steht in Kontrast zu William James' Hypothese, wonach dieselben Hirnregionen für Antizipation und Perzeption relevant sind (vgl. William James 1892). Verschiedene Studien scheinen diese Position zu stützen (vgl. Carlsson et al. 2000; Breiter et al 2001; Bush et al. 2000; Phan et al. 2002). Die Daten der vorliegenden Studie legen dagegen nahe, dass es weniger oder sogar keine tonische Vor-Aktivierung in den Bereichen der emotionalen Bildprozessierung gibt. Anstelle der antizipatorischen Aktivität im emotionalen perzeptiven Netzwerk wurde die Aktivierung eines separaten Netzwerks während der emotionalen Aufmerksamkeit gemessen. Zusätzlich spiegeln die Ergebnisse getrennter Aktivierungsmuster während verschiedener Phasen des Paradigmas möglicherweise einen Unterschied zwischen antizipatorischen und perzeptiven Komponenten der emotionalen Stimulus-Prozessierung (vgl. Ploghaus et al. 1999; Knutson et al. 2001; O'Doherty et al 2002; Knutson et al. 2003; alle zitiert nach: Bermpohl et al. 2006, 596). Diese abweichenden Ergebnisse der Studien sind gegebenenfalls auf das unterschiedliche Studiendesign mit sicherer bzw. unsicherer emotionaler Aufmerksamkeit zurückzuführen. Insofern zeigt die vorliegende Untersuchung zumindest, dass unsichere emotionale Aufmerksamkeit Hirnregionen (SAC, CMA, POS) umfasst, die von einem emotionalen Wahrnehmungsnetzwerk unterschieden werden müssen (vgl. Bermpohl et al. 2006, 596).




Grafik 12: Unterscheidung zwischen Netzwerken, die während emotionaler Aufmerksamkeit und emotionaler Bildwahrnehmung aktiv waren. Die Kontraste „emotionale Aufmerksamkeit > neutrale Aufmerksamkeit" (rot) und „emotionale Bildwahrnehmung > neutrale Bildwahrnehmung" (grün) wurden in einem so genannten gläsernen Gehirn überlagert. Die Farbe gelb wurde genutzt, um überlappende Kontraste darzustellen. Insgesamt zeigen die drei Projektionen, dass die zwei Kontraste unterschiedliche neuronale Netzwerke umfassen (vgl. Bermpohl et al. 2006, 593).

Die Baselinevergleiche zeigten drei verschiedene Aktivitätsmuster innerhalb des Paradigmas. Zunächst zeigte der SAC unterschiedliche Aktivität während Aufmerksamkeit unter der emotionalen Kondition (Interaktion zwischen Aufmerksamkeit und Emotion). Während der emotionalen Aufmerksamkeit wurde im Vergleich sowohl zur neutralen Aufmerksamkeit als auch zur emotionalen Wahrnehmung größere Aktivität beobachtet. Kein Unterschied wurde zwischen der emotionalen und der neutralen Wahrnehmung gefunden. Daneben zeigten der DLPFC und der LOC unterschiedliche Aktivität während der Bildwahrnehmung unter der emotionalen Kondition (Interaktion zwischen Wahrnehmung und Emotion). Während der emotionalen Bildwahrnehmung wurde im Vergleich sowohl zur neutralen Bildwahrnehmung als auch zur emotionalen Aufmerksamkeit größere Aktivität beobachtet. Kein Unterschied wurde dagegen zwischen der emotionalen und der neutralen Aufmerksamkeit beobachtet. Schließlich zeigte die Amygdala, vergleichbar mit dem DLPFC und dem LOC, unterschiedliche Aktivität während der emotionalen Bildwahrnehmung. Diese Region zeigte zusätzlich einen spezifischen Effekt der emotionalen Aufmerksamkeit während der Phase der Bildwahrnehmung. Im Unterschied zur neutralen Bildwahrnehmung konnte bei der emotionalen Bildwahrnehmung unter vorhergehender emotionaler Aufmerksamkeit größere Aktivität beobachtet werden. Während der Phase der emotionalen Aufmerksamkeit selbst konnte keine ansteigende Aktivität beobachtet werden. Insofern verweisen auch im Baselinevergleich die Ergebnisse auf unterschiedliche Netzwerke für emotionale Aufmerksamkeit und emotionale Bildwahrnehmung. Zusätzlich scheinen unterschiedliche Hirnregionen während unterschiedlicher Phasen der emotionalen Bildwahrnehmung unter dem emotionalen Aufmerksamkeitseffekt beteiligt zu sein: Der SAC zeigt diesen Effekt während der

Aufmerksamkeitsphase und die Amygdala während der Phase der Bildwahrnehmung, während DLPFC und LOC nicht von der emotionalen Aufmerksamkeit direkt betroffen sind.[139]

Die Dekorrelation der aufmerksamkeits- und bildbezogenen BOLD-Aktivität konnte insofern aufgrund des Studiendesigns nur teilweise gewährleistet werden. Somit ergibt sich die grundsätzliche Möglichkeit, dass sich die im Kontrast „emotionale > neutrale Aufmerksamkeit" beobachtete SAC-Aktivität auf die Präsentation erwarteter emotionaler Bilder anstelle der emotionalen Aufmerksamkeitsperiode bezieht (vgl. Bermpohl et al. 2006, 596).[140] Die Ergebnisse der Studie sprechen allerdings gegen eine solche Möglichkeit. Erstens wurden unterschiedliche Netzwerke für die Aufmerksamkeit und die Perzeption emotionaler Stimuli gefunden. Zweitens zeigt das zeitliche Histogramm, dass die SAC-Aktivierung von der emotionalen Aufmerksamkeit induziert wurde, bevor das nachfolgende Bild erschien. Drittens zeigten die Vergleiche mit der Baseline einen Trend in Richtung geringerer und nicht stärkerer Signal-Intensität im SAC hinsichtlich der erwarteten emotionalen Bilder im Vergleich zu den unerwarteten emotionalen Bildern oder den erwarteten neutralen Bildern. Insgesamt legen die Ergebnisse der Studie somit nahe, dass die SAC-Aktivierung in Verbindung mit der emotionalen Aufmerksamkeit nicht wesentlich durch die nachfolgende Bildphase beeinflusst wurde.

Das größte Aktivitätscluster im Zusammenhang mit der emotionalen Aufmerksamkeit im Vergleich zur neutralen Aufmerksamkeit wurde im SAC gemessen. Grundsätzlich gilt diese Region aufgrund verschiedener Läsionsstudien und Studien mit bildgebenden Verfahren als multiintegrative Struktur, die eine Vielzahl an affektiven, kognitiven und motorischen Prozessen in Beziehung zu adaptivem Verhalten umfasst (vgl. Devinsky et al. 1995; Paus 2001). Die Ergebnisse der Studie unterstützen diese Einschätzung, da sie den antizipatorischen Aspekt dieses Prozesses herausstellen (vgl. Bermpohl et al. 2006, 597).

Obwohl der SAC als „kognitive Einheit" im anterioren Cingulum gilt (vgl. Devinsky et al. 1995, Bush et al. 2000), wurden auch verschiedene affektive Funktionen mit dieser Region in Verbindung gebracht: beispielsweise emotionale Aufmerksamkeit ( vgl. z. B. Vuilleumier et al. 2001), Belohnung (vgl. Breiter et al. 1997; Bush et al. 2002) oder Schmerz (vgl. Rainville et al. 1997; Becerra et al. 2001; Singer et al. 2004). Die Ergebnisse dieser Studie legen jedoch nahe, dass

[139] Im Unterschied dazu haben Studien über aversive und appetitive Konditionierung Aktivität der Amygdala mit Bezug auf die konditionierten Stimuli beobachtet (vgl. Buchel et al. 1998; LaBar et al. 1998; Parkinson et al. 2000; Gottfried et al. 2002). Dieser Unterschied in den Ergebnissen kann auf die unterschiedliche Valenz der Erwartungssignale (spezifisch bzw. unspezifisch) zurückgeführt werden (vgl. Bermpohl et al. 2006, 596).

[140] Grundsätzlich besteht für Aufmerksamkeitsstudien eine methodische Herausforderung darin, die BOLD-Aktivität der Signale und der Zielbilder zu unterscheiden. Die Tendenz der ‚Willkür' der Signale für emotionale Aufmerksamkeit war hinsichtlich zweier Aspekte des Paradigmas gegeben: 1) Es wurden mit den Pfeilen unbestimmte emotionale Aufmerksamkeitssignale eingesetzt, die an sich einen möglichen Grad an Unbestimmtheit enthalten. 2) Dem Fixierungskreuz folgten in 25 % der Proben direkt emotionale Bilder als Kontrollkondition mit Bildern ohne vorhergehende emotionale Aufmerksamkeit.

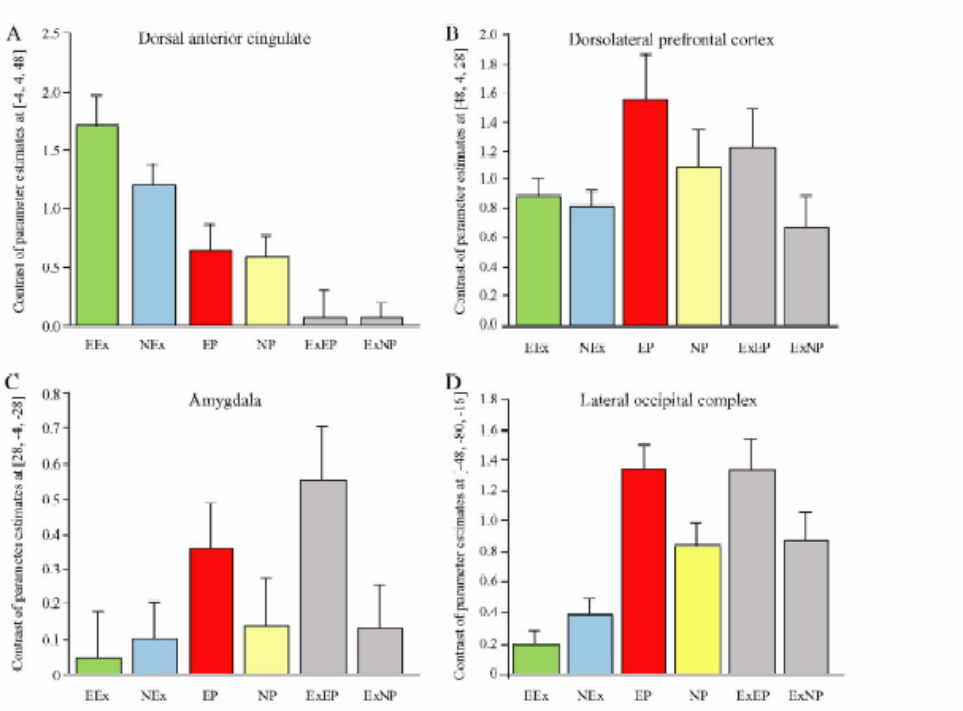
die Prozesse, die durch den SAC vermittelt werden, unabhängig von affektiven Stimuli sind – die bloße Aufmerksamkeit emotionaler Bilder verursachte die SAC-Aktivierung. Zudem war die Reaktion während der Aufmerksamkeitsphase deutlich stärker, als während der eigentlichen Perzeption der emotionalen Bilder. Letztere unterschied sich nicht von der Perzeption neutraler Bilder, in der Tendenz wurde sogar eine geringere SAC-Aktivierung verursacht, als während der neutralen Aufmerksamkeit. Damit unterstützen diese Ergebnisse den antizipatorischen Charakter der Funktion des SAC. Sie lassen eher eine Rolle des SAC bei vorhergehender emotionaler Aufmerksamkeit bzw. Erwartung vermuten als bei Aufmerksamkeitsprozessen, die die tatsächliche Präsenz affektiver Stimuli benötigen.

Aktivität des SAC, der CMA oder des POS wurde bereits in verschiedenen neurowissenschaftlichen Studien hinsichtlich antizipatorischer Angst (vgl. Chua et al. 1999), antizipatorischer Erregung (Crchley et al. 2001) sowie der Erwartung von Belohnung (vgl. Kirsch et al. 2003) und Schmerz (vgl. z. B. Porro et al. 2002; Jensen et al. 2003; Porro et al. 2003) beobachtet. Obwohl diese Studien nicht ausdrücklich die Aufmerksamkeit emotionaler Stimuli untersuchen, scheinen sie diesen Aspekt zumindest implizit mit einzubeziehen. In der vorliegenden Untersuchung konnte gezeigt werden, dass ein antizipatorischer Prozess nicht zwangsläufig klassische Schmerz- oder Belohnungsstimuli voraussetzt, um Aktivität in den Hirnregionen SAC, CMA und im POS hervorzurufen. Diese Regionen werden unter emotionaler Aufmerksamkeit unabhängig von Belohnung oder Schmerz aktiviert. Zudem wurde aufgrund der Datenanalyse deutlich, dass es nicht nur einen Signalanstieg im SAC während der emotionalen, sondern auch während der neutralen Aufmerksamkeit gab. Daraus ergibt sich die Annahme, dass die unter der Aufmerksamkeitskondition gezeigten Pfeile auch einen nicht-spezifischen Aufmerksamkeitseffekt im SAC hervorrufen. Da das Signal während der emotionalen Aufmerksamkeit stärker ist als während der neutralen Aufmerksamkeit, kann geschlossen werden, dass sowohl unspezifische Aufmerksamkeitseffekte, zum Beispiel durch den Hintergrund, als auch spezifische emotionale Aufmerksamkeit zur Aktivierung während der emotionalen Aufmerksamkeit beitragen (vgl. Bermpohl et al. 2006, 597).[141]

Neben affektiven und kognitiven Prozessen scheint der SAC – speziell sein caudaler Teil, das CMA – implizit bei der Prozessierung von motorischen Reaktionen auf behavioural relevante Stimuli zu sein. Aufgrund seiner dichten Verbindung zu primären und sekundären motorischen

[141] Allerdings kann die beobachtete Aktivierung des SAC nicht einfach durch einen nicht-spezifischen Erregungseffekt erklärt werden. Dann müsste man annehmen, dass ein grundsätzlicher Erregungseffekt nicht nur im Aufmerksamkeitskontrast, sondern auch im Wahrnehmungskontrast des Paradigmas Aktivität verursachen würde (vgl. Bermpohl et al. 2006, 27). Aufgrund des Vergleichs „emotionaler Bildwahrnehmung > neutraler Bildwahrnehmung", der keine unterschiedliche Aktivität im SAC zeigte, wird im Rahmen der Studie geschlossen, dass die Aktivität im SAC während emotionaler Aufmerksamkeit keinen generellen Erregungseffekt darstellt (vgl. Bermpohl et al. 2006, 597).

Regionen scheint die Region geeignet, affektive und kognitive Informationen in Handlung zu übersetzen (vgl. Paus 2001). Es wurde gezeigt, dass Aktivität in der CMA nicht Handlungsausübung per se spiegelt, sondern eher den antizipatorischen Zustand, in welchem eine Person sich befindet, um als Reaktion auf einen motivierenden Stimulus eine Handlung auszuwählen (vgl. Woldorff et al. 1999). In der vorliegenden Studie wurde die CMA während der emotionalen Aufmerksamkeit aktiviert, im Vergleich zur neutralen Aufmerksamkeit und emotionalen Stimuluswahrnehmung. Hieraus lässt sich schließen, dass die Aufmerksamkeit emotionaler Bilder auch einen Zustand der Bereitschaft für motorische Reaktion umfasst. Diese Annahme wird aus der beobachteten Co-Aktivierung der SMA und prä-SMA gestützt, die grundsätzlich in solche Bereitschaft und Ausführung von Bewegung involviert sind.




Grafik 13: Die Wirkung in den vier Kernregionen (A) anteriores Cingulum, (B) dorsolateraler präfrontaler Cortex, (C) Amygdala und (D) lateraler occipitaler Komplex. Die Balken repräsentieren den Vergleich zwischen den verschiedenen experimentellen Konditionen und der Baseline. Die Farbsymbolik für die verschiedenen Konditionen entspricht der Zuordnung in Grafik 11. Die Peakvoxel wurden für jede Kernregion bestimmt. EEx: emotionale Aufmerksamkeit, NEx: neutrale Aufmerksamkeit, EP: emotionale Bildwahrnehmung ohne vorhergehende Aufmerksamkeit, NP: neutrale Bildwahrnehmung ohne vorhergehende Aufmerksamkeit, ExEP: emotionale Bildwahrnehmung mit vorhergehender Aufmerksamkeit, ExNP: neutrale Bildwahrnehmung mit vorhergehender Aufmerksamkeit (vgl. Bermpohl et al. 2006, 595).

Neben dem SAC und der SMA wurde der parieto-occipitale Sulcus (POS, einschließlich der medialen Bereiche BA 7 und BA 19, bis hin zu BA 31) in dem Kontrast „(emotionale Aufmerksamkeit > neutrale Aufmerksamkeit) > (emotionale Wahrnehmung > neutrale Wahrnehmung)“ aktiviert. Diese Beobachtung steht in Einklang mit verschiedenen Studien, die beispielsweise die Aktivierung dieser Region während der Aufmerksamkeit von Schmerz untersucht haben (vgl. Buchel et al. 1998; Porro et al. 2003). Das POS kann als anteriorer Teil des

dorsal visuellen Pfades gelten, der von frühen visuellen Regionen hin zum posterioren parietalen Cortex sendet. Wie die CMA wird dieser dorsale Verlauf mit der Prozessierung von Handlung in Verbindung gebracht (vgl. Goodale/Milner 1992; Goodale/Westwood 2004). Insbesondere vermittelt er die benötigte sensorimotorische Transformation für visuellorientierte Handlung. Auch in der vorliegenden Studie wurde der POS während der Aufmerksamkeitsphase aktiviert, wobei keine visuellorientierte Handlung involviert war. Hieraus ergibt sich die Annahme, dass die Aufmerksamkeit motivational relevanter Stimuli einen Zustand der Handlungsbereitschaft implizieren kann. Dies kann antizipatorische Aktivierung im dorsalen Verlauf sogar in Abwesenheit handlungsorientierter visueller Stimulation verursachen. Diese Ergebnisse verweisen zusammengenommen auf den antizipatorischen Charakter der Funktion des SAC. Auch unter Verweis auf vorherige Studien kann angenommen werden, dass diese multi-integrative Region in emotionale Aufmerksamkeit und seinen begleitenden Zustand der Bereitschaft für motorische und reflexhafte Reaktionen in besonders emotionalen Situationen involviert ist (vgl. Bermpohl et al. 2006, 598).

***Quintessenz emotionale Aufmerksamkeit***

- *visuelle oder taktile Antizipation sowie Antizipation von Schmerzen oder Belohnung feststellbar*
- *Messung neuronaler Korrelate zum subjektiven Erleben während der Aufmerksamkeit und der Wahrnehmung emotionaler Bildstimuli*
- *messbare Korrelation neuronaler Aktivität im SAC, in der SMA und im POS während Aufmerksamkeitsphase von Bildstimuli mit unbestimmtem emotionalem Inhalt*
- *Abgrenzung dieses Netzwerks der emotionalen Aufmerksamkeit von spezifischen Regionen für emotionale Bildwahrnehmung*
- *Annahme: Abgrenzung spiegelt Unterteilung zwischen antizipatorischen und perzeptiven Bestandteilen emotionaler Stimulus-Prozessierung wider*

## 5.2 Das Paradigma der emotionalen Beurteilung[142]

Emotionen werden üblicherweise als ein komplexer Prozess aufgefasst, der verschiedene Dimensionen umfasst (vgl. Russel 1980; Lang et al. 1993; Barrett/Russel 1999). Unter anderem umfassen diese Dimensionen die Valenz, die Intensität und das Wiedererkennen eines emotionalen Stimulus. Die Valenz bezieht sich auf die subjektive Erfahrung einer affektiven Qualität des emotionalen Stimulus, unabhängig von seiner sensorischen Modalität (Russell 1980; Lang et al. 1999; Anderson et al. 2003a; Anderson et al.2003b). Beispielsweise wird eine Emotion in einem Kontinuum zwischen positiv und negativ erfahren. Die Intensität bezieht sich darauf, wie stark eine Emotion erfahren wird. Zudem kann eine Emotion innerhalb eines Kontinuums je nach Stärke oder Häufigkeit als ruhig oder aufregend erfahren werden. Das Wiedererkennen umfasst schließlich den Vergleich mit und die Erinnerung an eine vergangene Emotion. Es scheint eine wichtige Strategie zu sein, um den affektiven Gehalt eines emotionalen Stimulus und seiner herausragender Bedeutung für den Probanden zu ermitteln (Mayberg et al. 1999; Lee et al. 2002).

Verschiedene Studien haben bereits gezeigt, dass sich Intensität und Valenz nicht nur psychologisch, sondern auch in ihrer neuronalen Zusammensetzung unterscheiden (vgl. Anderson et al. 2003; Small et al. 2003; Anders et al. 2004; Dolcos et al. 2004). Emotionale Valenz, hervorgerufen durch gustatorische, olfaktorische oder visuelle emotionale Stimuli, wird mit neuronaler Aktivität im Ventromedialen Präfrontalen Cortex (VPMFC) in Verbindung gebracht (vgl. Anderson et al. 2003; Small et al. 2003; Dolcos et al. 2004). Andere Regionen, die mit Valenz zusammenhängen, sind der laterale präfrontale Cortex einschließlich des Ventro- und Dorsolateralen Präfrontalen Cortex (VLPFC, DLPFC) (vgl. Dolcos et al. 2004). Intensität hängt mit neuronaler Aktivität in der Amygdala (vgl. Anderson et al. 2003b; Small et al. 2003) und dem Dorsomedialen Präfrontalen Cortex (DMPFC) zusammen (vgl. Anders et al. 2004; Dolcos et al. 2004). Wiedererkennen schließlich wird in Verbindung mit neuronaler Aktivität im Pericenualen Anterioren Cingulären Cortex (PACC) gebracht. All diese Beobachtungen lassen vermuten, dass verschiedene präfrontale cortikale Regionen verschiedene Dimensionen der emotionalen Prozessierung vermitteln.

Dennoch ist es eine verwirrende Tatsache, dass dieselben präfrontalen cortikalen Regionen auch in kognitive Prozessierung wie z. B. Urteilen oder vorhergehende Aufmerksamkeit verbunden sind. Wenn man verschiedene Ergebnisse hinsichtlich emotionaler und kognitiver Prozessierung vergleicht, scheint eine Überlappung im medialen und im lateralen präfrontalen Cortex zu bestehen (vgl. Nakamura et al. 1999; Ochsner et al. 2002; Lange et al. 2003; Raichle et al. 2001;

[142] Das folgende Kapitel ist eine zusammenfassende Übersetzung der Studie „Segregated neural representation of distinct emotion dimensions in the prefrontal cortex – an fMRI study", die bereits in der Fachzeitschrift NeuroImage veröffentlicht wurde (vgl. Grimm et al. 2006).

Goel and Dolan 2003a; Northoff et al. 2004). Hieraus ergibt sich das Problem, wie emotionale Prozessierung und die damit verbundene kognitive Prozessierung voneinander getrennt werden können. Dies ist notwendig, um die spezifische Korrelation für die emotionalen Dimensionen und die veränderte Aktivität im präfrontalen Cortex zu erhalten. Ziel dieser Studie war es deshalb einerseits, die veränderte Aktivität im präfrontalen Cortex während emotionaler Prozessierung aufzuzeigen, und von der damit verbundenen kognitiven Prozessierung zu trennen. Hierfür wurde die Wahrnehmung mit der Beurteilung emotionaler Bilder verglichen. Um die vorhergehende Aufmerksamkeit zu kontrollieren, wurden die Bildwahrnehmung und die Bildbeurteilung unter einer erwarteten und unerwarteten Kondition verglichen. Die zweite Zielsetzung dieser Studie betraf die Korrelation dieser Ergebnisse mit den Ergebnissen der subjektiven Bewertung der emotionalen Dimensionen Valenz, Intensität und Wiedererkennen durch die Probanden. Hierfür wurden die während der Bildwahrnehmung/-beurteilung erhaltenen T-Werte mit der Post-Scanning Bewertung der Probanden korreliert. Aufgrund bisheriger Untersuchungen wurde eine getrennte neuronale Repräsentation der emotionalen Dimensionen in verschiedenen präfrontalen cortikalen Regionen erwartet. Einerseits wurde angenommen, dass neuronale Aktivität im VMPFC in Zusammenhang mit Valenz während des unerwarteten Bildsehens steht. Außerdem wurde vermutet, dass die Aktivität im VLPFC / DLPFC mit der Valenz und Intensität während der unerwarteten Bildbeurteilung korreliert. Schließlich wurde angenommen, dass das Wiedererkennen der emotionalen Bilder mit Aktivitätsveränderung des PACC während des unerwarteten emotionalen Bildsehens zusammenhängt (vgl. Grimm et al. 2006, 325f.).

Die Ergebnisse der Studie zeigen, dass die emotionalen Dimensionen Valenz, Intensität und Wiedererkennen in verschiedenen präfrontalen cortikalen Regionen getrennt neuronal repräsentiert werden. Außerdem wurde eine inverse Korrelation zwischen medialen und lateralen präfrontalen cortikalen Regionen für die Dimension der Valenz während des Bildsehens und der Bildbeurteilung von emotionalen Bildern beobachtet. Insgesamt verweisen die Ergebnisse der Studie auf getrennte neuronale Repräsentation der verschiedenen emotionalen Dimensionen in verschiedenen cortikalen Bereichen.

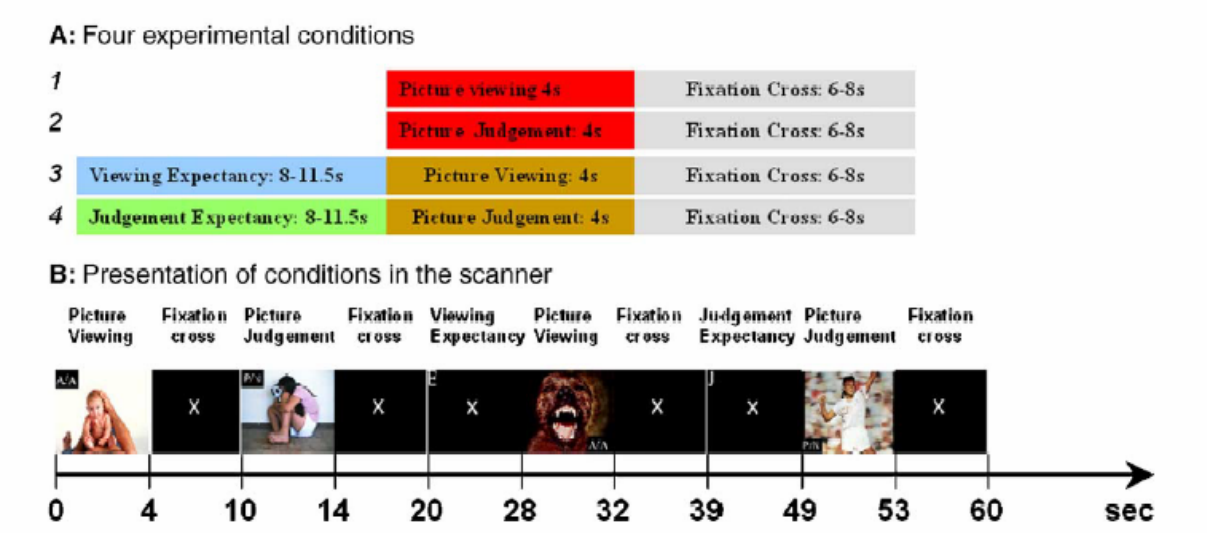
### 5.2.1 Untersuchungsmethode und Datensammlung

Auch in der zweiten Studie wurden den Probanden Bilder des International Affective Picture Systems (IAPS, Lang et al. 1999) gezeigt. Es wurden 29 gesunde Probanden (21 Frauen, 8 Männer) im Alter von 28 bis 56 Jahren ohne bekannte psychiatrische, neurologische oder sonstige Erkrankung untersucht. Für den Versuch wurden Bilder mit positiver (IAPS-Standard 7,32 ± 2,06) und negativer (IAPS-Standard 2,24 ± 2,67) Valenz ausgewählt. Die Bilder wurden sowohl

für die gesamte Versuchsgruppe als auch für jeden einzelnen Probanden hinsichtlich der zwei Valenzkategorien ausgeglichen. Die verwendeten Bilder wurden in zwei Gruppen unterteilt, eine für die Bildwahrnehmung und eine für die Bildbeurteilung. Die IAPS Bilder wurden für eine Dauer von vier Sekunden gezeigt. Bei der Hälfte der Bilder mussten die Probanden beurteilen, ob die Bilder positiven oder negativen Inhalt vermittelten; diese Aufforderung wurde durch die Buchstaben „P/N" in einer der Bildecken induziert. Im Falle einer positiven Beurteilung sollte der rechte Knopf, im Falle einer negativen der linke gedrückt werden. Die Reaktionszeit und die Beurteilung wurden aufgenommen. Bei der anderen Hälfte der Bilder mussten die Probanden lediglich passiv die Bilder sehen; diese Aufforderung wurde durch die Buchstaaben „A/A" in einer der Bildecken induziert. Als Bewegungskontrolle sollten die Probanden in diesem Fall beide Knöpfe drücken. Eine nicht-emotionale, neutrale Kontroll-Kondition wurde in dieser weiteren Studie nicht durchgeführt, da der Fokus des Paradigmas im Vergleich zwischen der Wahrnehmung und der Beurteilung emotionaler Bilder lag (vgl. Grimm et al. 2006, 326). Ferner wurden 50 % der Bilder in einer unerwarteten Kondition gesehen und beurteilt; ihnen ging als Baseline lediglich ein Fixierungskreuz voran. Den anderen 50 % der Bilder ging eine Aufmerksamkeitsphase von 8-11,5 Sekunden voran, die die Aufgabenstellung hinsichtlich des nachfolgenden Bildes bestimmte. Die vorhergehende Aufmerksamkeitsphase diente als Kontrolle der vorhergehenden Aufmerksamkeit auf das Bildsehen bzw. die Bildbeurteilung. Diese Phase wurde durch die Präsentation eines weißen Fixierungskreuzes vor dunklem Hintergrund und einem Buchstaben in einer Bildecke induziert: „J" induziert die Aufmerksamkeit für eine nachfolgende Beurteilungsaufgabe, der Buchstabe „E" induzierte die Aufmerksamkeit für das nachfolgende Bildsehen – die bloße Wahrnehmung (vgl. Kastner et al. 1999; Driver/Frith 2000).

Nach der Bildpräsentation wurde ein Fixierungskreuz für 6-8 Sekunden gezeigt, um den Probanden ein Erholungsintervall von den emotionalen Stimuli zu ermöglichen und als Baseline-Kondition zwischen Aktivierung und Deaktivierung zu unterscheiden (vgl. Stark/Squire 2001). Insgesamt wurden 159 Aufgaben in sechs Durchläufen präsentiert. Die verschiedenen IAPS Bildtypen und Seh-/Beurteilungsaufgaben wurden pseudorandomisiert in sechs Durchläufen präsentiert. Vor der Durchführung des Experiments wurden die Probanden durch einen Testlauf mit dem Paradigma vertraut gemacht (vgl. Grimm et al. 2006, 326). Es wurde ferner die Reaktionszeit zwischen dem Erscheinen eines Bildes und dem folgenden Knopfdruck gemessen. Die Messung erfolgte für die Phasen der Bildwahrnehmung und der Bildbewertung sowohl mit als ach ohne vorhergehende Aufmerksamkeitsphase getrennt. Die im Scanner gefällten positiven bzw. negativen Beurteilungen wurden aufgezeichnet und mit den Ergebnissen der Post-hoc-Evaluation der Valenz verglichen (vgl. ebd.) Im Unterschied zu dem vorhergehenden Paradigma „emotionale Aufmerksamkeit" wurde die subjektive Bewertung bzw. Skalierung der Bilder durch die Probanden direkt im Anschluss an die fMRI-Sitzung durchgeführt. Jedem Bild folgte eine Aufgabenpha-

se, in der die emotionale Intensität und die Valenz bewertet wurden sowie ein Wiedererkennungstest. Im Rahmen einer visuellen Analogskala wurde die Valenz in einem Kontinuum von „sehr negativ“ (1) bis „sehr positiv“ (9), die emotionale Intensität in einem Kontinuum von „gering“ (1) bis „hoch“ (9) sowie die Wiedererkennung von „definitiv nicht erinnert“ (1) bis „definitiv erinnert“ (9) gemessen[143]. Die Post-Scanning Messung der Valenz wurde in entweder positiv oder negativ klassifiziert und anschließend durch den T-Test mit der intra-Scanning-Messung verglichen. Für jeden Probanden wurde abschließend der Durchschnitt der drei Bewertungen berechnet.




Grafik 14: (A) In dem fMRT-Paradigma wurden die emotionale Bildwahrnehmung und die emotionale Bildbeurteilung unterschieden. Die Konditionen (1) und (2) zeigen die Bildwahrnehmung und -beurteilung ohne vorhergehende Aufmerksamkeitsphase. Ihnen folgte bzw. ging ein Fixierungskreuz als Baselinekondition voran. Die Konditionen (3) und (4) zeigen die Bildwahrnehmung und -beurteilung mit vorhergehender Aufmerksamkeitsphase. Die Aufmerksamkeitsphase induzierte entweder Bildwahrnehmung oder Bildbeurteilung, wobei offen blieb, ob ein positives oder negatives Bild folgte. (B) In dem Beispiel einer Stimulus- bzw. Bildphase werden die Konditionen Bildwahrnehmung und Bildbeurteilung ohne vorhergehende Aufmerksamkeit sowie Bildwahrnehmung und Bildbeurteilung mit vorhergehender Aufmerksamkeit dargestellt. In allen Fällen wurde die Aufmerksamkeitsphase durch ein Fixierungskreuz und einen Buchstaben induziert. Buchstabe E kündigte die nachfolgende Bildwahrnehmung an; Buchstabe J kündigte die nachfolgende Bildbeurteilung an. Bilder unter der Beurteilungskondition wurden mit den Buchstaben P/N markiert, damit die Probanden eine positive oder negative Beurteilung durchführten (vgl. Grimm et al. 2006, 328).

Die Messungen wurden mit einem Philips Intera 3 T Ganzkörper MR-Einheit, die mit einem Achtkanal Philips SENSE Kopfkeule ausgestattet ist, durchgeführt. Die funktionalen Zeitserien wurden mit einer „sensitivity encoded single-shot echo-planar Sequenz“ ermittelt (vgl. Pruessmann et al., 1999). Für die Datenanalyse wurden Cluster ausgewählt, die einen signifikanten Effekt ($P < 0{,}001$, uncorrected) von 10 oder mehr fortlaufenden Voxeln anzeigten. Für jeden Probanden wurde eine Designmatrix definiert, die das unerwartete und erwartete Bildsehen sowie die unerwartete und erwartete emotionale Bewertung als getrennte Ereignisse modulierte. Zusätz-

[143] Für die Messung der Wiedererkennung wurden zu einer Auswahl von 151 bereits während des Scans verwendeten Bildern zusätzlich 60 neue Bilder hinzugezogen, die ebenfalls hinsichtlich der Kriterien Valenz, Intensität und Dominanz ausgewählt wurden.

lich wurde die Baseline (6-8 sek.) als separates Ereignis in die Designmatrix mit einbezogen. Es wurde eine variable Dauer (6,0; 6,5; 7,0; 7,5; 8,0 sek.) der Baseline moduliert, die eine explizite Anwendung variabler Intervalle in der Datenanalyse ermöglicht. Nach der Berechnung aller Modellparameter wurden spezifische Effekte durch die Anwendung geeigneter linearer Kontraste für jede Kondition getestet, die in einer T-Statistik für jeden Voxel resultieren. Die fMRT-Analyse fokussierte dabei zwei Aspekte: Einerseits diente die Analyse dazu, die Regionen aufzuzeigen, die speziell mit Bildsehen zusammenhängen, unabhängig von kognitiver Prozessierung wie z.B. Beurteilung. Andererseits wurden die erhaltenen Daten mit den subjektiven Post-Scanning-Bewertungen der emotionalen Dimensionen Valenz, Intensität und Wiedererkennung korreliert. Die Daten wurden hinsichtlich der Effekte der Bildwahrnehmung analysiert. Bildwahrnehmung und Bildbewertung unter der unerwarteten Kondition wurden miteinander verglichen, um die Effekte der Bildbetrachtung aufzuzeigen, während die Effekte der Beurteilung kontrolliert wurden. In einem nächsten Schritt wurden die Bildbetrachtung und -beurteilung innerhalb der Aufmerksamkeitskondition verglichen, um die Aufmerksamkeitseffekte zu kontrollieren (vgl. Grimm et al. 2006, 328). Um zu bestimmen, ob bei den Probanden überlappende oder abtrennbare neuronale Netzwerke während der unerwarteten Bildbetrachtung und der erwarteten Bildbetrachtung aktiviert wurden, wurden über eine Random-Effekt-Konjunktionsanalyse der beiden Kontraste die jeweils aktiven Regionen für beide Konditionen bestimmt (vgl. Price/Friston 1997). Der exklusive Maskierungskontrast wurde auf die Kernkondition, die unerwartete Bildwahrnehmung angewendet, um Voxel zu eliminieren, die unter der erwarteten Kondition aktiviert waren. Der Kontrast zwischen unerwarteter Bildwahrnehmung und unerwarteter Bildbeurteilung wurde mit dem Kontrast der erwarteten Bildwahrnehmung und der erwarteten Bildbeurteilung abgedeckt. Der Signifikanzlevel der Maskierungen wurde bei $P < 0,05$ angesetzt. Schließlich wurde der Durchschnittswert jeder funktionalen Signalveränderung der relevanten Voxel über die 29 Probanden hinweg für jede einzelne der vier Konditionen getrennt bestimmt. Die Kernvoxel ergaben sich aus dem Kontrast zwischen „unerwarteter Bildwahrnehmung > unerwarteter Bildbeurteilung“ und „erwarteter Bildwahrnehmung > erwarteter Bildbeurteilung“.

In einem weiteren Analyseschritt wurden die Werte der individuellen Bildquoten für Valenz, Intensität und Wiedererkennung als ein Regressor in die Designmatrix für alle 29 Probanden miteinbezogen. Durch diese Korrelation konnte die Relation zwischen den T-Werten der Bildwahrnehmung bzw. Bildbeurteilung und den Bewertungen der Probanden für emotionale Valenz, Intensität und Wiedererkennung durchgeführt werden. Der Grenzwert für eine signifikante Korrelation wurde bei $P < 0,001$ uncorrected angesetzt. Um schließlich den valenzspezifischen Effekt für die Korrelationen zu ermitteln, wurde eine separate Korrelationsanalyse für positive und negative Bildwerte mit T-Werten für unerwartete Bildwahrnehmung, erwartete Bildwahrnehmung, unerwartete Bildbeurteilung und erwartete Bildbeurteilung durchgeführt. Es wur-

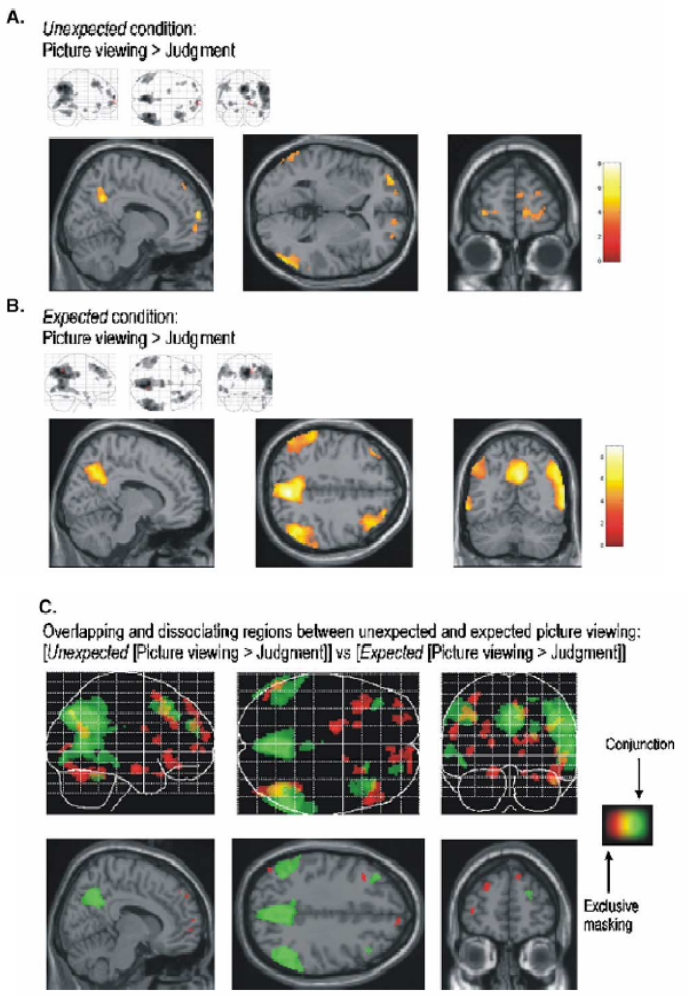
de jedoch kein valenzspezifischer Effekt für Intensität oder Wiedererkennung beobachtet. Deshalb wurde die valenzspezifische Korrelation lediglich für Valenz durchgeführt (vgl. Grimm et al. 2006, 330).

### 5.2.2 Unterscheidung neuronaler Repräsentationen

Der Knopfdruck während des fMRT-Scans diente dazu, die Entscheidungen über die emotionalen Bilder zu prüfen, wobei als Kontrolle beide Knöpfe beim bloßen Bildsehen ohne Entscheidung gedrückt wurden. Die Reaktionszeiten zeigten einen marginalen signifikanten Effekt bei den vier Konditionen: Die Reaktionszeiten während unerwarteten Bildsehens (1,58 +- 0.58 sek) waren signifikant schwächer als während unerwarteter Bildbeurteilung (1,67 +- 0,79 sek). Die Reaktionszeiten während erwarteten Bildsehens (1,57 +- 1,62 sek) waren ebenfalls signifikant schwächer als während erwarteter Bildbeurteilung (1,62 +- 0,61 sek). Außerdem wurde eine signifikante Differenz der Reaktionszeiten zwischen positiven (1,53 +- 1,51 sek) und negativen Bildern (1,64 +- 0,62 sek) gemessen. Die Post-Scanning Sitzung erlaubte die Untersuchung der Wirkung des Bildsehens und -beurteilens auf die subjektive Bewertung der emotionalen Valenz und Intensität. Die Bewertung der Valenz zeigte, dass die in dem Paradigma verwendeten negativen und positiven Bilder als solche erlebt wurden. Zusätzlich zeigten die Valenz-Bewertungen einen signifikanten Unterschied zwischen den Bildern in den vier Konditionen: Die Bewertung für erwartetes Bildsehen (5,24 +- 3,43 sek) waren positiver als für erwartete Bildbeurteilung (4,71 +- 3,46 sek). Obwohl kein signifikanter Level erreicht wurde, konnte ein ähnliches Ergebnis unter der unerwarteten Kondition beobachtet werden: unerwartete passiv gesehene Bilder wurden positiver bewertet als unerwartete beurteilte Bilder (vgl. Grimm et al. 2006, 330). Die Bewertung der Intensität zeigte, dass die in dem Paradigma als negativ und positiv klassifizierten Bilder mit derselben Intensität erlebt wurden. Die meisten Intensitäts-Bewertungen für positive und negative Bilder lagen bei 7,65 (SD 1,81) und 7,64 (SD 1,87). Im Unterschied zur Valenz zeigten die Intensitätsbewertungen keine signifikante Differenz zwischen den Bildern unter den vier Konditionen im fMRT. Die Bewertung der Wiedererkennung zeigte eine signifikante Differenz zwischen den Bildern unter den vier Konditionen. Erwartete beurteilte Bilder wurden stärker erinnert als lediglich erwartete wahrgenommene Bilder. Allerdings wurden weder bei unerwarteter Bildwahrnehmung und Bildbeurteilung, noch bei positiven und negativen Bildern signifikante Unterschiede bei der Bewertung der Wiedererkennung beobachtet (vgl. ebd.).

Um die Effekte des Bildsehens in der unerwarteten Kondition zu messen, wurde die Hirnaktivität während der unerwarteten Wahrnehmung emotionaler Bilder mit der Hirnaktivität während der unerwarteten Bildbeurteilung verglichen. Der Kontrast zeigte verschiedene Aktivität

besonders im VMPFC, DMPFC und im posterior-cingulären Cortex/Praecuneus. Weitere Regionen waren aktiv, einschließlich der bilateralen Amygdala, dem bilateralen Dorso- und Ventrolateralen Präfrontalen Cortex (DLPFC, VLPFC), dem bilateralen parietalen Cortex und der rechten anterioren Insula. Der umgekehrte Kontrast zwischen unerwarteter Bildbeurteilung und unerwarteter Bildwahrnehmung zeigte Aktivität im dorsalen anterioren cingulären Cortex, im medialen prämotorischen Cortex und im supplementärmotorischen Areal (SMA), im bilateralen DLPFC, im bilateralen Thalamus, im rechten und im medialen occipitalen Cortex sowie im Cerebellum (vgl. ebd., 330f.).




Grafik 15: (A) Unerwartete Bildwahrnehmung: Gezeigt wird die maximale Intensität der signifikanten Aktivierungscluster sowie die statistischen übereinander gelegten parametrischen Karten für ein normalisiertes Gehirn eines Probanden mit den Koordinaten [12, 62, -2]. (B) Erwartete Bildwahrnehmung: Gezeigt wird die maximale Intensität der signifikanten Aktivierungscluster sowie die statistischen übereinander gelegten parametrischen Karten für ein normalisiertes Gehirn eines Probanden mit den Koordinaten [12, -62, 42]. (C) Überlappende (grün) und getrennte (rot) Regionen in Aktivierungsclustern innerhalb der erwarteten wie unerwarteten Bildwahrnehmung wird für die maximale Intensität sowie für standardisierte Hirnregionen mit den Koordinaten [10, 54, 38] gezeigt (vgl. Grimm et al. 2006, 331).

Um die Effekte des Bildsehens in der erwarteten Kondition zu messen, wurde die Hirnaktivität während des erwarteten Bildsehens mit der Hirnaktivität während der erwarteten Bildbeurteilung verglichen. Der Kontrast zeigte verschiedene Aktivität besonders im Supragenualen Anterioren

Cingulären Cortex (SAC) und im Posterio ren Cingulären Cortex bzw. Präcuneus (PACC). Im Unterschied zur unerwarteten Kondition wurde keine Aktivität mehr in Regionen wie dem VMPFC und dem DMPFC beobachtet. Weitere Regionen waren aktiv, einschließlich dem bilateralen DLPFC und VLPFC, dem bilateralen parietalen Cortex, dem rechten temporalen Lobulus, dem Mittelhirn und der rechten Amygdala. Der konverse Kontrast zwischen erwarteter Bildbeurteilung und erwarteter Bildwahrnehmung zeigte Aktivität im bilateralen DLPFC, in der bilateralen anterioren Insula sowie im bilateralen occipitalen Cortex (vgl. ebd., 331f.).

In einem nächsten Schritt wurde untersucht, ob bei den Probanden überlappende oder getrennte neuronale Netzwerke während des unerwarteten und erwarteten Bildsehens aktiviert waren. Zu diesem Zweck wurde eine Konjunktionsanalyse zwischen dem Kontrast „(unerwartete Bildwahrnehmung > erwartete Bildwahrnehmung) > (unerwartete Bildbeurteilung > erwartete Bildbeurteilung)" durchgeführt. Die Konjunktionsanalyse zeigte überlappende Voxel bei $P < 0.05$ FDR-corrected im posterioren cingulären Cortex bzw. Praecuneus, dem bilateralen DLPFC, dem rechten VLPFC und dem bilateralen parietalen Cortex. Keine überlappenden Voxel wurden in anterioren medialen cortikalen Regionen, einschließlich VMPFC, SAC und DMPFC beobachtet. Diese Trennung zwischen den anterioren medialen cortikalen Regionen wurde zusätzlich durch den Maskierungskontrast „(unerwartete Bildwahrnehmung > unerwartete Bildbeurteilung) > (erwartete Bildwahrnehmung > erwartete Bildbeurteilung)" bestätigt (vgl. ebd., 332).

Um die behaviourale Relevanz zu untersuchen, wurden die fMRT-Ergebnisse mit den behaviouralen Daten korreliert. Emotionale Valenz zeigte eine positive Korrelation mit den T-Werten im VMPFC während des unerwarteten Bildsehens: je positiver die Valenz, desto höher war der T-Wert im VMPFC. Um den Einfluss der Aufmerksamkeit auf die Korrelation zu untersuchen, wurde dieselbe Analyse mit dem Kontrast „erwartete Bildwahrnehmung > erwartete Bildbeurteilung" durchgeführt. Diese Analyse zeigte keine signifikante Korrelation mehr im VMPFC, sondern nur im linken DLPFC. Anschließend wurde eine getrennte Korrelationsanalyse zwischen positiv und negativ bewerteten Bildern während der unerwarteten Bildwahrnehmung und der erwarteten Bildwahrnehmung durchgeführt. Positive Bilder zeigten unter beiden Konditionen eine signifikante Korrelation im DLPFC. Negative Bilder zeigten eine signifikante Korrelation mit dem T-Wert in der rechten Amygdala während unerwarteter Bildwahrnehmung; während der erwarteten Bildwahrnehmung wurde jedoch keine signifikante Korrelation beobachtet. Weder für die unerwartete noch die erwartete Bildwahrnehmung konnte eine signifikante Korrelation im VMPFC oder einer anderen medialen cortikalen Region gemessen werden. Um die Valenz-Bewertung mit Hirnaktivität während der Beurteilung zu vergleichen, wurden die subjektiven Bewertungen mit dem Kontrast „unerwartete Bildbewertung > unerwartete Bildwahrnehmung" korreliert: zwar konnte eine signifikante Korrelation zwischen der emotionalen Valenz und den T-Werten im bilateralen DLPFC gezeigt werden, aber nicht im VMPFC oder einer ande-

ren medialen cortikalen Region. Dagegen wurde eine negative Korrelation beobachtet: je positiver die emotionale Valenz bewertet wurde, desto geringer war der T-Wert im bilateralen DLPFC. Keine signifikante Korrelation konnte dagegen für die Phase der erwarteten Bildbeurteilung beobachtet werden. Schließlich wurde eine getrennte Korrelationsanalyse zwischen positiv und negativ bewerteten Bildern während der Beurteilung durchgeführt. Positive Bilder korrelierten signifikant negativ mit T-Werten im linken DLPFC, während negative Bilder signifikant negativ mit T-Werten im rechten DLPFC bzw. VLPFC korrelierten. Auch hierbei konnte keine signifikante Korrelation für die Phase der erwarteten Bildbeurteilung beobachtet werden. Im Unterschied zu der Valenz korrelierte die Intensitäts-Bewertung zwar nicht signifikant mit einer medialen präfrontalen Region während der unerwarteten Bildwahrnehmung oder Bildbeurteilung, aber sie korrelierte signifikant mit den T-Werten im medialen Cerebellum während der unerwarteten Bildwahrnehmung. Während der Phase der unerwarteten Bildbeurteilung zeigte die Intensität eine signifikante positive Korrelation mit den T-Werten im rechten VLPFC: je intensiver die Emotion erlebt wurde, desto höher war der T-Wert im rechten VLPFC. Eine signifikante Korrelation der Intensität mit neuronaler Aktivität während der Phase der erwarteten Bildwahrnehmung wurde nicht gefunden, dagegen korrelierte Intensität signifikant mit den T-Werten im Nucleus accumbens und DMPFC während der Phase der erwarteten Bildbeurteilung. Die Wiedererkennung schließlich korrelierte signifikant mit den T-Werten im PACC und im rechten DLPFC während der unerwarteten Bildwahrnehmung. Daneben wurde eine negative Korrelation beobachtet: je weniger Gegenstände wiedererkannt wurden, desto höher waren die T-Werte im PACC. Eine signifikante Korrelation während der unerwarteten Bildbeurteilung wurde dagegen nicht beobachtet. Während die Phase der erwarteten Bildwahrnehmung eine signifikante Korrelation der Wiedererkennung mit den T-Werten im rechten DLPFC und der rechten Amygdala zeigte, konnte für die Phase der erwarteten Bildbeurteilung eine signifikante Korrelation der Wiedererkennung mit dem SAC und dem linken parietalen Cortex festgestellt werden (vgl. ebd., 333f.).

### 5.2.3 Diskussion der Ergebnisse

Innerhalb der vorliegenden Studie wurde die Beteiligung präfrontaler cortikaler Regionen bei der Prozessierung emotionaler Dimensionen untersucht. In einem ersten Schritt wurden verschiedene präfrontale cortikale Regionen während der Wahrnehmung emotionaler Bilder unabhängig von kognitiver Prozessierung wie Beurteilung oder vorhergehender Aufmerksamkeit beobachtet. Diese Ergebnisse wurden dann mit den Post-Scanning-Ratings der Probanden zu den unterschiedlichen emotionalen Dimensionen Valenz, Intensität und Wiedererkennung korreliert. Die Ergebnisse bestätigen die Eingangshypothese und zeigen getrennte neuronale Repräsentation für

Valenz, Intensität und Wiedererkennen in verschiedenen präfrontalen cortikalen Regionen. Außerdem wurde eine konverse Korrelation bei der Valenz mit neuronaler Aktivität in medialen und lateralen präfrontalen cortikalen Regionen gemessen. Diese Ergebnisse unterstützen verschiedene neuere Beobachtungen einer reziproken Modulation zwischen diesen Regionen während emotional-kognitiver Interaktion. Sie sind außerdem konsistent mit psychologischen Theorien einer zugrunde liegenden emotionalen Struktur, wonach emotionale Prozessierung ein Ergebnis der Aktivität verschiedener Dimensionen ist (vgl. Russel 1980; Lang et al. 1993; Barrett/Russell 1999).

Aktivierung im VMPFC korrelierte positiv mit Valenz während unerwarteter Bildwahrnehmung. Je positiver die – sowohl negativen als auch positiven – emotionalen Bilder bewertet wurden, desto mehr Aktivierung wurde im VMPFC beobachtet. Dieses Ergebnis wird prinzipiell durch eine ähnlichen Studie gestützt, bei der die positive Relation zwischen der Valenz olfaktorischer bzw. gustatorischer Stimuli und neuronaler Aktivität im medialen orbitofrontalen Cortex beobachtet wurde (vgl. Anderson et al. 2003b; Small et al 2003). Eine weitere Studie konnte jedoch keine Korrelation der neuronalen Aktivität im VMPFC mit Valenz feststellen, was möglicherweise daran liegt, dass hierbei psychophysiologische Parameter miteinbezogen wurden (vgl. z. B. Anders et al. 2004). Dies lässt vermuten, dass der VMPFC emotionale Valenz unabhängig von psychophysiologischer Prozessierung repräsentiert. So wird eine Korrelation zwischen dem VMPFC und der Valenz unter einer emotionalen Kondition gezeigt. Diese scheint unabhängig von der sensorischen Modalität, in welcher die Emotion induziert wird, zu bestehen, da sie bei verschiedenen sensorischen Modalitäten auftritt. Die Ergebnisse ergänzen aufgrund der Beobachtung der kognitiven Prozessierung von Beurteilung und vorhergehender Aufmerksamkeit bisherige Studien, die mit der Prozessierung von emotionaler Valenz in Verbindung stehen. Weder während der erwarteten Bildwahrnehmung noch während der unerwarteten und erwarteten Bildbeurteilung wurde eine signifikante Korrelation zwischen dem VMPFC und der Valenz gemessen. Daraus lässt sich schließen, dass eine Valenz-abhängige Aktivierung des VMPFC nicht auf die damit verbundenen kognitiven Prozesse wie Beurteilung oder Aufmerksamkeit zurückgeführt werden kann. Neuronale Aktivität scheint speziell emotionale Valenz unabhängig von kognitiver Prozessierung zu kodieren. Vergleichbare Metaanalysen zeigten, dass emotionale Aufgaben ohne kognitive Anforderung ebenfalls Aktivität im VMPFC induzierten (vgl. Duncan/Owen 2000; Pan et al. 2002; Steele/Lawrie 2004).

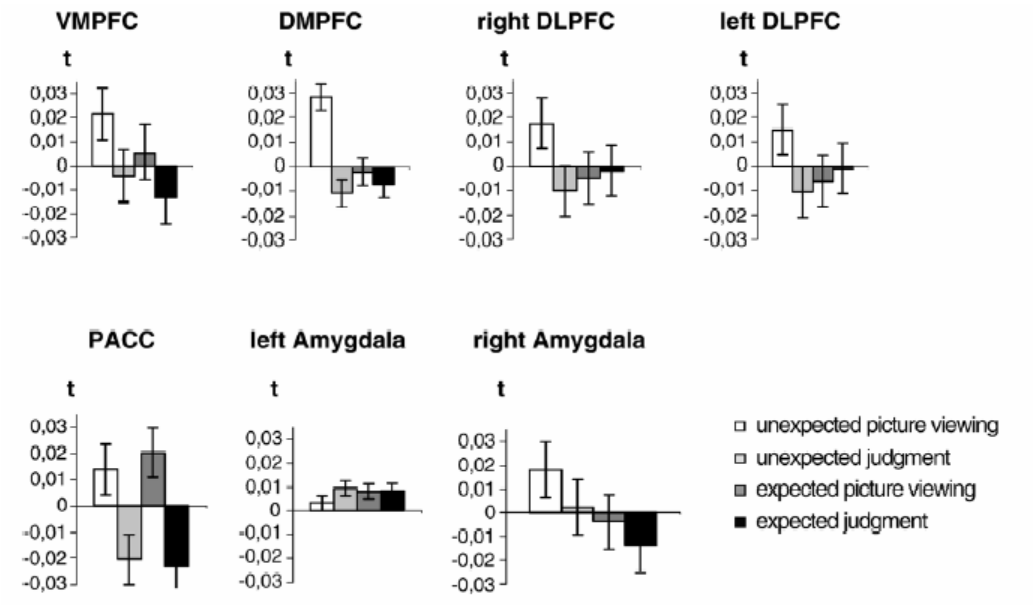
Allerdings konnte keine spezifische regionale Korrelation für positive wie negative emotionale Valenz mit dem VMPFC während der Bildwahrnehmung gemessen werden. Dieses Ergebnis wird durch vergleichbare Metaanalysen gestützt, die die Beteiligung des VMPFC unabhängig von positiver oder negativer Emotion beobachteten (vgl. Phan et al. 2002; Murphy et al. 2003; Steele/Lawrie 2004). Diesen Ergebnissen zufolge scheint neuronale Aktivität im VMPFC speziell den

intrinsischen affektiven Wert bzw. die Valenz emotionaler Stimuli zu kodieren unabhängig davon, ob sie positiv oder negativ sind. Diese Spekulation konnte derzeit nicht bestätigt werden, auch wenn verschiedene Läsionsstudien diese Hypothese zu unterstützen scheinen (vgl. z. B. Hornak et al. 2003; Berlin et al. 2004). Zusammengenommen scheinen die Ergebnisse nahezulegen, dass mediale Regionen im orbitofrontalen und ventralen präfrontalen Cortex spezifisch den Aspekt der subjektiven Erfahrung der emotionalen Prozessierung vermitteln (vgl. Grimm et al. 2006, 335).

In Ergänzung zum VMPFC wurde die Valenz mit Signalveränderungen im bilateralen DLPFC korreliert. Im Unterschied zum VMPFC wurde diese Korrelation nicht während des (unerwarteten) Bildsehens, sondern während der (unerwarteten) Bildbeurteilung beobachtet. Diese Relation wird auch in verschiedenen Studien bezüglich des DLPFC und emotionaler Beurteilung verglichen mit nicht-emotionaler Beurteilung beobachtet (vgl. Nakamura et al. 1999; Hariri et al 2000; Gorno-Tempini et al. 2001; Keigthley et al. 2003; Lange et al. 2003; Northoff 2004). Die Ergebnisse zeigten eine parametrische Abhängigkeit der neuronalen Aktivität im DLPFC von emotionaler Valenz. Aufgrund der Bildbeurteilung innerhalb des Paradigmas konnte gezeigt werden, dass Valenz nicht nur mit dem VMPFC zusammenhängt, sondern auch mit dem DLPFC (vgl. Dolcos et al. 2004). Die Beurteilung emotionaler Bilder scheint von einer wachsenden Aufmerksamkeit begleitet zu werden, die ebenfalls mit dem DLPFC in Verbindung steht (vgl. Taylor et al. 2003). Allerdings bleibt unklar, ob die Korrelation zwischen Valenz und dem DLPFC mit der Beurteilung oder der wachsenden Aufmerksamkeit zusammenhängt. Die Tatsache, dass keine signifikante Korrelation zwischen dem DLPFC und der Valenz während erwarteter Beurteilung beobachtet wurde, ist ein Hinweis darauf, dass die Valenz-abhängige Aktivität im DLPFC eher mit Beurteilung als mit vorhergehender Aufmerksamkeit zusammenhängt.

Psychologisch beinhaltet die Valenz eines Stimulus einen affektiven und einen evaluativen Aspekt (Russell 1980; Barrett/Russell 1999). Der affektive Aspekt bezieht sich auf die Qualität des Stimulus, während der evaluative Aspekt die Klassifikation des affektiven Stimulus als negativ oder positiv betrifft. Die Ergebnisse der Studie unterstützen diese Unterscheidung in neuronaler Hinsicht: Der affektive Aspekt kann mit neuronaler Aktivität im VMPFC verbunden werden; dies zeigt sich in der Korrelation während des unerwarteten Bildsehens. Der evaluative Aspekt hingegen hängt eher von neuronaler Aktivität im DLPFC ab, wie die Korrelation während der unerwarteten Bildbeurteilung nahelegt. Entsprechend lassen die Ergebnisse vermuten, dass die psychologische Unterscheidung zwischen affektiven und evaluativen Aspekten der Valenz mit neuronaler Repräsentation in unterschiedlichen präfrontalen cortikalen Regionen, z. B. dem VMPFC und dem DLPFC, korrespondiert. Die funktionale Unterscheidung zwischen VMPFC und DLPFC wird durch die Ergebnissen der valenz-spezifischen Korrelation unterstrichen. Im Gegensatz zum VMPFC wurden positive und negative valenz-spezifische Korrelation im linken

und rechten DLPfC während unerwarteter Bildbeurteilung gemessen. Die Signalveränderungen im rechten DLPFC korrelierten signifikant mit negativ bewerteten Bildern, während die Signalveränderungen im linken DLPFC mit positiv bewerteten Bildern korrelierten. Diese Dichotomie wird in verschiedenen Studien bestätigt (vgl. Canli et al. 1998; Davidson/Irwin 1999; Murphy et al. 2003; Wagner et al. 2004). Positive Emotionen werden tendenziell in der linken Hemisphäre lateralisiert, während negative Emotionen mit der rechten Hemisphäre in Zusammenhang gebracht werden.




Grafik 16: Die Balkendiagramme repräsentieren die T-Werte im Durchschnitt für das Experiment mit 29 Probanden und jede der vier Konditionen („Bildwahrnehmung ohne Aufmerksamkeit", „Beurteilung ohne Aufmerksamkeit", „Bildwahrnehmung mit Aufmerksamkeit", „Beurteilung mit Aufmerksamkeit") für verschiedene Regionen während Bildwahrnehmung und Beurteilung unter unerwarteten und erwarteten Konditionen (vgl. Grimm et al. 2006, 332).

Zusätzlich wurde ein konverses Korrelationsmuster im VMPFC und im DLPFC beobachtet. Der VMPFC zeigte eine positive Korrelation zur Valenz, während neuronale Aktivität im DLPFC eher negativ korrelierte. Je positiver die emotionalen (positiven und negativen) Bilder bewertet wurden, desto weniger Aktivität wurde im bilateralen DLPFC gemessen. Diese konverse Korrelation zwischen VMPFC und DLPFC ist in Übereinstimmung mit der reziproken Modulation zwischen medialem und lateralem präfrontalem Cortex während emotional-kognitiver Interaktion. Reziproke Modulation kann als Aktivitätsveränderung in entgegengesetzter Richtung (z. B. Anstieg und Verringerung) in verschiedenen Regionen definiert werden. Es ist bekannt, dass emotionale Prozessierung zu Aktivitätsanstieg im VMPFC führt bei gleichzeitiger Aktivitätsverringerung im DLPFC, vice versa (vgl. Goel/Dolan 2003a; Northoff et al. 2004). Kognitive Aufgaben induzieren dagegen das umgekehrte Muster mit Aktivitätsanstieg im lateralen präfrontalen Cortex und Verringerung im medialen präfrontalen Cortex. Die vorliegende Studie spezifiziert die funktionalen Mechanismen der reziproken Modulation hinsichtlich der Valenz: Beide,

VMPFC und DLPFC, sind anscheinend empfindlich für die Modulation durch Valenz, wenn auch in unterschiedlicher Weise, nämlich affektiv und evaluativ. Beide Aspekte wiederum werden reziprok moduliert (vgl. Grimm et al. 2006, 336).

Es wurde eine signifikante Korrelation der Intensität mit Aktivitätsveränderung im rechten VLPFC beobachtet. Je intensiver der emotionale Stimulus erlebt wurde, desto mehr Aktivität wurde im VLPFC induziert. Diese Korrelation wurde nur bei der unerwarteten Beurteilung beobachtet, aber nicht bei erwarteter Beurteilung oder unerwartetem oder erwartetem Bildsehen. Die Ergebnisse zeigen, dass neuronale Aktivität im VLPFC in Korrelation mit Intensität sich speziell auf die Prozessierung von Beurteilung bezieht und nicht auf Aufmerksamkeit.

Beim VLPFC wurde bereits des öfteren Aktivität bei emotional-kognitiver Interaktion beobachtet (vgl. Drevets/Raichle 1998; Phan et al. 2002; Murphy et al. 2003; Northoff et al, 2004; Steele/Lawrie 2004). Eine jüngste Studie zeigte, dass neuronale Aktivität in dieser Region möglicherweise zur „down-Modulation" emotionaler Intensität bei Beurteilung beiträgt (vgl. Ochsner et al. 2002). Die Ergebnisse dieser Studie unterstützen diese Beobachtung, da die neuronale Aktivität im VLPFC spezifisch mit der Beurteilung und nicht mit der Aufmerksamkeit verbunden zu sein scheint. Darüber hinaus stützen die Resultate die Frage der Lateralisation: es wurde eine Korrelation für den rechten, aber nicht für den linken VLPFC beobachtet. Nur ersterer wird, im Unterschied zum letzteren, mit Selbst-Erkennung und Selbst-Bewusstsein in Zusammenhang gebracht (vgl. Keenan et al. 2000; Platek et al. 2004).

Außerdem wurde eine signifikante Beziehung zwischen emotionaler Intensität und neuronaler Aktivität im DMPFC beobachtet. Diese Tendenz zeigen auch andere Studienergebnisse (vgl. Anders et al. 2004; Dolcos et al. 2004). Die funktionale Rolle des DMPFC lässt sich nun insofern spezifizieren, als die neuronale Aktivität nur während der erwarteten Beurteilung, aber nicht während der unerwarteten Beurteilung mit Intensität korreliert. Neuronale Aktivität im DMPFC scheint emotionale Intensität hinsichtlich der Interaktion zwischen emotionaler Beurteilung und Aufmerksamkeit zu modulieren. Dieses Ergebnis wird durch Studien über emotionale Beurteilung und Aufmerksamkeit gestützt (vgl. Gusnard et al. 2001; Northoff et al. 2004). Dies legt auch die beobachtete parametrische Abhängigkeit des DMPFC von emotionaler Intensität während der Interaktion zwischen Beurteilung und Aufmerksamkeit nahe.

Schließlich wurde während der IAPS-Bildwahrnehmung auch Aktivität in der Amygdala beobachtet (vgl. auch Phan et al. 2002; Murphy et al. 2003; Steele and Lawrie 2004; Morris et al 1998a, b; Hariri et al. 2000, 2003; Taylor et al. 2003). Allerdings wurde keine signifikante Korrelation zwischen veränderten Signalen der Amygdala und der emotionalen Intensität gemessen. Dies steht in Widerspruch zu früheren Ergebnissen, die eine solche signifikante Korrelation berichteten (vgl. Small et al. 2003; Anderson et al. 2003b). Diese Diskrepanz kann auf sensorische Unterschiede der in den Studien verwendeten emotionalen Stimuli (Geruch/Geschmack; visu-

al/emotional) zurückgeführt werden. Entscheidend ist dabei, dass visual-emotionale Stimuli semantisch wesentlich reicher und differenzierter erscheinen als emotionale Stimuli, die mit Gerüchen oder Geschmack verbunden sind.

Es konnte eine signifikante Korrelation des Perigenualen Anterioren Cingularen Cortex (PACC) zwischen dem unerwarteten Bildsehen und der subjektiven Bildbewertung der Wiedererkennung der emotionalen Bilder beobachtet werden. Dieses Ergebnis stimmt auch mit einer Metaanalyse überein, die die Mitwirkung dieser Region in Studien mit emotionaler Erinnerung untersuchte (vgl. Phan et al. 2002). Eine weitere Studie, die Emotionen durch Erinnerungen oder persönliche Bilder mit affektivem Inhalt induzierte, zeigte ebenfalls Aktivität im PACC (vgl. Reiman et al. 1997). Die funktionale Rolle des PACC lässt sich nun insofern noch deutlicher spezifizieren, als seine neuronale Aktivität parametrisch von der Erinnerung emotionaler Ereignisse abhängt. Weiter spezifiziert, wurde sogar die negative Korrelation zwischen Signalschwankungen des PACC und der Wiedererkennung beobachtet: je schlechter die Probanden die emotionalen Stimuli erinnerten, desto mehr Aktivität war im PACC zu beobachten. Dementsprechend scheint geringe Erinnerung emotionaler Stimuli mit ansteigender Aktivität im PACC zusammenzuhängen.[144] Aufgrund der Ergebnisse kann vorsichtig die Hypothese entwickelt werden, dass, je besser der emotionale Stimulus encodiert wurde, desto mehr neuronale Aktivität im Perigenualen (und Supragenualen) Anterioren Cingulären (und Dorsolateralen Präfrontalen) Cortex induziert wurde und desto besser die Probanden später den emotionalen Stimulus wieder erkennen bzw. erinnern können (vgl. Grimm et al. 2006, 337).

Der PACC muss vom Supragenualen Anterioren Cingularen Cortex (SAC) unterschieden werden, der eher mit grundsätzlicher Aufmerksamkeit wie z. B. kognitivem Monitoring in Verbindung gebracht wird (vgl. Bush et al. 2000; Duncan/Owen 2000). So wurde in der vorliegenden Studie signifikante Korrelation des SAC mit Wiedererkennung während der aufmerksamen Beurteilung beobachtet. Bei unerwarteten Konditionen war dies entsprechend nicht der Fall. Dies legt nahe, dass der SAC im Gegensatz zum PACC kein emotionales Stimulus-Wiedererkennen per se vermittelt, sondern eher nicht-spezifische Aufmerksamkeitsprozessierung. Die Ergebnisse legen somit insgesamt die funktionale Trennung von PACC und SAC in der emotionalen Stimulus-Prozessierung nahe, wobei ersterer bei dem Wiedererkennen von Emotionen und letzterer bei der Aufmerksamkeit von Emotionen beteiligt ist (vgl. Grimm et al. 2006, 337).

[144] Diese Beobachtung stimmt auch mit der Hypothese von Mayberg (2003) überein, wonach die Funktion der „extremen Wiedererkennung“ mit dem PACC in Verbindung gebracht wird. Extreme Wiedererkennung setzt emotionale Wiedererkennung voraus, die innerhalb der vorliegenden Studie untersucht wurde. Extreme Wiedererkennung verändert sich bei Patienten mit Major Depression, da sie unfähig sind, zwischen emotional herausragenden Inhalten und nicht herausragenden zu unterscheiden. Diese Patienten können sich nicht an positive persönliche Ereignisse ohne starken emotionalen Inhalt erinnern. Zugleich zeigen die depressiven Patienten Hyperaktivität im PACC (vgl. Mayberg et al. 1999; Mayberg 2003).

***Quintessenz Bildbeurteilung***

- *vielfältige Aktivierungsbereiche in verschiedenen präfrontalen cortikalen Regionen bei emotionalem Bildsehen unabhängig von der damit verbundenen kognitiven Prozessierung*
- *Korrelation mit den subjektiven Bewertungen unterschiedlicher emotionaler Dimensionen wie emotionaler Valenz, Intensität und Wiedererkennen durch die Probanden zeigt Beteiligung unterschiedlicher präfrontaler cortikaler Regionen an der Prozessierung der unterschiedlichen Dimensionen emotionaler Stimuli:*
    - *Zusammenhang des VMPFC mit affektivem Aspekt der Valenz*
    - *Zusammenhang zwischen bilateralem DLPFC mit eher evaluativen Aspekten*
    - *Beteiligung des rechten VLPFC bei beurteilender Kontrolle emotionaler Intensität*
    - *Interaktion „Beurteilung und Aufmerksamkeit" bei emotionaler Intensität durch Vermittlung des DMPFC*
    - *Beteiligung des PACC bei Enkodierung und bzw. oder bei Erinnerung emotionaler Ereignisse*
    - *Vermittlung grundsätzlicher Aufmerksamkeitsprozessierung mit Dimensionen Valenz und Wiedererkennen emotionaler Stimuli durch den SAC*
- *getrennte Repräsentationen der verschiedenen emotionalen Dimensionen im präfrontalen Cortex*

## Kapitel 6

## Vergleich der Kriterien und Paradigmen

Der Vergleich bzw. die Diskussion der philosophischen Theoriesätze mit den neurowissenschaftlichen Beobachtungssätzen erfolgt vor dem Hintergrund der Prinzipien der Asymmetrie, der Bidirektionalität sowie der transdisziplinären Zirkularität (vgl. Northoff 2004, 27ff.). Aufgrund des epistemischen Ungleichgewichts zwischen logischen und natürlichen Bedingungen ist ein direktes Schlussfolgern zwischen philosophischen und neurowissenschaftlichen Aussagen irreführend. Lediglich aufgrund der vorhandenen Schnittmenge beider Bedingungsbereiche können philosophische Theorieaussagen und neurowissenschaftliche Hypothesen miteinander verknüpft werden. Als Ergebnis einer solchen Verknüpfung ist neben der klaren Abgrenzung oder aber der Übereinstimmung beider Bereiche auch eine Modifizierung im Sinne eines „definitorischen Wandels" bzw. einer „begrifflichen Weiterentwicklung" möglich (vgl. Kap. 1.3). Im Rahmen dieser verknüpfenden Diskussion beziehen sich die Theoriesätze auf die logischen Bedingungen sowie ontologische wie epistemische Annahmen. Zudem werden die empirischen Implikationen der theoretischen Annahmen herausgearbeitet. Die Beobachtungssätze beziehen sich auf die natürlichen Bedingungen sowie Beobachtungen der tatsächlichen Welt. Für sie werden die theoretischen Explikationen herausgearbeitet. In einem ersten Schritt wurde dieses in den beiden vorhergehenden Kapiteln 4 und 5 bereits durchgeführt: Für die Entwicklung der Kriterien für mentale Kausalität wurden die theoretischen Bedingungen und teilweise einige empirische Implikationen herausgearbeitet. Es wurden dabei unterschiedliche Positionen und Widersprüche verdeutlicht. Im Gegenzug wurden die empirischen Bedingungen und einige theoretische Explikationen der neurowissenschaftlichen Studien um das Paradigma zur emotionalen Aufmerksamkeit und das Paradigma zur Bildbeurteilung unter verschiedenen Konditionen dargestellt. Hierfür können beispielsweise verschiedene intra-mentale kausale Prozesse oder kausale Wechselwirkungen zwischen Gehirn und Umwelt abgeleitet werden.

Im folgenden Kapitel werden für den wechselseitigen Vergleich die jeweiligen theoretischen wie empirischen Implikationen und Explikationen weiter herausgearbeitet. Die ontologischen, epistemischen bzw. empirischen Annahmen des theoretischen wie des empirischen Bereichs werden analogisiert und in Reflexion der etablierten Voraussetzungen für mentale Kausalität wechselseitig aufeinander bezogen (vgl. Kap. 3). Es bestätigt sich die Tendenz, wonach der herkömmliche Begriff der mentalen Kausalität redefiniert und modifiziert werden sollte. Beispiel ist der veränderte Gesetzesbegriff für die vertikale wie horizontale Relation mentaler Ereignisse, der sich an das Modell von ceteris-paribus-Gesetzen anlehnt. Entscheidend ist hierbei die Erweiterung der Relation auf drei Komponenten, die neben der physischen und der subjektiv-

phänomenalen Ebene auch die Relation zu äußeren Ereignissen bzw. Umweltstimuli berücksichtigt. Die herkömmliche Abgrenzung externer Stimuli als extramentale Ereignisse wird in Frage gestellt. Sie sind schließlich wichtiger Bestandteil eines mentalen Ereignisses selbst. Dieser Umstand wird auch über den Begriff der „kausalen Relation" im Rahmen des Vergleichs mit dem Kriterium der Multikausalität deutlich (vgl. Kap. 6.2.3). Die Notwendigkeit zur begrifflichen Differenzierung für mentale Kausalität zeigt sich ferner in der Bestimmung verschiedener Inhaltstypen oder in der Differenzierung zwischen abstrakten und konkreten Inhalten. Es bestätigt sich die Anforderung wechselseitiger Modulation für mentale Kausalität, anstelle einer monokausalen Struktur (vgl. Kap. 6.3.2; 6.3.3). Dabei scheint auch eine begriffliche Binnendifferenzierung für neuronale mentale Ereignisse erforderlich, die schließlich die Bestimmung des Gehirns als intrinsisch-relationale Funktionseinheit begründet (vgl. Kap. 6.4.3; 6.4.4).

Schließlich wird auch in dem Vergleich zwischen den neurowissenschaftlichen Paradigmen und den Kriterien der Intentionalität und der Proportionalität deutlich, dass sich ein Modell der mentalen Kausalität von dem klassischen Modell einer rein physischen Wirkkausalität unterscheiden muss. Deutlich wird dies unter anderem aufgrund der begrifflichen Modifizierung intentional-biologischer Prozesse oder der Verbindungsfunktion mentaler Eigenschaften einerseits (vgl. Kap. 6.5), sowie aufgrund der Ausdifferenzierung des Abstraktionsgrades mentaler wie physischer Eigenschaften, der Notwendigkeit einer adäquaten Eigenschaftsinstanziierung vor dem Hintergrund eines epistemischen Mehrwerts mentaler wie physischer Eigenschaftsbeschreibungen andererseits (vgl. Kap. 6.6). Deshalb kann schließlich auf dieser Basis in Kapitel 7 schließlich mit der neurophilosophischen Hypothesenbildung ein eigenständiges Modell für mentale Kausalität entwickelt werden, das sich an den Ergebnissen dieser vergleichenden Diskussion orientiert.

## 6.1 Emotionale Aufmerksamkeit und Regularität

Wie dargelegt wurde, ist es grundsätzlich problematisch, mentale Kausalität in einer genau definierten Regularität einzufangen, da sie als „Diener zweier Herren" sowohl physischen als auch mentalen Ansprüchen gerecht werden muss. Eine rein physische Kausalität wird als linearer kausal-mechanistischer Prozess verstanden, der strikten Gesetzen unterliegt, während mentale bzw. psychologische Prozesse vielmehr Rationalitätskriterien oder allenfalls ceteris-paribus-Gesetzen unterliegen.

Auch in der neurowissenschaftlichen Forschung werden Gesetze für neuronale Aktivität, aber auch für psychologische Funktionen sowie für neuronale Korrelate für diese Funktionen und somit zumindest implizit psychophysische Gesetze vorausgesetzt. Dies zeigen auch exemplarisch die vorliegenden Studien über unterschiedliche Netzwerke für Aufmerksamkeit und Perzep-

tion emotionaler Stimuli sowie unterschiedliche neuronale Repräsentationen für verschiedene emotionale Dimensionen im menschlichen Gehirn (vgl. Kap. 5.2; 5.3). So ist es geradezu das Ziel solcher Studien, die gesetzmäßigen Funktionen des Gehirns aufzuzeigen. Es können dabei auf der empirischen Ebene drei Gesetzesarten unterschieden werden:

1. Gesetze für neuronale Aktivität
2. Gesetze für psychische Funktionen
3. Gesetze für Korrelationen (Korrelationsgesetze)

Aufgrund der philosophischen Diskussion erscheint es problematisch, für 2. und 3. strikte Gesetze aufzustellen, während eine gesetzmäßige rein neuronale Aktivität unter empirischen wie theoretischen Gesichtspunkten zunächst unproblematisch erscheint. Hier stimmen empirische Implikation und theoretische Explikation überein.

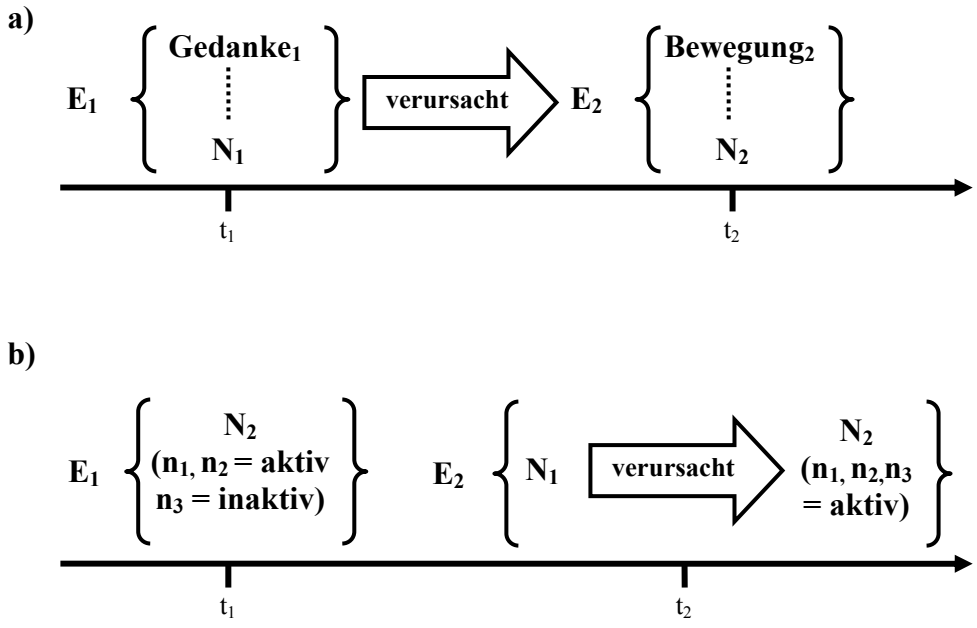
Im Folgenden wird untersucht, wie sich diese Situation unter neurowissenschaftlichen Bedingungen „in der Praxis" darstellt. Dabei wird deutlich, dass mögliche Gesetze für neuronale Aktivität gegebenenfalls weniger strikt ausfallen und mögliche Gesetze für psychologische Aktivität gegebenenfalls weniger vage. Dennoch scheint eine Reduzierung der psychologischen auf die neuronale Ebene nicht sinnvoll.

### 6.1.1 Gesetze für neuronale Aktivität

Aber auch für die Gesetze für neuronale Aktivität muss theoretisch genau definiert werden, auf welche Ursache hin welche Wirkung folgt. Nach Kim (2003) müsste es sich hierbei um den Fall der horizontalen Kausalität handeln. Ein neuronales Netzwerk wie z. B. das antizipatorische Netzwerk ist Ursache a) für ein anderes Netzwerk oder b) für eine spezifische Wirkung in einem anderen Netzwerk (vgl. Grafik 17). Auf der rein empirischen Ebene ist für das antizipatorische Netzwerk SAC, bestehend aus CMA, SMA und POS, zumindest der Fall b) darstellbar, wonach Aktivität in diesem Netzwerk als Ursache eine Verstärkung der anschließenden emotionalen Perzeption bzw. höhere Aktivität im emotionalen perzeptiven Netzwerk als Wirkung zur Folge hat. Wesentlich schwieriger ist es dagegen, Fall a) zu belegen, inwieweit das antizipatorische Netzwerk Ursache für ein anderes Netzwerk bzw. die Aktivierung eines anderen Netzwerks ist.

Dies hängt auch mit der spezifischen Funktion bzw. multi-integrativen Struktur des SAC zusammen, wonach emotional auffallende Ereignisse vermittelt und affektive, kognitive oder auch motorische Informationen und Prozesse in Handlung übersetzt werden. Rein empirisch betrachtet, scheint der SAC also eher eine vermittelnde Rolle in einer neuronalen Folge innezu-

haben. Die Aktivität des SAC ist somit notwendige Bedingung für verschiedene Handlungsarten. Es ginge jedoch zu weit, davon zu sprechen, der SAC sei Wirkursache für die Aktivierung z. B. von DLPFC und Amygdala. Als hinreichende Bedingung kommen möglicherweise andere neuronale Netzwerke oder auch externe Stimuli, wie z. B. die emotionalen bzw. neutralen Bilder in Frage. Diese scheinen als Ursachekandidaten insofern besser geeignet, als auch die Studie zeigt, dass das antizipatorische Netzwerk zwar eine Bereitschaft aufbaut, aber ein deutlich unterscheidbares emotional-perzeptives Netzwerk (bestehend aus Amygdala, Insula, medialem und lateralem präfrontalem Cortex, Putamen, Mittelhirn, Cerebellum, occipitotemporalen visuellen Regionen) in Korrelation zu den gezeigten emotionalen und neutralen Bildern Aktivität erkennen lässt. Zumindest für das antizipatorische Netzwerk erscheint es auf empirischer Ebene somit problematisch, dieses als alleinige Ursache für die Aktivierung weiterer Netzwerke und Hirnregionen zu definieren; vielmehr scheinen hier äußere Ursachen relevant zu sein.




Grafik 17: Ein neuronales Netzwerk ist a) zum Zeitpunkt $t_1$ Ursache für ein anderes neuronales Netzwerk zum Zeitpunkt $t_2$ in einem wirkmechanistischen Sinn, wonach Ereignis $E_1$ als alleinige Ursache die Bewegung von Ereignis $E_2$ verursacht. Ein neuronales Netzwerk ist b) zum Zeitpunkt $t_2$ Ursache für eine spezifische Wirkung in einem anderen Netzwerk in dem Sinn, dass die bereits vorhandene (Basis)aktivität des zweiten Netzwerks $E_2$ durch das erste $E_1$ moduliert und beeinflusst wird.

Sind dagegen andere Regionen als Ursache für die Aktivierung weiterer Regionen denkbar? Auf empirischer Seite wird zunächst zwischen internen und externen Stimuli unterschieden. Innerhalb der Abfolge der neuronalen Prozesse wird allerdings eher von temporaler statt kausaler Folge gesprochen. Zwar wird untersucht, welche Region zuerst aktiviert wird oder welche Areale miteinander verbunden werden, aber eine kausale Relation wird nicht als Begründung angeführt. Der Tenor ist meist deskriptiv mit einer vorsichtigen Deutung verbunden. Beispielsweise werden vi-

suelle Stimuli zunächst vom Auge über den Sehnerv an die Sehrinde und okzipitemporale Regionen geleitet und dort verarbeitet, ehe sie über die Aktivierung weiterer Hirnregionen bewusst werden. Andererseits ist auch denkbar, dass von einem inneren Stimulus eine temporale Folge initiiert wird und verschiedene Areale umfasst.

Hier stellt sich die Frage, wie genau ein solcher innerer Stimulus definiert werden kann, und was einen solchen inneren Stimulus selbst treibt bzw. hervorruft – oder etwa auslöst. Ein innerer Stimulus könnte auf psychologischer Ebene ein Wunsch oder eine Absicht sein (vgl. Kap. 4.5). Hierbei steht die Neurowissenschaft vor der Herausforderung, solche inneren Stimuli allein auf neuronaler Basis zu erkennen und zu messen; dennoch wird unter der Annahme einer idealen Neurologie angenommen, dass auch diese Prozesse (strikten) physikalischen Gesetzen unterliegen. Dabei wird deutlich, dass zwar durch die Messung der BOLD-Aktivität einzelne Regionen für spezifische Funktionen definiert werden können. Allerdings muss beachtet werden, dass jene Aktivität nicht bei allen Probanden gleich stark ist – dies müsste ein striktes Gesetz jedoch leisten. Geringe Abweichungen sind jedoch nicht zwingend für Regionen, aber zumindest für Netzwerke oder die Intensität der BOLD-Signale möglich.

Aus der neurowissenschaftlichen Perspektive stellt sich also die Frage, ob es sich tatsächlich um strikte kausale Gesetze handelt. Sowohl bei der Aktivität des antizipatorischen Netzwerks als auch bei der Aktivität der perzeptiven Regionen oder auch bei Aktivität motorischer Regionen z. B. aufgrund dispositionaler Initiativen ist es innerhalb der Neurowissenschaft eher üblich, diese Formen der Aktivität als Modulation, Repräsentation oder Prozessierung zu bezeichnen. Diese verschiedenen Funktionen werden in der philosophischen Diskussion – beispielsweise innerhalb des Funktionalismus oder teleologischer Theorien (vgl. Kap. 2.2; 4.5.1) – auch auf kausale Prozesse zurückgeführt. Für die neurowissenschaftliche Arbeit ist es aber allein auf der neuronalen Ebene fraglich, ob man sinnvoll davon sprechen kann, dass neuronale Netzwerke im Sinne einer monokausalen Relation einander verursachen. Dies muss kein Widerspruch zu stabilen physikalischen Gesetzen sein. So stellt sich überhaupt die Frage, wie genau neuronale Gesetze formuliert werden können. Denn wenn einige Regionen unterschiedliche neuronale Funktionen innehaben und mit unterschiedlichen weiteren Regionen vernetzt sind, sind Gesetze der Art, dass immer, wenn Region A aktiv ist, auch Aktivität von Region B folgt, nur eingeschränkt bei sehr konkreten Abläufen möglich. Vielmehr folgt auf Aktivität von Region A die Aktivität von Region B oder Region C etc. Zusätzliche Faktoren müssen unter diesen Umständen erfüllt werden, um genaue Aussagen zu treffen. Es ist die Frage, wie strikt allgemeine Gesetze über Neuronen bzw. neuronale Netzwerke in der Anwendung sein können und auf welcher Mikro- oder Makroebene diese gegebenenfalls anzusiedeln sind. Gegebenenfalls wäre also eine Modifizierung der Begriffe erforderlich, als von einer „kausalen Modulation“ oder einer „kausalen Prozessierung“ gesprochen werden kann. Dies müsste aber weitergehend geprüft werden.

### 6.1.2 Gesetze für psychische Funktionen

Was kann von neurowissenschaftlicher Seite über die zweite Gesetzesart ausgesagt werden? Grundsätzlich gilt, dass, um neurowissenschaftlich sinnvolle Studien zu entwickeln, das Vorhandensein psychischer Gesetze angenommen werden muss. Eine ideale Neurologie setzt implizit eine ideale Psychologie voraus. Eine Anomalie des Mentalen im Sinne Davidsons kann es demnach nicht geben. In der Praxis werden die Perspektiven im Rahmen des Versuchsaufbaus oder der Schlussfolgerungen ausgetauscht, weshalb die Studien somit für beide Bereiche interessant sind.

Mit der Feststellung auf philosophischer Seite, dass es keine strikten psychologischen Gesetze gibt, wird diese Praxis somit zunächst einmal in Frage gestellt. So wird in beiden vorliegenden Paradigmen ein Post-hoc-Test durchgeführt, mit dem die subjektive Wahrnehmung der Emotionen sowie eine Bewertung der Dimensionen Valenz, Intensität und Wiedererkennen erfasst werden sollen. Dabei wird ermittelt, dass es hinsichtlich der Valenz-Bewertung einen signifikanten Unterschied zwischen den Bildern in den vier Konditionen gibt, oder dass erwartete beurteilte Bilder signifikant stärker wiedererinnert werden als erwartete nicht-beurteilte Bilder. Grundsätzlich zeigt die Valenz-Bewertung, dass die verwendeten negativen, positiven und neutralen Bilder als solche wahrgenommen werden (vgl. Kap. 5.2.3). Dies zeigt, dass es offensichtlich auf der psychologischen Ebene zumindest beobachtbare Regularitäten gibt. Die emotionalen Dimensionen unterliegen ebenso wie die positive oder negative emotionale Aufmerksamkeit einer Regularität. So gibt es eindeutig Parameter für die Wiedererkennung. Auch die Folge „emotionale Aufmerksamkeit; emotionale Bildwahrnehmung“ oder grundsätzlich das Erkennen der IAPS-Bilder verweisen darauf. Es kann formuliert werden, dass „Kriegsbilder bei Probanden stärkere negative Emotionen bewirken als ein Foto mit einem Blitzeinschlag“. Ebenso kann formuliert werden, dass „ein Bild mit einem lachenden Baby stärkere positive Emotionen bewirkt als ein Bild mit einem Blumenfeld“. Solche Generalisierungen sind zutreffend, sie verweisen mindestens auf ceteris-paribus-Gesetze.

Wenn man Davidsons Einschätzung zugrunde legt, wonach insbesondere intentionale Ereignisse keinen strikten Gesetzen unterliegen, während die physische Realisierung mentaler Ereignisse wie z. B. Schmerzen gewährleistet bleibt, ließe sich feststellen, dass die angesprochenen Ergebnisse nicht zwingend in die Kategorie intentionaler Ereignisse fallen. Allerdings ist es sehr wohl möglich, dass die emotionalen Dimensionen Bestandteil eines intentionalen Ereignisses sind. Dieser Umstand verweist auf die Unterscheidung zwischen psychologischen und phänomenalen Eigenschaften bzw. die Frage der Möglichkeit „hybrider Eigenschaften“ (vgl. Kap. 4.1.1). Bei dem Ereignis der emotionalen Aufmerksamkeit oder der emotionalen Beurteilung ist es plausibel, von einer solchen Variante zu sprechen. Einerseits gibt es getrennte kognitive und

emotionale Netzwerke, andererseits sind diese miteinander verbunden: z. B. umfasst Valenz einen affektiven und einen evaluativen Aspekt (vgl. Kap. 5.2.3). Insofern ist Davidsons Unterscheidung nicht zutreffend. Das Problem der mentalen Kausalität stellt sich für phänomenale und psychologische Eigenschaften in ähnlicher Weise – zumindest hinsichtlich der Regularität. Searle hat in seiner detaillierten Analyse verschiedene Formen intentionaler Ereignisse aufgezeigt (vgl. Kap. 4.5.2). Danach ist es gut möglich, dass sich auch bei den im Rahmen der Studie durch die gezeigten Bilder induzierten Emotionen um intentionalen Gehalt in der Ausrichtungsart Geist-auf-Welt handelt, wozu schließlich Wahrnehmungen zählen. Zudem dürfte es sich bei den mit der Bildbeurteilung verbundenen Emotionen um einen intentionalen Gehalt derselben Ausrichtungsart handeln, da auch Überzeugungen zu dieser Art zählen (vgl. Searle 2001, 122; Kap. 4.5.2). Schließlich wurden im Beurteilungsparadigma auch Bewertungen aufgrund der gezeigten Bilder vorgenommen.

Inwiefern den Generalisierungen in der Psychologie ein Charakter strikter Gesetze zukommt, ist auf lange Sicht Aufgabe differenzierter Forschung. Wobei sicherlich die Vagheit und Unbestimmtheit subjektiver Zustände und psychologischer Be-griffe ein Hindernis darstellen. Aber auch hierfür ist es möglich und bereits üblich, psychologische Phänomene zu differenzieren und entsprechend zu klassifizieren. Es würde Neurowissenschaftler vor ein grundsätzliches methodisches Problem stellen, wenn sie aufgrund der insbesondere von Davidson aufgeführten Prämisse den „Kompass" für neurowissenschaftliche Forschung verlören. Dies gilt auch und möglicherweise gerade für intentionale Zustände. Schließlich ist es nicht möglich, allein aufgrund neuronaler Messungen subjektive Zustände zu bestimmen; die Aussagen der Probanden über ihre subjektiven Erlebnisse sind allein aus methodischen Gründen elementar für die Ergebnissicherung der Studien.

### 6.1.3 Korrelationsgesetze

Für die vorliegenden Paradigmen ist die Annahme einer gesetzesmäßigen psychophysischen Korrelation eine implizite Annahme des Versuchs bzw. des Versuchsaufbaus. Vom Design her werden empirische Bildgebungsstudien aus philosophischer Perspektive streng genommen „cross-level" bzw. vertikal entwickelt. Es findet ein Wechsel zwischen der physischen und der mentalen Ebene statt: (Aufgabenstellung → Gehirn → Umsetzung) Diese Folge beruht auf der Hintergrundannahme einer durchgehenden Korrelation. Manchmal ist die physische Ebene besser beschreibbar, manchmal die mentale. In der Analyse der Studien wird ebenfalls zwischen den Ebenen geschlossen. Beispielsweise wird geschlossen, dass die Aktivierung der Amygdala darauf zurück zuführen ist, dass der Proband Angst – als emotionales Ereignis – hatte. Eine (visuelle)

Aktivierung einer Region lässt wiederum ein entsprechendes Verhalten bzw. entsprechende neuronale Aktivität erwarten. Ferner kann der Versuch in psychologische und funktionelle Teilaspekte aufgeschlüsselt werden. Das „spezifisch Emotionale" ist dann nicht mehr die aktivierte Region, sondern der gemessene BOLD-effect bzw. die verstärkte Duchblutung der aktivierten Region.[145]

Aufgrund der „cross-level" Perspektive kann davon gesprochen werden, dass die Korrelationsgesetze im Gegensatz zu den Gesetzen 1. und 2. einen horizontalen und vertikalen Charakter haben. Kann deshalb neben der horizontalen Kausalität auf den jeweiligen mentalen bzw. physischen Ebenen auch eine vertikale Kausalität zwischen den Ebenen bestimmt werden? Diese Frage kann nur vorsichtig unter Verweis auf die vertikale Korrelation beantwortet werden. Dabei auf kausale Prozesse zu schließen, führt jedoch sehr weit, da alternativ auch eine Realisierungsrelation in Frage kommt.

Konkretere Korrelationen zwischen der neuronalen und der psychologischen Ebene konnten für die emotionalen Dimensionen Valenz (VPMFC, VLPFC, DLPFC), Intensität (DMPFC, VLPFC) und Wiedererkennen (PACC) beobachtet werden. Bei der eingehenden Untersuchung der emotionalen Valenz wurden weitere konkrete Korrelationen beobachtet. Beobachtet wurde eine positive Korrelation mit dem VMPFC während nicht erwarteten Bildsehens – je positiver die Valenz, desto höher der T-Wert im VMPFC. Positive Bilder bzw. positives Bilderleben zeigten für die erwartete wie unerwartete Kondition eine signifikante Korrelation im DLPFC, negative Bilder bzw. negatives Bilderleben zeigten zumindest für die unerwartete Kondition eine signifikante Korrelation mit dem T-Wert der rechten Amygdala. Für die Phase der Beurteilung konnte zumindest eine Korrelation zwischen dem Valenzwert und den T-Werten im bilateralen DLPFC gezeigt werden. Dabei handelt es sich interessanterweise um eine negative Korrelation – je positiver die emotionale Valenz betrachtet wurde, desto geringer war der T-Wert im bilateralen DLPFC (vgl. Kap.5.2.1).

Zudem spiegelt sich die psychologische Differenzierung der Valenz eines Stimulus in einen affektiven und einen evaluativen Aspekt (Russell 1980; Barrett/Russell 1999) in der neuronalen Aktivität des VMPFC und des DLPFC (vgl. Kap. 5.2.3). Das gilt nicht nur für die verschiedenen Korrelationen der emotionalen Valenz. Auch bei emotionaler Intensität werden verschiedene Korrelationen be-obachtet. So wurde eine signifikante Korrelation der Intensität mit Aktivitätsveränderung im rechten VLPFC beobachtet. Je intensiver der emotionale Stimulus erlebt wurde, desto mehr Aktivität wurde im VLPFC induziert. Diese Korrelation wurde nur bei der nicht-erwarteten Beurteilung beobachtet, aber nicht bei erwarteter Beurteilung oder nicht-erwarteter oder erwarteter Bildwahrnehmung. Die Ergebnisse zeigen, dass neuronale Aktivität im VLPFC in Korrelation mit Intensität sich speziell auf die Prozessierung von Beurteilung bezieht und nicht

[145] Interessant wäre es, zu untersuchen, welche Teilaspekte eher epiphänomenalen Charakter haben.

auf Aufmerksamkeit (vgl. ebd.). Außerdem wurde eine signifikante Beziehung zwischen emotionaler Intensität und neuronaler Aktivität im DMPFC beobachtet. Die funktionale Rolle des DMPFC lässt sich insofern spezifizieren, als die neuronale Aktivität nur während der erwarteten Beurteilung, aber nicht während der nicht-erwarteten Beurteilung mit Intensität korreliert (vgl. ebd.).

Die Neurowissenschaft geht zwar implizit von strikten physischen Gesetzen aus, die Erklärung und Prognose für Verhalten o. ä. ermöglichen. Da es jedoch viele Faktoren, wie z. B. Gene oder Umwelt für neuronale oder psychische Funktionen gibt, wird in der Praxis eine Wahrscheinlichkeit bzw. beschränkte Prognosefähigkeit angenommen.[146] Physische Eigenschaften sind zwar notwendig, aber nicht hinreichend für weitere mentale Ereignisse. Als Ursache sind auch externe Bilder, konkret die IAPS-Bilder, möglicherweise als gemeinsame Stimuli für physische und mentale Eigenschaften entscheidend. Somit stellt sich für die Neurowissenschaft die Frage, ob hinsichtlich der horizontalen wie auch der vertikalen Relation derzeit nicht ceteris-paribus-Gesetze der tatsächlichen methodisch bedingten Situation eher gerecht werden. Eine gesetzmäßige physische und mentale Beschreibung scheint in der neurowissenschaftlichen Praxis grundsätzlich notwendig zu sein. Dies zeigt sich in der Verknüpfung der fMRT-Daten mit den Post-hoc-Ergebnissen. Aus einer neurowissenschaftlichen Perspektive heraus müssen solche Verknüpfungen zumindest einen stabilen Gesetzescharakter haben. Inwiefern dieser strikt ist, muss jedoch vor dem Hintergrund der philosophisch-theoretischen Überlegungen offen bleiben. Damit scheint aber auch eine „Cross-Level-Kausalität" fraglich – in beide Richtungen! Auf der epistemischen Ebene ist es wichtig, beide Ebenen gleichberechtigt zu berücksichtigen, eine vollständige Reduktion in die eine oder andere Richtung wird jedoch ausgeschlossen – physische Gesetze können die psychologische Ebene nicht vollständig beschreiben. Dagegen scheinen beide Ebenen relevant zu sein, um eine Kausalität zwischen der Phase der emotionalen Aufmerksamkeit und der Phase der emotionalen Bildwahrnehmung zu beschreiben. Es bleibt jedoch problematisch, aus dieser epistemischen Situation heraus ontologische Schlüsse zu ziehen.

Alternativ zu einer Kausalität zwischen den Ebenen ermöglicht das Modell der Realisierungsrelation die gleichzeitige Instanziierung beider Eigenschaften. Darauf verweisen funktionalistische oder auch supervenienztheoretische Überlegungen (vgl. Kap. 2.2.1; Kap. 2.4.1). Zumindest müsste man auch aufgrund der theoretischen Anforderungen, Beispiel ist hier vor allem das Prinzip der Multirealisierbarkeit (vgl. Kap. 4.2; Kap. 6.2), dazu übergehen, anstelle impliziter Typen-Gesetze die Korrelation zwischen neuronaler Aktivität und subjektiver Wahrneh-

[146] Das Gehirn kann als reizoffenes Organ als störanfällig betrachtet werden. Danach hat eine kleine Reizänderung eine große Wirkung auf das Gehirn. Wie stark die Störanfälligkeit des Gehirns gegenüber äußeren oder auch inneren Einflüssen tatsächlich ist, bedarf weiterer Forschung.

mung als Tokenrelation zu diskutieren. Hierfür spricht auch das empirische Problem, dass es teilweise deutliche individuelle Abweichungen bzw. Extremwerte im Vergleich zu den neuronalen wie subjektiven Durchschnittswerten der Probanden gibt. Die Tokenrelation als metaphysische Voraussetzung könnte hierfür eine Erklärung liefern und ein Kompromiss zwischen der empirischen neurowissenschaftlichen Herausforderung und der philosophisch metaphysischen Forderung nach strikten Gesetzen sein. Hierbei könnte man sich an Davidsons Überlegungen für eine Tokenidentität orientieren, wobei gerade er deutlich macht, dass es keine strikten psychophysischen Gesetze gibt (vgl. Kap. 2.3.2). In Verbindung mit der Idee der Realisierung bzw. gleichzeitigen Instanziierung scheint dies nicht mehr hinderlich zu sein. Es sind die Eigenschaftsinstanziierungen kausal wirksam als konkrete, partikulare Veränderung innerhalb einer bestimmten Zeit. Dabei müssen möglicherweise auf der epistemischen Ebene immer mentale und physische Eigenschaften für eine Instanziierung berücksichtigt werden. Dieser Umstand muss auch auf neurowissenschaftliche Korrelationsgesetze übertragen werden.

### 6.1.4 Realisierungsrelation als Alternative

So wie verschiedene philosophische Theorien innerhalb der Leib-Seele-Diskussion einen impliziten (Eigenschafts-)Dualismus verfolgen, muss auch die Neurowissenschaft allein aus methodischen Gründen zwei verschiedene Bereiche, den neuronalen und den subjektiv phänomenalen Bereich, voraussetzen. Dem Versuch nach handelt es sich dabei streng genommen um eine Relation zwischen drei Komponenten: zwischen den Bildsignalen, der subjektiven Perspektive sowie den fMRT-Daten. Letztere werden üblicherweise mit dem tatsächlichen neuronalen Zustand identifiziert. Für die Phase der Aufmerksamkeit gilt, dass vertikal das Pfeilsignal, ein subjektiver Zustand der Aufmerksamkeit und ein neuronaler Zustand, der das antizipatorische Netzwerk um den SAC umfasst, miteinander korrelieren. Ebenso korrelieren während der Phase des Bildsehens das Bildsignal, ein neutraler oder emotionaler subjektiver Zustand sowie ein neuronaler Zustand, der z. B. das perzeptive Netzwerk umfasst, miteinander.

Die Variante der horizontalen Kausalität wäre nun zwischen der Aufmerksamkeitsphase und der Bildphase anzusiedeln, wobei sich über die Ebenen hinweg die Frage stellt, was als Ursache für welche Wirkung in Frage kommt. Die starke Physikalismusthese besagt, dass jede physische Wirkung auf eine physische Ursache zurückzuführen ist (vgl. Kap. 3.3). Dabei spielt auch eine Rolle, dass die Kausalrelation innerhalb eines geschlossenen Systems besteht. Demnach müsste die neuronale Aktivität während der Aufmerksamkeitsphase ursächlich für die neuronale Aktivität während der Bildphase sein. Ebenso müsste das subjektive Erleben der beiden Phasen kausal aufeinander folgen. Die meisten philosophischen Positionen lassen zu, dass die subjektive

Ebene vertikal (bzw. supervenient o. ä.) auf die neuronale Ebene zurückzuführen ist. Auf diese Weise wird die zeitliche Abfolge der Ereignisse auf beiden Ebenen mit einer physischen Fundierung erklärt. Die subjektive Ebene bleibt aber kausal impotent.

Der Versuch legt diesen Schluss nicht ohne weiteres nahe. So wird die Relation um eine dritte Komponente erweitert. Die Versuchsergebnisse machen nun deutlich, dass aus neurowissenschaftlicher Perspektive der Fokus insbesondere auf die externen Stimuli zu legen ist. Schließlich ist es z. B. in Anschluss an das Aufmerksamkeitssignal ein emotionales bzw. neutrales Bild, das die Aktivität des perzeptiven Netzwerks verursacht. Nun kann zwar vertreten werden, dass damit immer noch ein physisches Signal Ursache für die neuronale Aktivität und die subjektive Ebene kausal irrelevant ist. Diese Position ist insofern problematisch, als damit ein Ebenenwechsel bzw. eine Systemöffnung – nämlich von der neuronalen Ebene zur Bildebene – einhergeht, die eigentlich nach der Physikalismusthese vermieden werden sollte. Außerdem ist es problematisch, ein emotionales bzw. neutrales Bild lediglich als physischen Reiz zu definieren, da der Inhalt eines solchen Stimulus zwar aus verschiedenen physischen Komponenten besteht, die wiederum in sich zwischen Makro- und Mikroebene, zwischen supervenient und subvenient unterteilt werden müssten, aber eben aufgrund des subjektiven Erlebens – und damit auf der mentalen oder psychologischen Ebene – verstanden und verarbeitet wird (vgl. Kap. 6.6). Dies zeigen in den vorliegenden Studien die Bewertungen der emotionalen Dimensionen Valenz, Intensität und Wiedererkennen, die eben zunächst auf der psychologischen Ebene angesiedelt werden. Somit scheint aus neurowissenschaftlicher Sicht die subjektive Perspektive bzw. Ebene durchaus eine kausale Rolle zu spielen mit Einfluss auf die neuronale Ebene.

Außerdem scheint es – hinsichtlich der kausalen Frage – nun zwei Ursachen zu geben, die eine Wirkung auf das perzeptive Netzwerk, genauer gesagt: auf den neuronalen Zustand des Gehirns zu dem Zeitpunkt $t_2$ {Bildsehen} haben, wobei $t_2$ auf $t_1$ {Pfeilsignal} folgt. Die eine Ursache, das antizipatorische Netzwerk, ist eine neuronale, die andere Ursache, das Bildsignal, ist eine abstrakte. Dies steht in Konflikt zu der Voraussetzung, wonach eine Wirkung nicht mehrere individuell hinreichende Ursachen haben darf. Nun kann zwar weiter differenziert werden, dass z. B. die Aktivität des perzeptiven Netzwerks durch das Bildsignal verursacht wird und der vorangehende neuronale Zustand mit Aktivität im antizipatorischen Netzwerk lediglich einen Verstärker bildet. Allerdings kann ebenso in einem kontrafaktischen Konditional ausgesagt werden, dass jener neuronale Zustand zum Zeitpunkt $t_2$ mit Ausbleiben einer der beiden Ursachen ein anderer wäre bzw. als solcher nicht da wäre (sondern ein anderer) und somit beide Ursachen den Zustand kausal verursacht haben. Neurophilosophisch ist dabei interessant, dass hier die Ursache-Wirkung-Relation begrifflich und auch methodisch weiter hinterfragt werden muss. So stellt sich eben auch die Frage, wie ein neuronaler Zustand zu einem bestimmten Zeitpunkt zu bestimmen ist: lediglich über einige besonders aktive Hirnregionen oder Netzwerke, oder „globa-

ler“ in Bezug auf das gesamte Gehirn? Davon hängt wiederum ab, welche Ereignisse oder Stimuli als Ursache für eine Veränderung des neuronalen Zustands in Frage kommen.

Der Begriff der Korrelation ist in diesem Zusammenhang problematisch. Man kann nicht einfach schließen, je höher der T-Wert bzw. die Aktivität in einer Region ist, desto „mehr“ Identität der Eigenschaften bzw. Ereignisse liegt vor. Denn bei anderen Regionen kann eine negative Korrelation festgestellt werden. Beispiel ist die negative Korrelation für die emotionale Dimension Wiedererkennen – je weniger IAPS-Bilder erinnert wurden, desto höhere T-Werte wurden im PACC gemessen. Dementsprechend scheint geringe Erinnerung emotionaler Stimuli mit ansteigender Aktivität im PACC zusammenzuhängen.

Zusätzlich wurde ein konverses Korrelationsmuster im VMPFC und im DLPFC beobachtet. Der VMPFC zeigte eine positive Aktivität zur Valenz, während neuronale Aktivität im DLPFC eher negativ korrelierte. Je positiver die emotionalen Bilder (einschließlich positiver und negativer Bilder) bewertet wurden, desto weniger Aktivität wurde im bilateralen DLPFC gemessen. Diese konverse Korrelation zwischen VMPFC und DLPFC ist in Übereinstimmung mit der reziproken Modulation zwischen medialem und lateralem präfrontalem Cortex während emotional-kognitiver Interaktion. Emotionale Prozessierung führt zu Aktivitätsanstieg im VMPFC, bei gleichzeitiger Aktivitätsverringerung im DLPFC. Kognitive Aufgaben induzieren dagegen das umgekehrte Muster mit Aktivitätsanstieg im lateralen präfrontalen Cortex und Verringerung im medialen präfrontalen Cortex. VMPFC und DLPFC sind unterschiedlich spezifisch empfindlich für die Modulation durch Valenz (vgl. Kap. 5.2.3). Die affektive oder evaluative Modulation erfolgt reziprok. Hinsichtlich der Realisierung ist die reziproke Modulation insofern interessant, als die Realisierung eines mentalen Zustands offensichtlich ein Prozess ist, der nicht nur der Aktivität eines neuronalen Areals bedarf, sondern der gleichzeitigen Verringerung eines anderen Areals. Es ist im Gehirn ein üblicher Vorgang, dass die Aktivität in verschiedene Areale verlagert wird. Dennoch macht dieser Umstand deutlich, dass ein Hirnzustand – und möglicherweise auch ein mentaler Zustand – nicht allein positiv zu bestimmen ist, sondern auch negativ durch die Reduzierung der Aktivität in anderen Regionen.

Diese Beispiele der negativen sowie der konversen Korrelation verdeutlichen, dass eine voreilige Identitätsannahme problematisch ist. Es ist nicht möglich zu prüfen, inwieweit Inhalt und neuronale Aktivität extensionsgleich sind. Die Neurowissenschaft untersucht lediglich die empirische Extension der neuronalen Korrelate. Ob diese extensionsgleich zu den psychologischen Funktionen sind, ist nicht zwingend. Die logische Unterscheidung zwischen einer offenen (mentalen) Menge und einer geschlossenen (physischen) Menge wird nicht gemacht. Idealerweise müsste vorausgesetzt werden, dass beide Mengen geschlossen und extensionsgleich sind. Beobachtete Korrelationen stützen sicherlich die Relation zwischen neuronaler Aktivität und subjek-

tivem Erleben. Allerdings muss dabei auch darauf hingewiesen werden, dass positive, negative und konverse Korrelationen nicht durchgängig ermittelt werden konnten (vgl. Kap. 5.2.2; Kap. 5.2.3). Dies lässt auf der empirischen Ebene zumindest Spielraum zu und verweist auf die Flexibilität und die starke Vernetzung der neuronalen Areale. Mitunter konnte eben kein Areal signifikant aus der gesamten neuronalen Aktivität isoliert werden. Ob es dabei um die Relation der Identität handelt, kann jedoch nicht bestimmt werden.

***Quintessenz Vergleich Regularität***

*mentale Kausalität impliziert drei Gesetzesarten*

- *neuronale Gesetze:*
- *Veränderung neuronaler Aktivität: Aktivität neuronaler Populationen notwendige Bedingung und externe Stimuli hinreichende Bedingung*
- *Grundfrage: biologische Prozesse nicht zwingend „kausal" unter „strikten Gesetzen"*

- *psychische Gesetze:*
- *praxisrelevant für neurowissenschaftliche Forschung*
- *feststellbare Regularität, mindestens auf Basis von ceteris-paribus-Gesetzen*
- *„hybride Eigenschaften": Verknüpfung emotionaler Dimensionen und Intentionalität hinsichtlich mentaler Kausalität*

- *Korrelationsgesetze:*
- *empirische-methodische Hintergrundannahme als ceteris-paribus-Gesetz*
- *keine Beschreibung der psychischen Ebene durch physische Gesetze*
- *Modell der Realisierung entspricht Tokenrelation: Instanziierung von Ereignissen*
- *Drei-Komponenten-Relation führt zu Systemöffnung: Erste-Person-Perspektive, neuronale Aktivität, externe Stimuli*
- *Grundfrage: was ist ein neuronaler Zustand?*

## 6.2 Emotionale Aufmerksamkeit und Multikausalität

Das Argument der Multirealisierbarkeit wendet sich in erster Linie gegen Reduktion und Identität mentaler und physischer Eigenschaften. Auch wenn beispielsweise Kim entgegen der ursprünglichen Stoßrichtung das Argument dazu verwendet, Reduktion zu begründen (vgl. Kim 1999b;

Kap. 4.2.2), kann das Argument in seiner funktionalistischen Ausprägung als valide gelten, da es sich schlüssig gegen einen (artenspezifischen) Chauvinismus wendet, und diese Position für verschiedene logisch mögliche Welten begründet wird. Zwangsläufig wird Multirealisierbarkeit damit auch für mentale Kausalität relevant, da sowohl Ursache als auch Wirkung und somit schließlich die Relation in sich multirealisierbar bzw. multikausal ist (vgl. Kap. 4.2.2).

Multikausalität ist insofern ein Kriterium für mentale Kausalität, als a) multirealisierte Eigenschaft und Realisierer in Relation zueinander stehen und b) auf der Realisierer- wie auch auf der Eigenschafts-Ebene verschiedene Ursachen für eine Eigenschaft in Frage kommen. Dies verdeutlich der entwickelte Grundsatz, wonach eine

> […] Eigenschaft F in einem System A als Ursache Eigenschaft G oder Eigenschaft H oder Eigenschaft I usw. haben kann, wobei die Eigenschaften G, H, I usw. für F hinreichend sein müssen, auch wenn sie selbst durch eine bestimmte Zahl variabler Eigenschaften realisiert sind (vgl. ebd.).

Anhand der vorliegenden Paradigmen wird nun diskutiert, inwiefern eine vertikale oder eine horizontale Multikausalität bestimmt werden kann. Es sollte also gezeigt werden, dass auf der psychologischen Ebene, auf der neuronalen Ebene und für die Korrelation multirealisierbare bzw. multikausale Ursachen bestimmbar sind.[147] Zudem sollte für die Korrelation gezeigt werden, inwiefern die miteinander korrelierenden Ereignisse in einer multiplen kausalen Relation zueinander stehen bzw. ob eine zeitliche Korrelation auch mit Multikausalität zusammenfällt.

### 6.2.1 Multikausalität auf der psychologischen Ebene

Auf der (horizontalen) psychologischen Ebene kann eine Multikausalität z. B. an den emotionalen Dimensionen festgemacht werden. Diese werden durch verschiedene emotionale Bilder ausgelöst bzw. verursacht, werden aber im subjektiven Erleben als konstant positiv, negativ oder neutral bewertet. Dabei können unterschiedliche positive Bilder wie ein lachendes Kind oder ein Liebespaar positive Emotionen auslösen; unterschiedliche negative Bilder wie eine Gewaltdarstellung oder ein Unfall können negative Emotion auslösen; und unterschiedliche neutrale Bilder wie ein Stuhl oder ein Tisch können eine neutrale Haltung auslösen.

[147] Streng genommen ist es nicht notwendig zu zeigen, dass auf neuronaler Ebene Neuronen, Areale o. ä. multirealisierbare Ursachen haben, da sie zunächst lediglich die physische Ebene, also eine rein physische Kausalität beschreiben. Andererseits wurde gezeigt, dass für die Korrelation eben jene multiple Kausalität relevant sein kann, da ein Gedanke $G_2$ indirekt von der physischen Eigenschaft $P_1$ über die physische Eigenschaft $P_2$ verursacht werden kann, wobei $P_1$ und $P_2$ beide multirealisierbare Ursachen von $G_2$ sein könnten (vgl. Kap. 4.2.2, Fall 2c).

Hiergegen sind auf der theoretischen Ebene verschiedene Einwände möglich. Zum einen ist nicht eindeutig festgelegt, inwiefern die dargestellten Bilder als physische oder eher als psychologische Signale aufzufassen sind (vgl. Kap.4.3). Im Zweifel wird hierbei bereits eine Kausalrelation zwischen verschiedenen Ebenen beschrieben. Zum anderen erscheinen die emotionalen Dimensionen positiv, negativ und neutral recht grob, so dass es nicht verwundert, dass verschiedene Erlebnisse in einer Schnittmenge unter diese Begriffe subsumiert werden können. Schließlich löst eine Gewaltdarstellung ein anderes negatives Gefühl aus als eine Unfallfotografie. Andererseits sind auch Bilder vorstellbar, die sich weniger deutlich voneinander unterscheiden lassen, aber unterschiedlich sind, und bei denen die Grenzziehung im emotionalen Leben weniger deutlich ausfällt. Beispiele wären Fotografien mit demselben Bildmotiv, aber verschiedenen Protagonisten: verschiedene Babys, die lachen, oder verschiedene Unfallszenarien. Im subjektiven Erleben wird eine Unterscheidung hierbei – wenn überhaupt – äußerst nuanciert ausfallen. Man könnte also durchaus davon sprechen, dass unterschiedliche Bildstimuli ein und denselben emotionalen Erlebniszustand verursachen.

Umgekehrt kann auch beobachtet werden, dass dieselben Bilder bei unterschiedlichen Probanden zwar als positiv, negativ oder neutral eingestuft werden, aber in den Dimensionen der Valenz oder der Intensität unterschiedlich erlebt werden. Zwar zeigen die Ergebnisse des Post-hoc-Tests eine ähnliche Tendenz für alle Probanden, im Einzelfall gibt es jedoch Abweichungen und unterschiedliche Ausprägungen. Hinsichtlich der Gruppe der Probanden ist ebenfalls anzunehmen, dass trotz desselben Aufmerksamkeitsstimulus das subjektive Erleben der Aufmerksamkeitsphase, der Bildphase, der Beurteilungsphase sowie der Aufmerksamkeitseffekt individuell variieren. Im Ergebnis werden diese unterschiedlichen Werte in der neurowissenschaftlichen Praxis gemittelt, um eine übergreifende Tendenz festzustellen. Insofern scheint das Erleben eines bestimmten Bildes mit bestimmten emotionalen Dimensionen zusammenzuhängen. Auf einer epistemischen bzw. ontologischen Ebene ist dagegen eine Einzelfallbetrachtung durchaus sinnvoll. Es ist sehr wohl ein Unterschied, ob man einen Durchschnittswert betrachtet oder individuelle, von einander abweichende Relationen. Schließlich verweisen letztere auf den Spielraum innerhalb der Relation von Ursache und Wirkung (vgl. Kap. 4.2). Hier wird ein Dilemma deutlich, da einerseits verschiedene Ursachen (Bilder) eine Wirkung (positiv, negativ, neutral) erzielen, jedoch andererseits auch eine Ursache (z. B. ein positives Bild) bei verschiedenen Probanden verschiedene Wirkungen (z. B. unterschiedlich intensiv) erzielen können. Während der erste Fall die Multikausalität betrifft, könnte der letztere als „inverse Multikausalität“ bezeichnet werden.

Beide Fälle zeigen, dass auf theoretischer Seite eine begriffliche Schärfung erfolgen müsste. Die Bewertung der Bilder als positiv, negativ und neutral erfolgt in ihren emotionalen Dimensionen äußerst individuell. Andererseits wird deutlich, dass ein emotionaler Zustand zu einem

bestimmten Zeitpunkt äußerst individuell und spezifisch ist. Es dürfte schwierig zu beurteilen sein, ob ein Proband sich zwei Mal in genau demselben psychologischen (oder auch biologischen) Zustand befindet. Auf dieses Problem der begrifflichen Unschärfe weist auch Papineau hin (vgl. Papineau. 2004, 225). Psychologische Begriffe bleiben in letzter Konsequenz vage und somit schwerlich vollständig zu klassifizieren.

Ferner kann auch a) die Relation (emotionale) Aufmerksamkeit und (emotionales) Bildsehen oder auch b) die Relation (emotionales) Bildsehen und (emotionale) Bildbeurteilung als Variante der Multikausalität auf psychologischer Ebene betrachtet werden. Bei a) handelt es sich für den einzelnen Probanden zunächst einmal um die umgekehrte Relation, da auf denselben Aufmerksamkeitsstimulus – das Fixierungskreuz – verschiedene Bildstimuli bzw. verschiedene subjektive emotionale Zustände folgen, während bei b) verschiedenen Bildern verschiedene Beurteilungen folgen.

Es ist eine mögliche Aufgabe, empirisch festzustellen, ob es sich in der Aufmerksamkeitsphase tatsächlich um ein und denselben psychologischen (oder auch neuronalen) Zustand bei den einzelnen Probanden handelt. Immerhin ist es gut möglich, dass der subjektive Eindruck (wie auch die neuronale Messung) der Probanden trotz desselben Aufmerksamkeitsstimulus variiert. Ebenso wäre zu prüfen, ob das (emotionale) Bildsehen der einzelnen Probanden bei ein und demselben Bild im subjektiven Eindruck variiert und ob ferner die gesteigerte Aufmerksamkeit – also der Aufmerksamkeitseffekt – im Einzelfall variiert oder konstant ist. Dasselbe gilt auch für den Fall der (emotionalen) Bildbeurteilung. Auch wenn es derzeit empirisch eher offen zu sein scheint, ob es innerhalb der einzelnen Relationen eine Form der Multikausalität gibt, sprechen die genannten Zusammenhänge eher dafür.

### 6.2.2 Multikausalität auf der neuronalen Ebene

Bei der Betrachtung Multikausalität auf neuronaler Ebene stößt man parallel zur psychologischen Ebene einmal auf Hirnareale bzw. Netzwerke, die durch verschiedene vorangehende Hirnareale bzw. Netzwerke aktiviert werden können sowie auf den umgekehrten Fall der „inversen Multikausalität", dass ein und dasselbe Netzwerk verschiedene weitere Netzwerke beeinflusst. Dieser Umstand ist auf die Interkonnektivität, die umfassende Vernetzung des Gehirns, zurückzuführen. Für die vorliegenden Paradigmen kann dabei einmal auf die Region des SAC hingewiesen werden, der eine multi-integrative Funktion zukommt. Diese Region umfasst eine Vielzahl an affektiven, kognitiven und motorischen Prozessen in Beziehung zu adaptivem Verhalten (vgl. Kap. 5.1.2). Das bedeutet, dass sie einerseits in Zusammenhang mit unterschiedlichen Hirnarealen,

z. B. dem motorischen Cortex oder dem präfrontalen Cortex aktiviert wird. Dies könnte als horizontaler Version der Multikausalität auf neuronaler Ebene ausgelegt werden.

Zudem wurde aufgrund der Datenanalyse deutlich, dass es nicht nur einen Signalanstieg im SAC während der Phase der emotionalen Aufmerksamkeit gab, sondern auch während der Phase der neutralen Aufmerksamkeit. Daraus ergibt sich die neurowissenschaftliche Annahme, dass auch die unter der Aufmerksamkeitskondition gezeigten Pfeile einen nicht-spezifischen Aufmerksamkeitseffekt im SAC hervorrufen. Das Signal ist zwar während der emotionalen Aufmerksamkeit stärker als während der neutralen Aufmerksamkeit. Dennoch handelt es sich auch hierbei um zwei verschiedene Ursachen, die kausal die Aktivierung des SAC bewirken: eine nicht-spezifische und eine spezifisch emotionale Aufmerksamkeit. Im SAC werden beide Konditionen prozessiert.[148] Außerdem gibt die Region des SAC wiederum Signale an verschiedene Regionen weiter und löst somit ursächlich Aktivität in diesen aus. So umfasst die Aufmerksamkeit emotionaler Bilder auch einen Zustand der Bereitschaft für motorische Reaktion. Die eine Ursache „Aktivität im SAC" kann also vielerlei Wirkung, nämlich „Aktivität in verschiedenen Regionen", haben.[149]

Den Hirnregionen kommt eine funktionsspezifische Stabilität zu. So kann die SAC-Region mit CMA, SMA und POS als antizipatorisches Netzwerk von einem perzeptiven Netzwerk mit Amygdala, Insula, medialem und lateralem präfrontalem Cortex, Putamen, Mittelhirn, Cerebellum und occipitotemporalen visuellen Regionen unterschieden werden. Hierfür scheint auch die getrennte neuronale Repräsentation der emotionalen Dimension, wie auch die Unterscheidung zwischen kognitiven und emotionalen Prozessen im Zusammenhang mit emotionaler Valenz zu sprechen. So kann für die Regionen VMPFC und DLPFC ein konverses Korrelationsmuster beobachtet werden, das auf eine funktionale Unterscheidung zwischen beiden Regionen zurückzuführen ist (vgl. Kap. 5.2.3).

Solche Beobachtungen sprechen zunächst dafür, dass auf empirischer Ebene innerhalb der neuronalen Kondition keine Multirealisierung oder gar Multikausalität vorhanden ist. Es sind konkrete Regionen, denen eine spezifische Funktion zukommt. Die Idee der Eigenschaftsvariabilität kann aber insofern erhalten werden, als die individuellen Unterschiede betont werden. Der plastische Aufbau des Gehirns bzw. die Verfeinerung der Netze ist z. B. von individuellen sensorischen Erfahrungen abhängig. Wenn genau dasselbe IAPS-Bild zwei Personen A und B über einen längeren Zeitraum täglich gezeigt wird, werden die Gehirne beider Personen vor dem Hin-

[148] Ein weiteres Beispiel aus der Aufmerksamkeitsstudie ist die Aktivierung des POS, der von den frühen visuellen Regionen hin zum posterioren parietalen Cortex sendet. Bei der vorliegenden Studie wurde der POS auch während der Aufmerksamkeitsphase aktiviert, wobei keine visuell-orientierte Handlung involviert war. In Ergänzung zu früheren Studien folgt, dass antizipatorische Aktivierung im dorsalen Verlauf in Anwesenheit und in Abwesenheit handlungsorientierter visueller Stimulation verursacht wird (vgl. Kap. 5.1.2).

[149] Diese Tatsache macht es schwer, allein aufgrund der Daten aus bildgebenden Verfahren auf den Zustand des Probanden zu schließen, zumal es sich bei den dichten Verbindungen des SAC zu primären und sekundären motorischen Regionen eher um eine wechselseitige und weniger um eine einseitig lineare Verbindung handelt.

tergrund ihrer weiteren persönlichen Entwicklung im Detail unterschiedliche Strukturen ausbilden. Außerdem werden sich die Gehirnstrukturen von A und von B selbst innerhalb des Zeitraumes unterscheiden. Hier wäre zu untersuchen, ob damit auch Konsequenzen für die Funktion der neuronalen Regionen verbunden sind. Das philosophische Argument der möglichen Welten erweitert die Perspektive: So gibt es auf der empirischen Seite mit den verschiedenen Hirnarealen ganz verschiedene „Arten" von Ursachen, während auf philosophischer Seite derselbe Ursachetyp durch eine andere Materieart (Neuronen, Silizium, grüner Schleim etc.) „ausgetauscht" wird. So gesehen muss auch hinsichtlich der Materieart „Neuronen" nicht lediglich ein „Standardgehirn", sondern gerade der Unterschied individueller menschlicher Gehirne betrachtet werden. So sind über verschiedene Probanden hinweg auch verschiedene Ausprägungen der Hirnregionen beobachtbar. Aber auch diese bestehen letztlich aus derselben Materie, aus Eiweiß. Für eine weitergehende Differenzierung ist hier die molekulare Neurowissenschaft gefordert (vgl. Kap. 4.4.3).[150]

Um die These der Multirealisierung bzw. Multikausalität zumindest bedingt für die neuronale Ebene zu stützen, wäre es insbesondere von Interesse, Randbereiche bzw. Extreme zu betrachten, wie z. B. Gehirne, in denen aufgrund von Läsionen oder anderer Veränderungen ein Hirnareal eine weitere oder gar eine neue Funktion übernimmt und somit in eine Relation zu ganz anderen Hirnregionen tritt. Ebenso kann abermals auf den inzwischen durchgeführten Einsatz von Mikrochips hingewiesen werden, die z. B. bei Parkinsonpatienten zumindest zeitweilig erfolgreich implantiert werden können. Für die mentale Kausalität ist dabei entscheidend, dass einmal unterschiedliche Materie eine kausale Funktion im neuronalen Netzwerk einnehmen kann und außerdem, wie oben beschrieben, unterschiedliche Regionen Ursachen für Aktivität in weiteren Regionen sein kann. Dabei wird im letzteren Fall der philosophische Begriff der Multikausalität insofern modifiziert (und möglicherweise abgeschwächt), als auf der horizontalen Ebene eine Wirkung verschiedene Ursachen haben kann. So kann auch das Herunterfallen eines Dachziegels verschiedene Ursachen (z. B. Wind, Dachdecker) haben. Ebenso scheinen die stabilen Strukturen im Gehirn flexibel genug, auf neue Ereignisse und Stimuli zu reagieren.

[150] Der Bereich der molekularen Neurowissenschaft setzt sich zwar kaum direkt mit kognitiven Strukturen auseinander. Über Verfahren wie Liquoranalyse, histologische Hirnuntersuchungen, Zellbiologie, Protein- und DNA-Analysen wird die molekulare Struktur des Gehirns untersucht. Unter anderem sollen Genmuster für bestimmte kognitive Dispositionen, wie z. B. Depression, ermittelt werden. Die molekulare Neurowissenschaft könnte prüfen, inwiefern eine vertikale Multirealisierung im Sinne der Schichtenwelt gegeben ist.

### 6.2.3 Multikausalität innerhalb der Korrelation

Schließlich muss das Kriterium der Multikausalität für mentale Kausalität auch hinsichtlich der Korrelation der psychologischen und der neuronalen Ebene diskutiert werden. Naturgemäß werden die oben dargestellten Kernpunkte der beiden Ebenen auch die Korrelation betreffen. Schließlich sind beide horizontalen Ebenen aufgrund der vertikalen Verknüpfung von Bedeutung für die Art und Weise, in der die kausale Relation möglich ist.

Tatsächlich sind verschiedene Korrelationen beobachtbar, die grundsätzlich auf eine Realisierung schließen lassen: SAC, CMA, POS sind während der Aufmerksamkeitsphase aktiv, ebenso wie Amygdala, Insula, DLPFC, Cerebellum und occipitotemporale Regionen während der Bildphasen aktiv sind. Erneut ist hierbei die SAC-Region hervorzuheben, da sie sowohl während der emotionalen als auch während der neutralen Aufmerksamkeit aktiv ist. Hier könnte gegebenenfalls sogar von einer „inversen Multikausalität" gesprochen werden – eine Region prozessiert bzw. realisiert zwei verschiedene mentale Zustände. Hier stellt sich die Frage, ob ein physischer Zustand nicht mit mehreren mentalen Zuständen korrelieren kann. Dies zeigt sich auch in der bereits erwähnten multiintegrativen Funktion der Region. Hierbei muss allerdings auf die Schwierigkeit hingewiesen werden, auf psychologischer Ebene die beiden mentalen Zustände genau zu differenzieren. Bei der emotionalen wie auch der neutralen Aufmerksamkeit handelt es sich möglicherweise lediglich um ein Konstrukt, das zwar auf neuronaler Ebene darstellbar ist, sich aber nicht im subjektiven Erleben der Probanden widerspiegelt. Schließlich werden diese Zustände methodisch vor allem durch Kontrastierung berechnet (vgl. Kap. 5.1.1).[151]

Ein weiteres Beispiel für die (inverse) Multikausalität ist die reziproke Modulation von VMPFC und DLPFC während der emotionalen Valenz: emotionale Prozessierung führt zu Aktivitätsanstieg im VMPFC bei gleichzeitiger Aktivitätsverringerung im DLPFC, während kognitive Prozessierung das umgekehrte Aktivitätsmuster induziert. Zudem überrascht hierbei sicherlich wenig, dass ein Gedanke bzw. ein subjektiv-mentaler Zustand vielfach realisiert wird, nämlich durch das Zusammenspiel vieler verschiedener Hirnareale. Insofern ist ein mentaler Zustand in einem sehr einfachen Sinne tatsächlich „multirealisiert".

Die verschiedenen mitunter sehr detaillierten Korrelationen sprechen dagegen für eine stabile Realisierung, weniger für die vertikale Multirealisierbarkeit. Korrelationen zwischen der neuronalen und der psychologischen Ebene konnten für die emotionalen Dimensionen Valenz (VPMFC,

[151] Beispielsweise wurde, um den Aufmerksamkeitseffekt bei der emotionalen Kondition zu ermitteln, zuvor die neuronale Aktivität bei erwarteten emotionalen Bildern mit der neuronalen Aktivität bei unerwarteten emotionalen Bildern kontrastiert. Dieser Kontrast zeigte signifikante Aktivität im medialen und lateralen präfrontalen Cortex, im parahippocampalen Gyrus, im Putamen, im Thalamus, im dorsalen Mittelhirn, im temporalen Lobulus, im Cerebellum und in occipitotemporalen visuellen Regionen (vgl. Kap. 5.1.1).

VLPFC, DLPFC), Intensität (VLPFC, DMPFC) und Wiedererkennen (PACC) beobachtet werden. Diese und andere signifikante und spezifizierte Korrelationen müssen jedoch nicht als empirischer Hinweis pro mental-physische Identität und sogar kontra Multirealisierbarkeit aufgefasst werden. Zwar scheint eine spezifische psychologische Gruppe mit genau einer weiteren neuronalen Gruppe zu korrelieren. Allerdings kann eingewendet werden, dass solche Korrelationen lediglich vereinzelt und nicht für alle Phasen des Paradigmas gemessen wurden. Außerdem wurden sowohl positive, als auch negative sowie konverse Korrelationen gemessen. Dabei zeigt sich wiederholt, dass die neuronale Aktivität in einen größeren Gesamtzusammenhang einzuordnen ist. Es ist nicht eine einzige Region, die einen Gedanken prozessiert, sondern das gesamte Gehirn. In dieser Region wird zumindest verstärkte neuronale Aktivität verzeichnet, sie scheint somit notwendig für die Prozessierung zu sein. Aufgrund individueller Unterschiede kann es jedoch zu abweichender und damit unterschiedlicher Realisierung kommen (vgl. Kap. 6.1.4). Außerdem können – wie im vorhergehenden Abschnitt erörtert – neuronale Funktionen inzwischen tatsächlich, wenn auch eingeschränkt, von Siliziumchips ausgeführt werden – was eben nicht nur logisch mögliche Welten betrifft. Das empirische Hauptargument pro Multirealisierbarkeit knüpft daran an und verweist auf die immense Plastizität des Gehirns, welches potenziell in der Lage ist, nach der operativen Entfernung einer gesamten Hirnhälfte (!) zahlreiche Funktionen durch die übrig gebliebene Gehirnhälfte zu übernehmen.[152]

Andererseits kann eingewendet werden, dass derselbe physische Typus – nämlich das menschliche Hirn, also Eiweiß etc. – betroffen ist und insofern dieselben physischen und mentalen Typen miteinander korrelieren, was nicht gegen eine Typenidentität spricht. Damit bewegt man sich jedoch auf einer ganz anderen Ereignis- bzw. Eigenschaftsebene.

In der philosophischen Diskussion wird in der Argumentation pro Multirealisierbarkeit auf neurowissenschaftliche Erkenntnisse oder pathologische Befunde verwiesen, um das Argument empirisch zu stützen. Stichwort ist hier insbesondere die Plastizität des Gehirns (vgl. Kap. 4.2). Tatsächlich lassen sich aus der neurowissenschaftlichen Perspektive auf empirischer Ebene einige Hinweise sammeln, die für Multirealisierbarkeit auf der empirischen Ebene sprechen. Allerdings scheint Multirealisierbarkeit zunächst in einem Widerspruch zur neurowissenschaftlichen Arbeit zu stehen, da allein methodisch davon ausgegangen werden muss, dass genau einem durchschnittlichen psychologischen Zustand ein durchschnittlicher neuronaler Zustand entspricht.

[152] Mithilfe von TMS kann geprüft werden, ob eine Region notwendig für eine Handlung ist. Interessant ist die Kompensationsfähigkeit des Gehirns; so kann das Areal für Sehen bei Blinden Tastfunktionen übernehmen. Ferner sind auch umliegende Nervenbahnen in die neuronalen Prozesse miteingebunden; hier ist schwer zu ermitteln, welche regionale Aktivität notwendig für Verhalten ist. Als hinreichend könnte das Hinzukommen eines verhaltensspezifischen neuronalen Prozesses angenommen werden.

Mögliche Formen der Multirealisierbarkeit müssten demnach subtil verlaufen und sich eventuell von der philosophischen Annahme der Multirealisierbarkeit unterscheiden.

Die theoretisch-logische Annahme einer Multirealisierbarkeit des Mentalen konfrontiert uns zudem mit diversen empirischen Problemen. Zum einen ist es derzeit faktisch (noch) nicht möglich, psychologische Funktionen und neuronale Korrelate in der Extension genau zu identifizieren, auch wenn diese Identifizierung für eine ideale Neurowissenschaft als möglich unterstellt wird. Wenn wir darüber hinaus aber in Betracht ziehen, dass auch die neuronalen Korrelate von dem Gesamtzustand des „Systems Mensch" im Fluss der Zeit abhängen, also von Stoffwechselprozessen, hormonellen Schwankungen u. a., ist es problematisch, identische neuronale Korrelate für die mentalen Zustände anzunehmen. Dass heißt, die Proposition einer Person „Auto" ist zu jedem Zeitpunkt anders bzw. zu keinen zwei Zeitpunkten identisch, weder psychologisch, noch neuronal. Für ein System S gilt demnach $M_{t1} \neq M_{t2}$. Damit ergeben sich aber für die physischen Zustände dieselben Probleme, die mentalen Zuständen zugeschrieben werden: sie sind nicht mehr generalisierbar, erklärbar und vorhersehbar.

Dem Problem der ungleichen Extension versucht Kim durch die Annahme disjunkter Eigenschaften zu begegnen (vgl. Kap. 4.2.1). Abgesehen davon, dass disjunkte Eigenschaften keine natürlichen Arten sind, scheint auch diese Variante keine genaue Erklärung oder Prognose zu sichern. In sich bliebe eine Eigenschaft, die aus einer disjunkten Menge besteht, prinzipiell heterogen und variabel. Strikte physische Gesetze für neuronale Aktivität sind auf diese Weise genauso fragwürdig, wie strikte psychische Gesetze zu Gedanken. Auch wenn für emotionale Aufmerksamkeit ein neuronales Netzwerk identifiziert werden kann, bleibt dieses Netzwerk zumindest im Detail in seiner Anordnung variabel im Unterschied von Proband zu Proband. Und auch innerhalb desselben Systems bzw. desselben Probanden kann das neuronale Netzwerk für emotionale Aufmerksamkeit im Detail variieren und sich verändern. Mit der Idee disjunkter Eigenschaften ist Kim näher an Davidsons Tokenrelation, als intendiert. Kim hat jedoch lediglich physische Eigenschaften als disjunkte Eigenschaften im Blick. Schließlich wären konsequenterweise letztlich beide Ebenen disjunkt, die mentale und die physische Ebene, da nicht nur physische Eigenschaften, sondern auch mentale Eigenschaften variabel sind – so ist das Gehirn, wie der Gedanke unter holistischen Gesichtspunkten von Mensch zu Mensch verschieden.

An dieser Stelle kann erneut die Frage gestellt werden, auf welcher Ebene Realisierung bzw. Multirealisierung diskutiert werden muss. Dass die Regionen empirisch voneinander abgrenzbar sind, ist kein zwingendes Argument gegen (metaphysische) Multirealisierbarkeit. Empirisch müsste dagegen gezeigt werden, dass genau dieselben Synapsen respektive Neuronen die entsprechenden Gedanken realisieren. Das erfordert ein konsequenter Identitätsbegriff. Andererseits kann sogar auf der empirischen Ebene selbst eingewendet werden, dass letztlich eine Synapse eine Synapse ist, sprich alle Synapsen bzw. Neuronen sind mehr oder weniger gleich und können des-

halb jede Funktion erfüllen. Das erklärt auch die Plastizität des Gehirns. Dies scheint aber in Widerspruch oder zumindest in einer Spannung zur Spezifizierung der Hirnareale oder bestimmter Netzwerke zu stehen. Zumindest wirft dies Fragen bezüglich ihrer Ausdifferenzierung auf: was unterscheidet Netzwerke voneinander? Zudem stellt sich damit die Frage der inversen Multirealisierbarkeit. Wenn alle Neuronen gleich sind, warum können sie verschiedene Funktionen erfüllen? Dies stünde auch im Widerspruch zu einem philosophischen Identitätsbegriff. Dieser müsste gegebenenfalls modifiziert werden, um speziell für psychoneuronale Identität zu gelten. Einen möglichen Ausweg bietet auch hier die Konzentration auf die Instanziierung, auf die Ereignisebene, als eigentliche Ebene kausalen Geschehens (vgl. Kap. 4.2.2). Eine Ereigniskausalität scheint über die mentale wie physische Eigenschaftsvariabilität hinweg eine relative Stabilität zu sichern.

Dabei kann hinterfragt werden, wer eigentlich wen realisiert bzw. verursacht. Hinsichtlich der Variante der Cross-Level-Kausalität wurde bereits festgestellt, dass sie insofern problematisch ist, als einerseits die physischen Eigenschaften als Realisierer und/oder Ursachen in eine doppelte Verantwortung, und andererseits die mentalen Eigenschaften als Realisierungen und/oder Wirkungen in eine doppelte Abhängigkeit geraten (vgl. Kap. 4.2.2). Hier zeigt die Abgrenzung des Netzwerkes der emotionalen Aufmerksamkeit von dem Netzwerk der emotionalen Bildwahrnehmung, dass empirisch eine Unterteilung zwischen antizipatorischen und perzeptiven Bestandteilen der Stimulus-Prozessierung gemacht werden kann (vgl. Kap. 5.1.3). Ebenso scheint es getrennte neuronale Repräsentationen für emotionale Dimensionen zu geben (vgl. Kap. 5.2.3). Dies legt die Vermutung nahe, dass unterschiedliche mentale Eigenschaften durch getrennte physische Eigenschaften realisiert werden. Hinsichtlich der kausalen Frage muss aber zusätzlich die Relation zwischen emotionaler Aufmerksamkeit und emotionaler Bildwahrnehmung oder Bildbeurteilung betrachtet werden. Hierbei wird deutlich, dass durchaus ein kausaler Zusammenhang zwischen den verschiedenen Netzwerken besteht. Beispielsweise konnte für die Region Amygdala ein spezifischer Effekt der emotionalen Aufmerksamkeit während der Phase der Bildwahrnehmung gezeigt werden: unter der Aufmerksamkeitskondition wurde größere Aktivität gemessen als bei neutraler Bildwahrnehmung (vgl. Kap. 5.1.3). Somit ist es also durchaus möglich, eindeutige kausale Relationen zwischen verschiedenen Netzwerken darzustellen. Andererseits macht diese Unterscheidung zwischen separaten Netzwerken, dass eine Erklärung der mental-physischen Korrelation allein auf der Basis der physischen Eigenschaften bzw. neuronaler Aktivität nicht möglich ist. Dies geht aus dem Umstand hervor, dass es wenig oder keine verstärkende Voraktivierung im perzeptiven Netzwerk vor der emotionalen Bildprozessierung gibt (vgl. ebd.). Schließlich verursacht der Zustand der emotionalen Aufmerksamkeit nicht allein den Zustand der verstärkten emotionalen Bildwahrnehmung; letzterer wird wesentlich unabhängig prozessiert und durch externe Stimuli verursacht. Insofern ist es problematisch, mentale und physische Eigen-

schaften losgelöst von ihrer Instanziierung, dem gesamten Ereignis, auf ihre kausale Wirkung hin zu prüfen.

In der Diskussion um das Kriterium der Relation wurde bereits dargestellt, dass nicht nur eine unidirektionale oder bidirektionale Relation zwischen mentalen und physischen Eigenschaften bzw. zwischen der psychologischen und neuronalen Ebene zu betrachten ist, sondern darüber hinaus eine äußere Ebene, auf der im Rahmen des Versuchs die visuellen Stimuli anzusiedeln sind, hinzugezogen werden muss. Die Beobachtung zeigt, dass einige Regionen, wie z. B. der SAC, multiintegrativ sind.[153] Andere Regionen wie VMPFC, DLPFC, VLPFC oder DMPFC sind zwar stärker in ihrer funktionalen Rolle spezifiziert. Sie sind aber auch bei unterschiedlichen neuronalen Prozessen eingebunden und nicht nur bei exakt identischen Prozessen aktiv. So scheint neuronale Aktivität im DMPFC emotionale Intensität hinsichtlich der Interaktion zwischen emotionaler Beurteilung und Aufmerksamkeit zu modulieren. Diese Aktivität betrifft aber verschiedene extramentale Fälle von emotionaler Beurteilung und Aufmerksamkeit. Sowohl positive als auch negative oder sogar inhaltliche Variation scheint hier möglich. Auch die funktionale Rolle des PACC lässt sich insofern spezifizieren, als seine neuronale Aktivität parametrisch von dem Wiedererkennen emotional relevanter äußerer Ereignisse abhängt. Unter funktionalistischen Gesichtspunkten würde man hier von einer funktionalen Rolle sprechen, die jedoch stark von äußeren Bedingungen abhängt (vgl. Kap. 2.2.1; Kap. 2.2.5). Ebenso sind es auf einer allgemeinen Ebene unterschiedliche emotionale extramentale Bilder, die durch das perzeptive Netzwerk, aber auch durch das antizipatorische Netzwerk prozessiert werden. Somit kann ausgesagt werden, dass Aktivität im perzeptiven Netzwerk durch unterschiedliche visuelle Ereignisse ausgelöst bzw. „realisiert" wird. Selbiges könnte für die oben genannten Fälle der „inversen Multikausalität" gelten. Die physischen Eigenschaften scheinen multikausal verursacht zu sein.

Dieser Gedanke ist zunächst einmal befremdlich, da doch üblicherweise die Neuronen als genuin physisch als potenzielle Ursachen oder Realisierer in Frage kommen. Unter der Voraussetzung, dass alles in dieser Welt physisch ist, kann der Kreis der Verursacher nun aber auch auf äußere Stimuli erweitert werden. Schließlich kann neurowissenschaftlich untersucht werden (und wird untersucht), inwieweit Lernen eine psychologische, wie neuronale „Anpassung" an Umweltereignisse darstellt (vgl. z. B. Dretske 1995) Wenn sich neuronale Verknüpfungen aufgrund wiederholter Erfahrungen, also dem wiederholten Kontakt mit externen Stimuli, aufbauen und stabilisieren, ist es plausibel, hierbei von einer „kausalen Realisierung" durch die Stimuli zu sprechen. Der Begriff der Multikausalität erscheint auf der empirischen Ebene somit auch einen bidirektionalen Geltungsbereich zu beschreiben.

[153] Umgekehrt gibt es beim Netzwerk für emotionale Aufmerksamkeit auch eine Schnittmenge aus dem antizipatorischen und dem perzeptiven Netzwerk. Ebenfalls wird bei emotionaler und kognitiver Prozessierung eine Überlappung im medialen und im lateralen Präfrontalen Cortex beobachtet. Insofern haben die Netzwerke eine gemeinsame Rolle.

***Quintessenz Vergleich Multikausalität***

- *empirische Bestätigung der theoretischen Position: auf psychologischer wie neuronaler Ebene verschiedene Ursachen der Eigenschaften möglich*

*psychologische Ebene:*

- *„grobe" Multikausalität bei emotionalen Dimensionen; „nuancierte" Multikausalität des subjektiven Erlebens*
- *inverse Multikausalität: einem Ursacheereignis folgen mehrere mögliche subjektive Erlebniszustände*
- *Notwendigkeit der Einzelfallbetrachtung als Folge begrifflicher Unschärfe: Identifizierung übertragbarer mentaler Zustände problematisch*

*neuronale Ebene:*

- *Multikausalität und inverse Multikausalität: multiintegrative neuronale Populationen (z. B. SAC)*
- *relevante individuelle Unterschiede trotz funktionsspezifischer Stabilität*
- *Einordnung neuronaler Aktivität in größeren, gehirnumfassenden Gesamtzusammenhang stützt Multirealisierbarkeit*

- *konsequenter Identitätsbegriff empirisch nicht haltbar; für mentale Kausalität Auflösung über Ereignisebene möglich*
- *Erklärung des mental-physischen Ereignisses allein auf physischer Ebene nicht möglich*
- *Zusätzliche Multikausalität durch äußere Ebene als externe Stimuli*
- *„Kausale Realisierung" als neuronale Verknüpfung durch externe Stimuli*

## 6.3 Emotionale Aufmerksamkeit und Inhalt

Ein zentrales mit der Frage der mentalen Kausalität verbundenes Problem ist zu erklären, wie geistiger Inhalt kausal in einer physischen Welt z. B. im Rahmen einer Handlung wirksam sein kann. Es ist die Frage, wie mentale Zustände die Zustände und Ereignisse in der Welt repräsentieren bzw. wie sich eine geistige Welt an sich darstellen lässt. Eine schlichte Identifizierung der neuronalen Eigenschaften mit den Inhaltseigenschaften führt nicht zum gewünschten Erfolg.[154] Neuronale Prozesse scheinen per se nicht empfindsam für (semantische) Inhalte zu sein, es bedarf weiterer Klärung, wie typischerweise abstrakter Inhalt naturalisiert bzw. auf neuronale Prozesse zurückgeführt werden kann. Dieses Problem wird im Vergleich zwischen der empirischen Studie zur Emotionalen Aufmerksamkeit und der philosophischen Position der Language of Thougth-Hypothese zur Repräsentation von Inhalt übergreifend diskutiert. Dabei wird auch untersucht, inwiefern die Annahme einer syntaktisch strukturierten Gedankensprache bzw. die Annahme des Denkens als Prozess kausaler Zeichensetzung durch die neurowissenschaftlichen Versuche unterstützt werden kann. Das Problem der speziell „semantischen Repräsentation" erfährt durch die Unterscheidung zwischen „weitem" und „engem" Inhalt eine Zuspitzung. Die folgende Diskussion zeigt, dass diese Zuspitzung weder zwingend noch zielführend ist. Hier scheint weitere Begriffsarbeit erforderlich. Das Projekt der Naturalisierung, so es gelingen soll, muss die mit Inhalt verbundenen Phänomene ernst nehmen.

### 6.3.1 Grenzen repräsentationaler Theorien

Inhaltseigenschaften scheinen fundamental verschieden von biologischen oder physischen Eigenschaften zu sein. Es stellt sich somit die Frage, welchen Platz Inhaltseigenschaften überhaupt in der physischen Welt einnehmen können – sollten sie nicht naturalisiert werden können. Verschiedene Theorien, wie z. B. kausale, informationstheoretische oder teleologische Ansätze, diskutieren die Möglichkeit der Naturalisierung mentaler Repräsentation (vgl. Kap. 4.3.1; Kap. 4.5.1). Im Folgenden soll die Language of Thought-Hypothese zwar nicht im Einzelnen auf ihre empirische Validität hin überprüft werden. Vielmehr bietet es sich an, anhand der vorliegenden Studien zu diskutieren, inwieweit das Projekt einer Naturalisierung von Inhalt überhaupt sinnvoll sein kann.

Üblicherweise wird im Rahmen der LOT-Hypothese angenommen, dass der Inhalt eines Gedankens durch den Inhalt seiner Begriffe und deren syntaktische Rollen bestimmt wird. Die

[154] Dies hat schließlich auch zur Hinwendung von der Identitätstheorie zum Funktionalismus geführt (vgl. Kap. 2.1.3; Kap. 2.2).

Inhalte der Begriffsrepräsentationen werden durch Kausalrelationen gebildet (vgl. Fodor 1987, 1990b). Auf die Studien übertragen, sind mögliche Begriffe an die gezeigten Bilder und Stimuli gekoppelt. Dem Symbol bzw. Begriff „Pfeil" entspricht eine neuronale Repräsentation. Gleiches gilt auch für die auf den IAPS-Bildern gezeigten Gegenstände, Personen oder sogar Situationen. Für einfache Wahrnehmungen für Begriffe wie z. B. „Hund", „Baum" oder „Auto" scheint die Kausaltheorie zunächst plausibel zu sein: ein bestimmtes Ereignis in der Welt wird im Gehirn des Menschen neuronal repräsentiert (vgl. Kap. 4.3.1).

Auch für den Versuch ist vorstellbar, dass das neutrale IAPS-Bild „Stuhl" im Probanden ein vorhandenes neuronales Aktivitätsmuster für „Stuhl" hervorruft. Da die Probanden sich allesamt in einem mittleren Alter befanden, kann davon ausgegangen werden, dass ein entsprechendes Aktivitätsmuster für „Stuhl" bei ihnen bereits vorhanden war. Allerdings waren ihnen das gezeigte Bild und somit auch der gezeigte „Stuhl" selbst unbekannt. Das trifft auch auf andere Bilder zu. Die Probanden waren also in der Lage, neue Situationen und Gegenstände zu repräsentieren, gegebenenfalls durch Kombination bekannter Formen o. ä. auf eine im Konkreten unbekannte Szenerie. Insofern scheint ein einfaches Kausalitätsmodell nicht ausreichend, den Prozess der Repräsentation von Inhalt zu beschreiben. Jenseits der Aktivierung der vorhandenen Repräsentation erfolgen auch neue Verknüpfungen, da der Stuhl in einem Raum steht und aus einem bestimmten Material in einer spezifischen Form beschaffen ist. Diese Herausforderung einer endlosen Produktivität versucht die LOT-Hypothese mit dem Verweis auf eine kombinatorische Syntax abzudecken (vgl. Kap. 4.3.1). Zusätzlich kann angenommen werden, dass sich in unserem Gehirn „neuronale Prototypen" befinden (vgl. Churchland/Sejnowski 1992), die auf neue Objekte und Kontexte bezogen werden können. Damit müssen neuronale Repräsentationen mehr leisten, als einfache Abbilder der Objekte und Ereignisse in der Welt zu sein. Diesen Umstand zeigen bereits die theoretischen Probleme der kausalen Tiefe, der kausalen Breite oder der Fehlrepräsentation, die mit der Repräsentationstheorie verbunden sind (vgl. Kap. 4.3.3).

Eine syntaktische Struktur des Gehirns scheint zwar grundsätzlich notwendig, um diese verschiedenen Verknüpfungen für eine Repräsentation zu gewährleisten. Eine solche Struktur kann für die vorliegenden neurowissenschaftlichen Versuche insofern beschrieben werden, als die Unterscheidung zwischen positiven und negativen Bildern für die emotionale Dimension der Valenz mit getrennter neuronaler Aktivität im linken DLPFC sowie im rechten DLPFC einhergeht. Ebenso konnten negative Korrelationen für die Dimension der Wiedererkennung im PACC bzw. eine konverse Korrelation für die Valenz beobachtet werden – die Unterscheidung zwischen kognitiven und emotionalen Aufgaben spiegelt sich auf der neuronalen Ebene in der jeweiligen Aktivität im VMPFC oder im DLPFC (vgl. Kap. 5.2.3). Grundsätzlich verweist auch die Differenzierung zwischen emotionalem Aufmerksamkeitsnetzwerk und emotionalem Wahrnehmungsnetzwerk auf verschiedene Strukturen (vgl. Kap. 5.1.3). Auch die Interaktion zwischen

emotionaler Beurteilung und Aufmerksamkeit ist als syntaktisches Element interpretierbar – die Interaktion wird durch neuronale Aktivität im DMPFC moduliert (vgl. Kap. 5.2.3). Doch es reicht nicht aus, diese strukturelle Aufteilung im Gehirn als Beleg für eine syntaktische Struktur einer mentalen Sprache zu bewerten. Die Annahme einer solchen Sprache geht über die Beobachtung spezifischer funktionaler Eigenschaften neuronaler Netze und Regionen hinaus. Wenn man die Idee einer mentalen Sprache, die in der syntaktischen wie semantischen Struktur einer Alltagssprache gleicht, wörtlich nimmt, müsste eine solche Struktur tatsächlich beobachtbar sein. Bei komplexeren Inhalten der IAPS-Bilder müsste entsprechend eine komplexere neuronale Verknüpfung zu beobachten sein, möglicherweise in Form einer Gedankensprache oder aber als „Maps".[155] Tatsächlich muss aber auf computationale Modelle ausgewichen werden, weil eine Gedankensprache im wörtlichen Sinne nicht beobachtbar ist. Ein wichtiger Grund hierfür mag sein, dass im Gehirn wesentlich Prozesse ablaufen, während ein formulierter Satz in seiner syntaktischen Struktur eine feste, starre Einheit bildet.

Hinsichtlich der semantischen Eigenschaften von Inhalt scheint eine syntaktische bzw. strukturelle Verknüpfung allein jedoch nicht hinreichend zu sein. Hier stellt sich die Frage, inwiefern kausale Prozesse für den (semantischen) Inhalt verantwortlich sein können. Dies ist die Zielsetzung der LOT-Hypothese; die kausale und syntaktische Kohärenz soll die semantische, normative Kohärenz sichern (vgl. Kap. 4.3.1). Es kann argumentiert werden, dass die Formen und Farben der Bilder neuronale Aktivität verursachen und parallel die verschiedenen Stimuli neuronal abgeglichen und verknüpft werden. Allerdings ist hiermit das Phänomen des mentalen Inhalts nicht ausreichend erklärt. Grundsätzlich steht die Annahme einer syntaktischen und semantischen Kohärenz auch im Widerspruch zur These des anomalen Monismus, der eine solche Kohärenz prinzipiell ablehnen muss (vgl. Kap. 2.3). Der semantische Gehalt und das aktive Interpretieren der Probanden positiver oder negativer emotionaler Bilder als solche bedürfen ggf. neuronaler Fähigkeiten, die über syntaktische Verknüpfungen hinausgehen. Wenn beispielsweise eine „Gewaltdarstellung" als unmoralisch bewertet wird, ist anzunehmen, dass verschiedene externe wie interne Verknüpfungen erfolgen. Der Lokalisationstheorie steht dabei inzwischen die Netzwerktheorie gegenüber, wobei sich beide Ansätze ergänzen. Man geht inzwischen nicht

[155] Rein empirisch ist es derzeit und womöglich auch auf längere Sicht nicht möglich, Anhaltspunkte zu finden, die Fodors „Language of Thought"-Position im Gegensatz zur Map-Theorie vollständig bestätigen, vice versa. Es ist zumindest vorstellbar, dass es eine Gedankensprache „Mentalese" gibt, die aufgrund ihrer Struktur, Systematik und Produktivität den neuronalen Hintergrund für die gesprochene Sprache und deren Inhalte darstellt (vgl. Fodor 1987; Kap. 4.3.1). Andererseits gibt es auch gute Gründe, die für die Map-Theorie sprechen, wie die Tatsache, dass sie auch subpersonalen inhaltlichen Repräsentationen, die nicht bewusst zugänglich sind, wie z. B. grobe visuelle Eindrücke, fassen (vgl. Marr 1982). Grundsätzlich müssen sich die beiden Theorien nicht ausschließen, verschiedene neuronale Funktionsweisen könnten sich möglicherweise komplementär ergänzen. Die Diskussion ist in diesem Punkt allerdings nicht allein empirisch zu lösen. Keine Partei nimmt an, dass tatsächliche Sätze oder Bilder im Gehirn existieren, es wird vielmehr auf ähnliche Strukturen hingewiesen, die Sätze oder Bilder repräsentieren. Diese vermuteten Ähnlichkeiten machen eine Entscheidung in dieser Debatte jedoch schwierig, da keine Kriterien bekannt sind, die die Repräsentation in der einen wie in der anderen Form rechtfertigen. Insofern ist auch eine empirische „Überprüfung" derzeit nicht möglich.

mehr davon aus, dass es einen „lokalen Gedächtnisort" im Gehirn gibt, in dem Erinnerung wie in einer Schublade hinterlegt ist. Vielmehr werden bestehende neuronale Verbindungen in aktuellen Situationen erneut aktiviert, weshalb Erinnerung stets ein gegenwärtiger Prozess ist (vgl. Roth 1996, 209f.; Papineau 2004, 106). Dennoch bedarf das Wissen um die einzelnen Inhalte auch eines physischen bzw. biologischen Speichers. Festzustellen, wie eine solche Sicherung einzelner Inhalte im Gehirn möglich ist, also inwiefern Neuronen oder neuronale Netze Träger von Informationen sein können, ist letztlich auch Aufgabe molekularer Neurowissenschaft. Sie müsste aufzeigen, auf welche Weise welche molekularen Strukturen es ermöglichen, Informationen zu speichern, wieder abzurufen, mit neuen Informationen zu vergleichen, etc.[156]

Biologistische Theorien (vgl. z. B. Millikan 1984, 2000; Dretske 1988, 1995; Searle 2001, Hornsby 1986, 1997; Cummins 2000) lassen hierfür prinzipiell den Spielraum, da sie sich im weiteren Sinne auf die Funktion des Organs „Gehirn" beziehen, die den Inhalt einer Repräsentation festlegt. So werden aufgrund biologischer Selektionsprozesse biologisch wichtige Verhaltensweisen und die damit verbundenen Inhalte über Generationen übertragen. Wie „die Biologie" dies leistet, können diese theoretischen Ansätze jedoch auch nicht darstellen. Es wird letztlich auch nur auf Korrelationen (vgl. Dretske 1988), biologische Normen und Isomorphiebeziehungen (Millikan 2000; Cummins 2000) oder nicht näher beschreibbare intrinsische biologische Eigenschaften (Searle 2001) verwiesen (vgl. Kap. 4.5.1). Dabei werden mögliche theoretische Voraussetzungen von Repräsentation und mentalem Inhalt aufbereitet, aber auch die theoretischen Vorbehalte bleiben bestehen. Das eigentliche biologische Phänomen, wie Inhalt im Gehirn gespeichert, produziert und subjektiv erlebt wird, können diese Ansätze jedoch nicht erklären.

Biologistische Positionen können zumindest den Umstand akzentuieren, dass über syntaktische Strukturen hinaus die tatsächliche physische Materie für die Funktion der Repräsentation relevant ist. In der Konsequenz muss eine plausible Theorie der Repräsentation diesem Umstand konzeptionell gerecht werden. Dabei kann auch gefragt werden, inwiefern modale Gedankenexperimente, in denen Robotern oder Marsmenschen mentale Inhalte zugesprochen werden, überhaupt sinnvoll sind. Wenn Inhalt nicht nur von der Struktur eines Systems abhängig ist, sondern auch von dessen stofflicher Zusammensetzung, kann bezweifelt werden, dass Roboter mit Silizium-Schaltkreisen oder Marsmenschen mit grünem Schleim tatsächlich sinnvolle Denkmöglichkeiten darstellen. Bis zum Gegenbeweis kann vertreten werden, dass Lebewesen oder Systeme ohne die entsprechende biologische Molekülstruktur nicht zu (selbständigen) Gedanken fähig

[156] Auch auf die Probleme der kausalen Tiefe, der kausalen Breite oder der Fehlrepräsentation ließen sich bei anhaltendem Fortschritt der molekularen Neurowissenschaft empirische Antworten erwarten, zumal die Kausalrelation auf diese Probleme allein keine Antwort bietet (vgl. Kap. 4.3.3). So wäre zu erwarten, dass sowohl Ursachen als auch Wirkungen neuronaler Repräsentation genau zu bestimmen wären. Hinsichtlich der Fehlrepräsentation könnte gezeigt werden, aufgrund welcher Parameter neuronale „Fehlschaltungen" erfolgen.

sind, die sie zudem subjektiv erleben.[157] Natürlich sind auch hier Abstufungen denkbar, dass beispielsweise klug konstruierte Roboter aufgrund ihrer Ausstattung mit Siliziumchips eher für die Informationsverarbeitung geeignet sind als wie auch immer gearteter Schleim. So ist auch objektiv beobachtbar, dass sich einige Metalle als Leiter von Elektrizität besser eignen als andere. Ebenso gibt es entscheidende Qualitätsunterschiede in der Speicherung von elektrischer Energie zwischen herkömmlichen Batterien oder Akkumulatoren.

### 6.3.2 Empirische Validität und enger Inhalt

Neben dem *Wie?* der mentalen Repräsentation ist auch das *Was?* fragwürdig geworden. Dies ergibt sich aus dem Umstand, dass in der dualistischen Unterscheidung zwischen „mentaler Außenwelt" und „physischer Gehirnwelt" – diese Unterscheidung wird in der Frage des Inhalts letztlich von physikalistischen Positionen getroffen – der externe, abstrakte, mentale Inhalt irgendwie mit den internen, konkreten, physischen Neuronen verknüpft werden muss. Die Unterscheidung zwischen internalen und externalen Eigenschaften ist gewissermaßen ein Spezialfall des „explanatory gap" (vgl. Kap. 3.2.4). Verschiedene Gedankenexperimente legen nahe, dass der Inhalt von Gedanken auf sowohl physischen wie sozialen externalen Faktoren beruht (vgl. Putnam 1975c; Burge 1989; Kap. 4.3.2).[158] Aufgrund dieser Anforderungen steht Inhalt in einem Spannungsfeld, da er einerseits Außermentales repräsentiert, andererseits kausal wirksam sein soll. Ein Lösungsvorschlag aus diesem Dilemma ist, zwei Arten von Inhalt, weiten Inhalt und engen Inhalt, anzunehmen (vgl. Kap. 4.3.2).

Dies verweist zunächst auf ein grundsätzliches epistemisches wie auch empirisches Problem bei der „Beobachtung" von Inhalt, was sich für die Naturalisierung von Inhalt ergibt – die Frage der Perspektive. Die Probanden haben einen privilegierten Zugang zu ihrem subjektiven Erleben, da nur sie tatsächlich ihre Gedanken erleben und darüber berichten können. Rein empirisch scheint derzeit insofern die Position der Internalisten gestärkt, als ausschließlich die Probanden selbst wissen, in welchem mentalen Zustand sie sich befinden, bzw. welchen Gedanken sie gerade haben (vgl. ebd.). Basierend auf den BOLD-Messungen im fMRT kann immerhin gezeigt

[157] Diese Aussage ist allein deshalb problematisch, da wir ebenso wenig Wissen über die Erlebnisperspektive von Fledermäusen wie von Computern haben. Selbst wenn ein Roboter äußert: „Ich habe Gefühle und eine reiche Gedankenwelt" kann dies kategorisch und dogmatisch bestritten werden mit dem Hinweis, dass man nicht wisse, wie es ist, ein Roboter zu sein, und dementsprechend keinen Vergleich zwischen Robotergedanken und Menschengedanken möglich ist. Andererseits ist auch denkbar, dass Roboter einen weiteren oder anderen logischen Rahmen haben als das menschliche Gehirn.

[158] Insofern die Inhalte mentaler Eigenschaften Zustände oder Ereignisse in der Welt betreffen, ihre kausale Kraft jedoch lediglich auf ihre internen Charakteristika zurückzuführen sein sollen, berührt das inhaltliche Kriterium offensichtlich das Kriterium der extrinsischen Relation.

werden, welche Hirnregionen korrelieren und hohe Aktivität verzeichnen; ein konkreter Gedanke ist auf diesem Weg jedoch nicht feststellbar.[159]

Allerdings legen die bereits genannten Korrelationen, z. B. zwischen den emotionalen Dimensionen Valenz, Intensität, Wiedererkennen und paralleler neuronaler Aktivität in VPMFC, DMPFC oder PACC nahe, dass eine inhaltliche Beziehung zwischen den Bildern und der neuronalen Aktivität besteht. Es konnte gezeigt werden, dass, je positiver im behaviouralen Test die emotionale Dimension Valenz bewertet wurde, desto geringer der T-Wert im bilateralen DLPFC war. (Hier scheint es einen Aktivitätswechsel zwischen linkem DLPFC bei positiv bewerteten Bildern und rechtem DLPFC sowie VLPFC bei negativ bewerteten Bildern zu geben.) Zudem konnten verschiedene parametrische Abhängigkeiten, z. B. der neuronalen Akivität im DLPFC von der emotionalen Valenz bei nicht-erwarteter Bildbeurteilung, beobachtet werden. Kann dies als Bestätigung für Externalität interpretiert werden? Immerhin verweisen diese Korrelationen auf den Zusammenhang zwischen neuronaler Aktivität und Außenwelt. So zeigt das Histogramm des zeitlichen Verlaufs im Rahmen der Aufmerksamkeitsstudie, dass die externalen Stimuli und die internale neuronale Aktivität zeitlich aufeinander folgen – die SAC-Aktivierung wird von der emotionalen Aufmerksamkeit induziert, bevor das nachfolgende Bild erscheint, welches wiederum die Aktivität des perzeptiven Netzwerkes induziert (vgl. Kap. 5.1.2). Die externen Stimuli scheinen somit Auslöser für vielfältige neuronale Aktivität bzw. zahlreiche internale Prozesse zu sein.

Kann nun, als Reaktion auf den externalen Inhalt, auf der psychologischen und auf der neuronalen Ebene die Extension eines „engen" bzw. „weiten" Inhalts bestimmt werden? Hinsichtlich der Studien erscheint es auf beiden Ebenen problematisch, beispielsweise den engen Inhalt des Bildsehens oder der emotionalen Aufmerksamkeit von einem weiten Inhalt zu unterscheiden. Die subjektive Wahrnehmung scheint, wie oben bereits geschildert, vorrangig einer außermentalen, normativen Realität zu unterliegen. Insbesondere das Verstehen und Bewerten der Bilder als positiv oder negativ ist ohne eine Sprachgemeinschaft nicht zu leisten (vgl. Burge 1986, 1989). Aber auch der psychologische Aufbau einer emotionalen Aufmerksamkeit verweist auf die direkte Umgebung, die schließlich die entsprechenden Stimuli mit Bedeutung bzw. Inhalt unmittelbar vor der Versuchsdurchführung versehen hat. Allen Probanden wurde der Versuch im Vorfeld erklärt. Dieser „weite" Inhalt ist Voraussetzung für die Durchführung des Versuchs. Warum ein „enger" Inhalt unterschieden werden soll, erscheint empirisch nicht plausibel. Schließlich würde sich die theoretische Frage ergeben, ob bei den unterschiedlichen Probanden im „engen" Inhalt

[159] Genau dies wäre die erwartete empirische Beweisführung, dass ein neuronaler Zustand mit dem semantischen Inhalt eines Gedankens identisch ist: allein auf der Basis der neuronalen Aktivität aus der Dritte-Person-Perspektive den subjektiven Inhalt „zu übersetzen".

Abweichungen möglich wären, die unter Umständen zu fehlerhaftem Verhalten geführt hätten. Im subjektiven Erleben kann zumindest kein „enger" von einem „weiten" Inhalt getrennt werden.

Das Modell des „engen" Inhalts ist als Versuch zu verstehen, Inhalt, der eben aufgrund seiner propositionalen Eigenschaften (psychologischen) Wahrheitsbedingungen unterliegt, zu naturalisieren. Aus diesem Grunde wurde ein Modell entwickelt, wonach weite Inhaltseigenschaften von engen Inhaltseigenschaften „encodiert" werden (vgl. Jackson 1998, 99; Fodor 1987). Letztere sind intrinsisch und somit kausal wirksam. Es stellt sich nun die empirische Frage, wie diese „Encodierung" des weiten durch den engen Inhalt neuronal funktionieren soll. Kann „weiter" Inhalt, wenn er keine intrinsischen Eigenschaften hat, überhaupt neuronal repräsentiert werden? Wo und wie soll die Encodierung des Inhalts stattfinden? Rein physiologisch ist auf der empirischen Ebene eine zeitliche und möglicherweise kausale Folge z. B. von den Augen, dem Sehnerv über das Sehzentrum hin zu weiteren Hirnregionen bei einer visuellen Wahrnehmung beobachtbar (vgl. Roth 1996, 98, Singer 2002c, 147f.). Stimuli werden dabei grundsätzlich neuronal „prozessiert" und „moduliert". Man kann somit auch von einer „Encodierung" etwaiger Stimuli sprechen, aber nicht zwangsläufig von einer Encodierung eines „weiten" in einen „engen" Inhalt.

„Weiter" Inhalt, der encodiert würde, dürfte überhaupt nicht neuronal oder auch in der subjektiven Wahrnehmung in Erscheinung treten bzw. bewusst werden. Die Wahrnehmung müsste sich von vornherein auf einen „engen" Inhalt beschränken, wenn „weiter" Inhalt nicht intrinsisch ist bzw. nicht neuronal prozessiert wird. „Weiten" Inhalt dürfte es unter solchen Umständen gar nicht geben. Am Beispiel der vorliegenden Paradigmen kann nicht aufgezeigt werden, inwiefern hier ein „weiter" Inhalt von einem „engen" Inhalt unterschieden werden soll. Allenfalls ist ein Verlauf im perzeptiven Netzwerk erkennbar und die Beteiligung unterschiedlicher Regionen z. B. von occipitotemporalen visuellen Regionen über das Mittelhirn hin zum medialen und lateralen präfrontalen Cortex. Grundsätzlich müssten hierbei für einen „engen" Inhalt Kernregionen definiert werden. Die verwendeten IAPS-Bilder liefern einfach einen Inhalt bzw. eine Inhaltsmenge. Eine Unterscheidung zwischen weit und eng scheint dagegen selbst innerhalb dieser Inhaltsmenge auf einer abstrakten Ebene zu bleiben (vgl. Kap. 4.6) und verfehlt somit das Ziel, die Kluft zwischen internalen und externalen Eigenschaften zu überwinden. Denn auch wenn unsere Wahrnehmung aus den unzähligen Eindrücken, die sich uns in der Umwelt aufdrägen, nur einen Bruchteil tatsächlich zu Bewusstsein bringt, sind diese wahrgenommenen Inhalte noch immer abstrakt und damit auch „weit". Ketzerisch kann gefragt werden, wie „eng" denn so ein Inhalt sein darf, damit wir ihn neuronal verarbeiten können – das Modell läuft letztlich auf einen Regress hinaus.

Dagegen scheint eine Differenzierung zwischen äußerem Inhalt im Sinne der Konvention einer Sprachgemeinschaft hinsichtlich ihrer Begriffe, und einem inneren Inhalt, der die denotative

Bedeutung mit einer subjektiven konotativen Bedeutung anreichert, möglich. Während der äußere Inhalt, die denotative Bedeutung als lexikalisch-definitorischer Begriff das intersubjektive Verstehen, z. B. des Wortes „Auto", garantiert, umfasst der innere Inhalt, die konotative Bedeutung, individuelle gedankliche Assoziationen, wie z. B. „mein erstes Auto war ein blauer Golf" oder „ein Autohersteller ist mein Arbeitgeber" oder „jeden Morgen stehe ich im Stau". Ohne den äußeren Inhalt wäre, wie gesagt, keine Interaktion zwischen Mensch und Umwelt möglich. Als weiter Inhalt muss er deshalb kausal relevant und wirksam sein. Aber auch über diese begriffliche Paarung kann nicht dargestellt werden, *wie* Inhalt repräsentiert werden kann, sie verbleibt auf einer abstrakten Ebene. Allerdings kann zumindest über das *Was* festgehalten werden: Gegenstand ist der gesamte (abstrakte) Inhalt, die Unterscheidung zwischen „engem" und „weitem" Inhalt ist nicht zielführend. Dieser umfasst sowohl rein innere Prozesse im Gehirn als auch Ereignisse außerhalb des Gehirns, die in der Kombination aktuelle inhaltliche Zustände bewirken.

### 6.3.3 Naturalisierung ohne Vereinfachung

Wie kann das Projekt der Naturalisierung hinsichtlich des *Was?* der Repräsentation in Angriff genommen werden? Hierfür scheint einerseits weitere Begriffsarbeit auf der subjektiv mentalen Ebene notwendig, um genau zu bestimmen, was mit Inhalt gemeint ist und wie Inhalt weiter differenziert werden kann: wie kann Inhalt denotativ und konotativ, intern und extern dargestellt werden? Gibt es neuronale Entsprechungen für die abstrakten Kategorisierungen? Dabei muss für die Bestimmung von Inhalt nicht zwangsläufig zwischen Innen und Außen getrennt werden, Neuronen müssen nicht gegen Abstraktes ausgespielt werden. Es empfiehlt sich, bei dem Begriff zunächst theoretische Vorarbeit zu leisten, ehe man ihn der Neurowissenschaft überantwortet.

In den vorliegenden Studien können verschiedene Inhaltstypen bestimmt werden. Zum einen gibt es a) den „klassischen Inhalt" der IAPS-Bilder, der eine emotionale oder neutrale Situation vermittelt, wobei innerhalb dieser Bilder weiter unterschieden werden kann zwischen dem abstrakten Inhalt, den ein solches Bild als „Szene" darstellt und einzelnen konkreten Inhalten, solange Gegenstände oder Personen auf den Bildern identifiziert werden. Zum anderen gibt es b) den erlernten Inhalt der Bildsignale „Neutraler Pfeil", „Aufmerksamkeitspfeil" und „Fixierungskreuz". Hierbei handelt es sich um bereits bekannte geometrische Figuren, die in Verbindung mit der Aufgabenstellung des Versuchs für die Probanden einen unmittelbaren, neu erlernten Inhalt erhalten. (Die Aufgabenstellung an die Probanden ist ein Inhalt, der den gesamten Versuch umfasst.) Schließlich müssen neben diesen versuchspezifischen c) weitere Inhalte angenommen werden, die der Gedankenwelt der Probanden zuzuordnen sind. Diese können sich mehr oder

weniger direkt auf den Versuch und die gezeigten Bilder beziehen und mit (persönlichen) Erinnerungen verbunden sein. [160] Es können also drei Inhaltstypen unterschieden werden:

a) aktueller Inhalt (abstrakt, konkret)
b) erlernter Inhalt (langfristig, kurzfristig)
c) assoziativer Inhalt (individuell)

Die drei Inhaltstypen schließen sich nicht grundsätzlich aus. Beispielsweise kann aktueller Inhalt durchaus erlernt sein, vice versa. Zudem zeigt sich von a) bis c) eine Abstufung des Inhalts hinsichtlich externaler und internaler Relation. Bei a) aktuellem Inhalt scheint es zunächst plausibel zu sein, dass sich der mentale Inhalt bzw. das subjektive Erleben des Probanden nach den extramentalen Bildern richtet. Allerdings handelt es sich dabei weniger um begriffliche Inhalte, vermutlich werden die gezeigten Situationen auch aufgrund der zeitlichen Begrenzung eher intuitiv hauptsächlich visuell in einem vorsprachlichen Raum erfasst. Dies scheint zusätzlich dafür zu sprechen, dass der mentale Inhalt dieser Eindrücke durch die extramentalen Bilder geprägt und hauptsächlich external ist. Doch auch hierbei wird der Proband bei der Wahrnehmung der Bilder auf bekannte Größen zurückgreifen und mit diesen vergleichen. Es handelt sich dabei um b) bereits erlernte Inhalte. Innerhalb der Versuche wird dabei die mit Bildsignalen verknüpfte Aufgabenstellung umgesetzt. Welchen Status diese mentalen Inhalte einnehmen, ist tatsächlich strittig. Während sie erlernt werden, müssen sie im Grunde wesentlich external sein. Lernen wird neuronal als eine „Prägung" des Gehirns durch äußere Ereignisse aufgefasst. Zu einem späteren Zeitpunkt, z. B. bei der Durchführung des Versuchs, kann diesen erlernten Inhalten aber auch ein internaler Status zufallen, da sie gegebenenfalls nun zum festen Repertoire des Probanden bzw. seines Gehirns gehören. Diesem etablierten Repertoire ist schließlich c) der individuelle assoziative Inhalt zuzuordnen. Es kann sich dabei im extremen Fall um eine sehr starke, „festgefahrene" Meinung o. ä. handeln, die auch durch äußere Einflüsse nur schwer bzw. gar nicht zu beeinflussen ist.

Für diese drei subjektiv-phänomenalen Typen müsste nun eine neuronale Entsprechung ermittelt werden. Tatsächlich konnte innerhalb der Studien während unterschiedlicher Phasen auch neuronale Aktivität unterschiedlicher Areale beobachtet werden (vgl. 5.1.2.; 5.2.2.). Allerdings entziehen sich solche assoziativen Gedanken bzw. Inhalte wie in c) der Kontrolle der Versuchsleitung und sind nicht über Tomografen o. ä. aus der Dritte-Person-Perspektive einsehbar, sondern nur den Probanden selbst über die Erste-Person-Perspektive verfügbar. Letztlich

[160] Eine Probandin äußerte, dass sie sich während der Versuchsdurchführung unweigerlich an den Roman „A Clockwork Orange" erinnert fühlte, da sie einige emotional sehr aggressive Bilder, wie z. B. eine Operation oder verwundete Menschen, ansehen musste bzw. sollte. In dem Roman wird der Protagonist Alex u. a. mit Kinobildern konditioniert, damit aus ihm ein friedvoller Mensch wird (vgl. Burgess 1975).

können alle drei Typen in unterschiedlicher Gewichtung internal wie external interpretiert werden. Grundsätzlich scheint sich auf der subjektiv mentalen Ebene die Gedankenwelt potenziell kontinuierlich zu verändern. Äußere Eindrücke werden mit gemachten Erfahrungen abgeglichen und erweitern den Erfahrungsschatz, das Meinungsbild einer Person. Aufgrund neurowissenschaftlicher Studien setzt sich die Erkenntnis durch, dass dem Lernprozess biologische Grenzen gesetzt sind. Mit der Ausbildung des Gehirns im jungen Erwachsenenalter sind die wesentlichen neuronalen Bahnen stabilisiert. Unter dem Stichwort „Neurogenese" wird aber erforscht, wie Lernen im fortgeschrittenen Alter unter neurobiologischen Aspekten möglich ist (vgl. Emsley/Mitchell/Kempermann/Macklis 2005; Kempermann 2006).

Es scheint bei der Prozessierung von Inhalt bzw. der Inhaltstypen auf das Zusammenspiel sämtlicher Hirnregionen anzukommen. Ein mentaler Zustand, wie der Gedanke „Das Bild zeigt einen Autounfall", müsste unter holistischen Gesichtspunkten dem Zustand des gesamten Gehirns zu einem bestimmten Zeitpunkt t entsprechen, wenn nicht sogar dem Zustand des gesamten Nervensystems oder des gesamten biologischen Systems. Hier stellt sich wieder die Frage der internalen Grenzziehung. Ein solcher mentaler Gesamtzustand umfasst auch Regionen, die nicht im Kern für den etwaigen Gedanken eine Rolle spielen. Von einer 1:1 Entsprechung der Bilder im Gehirn, zumal in verschiedenen Gehirnen der Patienten, sozusagen als Abdruck, kann nicht die Rede sein. Über die bloßen physischen Stimuli (Bildwellen) hinaus wird das Bild (aktiv) interpretiert – dies legt eher eine „wechselseitige Kausalität" nahe.

Allein hohe Aktivität in einer bestimmten Region scheint nicht hinreichend, um den Status mentalen Inhalts zu beschreiben. Der Inhalt komplexer Repräsentationen muss anders bestimmt werden, als der einfacher Begriffsrepräsentationen (vgl. Kap. 6.3.1). Dabei sollte die epistemische Differenzierung zwischen konkreten Inhalten, die direkt durch das Nervensystem induziert werden, und so genannten abstrakten, normativen Inhalten, für die andererseits auch internale Aktivität, wie z. B. Erinnern, notwendig ist, sich auch auf der neuronalen Ebene widerspiegeln. Es wird deutlich, dass ein einzelner Gedanke, der sich auf ein (externales) äußeres Ereignis bezieht, ein umfassendes Netz an (internaler) neuronaler Aktivität mit sich bringt. Die Vernetzung verschiedener Neuronenensemble für einen Gedanken setzt wiederum voraus, dass ein solches Netz für diesen Gedanken entwickelt wurde. Es müsste hinsichtlich des Modells der reziproken Modulation ein Modell für die Repräsentation von (mentalen) Inhalten entwickelt werden, dass a) nicht nur eine einseitige monokausale Repräsentation zwischen (empfangendem) Gehirn und (sendenden) Umweltstimuli berücksichtigt und b) die wechselseitige Modulation als internalen Prozess berücksichtigt.

Nicht alle neuronalen Prozesse, die für den Gedanken relevant sind, werden dem Subjekt selbst bewusst sein. Psychologische Erklärungen für mentale Repräsentation implizieren nicht, dass ihre Besitzer sich über diese (aktuell) bewusst sein müssen. Repräsentationen sind üblicherweise bewusst und unbewusst. Dies wirft die empirische Frage auf, welchen Zustand diese unbewussten Mentalese-Sätze oder gedanklichen Maps in Abgrenzung zu bewussten neuronalen Zuständen haben. Daneben können zusätzlich neuronale Prozesse unterschieden werden, die prinzipiell nicht bewusst werden (vgl. Evans 1982). An dieser Stelle soll lediglich darauf hingewiesen werden, dass der internale Inhalt auf unbewusste Dispositionen zurückzuführen ist, die aufgrund externaler Stimuli oder auch internaler Prozesse ausgelöst und bewusst werden können. Der empirische wie der ontologische Status solcher Dispositionen ist nicht eindeutig geklärt (vgl. Searle 2001; Kap. 4.5.2). Die Erinnerung an verschiedene bekannte Unfallbilder könnte unterhalb signifikanter neuronaler Aktivität und unbewusst verlaufen. Es werden somit bestehende Inhalte mit dem neuen externalen Inhalt abgeglichen. Das Gehirn befindet sich in einem Zustand der Überprüfung, die so auch dem Probanden nicht bewusst sein muss, was also seine Autorität hinsichtlich seiner Hirnvorgänge einschränkt – es sind eben nur die bewussten Vorgänge, zu denen er einen privilegierten Zugang genießt.

Das Problem bei der empirischen Interpretation für den Inhalt ist wiederum, dass nicht eindeutig bestimmt werden kann, was was bedingt. Natürlich lösen die Bilder verschiedene Reaktionen im Gehirn aus. Allerdings muss im Gehirn bereits eine (neurologische) Disposition bestehen, die Bilder zu erkennen und auf ihren Gehalt bzw. Inhalt hin einzuordnen. Hier muss also eine Bereitschaft vorhanden sein, das Bild im Rahmen seiner IAPS-Bewertung anzuerkennen. Denkbar wäre ein Kaspar Hauser, bei dem dies so nicht funktioniert. Eine solche Person, die nicht die Chance hatte, entsprechende neuronale Verknüpfungen zu entwickeln, dürfte nicht ad hoc in der Lage sein, komplexe Ereignisse oder Situationen zu verstehen oder entsprechende Ideen zu entwickeln.

Während konkrete Gegenstände (Äpfel, Züge, etc.) für diesen „tabula rasa"-Kasper Hauser vermutlich durch Zeigen relativ schnell erlernbar wären, wäre das Erklären komplexer, abstrakter Sachverhalte wesentlich schwieriger – wie auch das Verstehen. Eine solche Person wäre im Grunde nicht in der Lage, eine Liebesszene in einem Film als solche zu deuten, da die neuronale Vernetzung hierfür fehlt. Dabei geht man inzwischen davon aus, dass im präfrontalen Cortex bezüglich höherer kognitiver Funktionen eine Staffelung erfolgt. Einfachere Funktionen sind weiter hinten anzusiedeln; je komplexer die kognitive Funktion, desto weiter vorne kann sie möglicherweise im präfrontalen Cortex lokalisiert werden (vgl. Roth 2003, 480ff.). Bei einem Begriff wie „Liebe" ist dieser mit weiteren Regionen, wie z. B. emotionalen Regionen wie der Amygdala vernetzt. Ein „tabula rasa"-Kasper Hauser könnte eine solche Vernetzung nicht entwickeln. Aber auch bei „normalen" Personen scheint es einen Qualitätssprung zu geben zwischen der beob-

achtbaren neuronalen Vernetzung und dem subjektiven Erleben. Wie ist es möglich, subjektiv gedanklich einer konservativen oder einer liberalen Ideologie zu folgen, wenn neuronal lediglich eine Vernetzung und wechselnde Aktivitätspotenziale zu verzeichnen sind? Die derzeitigen philosophischen Modelle laufen im Grunde auf ein Modell hinaus, wonach eine lineare oder auch assoziative (subjektive) Gedankenfolge auf neuronaler Ebene eine sehr schnelle Abfolge neuronaler Vernetzungen mit den relevanten Arealen darstellt – gewissermaßen ein „legato-prestissimo" neuronaler Vernetzung. Unter dem Oberbegriff „Liebe" müssten in schneller Folge weitere Netzwerke assoziativ aktiv werden, die den Begriff individuell für die jeweilige Person ausfüllen.

***Quintessenz Vergleich Inhalt***

- *Plausibilität von Kausaltheorie (LOTH) für einfache Begriffe gegeben*
- *Problem: weniger Plausibilität von Kausaltheorie (LOTH) bei komplexen Begriffen*
- *Kombinatorische Syntax auf Kombinatorik neuronaler Prototypen übertragbar*
- *kein neuronal-struktureller Beleg für „Mentalese" möglich: Unterschied zwischen Prozess-Charakter des Gehirns und fester Struktur von Sätzen*

- *Bestimmung von Wahrnehmung als aktives Interpretieren*
- *Relevanz der Materie: Denkprozesse auf Basis des konkreten biologischen Speichers*
- *Korrelation neuronale Aktivität und Außenwelt betrifft subjektiven Inhalt*
- *Unterscheidung „weiter" Inhalt und „enger" Inhalt nicht zielführend; lediglich Enkodierung von Stimuli*
- *Mögliche Unterscheidung: „innerer Inhalt" (denotativ) und „äußerer Inhalt" (konotativ)*
- *Bestimmung dreier, extramentaler Inhaltstypen: aktuell, assoziativ, individuell*
- *aktives Interpretieren als wechselseitige Kausalität*
- *Frage der Lokalisation von Inhalt: Gehirnholismus beteiligt sämtliche Regionen*
- *Differenzierung: abstrakte versus konkrete Inhalte*
- *Modell der mentalen Kausalität nicht einseitig monokausal sondern auf Basis wechselseitiger Modulation (internaler Prozesse)*
- *Definition: Dispositionen als Begründung für internalen Inhalt, die aufgrund extramentaler Stimuli bzw. internaler Prozesse ausgelöst und bewusst werden können*
- *Zweifache Anforderung an Inhalt als Widerspruch zu Voraussetzungen für mentale Kausalität: Extramentales repräsentieren bei kausaler Wirksamkeit*

Die Frage nach dem *Wie?* der Repräsentation, der Konflikt zwischen external und internal, zwischen weitem Inhalt und engem Inhalt bleibt vorläufig bestehen. Neuronale Netze sind als Basis mentaler Repräsentation weitgehend akzeptiert, dennoch bleibt es zunächst ein (biologisches)

Rätsel, wie diese repräsentierten Inhalte auch (subjektiv) gedacht und erfahren werden können. Die Frage der Semantik ist noch nicht geklärt. Ferner sollen mentale Zustände, da sie extrinsisch sind, oder weiter Inhalt soll, da er external ist, keine kausale Wirkung haben. Andererseits muss vorher gefragt werden, wie nicht-intrinsische Zustände bzw. nicht-internaler Inhalt überhaupt repräsentiert werden kann. Wenn dies nicht möglich ist, erscheint es unplausibel, dass allein intrinsisch-kausale Zustände in Interaktion mit der Umwelt treten können. Dies scheint ein subjektiv erlebbares, wie auch neuronal beobachtbares Merkmal der Aktivität im Gehirn zu sein. Die zweifache Anforderung an Inhalt, dass er Extramentales repräsentiert, aber zugleich kausal wirksam sein soll, dieses Spannungsfeld ist unter den üblichen Voraussetzungen nicht aufzulösen.

### 6.4 Emotionale Aufmerksamkeit und Relationalität

Mentale Eigenschaften scheinen extrinsisch zu sein, was bedeutet, dass sie relational sind und sich auf etwas beziehen, das über ihren konkreten Objektbereich hinausgeht. Sie beziehen sich nicht bzw. nicht nur in Form, Gehalt oder innerer Struktur auf das Ereignis oder den Zustand, der diese Eigenschaft besitzt. Eine Beschreibung einer extrinsischen Eigenschaft bezieht sich entsprechend auf einen größeren Bereich, der auch – aber eben nicht nur – das zugehörige Ereignis bzw. den zugehörigen Zustand umfasst (vgl. Kap. 4.4). Für mentale Kausalität ist dabei ein Kernpro-blem, dass extrinsischen Eigenschaften grundsätzlich kausale Wirkung abgesprochen wird (vgl. Kap. 4.4.2).

Es fällt schwer, die Begriffe intrinsisch und extrinsisch neurowissenschaftlich zu fassen. Eine intrinsische Eigenschaft einfach als innere Eigenschaft zu beschreiben, scheint unpassend. Ferner ist unklar, was zu einer rein intrinsischen Eigenschaft zählen soll. Wären beispielsweise Blutzucker oder Hormone noch Bestandteil einer inneren Eigenschaft, da sie sich innerhalb des menschlichen Körpers befinden? Zudem kann gegen scheinbar intrinsische Eigenschaften eingewendet werden, dass sie möglicherweise durch zeitlich vorherige äußere Ereignisse individuiert werden. Doch ebenso schwierig ist es, extrinsische Eigenschaften neurowissenschaftlich zu bestimmen, da empirisch vor allem die inneren Verbindungen zwischen den neuronalen Arealen bestimmt werden. Allerdings scheint der Bezug zur Außenwelt bzw. äußeren Reizen für das Gehirn plausibel.

Für die Diskussion ist es insofern wichtig zu klären, inwiefern mentale Zustände – als psychologische und als neuronale Zustände – unter neurowissenschaftlichen Gesichtspunkten als extrinsisch oder intrinsisch gelten müssen. Die Begriffe „intrinsisch“ und „extrinsisch“ sind unscharf und können nicht eindeutig dem Bereich des subjektiven Erlebens oder dem neuronalen Bereich zugeordnet werden. Dabei wird deutlich, dass die klassische Dichotomie zwischen intrin-

sischen und extrinsischen Eigenschaften für den mentalen bzw. neuronalen Bereich nicht aufrechterhalten werden kann.

### 6.4.1 Extrinsische Eigenschaften der subjektiven Ebene

Hinsichtlich der empirischen Implikationen der neurowissenschaftlichen Studien scheint es auf der psychologischen Ebene selbstverständlich zu sein, dass die subjektiven Eindrücke der Probanden sich auf etwas außerhalb ihres Körpers beziehen und somit extrinsisch, relational sind. Schließlich beziehen sich die subjektiven Zustände der Probanden während der Aufmerksamkeitsphase oder auch während der Phase des Bildsehens auf äußere Ereignisse bzw. Stimuli. Sie sind relational zu anderen äußeren Ereignissen. Ihre Gedanken können als intentional eingestuft werden (vgl. Kap. 4.5; Kap. 6.5).

In diesem Zusammenhang kann auf die vergleichende Diskussion um die Regularität und mentale Kausalität (vgl. 6.1.4) verwiesen werden. Dabei wurde deutlich, dass auf empirischer Ebene nicht lediglich eine Relation zwischen zwei Komponenten zu betrachten ist, sondern eine Relation zwischen drei Komponenten – den Bildsignalen, der subjektiven Perspektive sowie den fMRT-Daten. In der Konsequenz kann nun gesagt werden, dass sich psychologische Zustände bzw. das subjektive Erleben auf die externen Bilder bezieht: die Probanden erleben abhängig von dem gezeigten positiven oder negativen Bild eine andere Emotion. Das Erleben bzw. Miterleben jedes einzelnen Bildes ist erkennbar auf das Bild selbst zurückzuführen. Also ist das Erleben der mentalen Zustände metaphysisch abhängig von äußeren Ereignissen und epistemisch nicht durch Beobachtung bestimmbar. Aus der Perspektive der Probanden können die Eigenschaften der psychologischen Zustände somit als extrinsische Eigenschaften angesehen werden (vgl. Kap. 4.1).

In Korrelation zu den gezeigten Bildern können auf der psychologischen Ebene bei den Probanden subjektive Bewertungen sowie die emotionalen Dimensionen gemessen werden, auf die die ebenfalls subjektive Beurteilung der Bilder durch die Probanden folgt. Umfasst dieser subjektive Zustand lediglich extrinsische oder auch intrinsische Eigenschaften? Die Anforderung für intrinsische Eigenschaften lautet, dass diese sich ausschließlich auf ihr Bezugsobjekt beziehen. Da sich die psychologischen Zustände der Probanden auf äußere Stimuli beziehen, könnte nun angenommen werden, dass sie zwangsläufig extrinsisch sein müssen. Die Bildwahrnehmung geht damit über so genannte innere Eigenschaften hinaus. Die entsprechende Inhaltseigenschaft kann also weder intrinsisch besessen werden, noch ist sie eine intrinsische Eigenschaft. Allerdings kann auch argumentiert werden, dass das Bezugsobjekt psychologischer Zustände bzw. des subjektiven Erlebens der äußere Stimulus bzw. das Bild ist. Ein subjektives Erleben des Bildes „Autounfall“

ist ohne das parallel gezeigte entsprechende Bild nicht denkbar. Insofern verhält sich das subjektive Erleben z. B. eines negativen Bildes intrinsisch zu dem Bild selbst – das Bild kann letztlich als das notwendige Bezugsobjekt betrachtet werden.[161]

Dafür, dass subjektive Zustände auch intrinsische Eigenschaften haben können, spricht, dass die emotionalen Dimensionen Valenz, Intensität und Wiedererkennen stark individuell geprägt sind. Die Valenz bezieht sich auf die subjektive Erfahrung einer affektiven Qualität des emotionalen Stimulus unabhängig von seiner sensorischen Modalität (vgl. Kap. 5.2.3). Ebenso bezieht sich Intensität darauf, wie stark eine Emotion erfahren wird, und Wiedererkennen umfasst den Vergleich mit vergangenen Emotionen. Die emotionalen Dimensionen sind somit zu einem gewichtigen Teil bereits in der Person der Probanden angelegt und werden nicht allein von außen bestimmt. In diesem Sinne können sie nicht metaphysisch intrinsisch sein, aber zumindest epistemisch (vgl. Kap. 4.3.1).

Dies macht deutlich, wie schwierig es ist, die Relation zwischen den Bildern und dem subjektiven Erleben als entweder ausschließlich intrinsische oder ausschließlich extrinsische Relationen zu bestimmen. Dabei werden allerdings auch definitorische Schwierigkeiten deutlich. Zum einen werden die physische und die mentale Ebene vermischt. Schließlich handelt es sich bei den Bildern möglicherweise – beispielsweise als Lichtwellen – um rein physische Stimuli, während das subjektive Erleben üblicherweise dem psychologischen bzw. mentalen Bereich zugeordnet wird. Zum anderen bedeutet relational, dass sich z. B. der Gedanke „positives Bild" nicht nur außerhalb des Kopfes, sondern in einem komplexen (psychologischen) Beziehungsgefüge befindet – wo aber soll sich das befinden, wenn nicht in dem Körper bzw. in der Person der Probanden?

### 6.4.2 Intrinsische Eigenschaften der neuronalen Ebene

Man könnte nun erwarten, dass die Verhältnisse unter Einbeziehung der neuronalen Ebene klarer werden. Schließlich wird gemeinhin angenommen, dass die kausale Kraft eines Objekts ausschließlich von den intrinsischen, nicht-relationalen Eigenschaften abhängt.[162] Üblicherweise wird argumentiert, dass mentale Eigenschaften, da sie extrinsisch sind, kausal wirkungslos bleiben. Im Umkehrschluss sind es allein die intrinsischen physischen Eigenschaften, die kausale Wirkung

[161] So handelt es sich bei den neutralen wie emotionalen Bildern oder auch bei den Aufmerksamkeitsstimuli per se um externe Signale, die streng genommen selbst die Eigenschaft haben, ein positives, negatives oder neutrales Bild bzw. ein Stimulus zu sein. Insofern sind die psychologischen Eigenschaften bereits (intrinsische) Eigenschaften der IAPS-Bilder – diese wurden gerade aufgrund ihrer Eigenschaften ausgewählt. Die Eigenschaft, ein positives, negatives oder neutrales Bild zu sein bzw. als solches wahrgenommen zu werden, scheint somit eine messbare Eigenschaft jener Bilder zu sein und kann somit als deren intrinsische Eigenschaft angesehen werden.

[162] Diese Intuition ergibt sich aus Stichs „principle of autonomy" – wonach Zustände und Prozesse von psychologischem Interesse auf aktuellen, internalen physischen Zuständen des Organismus supervenieren – und aus Fodors methodologischem Solipsismus (vgl. Fodor 1974; Stich 1983).

erzielen. Zudem übernehmen die physischen Eigenschaften die kausale Funktion vollständig, weshalb nicht alle Eigenschaften eines Objekts für kausale Wirkung notwendig sind (vgl. Kap. 3.4). In der Argumentation scheint es häufig so zu sein, dass Neuronen mit intrinsischen Eigenschaften identifiziert werden, falls die intrinsischen physischen Eigenschaften überhaupt näher bestimmt und nicht lediglich abstrakt diskutiert werden. Als nicht-relationale physische Eigenschaften sichern die Neuronen die kausale Wirksamkeit von Gedanken bzw. von neuronaler Aktivität überhaupt. Die Neuronen haben die physische Eigenschaft aufgrund ihrer selbst. Ihre Eigenschaften sind anscheinend durch Beobachtung bestimmbar. Damit erfüllen sie das metaphysische wie das epistemische Kriterium für intrinsische Eigenschaften (vgl. Kap. 4.3.1). Allerdings muss jemand, der in diese Richtung argumentiert, nicht annehmen, dass nur intrinsische Eigenschaften kausal wirksam sind. Intrinsische Eigenschaften sind keine essenziellen, sondern lediglich nicht-relationale Eigenschaften (vgl. Heil 1994, 159).

Daran anknüpfend muss untersucht werden, inwiefern Neuronen als essentiell intrinsisch gelten können. Die Anforderung an intrinsische Eigenschaften ist, dass keine weitere Beziehung als zu dem Bezugsobjekt vorliegen darf. Anders formuliert: „A thing has its intrinsic properties in virtue of the way that thing itself, and nothing else, is." (Lewis 1983b, 197; Kap. 4.3.1) Wie etwas ist, hängt vollständig von seinen intrinsischen Eigenschaften ab. Die Frage ist nun, ob Neuronen diese starken Anforderungen erfüllen können.

Die Relation zwischen drei Komponenten, auf die hinsichtlich der intrinsischen Relation auf der Ebene des subjektiven Erlebens bereits hingewiesen wurde, betrifft auch die neuronale Ebene. Danach beziehen sich auch die jeweiligen neuronalen Zustände, die durch das fMRT gemessen werden, auf die externen Stimuli und müssen folglich als relational betrachtet werden. Ihre neuronale Organisation ist abhängig von äußeren Ereignissen und damit metaphysisch extrinsisch (vgl. Kap. 6.3.1). Im Rahmen der vorliegenden Studien können hierfür zahlreiche Beispiele angegeben werden. So beziehen sich sowohl das antizipatorische als auch das perzeptive Netzwerk in der jeweiligen Phase auf die äußeren Bildstimuli. So wird signifikante Aktivität für die SAC-Region in Zusammenhang mit dem gezeigten Aufmerksamkeitsstimulus gemessen, entsprechende Aktivität des perzeptiven Netzwerkes im Zusammenhang mit den emotionalen und neutralen Bildstimuli. Des Weiteren kann gezeigt werden, dass der Signalanstieg im SAC und im PACC in der Kombination Aufmerksamkeitssignal und emotionales Bild intensiver ist, als in der Kombination neutrale Aufmerksamkeit und emotionales Bild (vgl. Kap. 5.1.3). Unter der letzten Kondition wird im Unterschied zur ersteren hauptsächlich Aktivität im VMPFC und DMPFC gemessen (vgl. Kap. 5.2.3). Ebenso kann, sogar noch feiner, Aktivität bei positiven Bildern im DLPFC und Aktivität bei negativen Bildern in der rechten Amygdala gezeigt werden, sowie Aktivität des linken DLPFC bei positiv beurteilten Bildern und Aktivität des rechten DLPFC und des VLPFC bei negativ beurteilten Bildern (vgl. Kap. 5.2.1). Dies sind nur einige auffällige Beispiele,

die darstellen, dass es offensichtlich einen relationalen Zusammenhang zwischen der internen neuronalen Prozessierung und den externen Stimuli gibt. Sie machen deutlich, dass sich die neuronale Aktivität auf externe Ereignisse bezieht. Damit ist es nicht einsichtig, dass Neuronen ausschließlich intrinsisch und nicht auch extrinsisch sein können.

Die hier dargestellten Beispiele machen deutlich, dass auf der empirischen Ebene eine einfache Gleichsetzung von „neuronal" und „intrinsisch" nicht zutreffend ist. Dagegen scheinen die Neuronen auch extrinsische Eigenschaften zu besitzen, wodurch die jeweilige Aktivierung bei positiven oder negativen Bildern möglich und plausibel wird. Als übliche Charakterisierung intrinsischer Eigenschaften werden Form, Größe, Masse oder interne Struktur genannt (vgl. Lewis 1983b, 197; Kap. 4.3.1). Natürlich lassen sich Neuronen und neuronale Netzwerke auf diese Weise bestimmen. Neuronen haben eine spezifische Form und auch ein Gewicht. Diese intrinsischen Eigenschaften scheinen sich tatsächlich nur auf die Neuronen selbst zu beziehen. Wie aber kann nun erklärt werden, dass einige Neuronen bei relationalen Ereignissen, wie positiven oder negativen Bildern, aktiviert werden? „The intrinsic properties of something depend only on that thing; whereas the extrinsic properties of something may depend, wholly or partly, on something else." (Lewis 1983b, 197) Die genannten und oben aufgeführten relationalen Ereignisse beziehen sich eben nicht nur auf die Neuronen selbst, sondern gehen darüber hinaus. Insofern ist es nicht plausibel anzunehmen, Neuronen wären bzw. hätten ausschließlich intrinsische Eigenschaften. Das Ereignis „neuronale Aktivität" muss selbst auch extrinsische Eigenschaften haben, damit sinnvoll eine physische Verbindung zu äußeren, relationalen Ereignissen bestehen kann. Ein Modell, wonach Neuronen als rein intrinsische Eigenschaften in einem Kausalgefüge durch äußere Stimuli verursacht werden und daraufhin weitere Wirkung erzielen, wird dagegen der Anforderung nicht gerecht, dass Neuronen bzw. neuronale Netzwerke selbst äußere Ereignisse repräsentieren und modulieren. Dies scheint eine essentielle Eigenschaft der neuronalen Netzwerke zu sein (vgl. Singer 2002d, 119).[163]

### 6.4.3 Binnendifferenzierung neuronaler Eigenschaften

Vor diesem Hintergrund ist es geboten, zwischen neuronalen Ereignissen und ihren Eigenschaften zu unterscheiden. Diese stehen, wie auch die psychologischen Ereignisse bzw. das subjektive Erleben, in Bezug zu Bildstimuli als externe Ereignisse. Dabei scheint es sinnvoll zu sein, auf der empirischen Ebene eine Binnendifferenzierung zwischen einzelnen Neuronen, einzelnen Regio-

[163] Alternativ wäre es auch denkbar zu sagen, das Bezugsobjekt der Neuronen ist das Bild, um somit einen exklusiven intrinsischen Bezugsrahmen zu bestimmen. Das würde aber die Begrifflichkeit „intrinsisch" und „extrinsisch" völlig aufheben.

nen und neuronalen Netzwerken vorzunehmen. So sind es insbesondere die neuronalen Netzwerke, im vorliegenden Falle das perzeptive, wie das antizipatorische, die in deutlicher Relation zu den externen Bildstimuli stehen. Aber auch einzelnen Regionen, wie z. B. der CMA, dem POS oder dem PACC, werden spezifische Funktionen zugeschrieben, die allerdings variieren, abhängig von dem jeweiligen Netzwerk und den weiteren Regionen, die unter einer bestimmten Kondition aktiviert werden. Beispielsweise gilt das POS als anteriorer Teil des dorsal visuellen Pfades, der von frühen visuellen Regionen hin zum posterioren parietalen Cortex sendet. Dies könnte noch als rein intrinsische Eigenschaft des POS gelten, die sich auf die interne Struktur bezieht. Wie die CMA und die SMA wird dieser dorsale Pfad mit der Prozessierung von Handlung in Verbindung gebracht (vgl. Kap. 5.1.3). Er vermittelt die benötigte sensorimotorische Transformation für visuell-orientierte Handlung. Hier stellt sich nun die Frage, ob diese Prozessierung gegebenenfalls relationaler Stimuli noch als rein intrinsische Eigenschaft zu bewerten ist. So war das POS in der vorliegenden Studie während der Aufmerksamkeitsphase aktiviert, wobei keine visuell orientierte Handlung involviert war. Üblicherweise ist dies allerdings der Fall. Hieraus folgt die Annahme, dass die Aufmerksamkeit motivational relevanter Stimuli einen Zustand der Handlungsbereitschaft implizieren kann – auch in Abwesenheit handlungsorientierter visueller Stimuli. Die Prozessierung von Handlung ist nun eine andere Funktion als die Prozessierung von Handlungsbereitschaft. Abhängig vom äußeren Kontext oder eingebunden in unterschiedliche Netzwerke scheint somit eine funktionale Variation einzelner neuronaler Regionen möglich zu sein.

Denkbar wäre eine Abstufung, wonach einzelnen Neuronen lediglich intrinsische Eigenschaften zukommen, während Neuronenensemble oder -netzwerke sich durch intrinsische wie extrinsische Eigenschaften auszeichnen. So wird für bestimmte Regionen oder Netzwerke eine funktionale Spezifizierung bestimmt, z. B. als „multiintegrativ“ (vgl. Kap. 5.1.3), wodurch diese Netzwerke einen relationalen Charakter erhalten. Welche funktionale Spezifizierung ein neuronales Netzwerk beschreibt, richtet sich nach Prozessen in der äußeren Umgebung und nicht allein danach, in welcher Hirnregion es sich befindet (vgl. Lewis 1986, 267). Es kann argumentiert werden, dass eben die funktionale bzw. kausale Rolle eine Hirnregion oder ein neuronales Netzwerk erst als dieses bzw. jenes auszeichnet. Im Gegensatz zur „kausalen Rolle einer Badewanne“ (vgl. Searle 1993, 10f.; Kap. 4.5.2) handelt es sich nicht lediglich um eine gesellschaftliche Konvention, sondern um eine echte biologische Funktion.

Hinsichtlich der Überlegung zu maximalen Eigenschaften könnte dabei die maximale neuronale Aktivität, die über den T-Wert ermittelt wird, die minimale neuronale Aktivität, die sich in der Baseline widerspiegelt, umfassen (vgl. Kap. 4.3.2). Das Gehirn (bzw. Neuronenensemble) verfügt über die maximale Eigenschaft, sämtliche mentalen Inhalte (einschließlich motorischer Eigenschaften o. ä.) zu prozessieren – als potenzielle Eigenschaft. Es können eben nicht nur vorhandene, sondern auch neue Inhalte prozessiert werden. Hiervon muss wiederum eine aktuelle

Funktion unterschieden werden, die sich aus einem konkreten Stimulus heraus entwickelt. Neuronen sind nun insofern grenzsensitiv, als es gerade die herausragende Funktion des Nervensystems ist, das Gehirn bzw. den Menschen mit der Umwelt und seiner Umgebung in „Kontakt" zu bringen. Problematisch ist dabei die Grenzziehung, auf welcher Ebene neuronalen Regionen oder Netzwerken auch extrinsische Eigenschaften zukommen. Möglicherweise ist ein Qualitätssprung auf Basis der Vernetzung der Regionen festzumachen, womit offensichtlich zugleich eine „höherstufige" nach außen gerichtete Funktion verbunden ist. Dabei müsste aber noch weiter bestimmt werden, ob das externe Bezugsobjekt sich jenseits des Körpers befindet oder auch ein Organ markieren kann, damit im mentalen Kontext von extrinsischen Eigenschaften gesprochen werden kann. Dass neuronale Netzwerke hierbei über extrinsische Eigenschaften verfügen, scheint aber notwendige Voraussetzung dafür zu sein, dass überhaupt eine Relation zu Ereignissen außerhalb des Gehirns möglich ist.

### 6.4.4 Wechselseitige Bedingung intrinsischer und extrinsischer Eigenschaften

Dem Umstand, dass das Gehirn ein biologisches Organ mit spezifischen Funktionen und eben nicht eine leblose Badewanne oder ein Stein ist, sollte auch hinsichtlich der Diskussion um intrinsische und extrinsische Eigenschaften Rechnung getragen werden. Das Gehirn ist ein essentiell reizverarbeitendes Organ; eine rein intrinsische Eigenschaftsbeschreibung wird diesem Umstand nicht gerecht. Indem es physische Stimuli bzw. psychologische Inhalte prozessiert, stellt es relationale Bezüge her. Diese wesentliche Funktion des Gehirns ist seine extrinsische Eigenschaft. Searle versucht, diese biologische Eigenschaft den physikalischen Voraussetzungen anzupassen, indem er vertritt, dass vermeintlich extrinsische Eigenschaften bzw. Zustände im Grunde intrinsisch sind. Deshalb setzt er den Begriff der „intrinsischen Intentionalität" als beobachterunabhängiges Merkmal des Geistes ein (vgl. Kap. 4.5). Tatsächlich ist es jedoch so, dass sowohl die subjektiven Erlebnisse als auch die neuronale Aktivität in Relation zu äußeren Ereignissen steht und ihre Eigenschaften über sie selbst hinaus gehen und beschrieben werden können (vgl. Kap. 5.1.3; Kap. 5.2.3). Insofern erscheint es nicht sinnvoll, extrinsische Eigenschaften einfach qua Definition zu intrinsischen Eigenschaften umzuwandeln. Vielmehr erscheint es angemessen zu akzeptieren, dass das Gehirn nicht ausschließlich als intrinsisches Organ beschrieben werden kann. Es ist vielmehr intrinsisch-relational.

Es ist problematisch, die kausale Kraft des Gehirns allein extrinsisch zu begründen. Das Gefühl der Angst, das Phänomen der Phobie sind Resultate externer und interner Stimuli. Bei der Phobie scheinen interessanterweise die externen Stimuli für die meisten Menschen nicht schlimm zu sein, die Phobie wird stark von internen Stimuli wie Gedanken oder düsteren Ideen bzw. „dys-

funktionalen Kognitionen“ bestimmt. Hieraus sollte jedoch nicht voreilig geschlossen werden, nur intrinsische Stimuli seien von Bedeutung. Immerhin müssen die Prozesse, die zur Phobie führen, vorher entwickelt worden sein – plausiblerweise durch externe Ereignisse. In verschiedenen neurowissenschaftlichen Studien wurde gezeigt, dass es für das Erkennen bzw. die Einordnung der Form, Größe oder Lage von Objekten notwendig ist, sie in Relation zum eigenen Körper oder zu anderen im Raum befindlichen Objekten zu setzen (vgl. Jacob 2005, 278).

Zu beachten ist dabei, dass die vorgenommene Messung wie auch die Messung der vorliegenden Studien den Querschnitt der Probandengehirne darstellt. Für diese Querschnittsmessung ist die Außenwelt möglicherweise gar nicht relevant, und es werden rein intrinsische Eigenschaften im Ergebnis dargestellt. Bei den einzelnen Probanden gibt es trotz der gleichen externen Stimuli Abweichungen z. B. in der neuronalen Aktivität. Dies muss nicht verwundern, da es sich nicht um „intrinsische Doppelgänger“ handelt, wie in den verschiedenen Gedankenexperimenten zur Externalität (vgl. Kap. 4.3.2). Auf der empirischen Ebene und unter Ausschluss möglicher Welten kann etwas anderes diskutiert werden. An sich müssten einzelne Neuronen in ihrem Bauplan „intrinsische Doppelgänger“ sein. In ihrer Anordnung und Vernetzung im Gehirn unterscheiden sie sich aber voneinander – einmal intern, innerhalb des Gehirns, da einige Neuronen Bestandteil verschiedener Regionen sind, und dann extern, im Vergleich zu anderen Gehirnen. Der Grund dafür ist sicherlich in der genetischen Veranlagung zu suchen (vgl. Singer 2002d, 113f.). Ein weiterer Grund ist aber der Einfluss der äußeren Umgebung, nach der sich das Gehirn während der Entwicklung in seiner Struktur ausrichtet. Man kann somit sagen, dass einige intrinsische Eigenschaften des Gehirns das Ergebnis der extrinsischen, relationalen Eigenschaft sind. So ist es sinnvoll anzunehmen, dass neuronale Netzwerke intrinsische wie extrinsische Eigenschaften haben. Die intrinsischen Eigenschaften ermöglichen ihnen die Teilhabe an physikalisch-kausalem Geschehen. Ohne die extrinsischen Eigenschaften wäre dieses potenzielle kausale Vermögen jedoch gewissermaßen blind – es gäbe weder Anlass noch Richtung oder Notwendigkeit zur Umsetzung bzw. zu kausaler Aktivität selbst. Zwar wird bezweifelt, dass die kausale Kraft einer Eigenschaft auf einer Relation beruhen kann, die über sie hinausgeht:

> According to externalism, contentful states are identified by reference to entities that lie outside the subject's body […] but the causal powers of a state must be intrinsically grounded; they cannot depend essentially upon relations to what lies quite elsewhere. (McGinn 1989, 133)

Dieser Aussage kann man, wie oben gezeigt, grundsätzlich widersprechen. Kausale Kraft muss nicht, auch wenn einiges dafür spricht, (allein) intrinsisch sein. Außerdem ist das Gehirn und sind die Prozesse im Gehirn wichtiger Teil der extrinsischen Relation; allein deshalb ist es nicht zutreffend, wie McGinn zu behaupten, die kausalen Kräfte eines extrinsischen Zustands

bzw. Ereignisses würden von Relationen, die irgendwo anders liegen, abhängen. Unter der Idee der Realisierungsrelation lässt sich schließlich die globale wie auch die lokale Definition intrinsischer Eigenschaften dahingehend interpretieren, dass ein mentales Ereignis M intrinsisch über eine intrinsisch-extrinsische Eigenschaft verfügt, da diese eine Realisierungsrelation zwischen mentaler und physischer Eigenschaft beschreibt. Bezogen auf die empirische Situation ist denkbar, dass Neuronen mentale Ereignisse prozessieren, da sie über mentale und physische Eigenschaften verfügen (vgl. Kap. 4.3.2).

***Quintessenz Vergleich Relationalität***

*subjektive Ebene nicht ausschließlich extrinsisch:*

- *abhängig von äußeren Ereignissen (metaphysisch)*
- *nicht durch Beobachtung bestimmbar (epistemisch)*
- *intrinsisch (z. B. emotionale Dimensionen)*

*neuronale Ebene nicht ausschließlich intrinsisch:*

- *Relation zwischen drei Komponenten*
- *externe Stimuli für Neuronen*
- *Gleichsetzung neuronal mit intrinsisch unzutreffend*
- *Ereignis „neuronale Aktivität" muss extrinsische Eigenschaft haben*

*Binnendifferenzierung erforderlich (Abstufung):*

- *zwischen neuronalen Ereignissen und ihren Eigenschaften*
- *zwischen Neuronen, Regionen, Netzwerken*
- *Netzwerke und Regionen mit extrinsischer Ausrichtung*

*intrinsische und extrinsische Eigenschaften bedingen sich wechselseitig:*

- *Reizverarbeitung als wesentliche, extrinsische Funktion des Gehirns*
- *intrinsische Eigenschaften als Ergebnis extrinsischer Eigenschaften*
- *intrinsische Eigenschaften ermöglichen Kausalität; extrinsische Eigenschaften geben die Zielrichtung an*
- *Gehirn mit neurobiologischen Eigenschaften als intrinsisch-relationale Funktionseinheit*

Man kann an dieser Stelle zusätzlich betonen, dass die Neuronen auch hinsichtlich ihrer physischen Eigenschaften und aufgrund ihrer Grenzsensitivität durchaus abhängig von äußeren, einwirkenden Kräften sind (vgl. Ellis 2001; Kap. 4.3.2). Wie gezeigt wurde, ist eine trennscharfe

Grenzziehung bei biologischen Prozessen bzw. Ereignissen aufgrund ihrer Funktion, die extrinsische Eigenschaften impliziert, in diesem Zusammenhang nicht zielführend. Die klassische Dichotomie zwischen extrinsischen und intrinsischen Eigenschaften kann zumindest hinsichtlich mentaler bzw. neuronaler Prozesse nicht aufrechterhalten werden. Dagegen scheint es vielmehr intrinsisch-relationale Eigenschaften zu geben.

Bei Dretske heißt es, dass die Neurowissenschaft das Gehirn als ein körperliches Organ untersucht, als einen Klumpen Materie mit verschiedenen intrinsischen (elektrischen und chemischen) Eigenschaften (und verschiedenen Relationen zu anderen inneren Organen). Die Psychologie untersucht dagegen das Gehirn als einen Konsumenten von Informationen über seine Umgebung, als einen Klumpen Materie mit verschiedenen extrinsischen Relationen zur Umwelt, auf die Handlung ausgerichtet wird. Dretske beschreibt das Gehirn deshalb als ein physisches Objekt mit zwei Eigenschaftsbereichen. Die intrinsischen Eigenschaften definieren es als ein spezielles neurobiologisches Organ, während die extrinsischen Eigenschaften die Relationen des Gehirns zum Rest des Universums umfassen, insofern es diesen Rest der Welt repräsentiert (vgl. Dretske 1994, 133). In Abwandlung zu Dretske lässt sich nun sagen, dass es gerade die neurobiologischen Eigenschaften sind, die dem Gehirn überhaupt intrinsisch-relationale Eigenschaften verleihen.

### 6.5 Emotionale Aufmerksamkeit und Intentionalität

In dem Phänomen der Intentionalität laufen die verschiedenen bereits diskutierten Kriterien extrinsische Relation und Inhalt zusammen. So wird Intentionalität u. a. bezeichnet als das mentale Vermögen zu repräsentieren bzw. die Spannung, mit der sich der menschliche Geist auf Objekte bezieht (vgl. Kap. 4.5). Die Schwierigkeiten und Möglichkeiten, die sich aus der Notwendigkeit neuronaler Repräsentation ergeben, wurden im vorhergehenden Kapitel (6.4.) eingehend diskutiert. Hierauf soll an dieser Stelle nur noch am Rande eingegangen werden.

Der Fokus der folgenden Diskussion soll auf der Behauptung liegen, dass intentionale Verursachung im Grunde eine systematische Verknüpfung des Repräsentationsvermögens mit den mentalen Kausalbeziehungen ist (vgl. Searle 2001, 127; Kap. 4.5.2). Diese systematische Verknüpfung beschreibt dabei ein Spannungsfeld zwischen normativer Repräsentation und physischer Kausalität. Es wird untersucht, inwiefern dieser Ansatz durch die empirischen Implikationen der vorliegenden neurowissenschaftlichen Studien unterstützt wird. Dabei wird deutlich, dass Intentionalität nicht nur ein spezifisches Merkmal subjektiv mentaler Phänomene, sondern auch für den neuronalen Bereich charakteristisch ist. Zentrale Eigenschaft in dem beschriebenen Spannungsfeld ist dabei die neuronale Disposition.

### 6.5.1 Intentionale Verursachung als bidirektionales Kausalitätsmodell

Wenn intentionale Zustände sich dadurch auszeichnen, dass sich ein mentaler Zustand auf ein zugehöriges intentionales Objekte bezieht, scheint Intentionalität grundsätzlich ein Begriff zu sein, der Zustände auf der Ebene des subjektiven Erlebens beschreibt. Grundsätzlich kann für die vorliegenden Studien die „primitive", nicht abgeleitete Version der Intentionalität veranschlagt werden, da hier eine direkte Bildbeobachtung bzw. -beurteilung vorgenommen wird. Die Version der abgeleiteten Intentionalität von Satzäußerungen liegt lediglich – wenn überhaupt – im Rahmen der Bewertung der emotionalen Dimensionen in der Post-hoc-Messung vor (vgl. Kap. 4.5). In der Betrachtung der vorliegenden Studienergebnisse kann hierfür das Erleben der emotionalen Bilder als positiv oder negativ als intentionaler Zustand aufgeführt werden. Die wahrgenommenen Bilder entsprechen dabei als konkrete physische Objekte in Raum und Zeit dem intentionalen Bezugsobjekt. Parallel entsprechen für den Zustand der emotionalen Aufmerksamkeit die Pfeilsignale zur Ankündigung eines emotionalen bzw. neutralen Bildes abstrakten intentionalen Objekten. Folglich beschreibt Intentionalität lediglich auf etwas andere Weise das Phänomen der extrinsischen Relation (vgl. Kap. 4.4). Dabei zielt der Begriff der Intentionalität insbesondere auf die Spannung zwischen Bewusstsein und Objekt.

Um diese Spannung zu naturalisieren und zu zeigen, dass auch physische Systeme intentionale Zustände enthalten können, wurden als teleologische Theorien Dretskes informationstheoretischer und Millikans konsumenten-orientierter Ansatz entwickelt. Die Validität dieser Ansätze muss in Bezug auf die vorliegenden neurowissenschaftlichen Paradigmen differenziert betrachtet werden. Nach der Indikatorrelation müsste für jeden Probanden als System S die Funktion der Indikation durch verschiedene neuronale Zustände $s_1$, $s_2$, … $s_n$ für die korrespondierenden Eigenschaften $f_1$, $f_2$, … $f_n$ eines Stimulus F gegeben sein. Um zu prüfen, inwiefern es sich dabei um eine nomologische Korrelation handelt, müsste derselbe Stimulus in mehrmaliger Wiederholung demselben Probanden gezeigt werden. Durch die Wiederholung ist allerdings eine Veränderung der neuronalen Aktivität durchaus möglich, die Indikatorrelation wäre damit nicht eindeutig und nicht nomologisch – dies würde den von Dretske aufgestellten Anforderungen widersprechen (vgl. Kap. 4.5.1). Ihm zufolge soll schließlich eine Indikatorrelation eine hundertprozentige Wahrscheinlichkeit aufweisen. Im Versuch ist beispielsweise die emotionale Aufmerksamkeit auf die senkrechten Pfeil-Stimuli festgelegt bzw. wird durch diese verursacht. Dennoch ist es möglich, dass das Phänomen der emotionalen Aufmerksamkeit sowohl auf der neuronalen Ebene in der Intensität als auch auf der subjektiven Ebene im Erleben variiert. Daneben wird umgekehrt während der Phase der emotionalen Bildwahrnehmung trotz unterschiedlicher Bildstimuli dasselbe Netzwerk bzw. dieselben Hirnregionen aktiviert. Beide Fälle stimmen insofern nur bedingt mit der Forderung nach Eindeutigkeit bzw. hundertprozentiger Wahrscheinlichkeit überein. Es

wäre zumindest eine weitere Differenzierung oder Eingrenzung der neuronalen Aktivität notwendig, um zu prüfen, welche Unterschiede in der Intensität oder in der Vernetzung zu anderen Hirnregionen zum Zeitpunkt der jeweiligen Bildprozessierung bestehen.

Auch hinsichtlich der Kausalrelation und des Abhängigkeitsverhältnisses gibt es mögliche Abweichungen von den Anforderungen der Indikatorrelation. So funktioniert zwar bei einfachen Kausalrelationen die Beschreibung, wonach G die Wirkung von F ist und entsprechend ein Stimulus neuronale Aktivität verursacht. Bei komplexen Relationen muss diese Beschreibung jedoch erweitert werden, da auch bei den vorliegenden Versuchen nicht nur eine unidirektionale Kausalität vom Stimulus zu den Neuronen berücksichtigt werden muss. Darüber hinaus ist einerseits das Vorwissen über die Versuchsanordnung und den -ablauf für die Probanden sehr kurzfristig erlernt. Sie sind aber nicht durch Konditionierung oder selektive Lernprozesse angeeignet – diese Art der schnellen Informationsaneignung wird durch die Indikatorrelation nicht berücksichtigt. Andererseits besteht ein Vorwissen über die negativen wie positiven Bildinhalte. Dieses Vorwissen wird aber in einem für die Probanden neuen Kontext neu verknüpft. Insofern entspricht es nicht mehr der hundertprozentigen, eindeutigen Wahrscheinlichkeit der Indikatorrelation. Dies wird auch im Abhängigkeitsverhältnis deutlich, das nicht, wie gefordert, asymmetrisch strikt ist, da beispielsweise für die emotionale Aufmerksamkeit wie für die emotionale Bildwahrnehmung eine entsprechende neuronale Aktivität oder subjektives Erleben auch ohne den exakten Stimulus als Abweichung möglich ist. So ist bei der emotionalen Bildwahrnehmung trotz verschiedener Bilder dasselbe Netzwerk aktiv (vgl. Kap. 5.1.2). Wenn S als neuronale Aktivität ein Indikator für F ist, muss überlegt werden, wie weit bzw. variabel F sein kann – beispielsweise als disjunkte Menge verschiedener Bildstimuli. Insofern bestätigt sich das „functional indeterminacy problem" durch die theoretischen Explikationen der neurowissenschaftlichen Studien (vgl. Kap. 4.5.1).

Bei Millikans konsumenten-orientiertem Ansatz sind insbesondere der Nutzungsbegriff sowie die Verknüpfung von intentionaler und evolutionärer Erklärung problematisch. Der Schwerpunkt wird bei dieser Theorie auf die Verarbeitung, die Funktion der Repräsentation gelegt. Damit ist der Nutzungsaspekt im Umweltkontext maßgeblich – der Interpret erhält wertvolle Informationen, die der Überlebenssicherung dienen. Im Rahmen eines naturalisierten Intentionalitätskonzeptes wie z. B. des Bienentanzes lässt sich diese Unterscheidung zwischen Informant und Interpret gut darstellen. Schwierig wird dies allerdings für kognitive Systeme. Hier wird lediglich auf die „innere Repräsentation" verwiesen, wobei Entschlüsselung und Nutzung im kognitiven System mehr oder weniger zusammenfallen müssen.

Wie aber kann dies im Detail beschrieben werden? Grundsätzlich können sämtliche Bildstimuli als produzierender Anteil der Repräsentation und der Mensch bzw. sein Gehirn als konsumierender Anteil der Repräsentation betrachtet werden. Doch Millikan scheint beide Bereiche, den produzierenden wie den konsumierenden Anteil, in das kognitive System bzw. das Gehirn zu

verlagern (vgl. Millikan 1989; Kap. 4.5.1). Das Gehirn konsumiert nur die Informationen, die es zuvor in Orientierung an die Umwelt produziert hat. Hinsichtlich der neuronalen Aktivität bei emotionaler Aufmerksamkeit oder emotionaler Bildwahrnehmung erscheint es jedoch problematisch, diese Unterscheidung nachzuvollziehen, sie ist neuronal nicht sichtbar. Ferner soll aus dem Konsumanteil der Repräsentation eine überlebenssichernde Nutzung folgen. Für die Unterscheidung zwischen positiven und negativen IAPS-Bildern innerhalb der Versuche scheint dieser evolutionäre Nutzungsbegriff theoretisch valide: Sie sichern das Überleben z. B. als Warnhinweis oder als Signal für eine sichere Umgebung o. ä. Andererseits kann für den Versuch bezweifelt werden, dass die Umsetzung der Funktion evolutionär zielführend ist. Schließlich geht es zunächst lediglich „per Knopfdruck" um die Erfüllung einer komplexen Versuchsanordnung.

In diesem Zusammenhang ist auch Millikans Überlegung einer „normalen Kondition" (vgl. ebd.) als grundlegendes Ereignis für den Inhalt einer Repräsentation problematisch. Natürlich ist die Annahme plausibel, dass neuronale Verknüpfungen bei Kindern in der Entwicklung sehr konkret durch Stimuli entwickelt werden. Andererseits scheint diese normale Kondition für die Bildwahrnehmung oder vor allem komplexe Repräsentationen nicht unbedingt bestimmbar zu sein. Zumal die Entstehung der normalen Kondition evolutionär verlaufen soll, dies aber für eher komplexe mentale Phänomene wie die emotionale Aufmerksamkeit oder die emotionale Bildwahrnehmung nur eingeschränkt begründet werden kann. Bei diesen beiden Zuständen kommt der emotionale Anteil sogar einer evolutionären Erklärungsstrategie entgegen, weil argumentiert werden kann, dass emotionale Wahrnehmungen und ihre Verstärkung sehr basal zu schnellen überlebenssichernden Reaktionen verhelfen können. Ist diese Erklärung aber für die heutige, tatsächliche Laborsituation hinreichend? Schließlich ist lediglich eine stärkere Konzentration und gegebenenfalls eine Bewertung der Bilder gefordert. Es ist zweifelhaft, dass anspruchsvolle, komplexe Begriffe und die Intentionalität mentaler Zustände nur über evolutionäre Tauglichkeit zu erklären sind.

Es könnte vermutet werden, dass lediglich aktive Bewusstseinszustände der Probanden wie z. B. der Wunsch, eine Tasse Kaffee zu trinken, zu intentionalen Zuständen zählen, während eher passive Bewusstseinszustände wie die Bildwahrnehmung nicht hinzugezählt werden können. Der Begriff der Intentionalität als Gerichtetheit ist jedoch weiter zu fassen. Searle unterscheidet deshalb auch zwischen der Geist-auf-Welt-Ausrichtung für Überzeugungen, Wahrnehmungen oder Erinnerungen und der Welt-auf-Geist-Ausrichtung für Wünsche und Absichten. Damit sollen systematisch Repräsentation und Kausalität miteinander verknüpft werden (vgl. Searle 2001, 122; vgl. Kap. 4.5.2). Die Phase der Bildwahrnehmung und die Phase der Bildbeurteilung sind auf der Ebene des subjektiven Erlebens als intentionale Gehalte der Geist-auf-Welt-Ausrichtung zuzuordnen, da das Bildsehen einer Wahrnehmung entspricht und die Bildbeurteilung einer Überzeugung. Diese Unterscheidung kann durch die theoretischen Explikationen durch die vorliegenden

Paradigmen bedingt unterstützt werden, da empirisch für die unterschiedlichen Funktionen auch unterschiedliche Hirnregionen zuständig sind. So wird sowohl in der Phase der Bildwahrnehmung als auch der Bildbeurteilung signifikante Aktivität im DLPFC gemessen, wobei während der unerwarteten Bildbeurteilung eine Lateralisierung beobachtet wurde: Der rechte DLPFC korreliert mit negativen Bildern und der linke DLPFC mit positiven (vgl. Kap. 5.2.2; Kap. 5.2.3).

In beiden Fällen richtet sich das subjektive Erleben auf äußere Ereignisse hin aus. Beide Fälle werden andererseits von jenen äußeren Ereignissen überhaupt erst initiiert. Mit dem Begriff der „intentionalen Verursachung" möchte Searle deshalb ein bidirektionales Kausalitätsmodell etablieren, dass den Anforderungen der Repräsentation und der Kausalität gerecht wird. Ursache und Wirkung sollen sich danach wechselseitig ergänzen und sind jeweils die Repräsentation des anderen (vgl. Kap. 4.5.2). Diese Relation wird deshalb als „kausale Selbstbezüglichkeit" beschrieben. Bei kognitiven Zuständen mit kausaler Selbstbezüglichkeit, wie z. B. Wahrnehmungen, ist zwar die Ausrichtung vom Typ Geist-auf-Welt, aber die Kausalität soll umgekehrt von der Welt auf den Geist verlaufen. Ein Beispiel dafür ist die Phase des Bildsehens, in der sich die neuronale Aktivität auf den Bildstimulus hin ausrichtet, wobei dieser den Vorgang eigentlich verursacht. Bei volitiven Zuständen mit kausaler Selbstbezüglichkeit, wie z. B. Absichten, ist dagegen die Ausrichtung vom Typ Welt-auf-Geist, aber die Kausalität soll vom Geist auf die Welt verlaufen. Für die vorliegenden Paradigmen könnte eingeschränkt die Beurteilung der Bilder hierfür als Beispiel gelten, wobei insofern eine Einschränkung getroffen werden muss, als aufgrund des Untersuchungsmodells dem mentalen Zustand ein Bildstimulus vorhergeht. Deshalb wurde die Bildbeurteilung bereits zur Geist-auf-Welt-Ausrichtung gezählt. Dennoch ist hiermit die Absicht verbunden, zwecks Beurteilung einen Knopf zu drücken, weshalb diese Absicht durchaus als volitiver Zustand bezeichnet werden kann.

Ein vielleicht extremes Beispiel für die kausale Selbstbezüglichkeit ist das Phänomen der emotionalen Aufmerksamkeit mit der Ausrichtung Geist-auf-Welt. Im Unterschied zur intentionalen Wahrnehmung bzw. dem reinen Bildsehen wird damit eine Erwartung auf ein kommendes Ereignis verbunden. Insofern wird die Ausrichtung des Geistes auf die Welt aufgrund der speziellen Aufmerksamkeit betont. Der kausale Prozess wird durch den Stimulus „Pfeil-Signal" ausgelöst; die geforderte normative Verknüpfung ergibt sich auf der subjektiven Ebene aus der von dem Probanden erlernten Aufgabenstellung. Entsprechend wird die emotionale Aufmerksamkeit auf der neuronalen Ebene durch das als „antizipatorisch" gekennzeichnete Netzwerk repräsentiert (vgl. Kap. 5.1.3). Diese antizipierende, emotionale Aufmerksamkeit kann als starke Form der Intentionalität angesehen werden. Dafür spricht auch, dass in Verbindung mit der emotionalen Aufmerksamkeit auch die funktionale Rolle des DMPFC hinsichtlich der Kondition „erwartete Bildbeurteilung" spezifiziert werden konnte. Unter dieser Kondition moduliert die Region die emotionale Intensität hinsichtlich der Interaktion von emotionaler Beurteilung und

Aufmerksamkeit (vgl. Kap. 5.2.3). Die Kausalität verläuft tatsächlich bidirektional, da einerseits ein äußeres Signal „aus der Welt" im subjektiven Erleben die gesteigerte Aufmerksamkeit verursacht, andererseits daran anschließend das nachfolgende Bild aufgrund der gesteigerten neuronalen Aktivität im antizipatorischen Netzwerk intensiv wahrgenommen wird.

Problematisch an der Bestimmung der Intentionalität ist wiederum unter neurowissenschaftlichen Gesichtspunkten, dass es sich als ein sehr allgemeines Kennzeichen von Bewusstsein wesentlich auf den Bereich des subjektiven Erlebens bezieht. Dabei scheint die Normierung dieser Ereignisse entscheidend für das Aufkommen und Funktionieren des intentionalen Prozesses zu sein. Hier stellt sich die Frage, wie diese Normierung im Rahmen repräsentationaler Prozesse, die intentionale Prozesse ja umfassen, neuronal geleistet werden kann. Schließlich scheint Intentionalität auch sehr individuell zu sein, und ist neuronal abhängig von jeweiligen emotionalen oder kognitiven oder auch motorischen Prozessen. Intentionalität einfach im präfrontalen Cortex anzusiedeln, wäre sicherlich viel zu allgemein, eine Bestimmung, die auf individuelle neuronale Verknüpfungen hinausläuft, dagegen zu spezifisch und auch zu beliebig. Möglicherweise wäre der dritte Weg, das Phänomen der Intentionalität zunächst für emotionale, kognitive und motorische Prozessierung zu unterscheiden, eine mögliche Lösung, an die eine vertiefende Bestimmung anknüpfen kann. Parallel zur emotionalen Aufmerksamkeit wären somit auch eine kognitive und eine motorische Aufmerksamkeit zu bestimmen.

### 6.5.2 Normativität als Herausforderung des neuronalen Bereichs

Intentionalen Zuständen wird zudem eine „innewohnende Normativität" bescheinigt (vgl. Searle 1993, 69f. ). In der Form propositionaler Ausdrücke unterliegen sie offensichtlich externen Wahrheitsbedingungen. Es kann eben ganz objektiv überprüft werden, ob es der Fall ist, dass New York südlich von Boston liegt oder dass ein Apfel grün ist und nicht blau. Eine Person, die anderes denkt, unterliegt zumindest einem Fehlschluss.[164]

Im Rahmen der vorliegenden Studien wurden mit den IAPS-Bildern tatsächlich Bilder gezeigt, die einem sehr geordneten normativen Rahmen unterliegen. Sie waren bereits vor der Untersuchung als „positiv", „negativ" bzw. „neutral" hinsichtlich des emotionalen Erlebens kate-

[164] Allerdings ist es ein Kennzeichen intentionaler Zustände, dass auch Fiktionales und Irrationales gedacht werden kann, weshalb solche irrationalen mentalen Zustände „weicheren" Wahrheitskriterien unterliegen, als so genannte realistische Gedanken. Das Einhorn ist ein Beispiel für eine solche intentionale Nichtexistenz, auch wenn es immerhin eine literarische Existenz aufweisen kann und dadurch wiederum einer gewissen literarischen Norm unterliegt. Ein vorgestelltes grünes Einhorn mit einem roten Höcker würde dieser literarischen Norm widersprechen, wäre aber zumindest eine sehr private intentionale Vorstellung (vgl. Kap. 4.5). Insbesondere die fiktionalen und irrationalen intentionalen Zustände verweisen darauf, dass diese sich nicht ausschließlich auf eine extramentale Welt beziehen. An die Neurowissenschaft würde sich hieraus die Frage ergeben, ob sich solche fiktionalen Zustände in Abgrenzung zu realistischen Zuständen auch neurologisch bestimmen lassen.

gorisiert. Dieser Bewertung folgte, wie aus dem Post-hoc-Test ersichtlich wurde, auch der wesentliche Teil der Probanden. Die durchschnittlichen Ergebnisse der Valenz-Einschätzung für die eingesetzten negativen, neutralen und positiven Bilder betrugen 1,81 (± 0.54, S.D.), 5,14 (± 0.30 S.D.), und 7,26 (± 0,73 S.D.). Insofern wird deutlich, dass auf der psychologischen Ebene das subjektive Erleben der Probanden gängigen Normen und den damit verbundenen Wahrheitsbedingungen folgt. Ein Unfallbild oder eine Kriegsbild gilt aufgrund gesellschaftlicher Normen als ein negatives Bild, während ein Babyfoto – möglicherweise aufgrund rudimentärer Instinkte, wie dem „Kindchenschema" – als ein positives Bild bewertet wird. Von Interesse sind hier aber auch die Abweichungen, die als Hinweis dafür angesehen werden können, dass intentionale Zustände neben äußeren normativen Bedingungen auch einer inneren Logik des Subjekts folgen. Hat beispielsweise ein Proband im Post-hoc-Test ein nach gesellschaftlichen Normen negatives Bild positiver bewertet als üblich, könnte dies auf eine sehr persönliche Erlebnis- und Ideengeschichte zurückzuführen sein. Ein scheinbar objektiver Fehler unterliegt also einer subjektiven Rationalität.

Auf neuronaler Ebene scheint es eine Entsprechung für diese Normierung der Bilder bzw. das subjektive Erleben der Probanden zu geben. Hierauf verweist zumindest die valenzspezifische Korrelation zwischen dem rechten DLPFC und negativen Bildern bzw. dem linken DLPFC und positiven Bildern (vgl. Kap. 5.2.3). Ferner ist unterschiedliche Aktivität bei den emotionalen im Vergleich zu den neutralen Bildern messbar. Das wird durch das spezifisch emotionale Wahrnehmungsnetzwerk deutlich. Daneben kann bei der Antizipation der neutralen Bildern kaum oder nur geringe Aktivität in der Region SAC beobachtet werden (vgl. Kap. 5.1.2; Kap. 5.1.3). Da der SAC emotional auffallende Ereignisse vermittelt, kann daraus geschlossen werden, dass eine Unterscheidung zwischen emotionalen und nicht-emotionalen Ereignissen bei den Probanden möglich ist – die Kontrollbedingung ist also gut gewählt.

Bedeutender hinsichtlich der Frage der Normierung ist aber die beobachtbare unterschiedliche Aktivität bei positiven Bildern im Vergleich zu negativen Bildern. So wurde eine Korrelation der positiven Bilder mit dem DLPFC unter der erwarteten und auch der nicht-erwarteten Bildkondition gemessen, sowie eine Korrelation der negativen Bilder mit der rechten Amygdala unter der nicht-erwarteten Bildkondition. Ebenso konnte unter der Beurteilungskondition eine signifikant negative Korrelation der positiv bewerteten Bilder mit dem linken DLPFC sowie eine negative Korrelation der negativ bewerteten Bilder mit dem rechten DLPFC und dem VLPFC gemessen werden (vgl. Kap. 5.2.3). Das subjektive Erleben der Bilder als „positiv" oder „negativ" korreliert also mit unterschiedlichen Hirnregionen. Diese Korrelationen können dahingehend interpretiert werden, dass soziale Normen eine neurologische Entsprechung haben.

Es stellt sich die grundsätzliche Frage, ob es für den theoretischen Begriff der Intentionalität und dem damit verbundenen Bezug zum Bereich der Normen eine konkrete neuronale Entsprechung geben kann, oder ob dieser Begriff nicht doch eher die allgemeine extrinsische Relation beschreibt. Auch die emotionalen Dimensionen Valenz, Intensität und Wiedererkennen sind mögliche intentionale Merkmale psychologischer Zustände, schließlich beziehen sie sich sehr direkt auf ein intentionales Objekt. Die Korrelationsanalyse zeigt, dass diese Dimensionen getrennt in verschiedenen Hirnregionen prozessiert werden (VPMFC, VLPFC, DLPFC für Valenz; VLPFC, DMPFC für Intensität; PACC für Wiedererkennen; vgl. Kap. 5.2.2). Möglicherweise sind es solche neuronalen Eigenschaften bzw. Netze, die in Kombination mit anderen Regionen die „intentionale Spannung" zwischen Subjekt und Objekt bewirken. Schließlich weisen die Ergebnisse darauf hin, dass es getrennte Repräsentationen der verschiedenen emotionalen Dimensionen im präfrontalen Cortex gibt, getrennt von kognitiver Prozessierung (vgl. Kap. 5.2.3). Hieraus ergibt sich die Frage, ob ggf. eine emotionale Intentionalität von einer kognitiven Intentionalität sowohl theoretisch als auch empirisch unterschieden werden muss. Das Phänomen der emotionalen Aufmerksamkeit könnte in diesem Zusammenhang ein Beispiel für emotionale Intentionalität sein.

Searle betont, dass die Kausaleigenschaften der Intentionalität[165] das Kausale mit dem Normativen verbinden. Allerdings muss bei einem biologischen Prozess vorsichtig mit dem Begriff „kausal" umgegangen werden. Schließlich ist die zeitliche Abfolge der Aktivierung verschiedener Hirnregionen kein Beleg für eine kausale Relation. Aus neurophysiologischer Sicht werden innerhalb des Gehirns zunächst einmal Signale prozessiert. Die Übermittlung von Signalen wird dabei gemeinhin als kausaler Prozess interpretiert.

Eindeutiger erscheint es dagegen, äußere Stimuli als Ursachen zu bestimmen und die gesamte neuronale Aktivität als Wirkung innerhalb einer kausalen Relation (vgl. Kap. 4.3.2; Kap. 4.4.1). Diese Ansicht kann neurowissenschaftlich unterstützt werden, weil sich auf der physiologischen Ebene ein kausaler Vorgang ereignet, von den Lichtwellen des Bildsignals ausgehend zur Retina des Auges über den Sehnerv und das Sehzentrum hin zu weiteren Hirnarealen, welche das Gesehene repräsentieren. Innerhalb dieser Areale haben sich neuronale Netzwerke ausgebildet, die möglicherweise kausal aktiviert werden (vgl. z. B. Roth 1996, 121ff.; Singer 2002c, 147f.). Auch der gängige Funktionalismus verbindet letztlich Repräsentation und Kausalität, wobei die semantischen Eigenschaften einer mentalen Repräsentation von ihrer funktionalen Rolle bestimmt werden. Die hinreichende Bedingung, eine semantische Eigenschaft zu haben, kann

[165] Searle verwendet den Begriff der „intrinsischen Intentionalität", um Intentionalität als biologisches Phänomen mit kausaler Kraft zu bestimmen. Hier wird dieser begrifflichen Setzung nicht gefolgt, da Intentionalität traditionell mit gutem Grund als extrinsische Relation verstanden wird. Im Unterschied zu Searle wird dagegen angenommen, dass mentale Zustände als extrinsisch-intrinsische Zustände dennoch kausal wirksam sein können (vgl. Kap. 4.5; Kap. 6.4.4).

danach in kausalen Begriffen bestimmt werden (Fodor 1994, 39; Kap. 2.2.1; Kap. 2.2.3). Dem kann zumindest hinsichtlich intentionaler Prozesse widersprochen werden; ohne die Funktion der Repräsentation würde lediglich ein kausaler Prozess ablaufen. Aufgrund dieser Funktion – eben in Hinblick auf die Normativität – muss dieser Prozess jedoch auch eine semantische Funktion umfassen. In der Folge werden „externe Wahrheitsbedingungen" nicht nur repräsentiert, sondern sie können auch Ursache der Repräsentation sein. Die Norm „positiv" bzw. „negativ" verursacht die entsprechende neuronale Aktivität. Die Kausalität darf demnach nicht von der Normativität getrennt werden, ansonsten ist das Phänomen der Intentionalität verloren.

### 6.5.3 Dispositionen als Ausdruck neuronaler Intentionalität

Wenn intentionale Zustände potenziell bewusst werden können, bedeutet das umgekehrt, dass sie auch unbewusst sein können. Sowohl bei der Welt-auf-Geist als auch der Geist-auf-Welt-Ausrichtung werden in wechselseitiger Ergänzung Ursache und Wirkung aufeinander bezogen. Abgesehen von dem Fall des Erlernens neuer externer Inhalte, also einer „Erstverursachung", umfasst die Relation der intentionalen Kausalität bekannte Umweltereignisse und entsprechend vorhandene Neuronenensembles. Diese stehen wiederum in Verbindung mit anderen intentionalen Zuständen (vgl. Kap. 4.5; Kap. 2.2). Verschiedene Überzeugungen sind Bestandteil einer Überzeugung. Wie aber kann diese Verbindung neuronal bestimmt werden, solange die intentionalen Zustände unbewusst sind? Ein intentionaler Zustand müsste streng genommen ständig aktivierbar bzw. vorhanden sein.

In den üblichen neurowissenschaftlichen Versuchsanordnungen werden die bewussten Handlungen der Probanden untersucht. Natürlich müssen ihnen unbewusste Dispositionen zugrunde liegen, woraus das aktuale Handeln und Empfinden resultiert. Allerdings ist es schwierig, innerhalb einer Versuchsanordnung zwischen bewusster und unbewusster Hirnaktivität zu unterscheiden. Ein Beispiel ist die Disposition zur Depression. Diese ist schwer darzustellen, da hiermit komplexe Eigenschaften bzw. Störungen der neuronalen Netzwerke verbunden sind. Man bemüht sich gegenwärtig, die komplexen Störungen der Netzwerke durch Aktivierungsparadigmen sichtbar zu machen. Depressive zeigen im strukturellen Bild keine Unterschiede zu Gesunden. Im resting state (ohne Aufgabe) zeigen sie im fMRT eine geringfügig erhöhte Grundaktivität im VMPFC und eine geringfügig verminderte Grundaktivität im DLPFC. Diese Veränderungen sind jedoch so geringfügig, dass sie nicht bei allen Depressiven zu messen sind. Dies gilt vermutlich erst recht für vermeintlich gesunde Probanden, die daraufhin untersucht werden sollen, ob sie eine Disposition zur Depression haben oder nicht. Durch Aktivierungsparadigmen wie z. B. die Aufmerksamkeitsstudie (vgl. Kap. 5.1) ist es eventuell möglich, den Defekt der De-

pressiven besser zu messen. Denn durch die Aktivierung bestimmter Hirnregionen lassen sich gegebenenfalls funktionelle Defizite darstellen. Auf diese Weise gewänne man ein Verfahren, auch bei vermeintlich gesunden Probanden die mögliche Disposition zur Depression festzustellen.

Abgesehen von Dispositionen zu einer Erkrankung, sind auch Dispositionen zu einem Gedanken, einer Idee oder einer Einstellung schwer zu fassen. In den gängigen neurowissenschaftlichen Versuchsanordnungen können nur aktuelle Gehirnzustände bzw. Einstellungen einer Person erfasst und dargestellt werden. Ein Bild von einem Krankenwagen ruft entsprechende Assoziationen hervor. Dass diese Person zudem Klavier spielt, aber Liszt verabscheut, kann durch das Krankenwagenbild nicht aktiviert oder sogar gemessen werden. Die Überlegung ist jedoch, wo die Disposition „Liszt verabscheuen" in der Zwischenzeit zu finden ist. Welchen ontologischen Status hat sie philosophisch und neurowissenschaftlich, wenn sie bei einer Person nicht aktuell repräsentiert wird?

Man kann erwidern, dass die Disposition „Liszt verabscheuen" nicht durch die Disposition „Krankenwagen-Bild" aktiviert und daher auch nicht als fMRT abgebildet wird. Aber die Disposition verschwindet nicht nach ihrem Abruf. Das ist ja das Wesen der Disposition, dass sie fortlaufend besteht, aber nur gelegentlich sichtbar wird, z. B. in einer depressiven Episode. Neurowissenschaftlich ist sie in der Zwischenzeit in strukturellen und funktionellen Eigenschaften des Gehirns zu finden. Wenn man gerade ein Krankenwagen-Bild zeigte, sieht man sie nicht. Wenn man ein Liszt-Bild zeigte, könnte man sie dagegen sehen. Vorhanden ist sie immer. Die depressive Disposition wird bislang nur sichtbar, wenn ein Mensch immer wieder depressiv erkrankt. Wenn man die richtigen Aufgabenstellungen und Messmethoden entwickelte, könnte man die Disposition aber auch in den vermeintlich gesunden Phasen dieses Menschen darstellen und messen. Denn auch in den scheinbar gesunden Phasen ist das Gehirn mit Depressions-Disposition strukturell und funktionell anders gestrickt als das ohne eine solche Disposition.

Eine Unterscheidung, wonach unbewusste Prozesse unterhalb des Signifikanzwertes bleiben, kann nicht getroffen werden. Dann würde man davon ausgehen, dass unbewusste Prozesse mit schwachen Signalen korrelieren. Das ist jedoch nicht der Fall. Der Unterschied zwischen bewusst und unbewusst scheint im Wesentlichen kein quantitativer zu sein im Sinne von: Unbewusste Angst führt zu geringer Aktivierung in der Amygdala, bewusste Angst führt zu starker Aktivierung in der Amygdala, und bei Überschreiten eines bestimmten Aktivitätsniveaus tritt der Übergang von unbewusst zu bewusst ein. Vielmehr scheinen bewusste Prozesse dadurch charakterisiert zu sein, dass zusätzlich weitere Regionen in das Konzert der Neuronen integriert werden. Unter diesem Aspekt wird auch die emotionale Aufmerksamkeit untersucht, um genau zu bestimmen, welche Regionen an einer verstärkten emotionalen Wahrnehmung beteiligt sind (vgl.

Kap. 5.1.2; Kap. 5.1.3). Unbewusste Dispositionen sind also empirisch schwer zu fassen. Da die neuronalen Netze aufgrund der Plastizität des Gehirns keine feste und eindeutige Dispositionszuschreibung haben, da diese „Dispositionsnetze" nur in aktuellen Situationen aktiviert sind, ist es schwierig, die entsprechenden Netze (die bestehen bleiben) mit ihren Dispositionen (die nach ihrem Abruf verschwinden) zu identifizieren: Dennoch sind sie ständig abrufbereit bzw. aktivierbar – abhängig von inneren wie äußeren Stimuli oder Ereignissen. Das Schema Geist-auf-Welt betrifft nicht nur die psychologische Ebene des subjektiven Erlebens, sondern auch die neuronale Aktivität selbst.

Intentionale Zustände können somit neuronal als Dispositionen beschrieben werden, die aufgrund externer Stimuli oder auch interner Prozesse ausgelöst und bewusst werden können. Empirisch müsste dabei zwischen unbewussten und bewussten Zuständen aufgrund der Aktivierung verschiedener relevanter Regionen zu unterscheiden sein, die einen bewussten intentionalen Zustand ausmachen. Ein intentionaler Zustand wird der Definition nach auf ein intentionales Objekt, ein Ereignis hin ausgerichtet. Empirisch kann dies anhand veränderter Hirnaktivität festgestellt werden, die ein Konzert der Neuronen ergibt. Ontologisch scheint dabei ein vorhandenes internes mentales Ereignis – der unbewusste Zustand – zusammen mit einem extramentalen Ereignis das externale mentale Ereignis, den bewussten intentionalen Zustand, zu verursachen.

Auf der subjektiven Ebene bleibt dennoch das Problem bestehen, dass einem unbewussten Zustand, der phänomenal noch nicht existiert, ein ontologischer Status verliehen werden muss. Hierin zeigt sich sehr deutlich der Unterschied zwischen physischen und mentalen epistemischen Möglichkeiten. Denn physisch ist ein unbewusster intentionaler Zustand aufgrund der strukturellen und funktionellen Zusammensetzung des Gehirns potenziell notwendig, aber nicht hinreichend, denn der „bewusste Gedanke" fehlt. Die physische subveniente Eigenschaft ist vorhanden, aber die mentale superveniente Eigenschaft nicht. Damit diese aber noch hinzukommt und aus dem potenziellen intentionalen Zustand ein bewusster, tatsächlicher wird, muss eine Veränderung erfolgen. Hier stellt sich nun die Frage, inwiefern die physische subveniente Eigenschaft *ohne* mentale supervenierende Eigenschaft dieselbe sein kann wie die physische subveniente Eigenschaft *mit* mentaler supervenierenden Eigenschaft. Scheinbar muss eine Schwelle überschritten werden, damit eine Disposition ein tatsächlicher Hirnzustand wird, bzw. aus einem unbewussten ein bewusster intentionaler Zustand. Wenn nun intentionale Kausalität aufgrund repräsentationaler Eigenschaften Normativität umfasst, muss diese Verbindung von Kausalität und Normativität sich in den neuronalen Dispositionen niederschlagen: sie sind Dispositionen mit einer Ausrichtung auf die normative Welt hin.

Beispielsweise kann für den subjektiven Zustand der emotionalen Aufmerksamkeit für die neuronale Aktivität ein bestimmter minimaler Grenzwert angenommen werden, ab welchem der subjektive Zustand instanziiert wird. Dieser Wert wird in der neurowissenschaftlichen Studie mit

dem signifikanten T-Wert beschrieben, an dem sich die für den Zustand entscheidende Aktivität festmachen lässt. Zu der Intensität der Aktivität kommt insofern noch die „Breite“ hinzu, als für einen bestimmten subjektiven Zustand bestimmte Areale zu einem Ensemble verknüpft werden müssen. Dabei müssen Neuronen, die im Rahmen der Wahrnehmung zuerst aktiviert sind, wie z. B. Detektoren, intensiv genug „feuern“, um eine Kaskade bzw. weitere Neuronen auszulösen. Andererseits müssen die Neuronenensembles entsprechende Signale erhalten, damit weitreichende Aktivität erfolgt – oder nicht. Es muss eben eine Disposition bestehen, nach der die Neuronen „auswählen“ können, ob eine Kaskade aktiviert wird oder nicht. Diese kann sich streng genommen nur auf die subjektiv mentale Ebene beziehen, denn auf dieser Ebene ist Normativität zunächst anzusiedeln. Es scheint, dass ohne diese abstrakte Ebene die Neuronen nicht geeignete normative Signale empfangen können, sondern lediglich rein physische Signale. Ohne die neuronalen Dispositionen scheint dagegen keine unbewusste Verarbeitung der erhaltenen Signale möglich zu sein. Insofern muss Intentionalität beide Bereiche umfassen, den subjektiv mentalen wie den neuronalen Bereich, und darf nicht als lediglich mentales Phänomen betrachtet werden.

***Quintessenz Vergleich Intentionalität***

- *bidirektionales Kausalitätsmodell*
- *Dretskes Indikatorrelation: in geforderter Strenge nicht erfüllbar*
- *Milikans teleosemantischer Ansatz: Verlagerung Produzenten-Konsumenten-Rolle in Gehirn problematisch; evolutionärer Begriff nur bedingt valide*
- *empirische Stützung von Geist-auf-Welt- und Welt-auf-Geist-Relation*
- *Emotionale Aufmerksamkeit als extremes Beispiel kausaler Selbstbezüglichkeit*
- *Begriffliche Modifizierung: Unterscheidung emotionaler, kognitiver und motorischer intentionaler Prozesse*
- *Entwicklung sozialer Normen mit neurologischer Entsprechung: äußere Ereignisse als intentionale Ursachen*
- *Intentionalität: Verknüpfung von Kausalität und Normativität*

- *Dispositionen als neuronale Intentionalität bzw. permanente Potenzialität (strukturelle und funktionale Hirneigenschaften)*
- *empirisch: Konzert der Neuronen nach Stimulus*
- *ontologisch: Verknüpfung von internalem mentalem Ereignis und externalem Ereignis*
- *Abhängigkeit physisch subvenienter Eigenschaften von supervenienten Eigenschaften: mentale Eigenschaften sichern Verbindung zwischen Umwelt und Gehirn*

### 6.6 Emotionale Aufmerksamkeit und Proportionalität

Mentale Kausalität soll als proportionale Relation erklärt werden. Grundlage ist das Prinzip der Proportionalität. Das bedeutet, dass die Beschreibungen der Ursachen und Wirkungen adäquat sind, wenn sie denselben Abstraktionsgrad oder aber dieselbe Detailebene beschreiben. Beispielsweise beschreibt die höhere Abstraktionsebene das Zerbrechen einer Fensterscheibe aufgrund des Schwungs, mit dem ein Gegenstand einschlägt, während auf der niedrigeren Abstraktionsebene weitere, detailliertere Eigenschaften die genaue Art und Weise des Zerbrechens beschreiben (vgl. Kap. 4.6). Das Prinzip der Proportionalität soll aber nicht nur auf der epistemischen Ebene, sondern auch auf der ontologischen Ebene gültig sein. Schließlich ändert sich mit dem Detailumfang einer Eigenschaft auch der Detailumfang der instanziierten Eigenschaft der Wirkung (vgl. Kap. 4.6.1). Der Vergleich der theoretischen Explikationen der neurowissenschaftlichen Studien mit den empirischen Implikationen des Prinzips der Proportionalität soll im Folgenden untersuchen, inwiefern der Anspruch der Proportionalität auf der empirischen Ebene erfüllt werden kann. In der Diskussion werden Schwierigkeiten deutlich, die sich aus der Übertragung des theoretischen Konzepts der Adäquatheit auf die empirische Situation ergeben.

#### 6.6.1 Korrelation zweier Erklärungsbereiche

Zunächst kann hinsichtlich der vorliegenden neurowissenschaftlichen Paradigmen festgestellt werden, dass sie auf der theoretischen Ebene tatsächlich zwei unterschiedliche Erklärungsbereiche miteinander verbinden. Einerseits wird mit dem bildgebenden Verfahren fMRT die neuronale Aktivität im Gehirn gemessen. Dieser Bereich wird üblicherweise als der physische Erklärungsbereich betrachtet (vgl. Kap. 6.2.2; Davidson 1970). Andererseits wird im Post-hoc-Test das subjektive Erleben der Probanden untersucht. Dieser Bereich wird üblicherweise als mentaler oder auch psychischer Erklärungsbereich betrachtet. Aufgrund des Studiendesigns wird in der Regel „cross-level“ zwischen beiden Erklärungsbereichen geschlossen, um somit eine Funktionsbeschreibung der im fMRT gemessenen neuronalen Aktivität zu ermitteln. In der Terminologie wird sehr vorsichtig von einer Korrelation gesprochen, womit beide Erklärungsbereiche prinzipiell aufrechterhalten werden. Andererseits wird implizit die Aufschlüsselung des neuronalen, als basalen Erklärungsbereichs vorangetrieben. Hieraus folgt nicht zwangsläufig eine Konkurrenz, allerdings werden beide Bereiche vermutlich langfristig als alternative Erklärungsbereiche angesehen.

Wenn mentale und physische Erklärungen nicht miteinander konkurrieren, sondern gleichberechtigt gelten, kann unterschieden werden, ob beide Erklärungsbereiche für sich essenzielle,

spezifische Aussagen treffen, die innerhalb des jeweiligen anderen Erklärungsbereiches nicht getroffen werden können (vgl. Burge 1999, 116; Kap. 4.7.1). Die bildgebende Neurowissenschaft macht unter anderem positive Aussagen über neuronale Aktivität. Abhängig von dem jeweiligen Verfahren (fMRT; PET, EEG) können aus der Dritte-Person-Perspektive Aussagen über das Gehirn und seine physischen Eigenschaften gemacht werden. Andererseits muss dabei auf die methodischen Limitationen der Neurowissenschaft verwiesen werden. Es ist fraglich, ob es jemals die technischen Mittel geben wird, tatsächlich alle subjektiven mentalen Phänomene zu messen und entsprechend zu beschreiben. Einige mentale Phänomene scheinen sich aufgrund der „Privatheit" oder aber aufgrund der „Vagheit" bzw. Unschärfe der mentalen Begriffe einer vollständigen neurowissenschaftlichen Bestimmung zu entziehen (vgl. Papineau 2004, 196ff.). Zusätzlich ist anzunehmen, dass zumindest im Detail ein inverses Problem für die Neurowissenschaft bestehen bleibt, und allein auf Basis technischer neurowissenschaftlicher Verfahren ohne subjektive Befragung keine vollständigen psychologischen Profile von Probanden erstellt werden können. Schließlich ist jedes Gehirn im Detail anders strukturiert.

Die Psychologie macht wiederum positive Aussagen über das subjektive Erleben. Mit unterschiedlichen Untersuchungsmethoden (z. B. Fragebogen, Skalierung etc.) wird versucht, die Erste-Person-Perspektive der Probanden bzw. die psychologisch-phänomenalen Ereignisse darzustellen. Andererseits gibt es auch eine offensichtliche epistemische Limitation im subjektiven, mentalen Bereich. Es ist uns nicht möglich, neuronale Aktivität im Gehirn direkt zu fühlen oder zu erleben. Dass signifikante Aktivität in der Amygdala zu verzeichnen ist, drückt sich für uns möglicherweise als Angstzustand aus, aber nicht als „Feuern" bestimmter Neuronen in bestimmten Regionen. Insofern kann zunächst rein methodisch festgehalten werden, dass die psychologische Ebene tatsächlich einen epistemischen Mehrwert liefert. Auf dieser Ebene wird Verhalten und die damit verbundene Kausalrelation in einer Art und Weise erklärt, die eine rein physische, neuronale Beschreibung nicht leisten kann (vgl. z. B. Menzies 2003, 196; Schröder 2003, 177; Kap. 4.6).

Damit gibt es auf der epistemischen Ebene zwei eigenständige Erklärungsbereiche. Keiner von beiden kann – beispielsweise im Sinne einer Reduktion – aufgegeben werden, da dies zu einem Verlust und einer Einschränkung des Erkenntnisraums führen würde. Aufgrund dieser epistemischen Situation kann auch kein ontologisches Primat für einen der beiden Bereiche folgen. Natürlich wird die bildgebende Neurowissenschaft vor dem Hintergrund des physikalischen Weltbilds (vgl. Kap. 3.3) gemeinhin favorisiert, aber eine schlichte Gleichsetzung von Neuronen und physischen Eigenschaften bei gleichzeitigem Ausschluss des subjektiven Erlebens vereinfacht die epistemische Ebene und verkürzt somit letztlich auch die ontologische Dimension.

### 6.6.2 Ausdifferenzierung des Abstraktionsgrades

Bei der Verknüpfung der theoretischen Explikationen mit möglichen empirischen Implikationen des Prinzips der Proportionalität muss bestimmt werden, welche Eigenschaften determinables und welche determinates entsprechen. Naheliegend wäre die Bestimmung subjektiv mentaler Eigenschaften als determinables bzw. als proportional zu höherstufigen abstrakten Eigenschaften. Beispiel wäre ein emotionales Bild „Babylächeln", das subjektiv als solches vom Probanden erkannt und erlebt wird. Parallel wäre die neuronale Aktivität als determinate bzw. als proportional weniger abstrakte Eigenschaft zu bestimmen:

a) neuronal/determinate: geringer Abstraktionsgrad
b) subjektiv mental/determinable: hoher Abstraktionsgrad

Hier stellt sich aber bereits die Frage, welche Eigenschaften auf der Bildseite als „weniger abstrakt" in Betracht kommen. Hierfür könnten gegebenenfalls verschiedene Farb- und Formeigenschaften des Bildes angegeben werden. Allerdings ist es die Frage, ob ein Babylächeln einen höheren Abstraktionsgrad besitzt, als Formen und Farben. Ein Babylächeln abstrakt zu nennen, scheint zunächst unpassend. Der Be-griff müsste modifiziert werden. Hier wäre es interessant zu untersuchen, wieweit sich eine solche theoretische Unterscheidung neuronal niederschlägt. Verschiedene computationale oder konnektionistische Modelle unterstützen das Prinzip der Proportionalität. Sie zeigen tatsächlich, dass einfache geometrische Formen von rekurrenten Systemen schneller erlernt bzw. dargestellt werden können, als komplexere Gegenstände (vgl. Churchland 1997, 143).

Andererseits lassen Kenntnisse insbesondere über den präfrontalen Cortex vermuten, dass im Gehirn selbst höherstufige kognitive Funktionen von niedrigstufigen unterschieden bzw. an anderer Stelle prozessiert werden. Damit muss, ausgehend von der neuronalen Aktivität, eine dreifache Unterscheidung getroffen werden, nämlich zwischen:

c) niedrigstufig neuronal/determinate: geringer Abstraktionsgrad
d) höherstufig neuronal/determinable bzw. determinate: hoher Abstraktionsgrad
e) höherstufig mental subjektiv/determinable: hoher Abstraktionsgrad

Die höherstufigen neuronalen Eigenschaften sind dabei einerseits mögliche determinables der niedrigstufigen neuronalen Eigenschaften und andererseits determinates der mentalen subjektiven Eigenschaften. Zusätzlich können noch

f) niedrigstufige mental subjektive Eigenschaften mit einem niedrigen Abstraktionsgrad

eingeführt werden, welche die oben genannten Farb- und Formeigenschaften umfassen würden. Diese wiederum wären determinates der höherstufigen mental subjektiven Eigenschaften und determinabels der niedrigstufigen neuronalen Eigenschaften, stünden aber in keiner direkten Beziehung zu den höherstufigen neuronalen Eigenschaften. In der Praxis kann gezeigt werden, dass niedrigstufige mentale Eigenschaften mit niedrigstufigen neuronalen Eigenschaften, höherstufige mentale Eigenschaften mit höherstufigen neuronalen Eigenschaften korrelieren. Es muss hier offensichtlich doppelt unterschieden werden: einmal vertikal zwischen zwei verschiedenen Erklärungsebenen (höherstufig versus niedrigstufig) sowie horizontal zwischen zwei verschiedenen Erklärungstypen (mental subjektiv versus neuronal). Während Schröder (2003) den Abstraktionsgrad cross-level ansetzt, wird hierbei zusätzlich intra-level unterschieden.

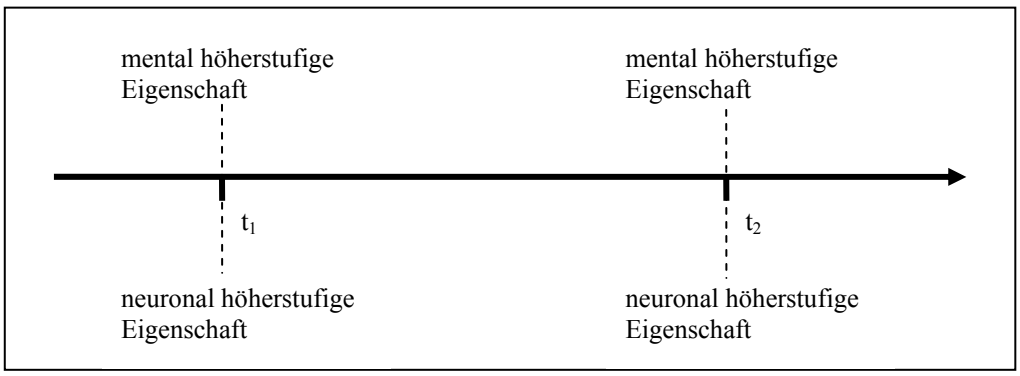
Dabei kann eingewendet werden, dass der cross-level-Vergleich nicht zulässig ist, solange zwischen zwei verschiedenen Erklärungstypen differenziert wird. Korrelationen stützen zwar die Verbindung beider Ebenen; dennoch sollte nicht vorschnell auf die Identität beider Bereiche geschlossen werden, weshalb die Frage der Proportionalität für mentale Kausalität virulent bleibt. Eine Eigenschaft verhält sich nur dann proportional zu einer anderen, wenn sie denselben Grad an Abstraktion erfüllt. Eine detaillierte Beschreibung des physischen Zustands, auf welchem die Intention zu trinken superveniert, soll nach Schröder irrelevant für die zu instanziierende Eigenschaft „trinken" sein. Die Angabe eines physischen Zustands bzw. spezifischer neuronaler Aktivität wäre ihm zufolge lediglich proportional sinnvoll, wenn die genauen Verbindungen, die das Trinken als Ereignis realisieren, bestimmt werden sollen. Unter der oben vorgenommenen doppelten Unterscheidung kann dagegen nur angenommen werden, dass in Korrelation zur instanziierten Eigenschaft „trinken" eine höherstufige neuronale Eigenschaft aktiviert wird. Diese ist jedoch nicht adäquate Ursache. Die adäquate Ursache muss vielmehr zeitlich der Wirkung vorangehen und ist in inneren (der Zustand „Durst") wie äußeren Stimuli (die Wahrnehmung eines Getränks) zu suchen.

Für das Bild „Babylächeln" muss wiederum eine dreifache proportionale Unterscheidung vorgenommen werden (vgl. Kap. 6.1.4) zwischen dem höherstufigen Abstraktionsgrad des Bildes, dem höherstufigen Abstraktionsgrad des subjektiven mentalen Erlebens und dem höherstufigen neuronalen Abstraktionsgrad. Damit sind wir wiederum bei der eingangs geschilderten Korrelation mentaler und physischer Erklärungstypen als alternative und darüber hinaus komplementäre Erklärungsbereiche angelangt.[166]

[166] Allerdings bleibt die Frage, wie entschieden werden soll, ob die höherstufige neuronale oder die höherstufige mentale Eigenschaft sich stärker proportional zu der höherstufigen Bildeigenschaft verhält. Die Antwort würde

Nach Schröder und Yablo müssten Eigenschaften der mentalen Erklärungsebene grundsätzlich determinables der Eigenschaften der physischen Erklärungsebene sein – ansonsten würde sein Prinzip der Proportionalität die kausale Kraft jener Eigenschaften nicht garantieren (vgl. Kap. 4.6.1). Ontologisch versucht Schröder über die Supervenienzrelation das Prinzip der Proportionalität zu stärken und die kausale Kraft mentaler Eigenschaften zu begründen. Damit ist auch eine Implementierungsrelation zwischen semantischen Eigenschaften und Aktivierungsvektoren als Modell möglich (vgl. Kap. 4.6.2). Entscheidend bleibt dabei, dass mentale Eigenschaften Bestandteil von Generalisierungen sein können, die nicht in physischen Begriffen beschrieben werden können (vgl. Marras 2003, 261).

In diese Richtung zielt Yablos Argumentation, wonach für physische Beschreibungen die Gefahr der Überspezifizierung besteht (vgl. Yablo 1997, 259; Kap. 4.6.1). Wichtiger noch ist die Perspektive, dass mentale Repräsentationen über Kriterien verfügen, die Beschreibung eines physischen Ereignisses in seiner Variabilität zu erfüllen. Mentale Eigenschaften umfassen verschiedene alternative neuronale Verknüpfungspfade und stellen sicher, dass eine Zielrepräsentation, wie z. B. ein positives IAPS-Bild, erreicht wird (vgl. Kap. 4.6.2). Das Argument beschreibt letztlich das Prinzip der Multirealisierbarkeit mentaler Eigenschaften. Die subjektive mentale Eigenschaft wäre auch mit einer anderen neuronalen Eigenschaft realisiert worden. Hinsichtlich der vorliegenden neurowissenschaftlichen Studien lässt sich lediglich sagen, dass dieselben Bilder von unterschiedlichen Probandengehirnen erkannt wurden. Es müsste darüber hinaus untersucht werden, ob auch bei denselben Probanden für dieselben Stimuli alternative Aktivierungsmuster erkennbar sind.




Grafik 18: Das Prinzip der Proportionalität bezieht sich auf die horizontale Ebene. Eine mental höherstufige Eigenschaft zum Zeitpunkt $t_1$ kann danach nur von einer (mental) höherstufigen Eigenschaft zu einem vorangegangenen Zeitpunkt $t_1$ adäquat verursacht werden. Die Korrelation der mental höherstufigen und neuronal höherstufigen Eigenschaften beschreibt dagegen eine vertikale Relation. Auch die neuronal höherstufige Eigenschaft zum Zeitpunkt $t_1$ und die mental höherstufige Eigenschaft $t_2$ verhalten sich nicht proportional zueinander, da unterschiedliche, komplementäre Erklärungsbereiche bzw. Eigenschaftsbereiche beschrieben werden.

hierbei vermutlich in der Terminologieentscheidung liegen, wonach die Terminologie der Bildeigenschaft „Babylächeln" vermutlich eher der Terminologie der subjektiv mentalen Eigenschaften „Babylächeln" entspräche.

### 6.6.3 Übertragungsmöglichkeiten für das Paradigma

Die emotionalen Dimensionen Valenz, Intensität und Wiedererkennen befinden sich auf einem anderen Abstraktionsgrad, als die entsprechenden prozessierenden neuronalen Regionen, mit denen jene korrelieren (vgl. Kap. 5.2.2; Kap. 5.2.3; Kap. 6.1.3; Kap. 6.1.4). Auffällig ist hier die Überlappung kognitiver und emotionaler Prozessierung bei den verschiedenen relevanten Hirnregionen. Zwar kann über den Aufbau der Studie hinsichtlich der neuronalen Aktivität weiter differenziert werden. Allerdings ist es fraglich, ob auf der Ebene des subjektiven Erlebens ebenfalls eine Überlappung vorhanden ist bzw. wahrgenommen wird. Diese Frage stellt sich auch bezüglich der übrigen Korrelationsanalysen. So können die Korrelationen der emotionalen Dimension Valenz äußerst differenziert dargestellt werden. Valenz korreliert mit dem VMPFC während unerwarteter Bildwahrnehmung, während der Bildbeurteilung jedoch nicht mit dem VMPFC, sondern mit dem bilateralen DLPFC. Die Intensität korreliert während unerwarteter Bildwahrnehmung signifikant mit dem medialen Cerebellum, während nicht-erwarteter Bildbeurteilung mit dem rechtem VLPFC und während erwarteter Bildbeurteilung mit dem Nucleus accumbens und dem DMPFC (vgl. Kap. 5.2.2). Auf der neuronalen Ebene wird hier wichtige Detailarbeit geleistet, allerdings bleibt die subjektive Ebene weiterhin wesentlich undifferenziert. Ein und derselbe subjektive mentale Zustand wird während verschiedener Konditionen signifikant von verschiedenen neuronalen Bereichen prozessiert.

Allerdings hat das Prinzip der Proportionalität nicht die einzelne Korrelation im Blick, sondern die kausale Relation zwischen zwei Korrelationen. Die neuronale Aktivität ist möglicherweise die passende subveniente Basis für das bewusste Erleben der Valenz oder Intensität. Andererseits erscheint z. B. die Aktivität im DLPFC während unerwarteter Bildwahrnehmung nicht adäquat für das folgende Ereignis der Bildbeurteilung zu sein. Hier lohnt auch die Betrachtung der Folge von dem Zustand der emotionalen Aufmerksamkeit zum Zustand der emotionalen Bildwahrnehmung auf der neuronalen wie subjektiven Ebene. Neuronal werden zwei separate Netzwerke aktiviert, wobei jedoch nicht davon gesprochen werden kann, dass das antizipatorische Netzwerk das Wahrnehmungsnetzwerk verursacht, da letzteres insbesondere durch externe Stimuli verursacht wird. Immerhin kann eine verstärkte Aktivität als Wirkung beobachtet werden. Im Gegensatz hierzu kann auf der subjektiven Ebene nicht genau differenziert werden zwischen dem Zustand der emotionalen Aufmerksamkeit und dem Zustand der gesteigerten emotionalen Wahrnehmung. Zwar lassen sich auch hier die Zustände anhand der gezeigten externen Stimuli unterteilen. Sie werden aber von den Probanden subjektiv als eine kontinuierliche Folge erlebt. Hinsichtlich der Frage der Proportionalität kann nun festgehalten werden, dass vor allem die Zielrepräsentation für den Zustand der emotionalen Bildwahrnehmung entscheidend ist. Dies kann aber nicht adäquat durch (abstrakte) physische Eigenschaften beschrieben werden – sie sind

zu spezifisch bzw. lediglich disjunkt (vgl. Kap. 4.6.1). Zusätzlich ist die Abweichung von neuronaler und subjektiver Beobachtung problematisch, da auch hierfür nicht deutlich ist, inwiefern die physischen Eigenschaften proportionale Ursache für das subjektive Erleben sein können.

Grundsätzlich kann festgestellt werden, dass innerhalb der Neurowissenschaft allein vom methodischen Interesse her vor allem die Detailebene untersucht wird. Schließlich liegt der Fokus darauf aufzuzeigen, welche Hirnregionen o. ä. welche Funktionen prozessieren. Die neuronale Detailebene und der damit verbundene niedrige Abstraktionsgrad werden „cross-level" mit dem höherstufigen Abstraktionsgrad der subjektiven mentalen Ebene korreliert. Wenn nun das Prinzip der Proportionalität auf die vorliegenden neurowissenschaftlichen Paradigmen übertragen wird, kann untersucht werden

a) was die adäquate Wirkung der Aufmerksamkeitsstimuli (Pfeile) ist,
b) was die adäquate Wirkung der emotionalen Stimuli (Bilder) ist.

Die Studien weichen insofern von Schröders Modell ab, als kein Impuls bzw. keine körperliche Handlung ausgeführt wird, sondern eine psychische Veränderung erfolgt. Zudem befinden sich die Bildstimuli als Ursache außerhalb des Körpers, während die Wirkung im Gehirn bzw. im subjektiven Erleben beobachtet wurde. Eingeschränkte Ausnahme ist die Beurteilung der Bilder durch die Probanden, da hierbei die Ursache in vorhandenen Dispositionen der Probanden zu suchen ist, auch wenn der Aufforderungsstimulus, das Bild per Knopfdruck als positiv oder negativ zu bestimmen, von Außen erfolgte. Insofern kann in Ergänzung diskutiert werden,

c) was die adäquate Wirkung der Beurteilung ist.

Für a) die Aufmerksamkeitsstimuli bzw. Pfeilsignale kommen neuronal das antizipatorische Netzwerk sowie subjektiv der Faktor emotionale Aufmerksamkeit in Betracht. Das antizipatorische Netzwerk ist auf der Detailebene mit geringerem Abstraktionsgrad anzusiedeln und der Faktor emotionale Aufmerksamkeit dagegen auf der Ebene mit höherem Abstraktionsgrad. Auf der epistemischen Ebene kann festgestellt werden, dass die Formulierung „emotionale Aufmerksamkeit" einen höheren Abstraktionsgrad besitzt als die konkrete neuronale Aktivität im SAC, einschließlich CMA, SMA und POS.

Auf welchem Abstraktionsgrad muss nun aber die Ursache dieser beiden unterschiedlichen Wirkungsbeschreibungen angesiedelt werden? Um die subjektive Eigenschaft „emotionale Aufmerksamkeit" zu erklären, soll die detaillierte Beschreibung des physischen Zustands, auf welchem der subjektive Zustand superveniert, irrelevant für die zu instanziierende Eigenschaft

sein (vgl. Schröder 2003; Kap. 4.6). Ursache soll eine Eigenschaft sein, die sich auf demselben hohen Abstraktionsgrad bewegt. Dies wäre z. B. die Instruktion „bei senkrechtem Pfeil folgt ein emotionales Bild“ (vgl. Kap. 5.1.1). Da das Pfeil-Signal dem Zustand der emotionalen Aufmerksamkeit zeitlich vorausgeht, wäre also auf der horizontalen Erklärungsebene die damit verbundene subjektiv erlebte Information Ursache für den subjektiven erlebten Zustand „Ich bin emotional aufmerksam“ bzw. „ich erwarte ein emotionales Bild“. Wenn dagegen die genaue Verbindung, die den Zustand der emotionalen Aufmerksamkeit realisiert, erklärt werden soll, dann ist nach Schröder die Angabe des physischen Zustands erforderlich. Dabei wird die Erklärungsebene im Sinne der determinables-determinates Relation vertikal nach unten verschoben. Prinzipiell handelt es sich hierbei um den komplementären Erklärungstypus. Hierbei muss allerdings für ein externales Bildsignal derselbe Abstraktions- bzw. Detailgrad dargestellt werden, wie für das neuronale Aktivitätsmuster.

Für b) die emotionalen Stimuli bzw. Bilder kommen neuronal das perzeptive Netzwerk sowie subjektiv der Faktor Emotion in Betracht. Das perzeptive Netzwerk ist auf der Detailebene mit geringerem Abstraktionsgrad anzusiedeln und der Faktor Emotion dagegen auf der Ebene mit höherem Abstraktionsgrad. Die Formulierung „emotionale Bildwahrnehmung“ besitzt einen höheren Abstraktionsgrad als die konkrete neuronale Aktivität von Amygdala, Insula, medialem und lateralem präfrontalem Cortex, Putamen, Mittelhirn, Cerebellum und okzipitalen visuellen Regionen. Der Abstraktionsgrad der Ursache dieser beiden unterschiedlichen Wirkungsbeschreibungen kann relativ einfach bestimmt werden: Er liegt zum einen im abstrakten, semantischen Inhalt der gezeigten Bilder („Babylächeln“, „Autounfall“). Zum anderen liegt er in den detaillierten Form- und Farbeigenschaften des Bildes.

Für c) die Bildbeurteilung ist die Wirkung recht eindeutig durch den Knopfdruck ja/nein und die korrelierende neuronale Aktivität zu bestimmen, denn der Beurteilung soll die motorische Handlung folgen. Auch hierfür gilt, dass der subjektive bewusste Knopfdruck als Wirkung adäquat zur subjektiven bewussten Beurteilung eines Bildes positiv bzw. negativ ist. Auf der Detailebene sind entsprechend die korrelierenden neuronalen Aktivitäten anzusiedeln: die höherstufige neuronale Aktivität parallel zur subjektiven Bewertung und die niedrigstufige neuronale Aktivität im motorischen Cortex parallel zur Handlung „Knopfdruck“ (vgl. Kap. 5.2.3). Allerdings ist die subjektive Bewertung eines Bildes selbst Wirkung eines ensprechenden Stimulus, der auf dem Bild als zusätzliches Signal zur Beurteilung auffordert. Dieses Beurteilungssignal kann ähnlich wie das Aufmerksamkeitssignal in die Detailebene der bloßen Form des Signals und in die höhere Abstraktionsebene der damit verbundenen Instruktion unterteilt werden. Für alle kausalen Relationen gilt dabei, dass Ursache und Wirkung nur auf der horizontalen Ebene adäquat sein können.

Die Verknüpfung der theoretischen Explikationen mit den empirischen Implikationen zeigt, dass auf der horizontalen Ebene eine proportionale kausale Linie verfolgt werden kann – separat für den entsprechenden Abstraktionsgrad des Eigenschaftsbereichs. Hiergegen kann jedoch eingewendet werden, dass das Prinzip lediglich für die Erklärungsebene sinnvoll übertragen werden kann. Es liefert zwar verschiedene Erklärungstypen, letztlich sind aber für die ontologische Ebene die physischen Eigenschaften relevant (vgl. Kap. 3.4). In einem ersten Schritt wird aber immerhin bestätigt, dass unter empirischen Gesichtspunkten weder ein Widerspruch zur Voraussetzung der kausalen Geschlossenheit noch zum Argument des kausalen Ausschlusses besteht – jedes mentale Ereignis hat mindestens eine physische Eigenschaft mit einer proportionalen physischen Ursache (vgl. Kap. 4.6.1).

### 6.6.4 Proportionale Kausalität und Eigenschaftspluralismus

Die Verknüpfung des Prinzips der Proportionalität mit den theoretischen Explikationen der neurowissenschaftlichen Studien hat gezeigt, dass es auf der epistemischen Ebene sehr wohl eine Unterscheidung zwischen der subjektiven und der neuronalen Ebene gibt. Dies bestätigt die Einschätzung, die sich aus der Problematisierung des Korrelationsbegriffs ergab (vgl. Kap. 6.1.4). Genauso wenig wie aus einem hohen T-Wert bzw. verstärkter Aktivität in einer Region auf eine Identität zwischen Bild und neuronalem Zustand geschlossen werden kann, kann allein auf dieser Basis die Adäquatheit zwischen Ursache und Wirkung festgestellt werden. Sicherlich liegt es aufgrund der empirischen Daten nahe zu vermuten, dass die verschiedenen Zustände oder Ereignisse in einer Relation zueinander stehen. Andererseits kann – entgegen einer möglichen Reduktion – der genaue Eigenschaftsbereich nicht bestimmt werden. So ist es möglich und beinahe selbstverständlich, dass sich trotz Überschneidung der Eigenschaftsbereich des Bildes und der neuronale Eigenschaftsbereich unterscheiden. Aufgrund dieser unterschiedlichen Eigenschaftsbereiche ist es fraglich, ob ein neuronaler Zustand mit einem Eigenschaftsbereich von hohem Detailgrad allein adäquate Wirkung eines Bildes mit einem Eigenschaftsbereich von hohem Abstraktionsgrad sein kann. Hier scheint ein subjektiver mentaler Zustand mit einem Eigenschaftsbereich von hohem Abstraktionsgrad geeigneter zu sein, den Anforderungen der Zielrepräsentation zu entsprechen.

Abgesehen davon, dass verschiedene zusätzliche Bedingungen erfüllt sein müssen – es muss überhaupt ein Bild gezeigt werden – , zeigen unter anderem die Korrelationsmessungen zwischen den fMRT-Daten und den emotionalen Dimensionen Intensität und Valenz, dass aus der empirischen Perspektive tatsächlich ein subjektiver mentaler Zustand der emotionalen Aufmerksamkeit instanziiert wird bzw. werden muss, ehe der verstärkte Zustand des emotionalen

Bildsehens erreicht wird (vgl. Kap. 5.2.3). Für die horizontale Ebene kann somit festgestellt werden, dass, bliebe der subjektive Zustand der emotionalen Aufmerksamkeit aus, auch ein gesteigertes emotionales Bildsehen nicht möglich wäre – weder auf der subjektiven noch auf der neuronalen Ebene. Für die ontologische Ebene kann gefolgert werden, dass das nachfolgende Ereignis zumindest ein anderes gewesen wäre.

Die Abhängigkeit einer Wirkeigenschaft richtet sich nach dem Detailumfang der ursächlichen Eigenschaft. Allerdings sind niedrigstufige Eigenschaften stets auch abhängig von höherstufigen Eigenschaften, was umgekehrt nicht gilt. Es handelt sich schließlich um eine asymmetrische Relation. Das antizipatorische Netzwerk oder das perzeptive Netzwerk werden nicht nur durch detaillierte Form- oder Farbstimuli instanziiert, sondern auch im Rahmen von Bildsignalen mit einem höheren Abstraktionsgrad. Dies zeigte sich in der Ausdifferenzierung des Abstraktionsgrades. Die höherstufige Abstraktionsebene umfasst als determinable immer auch die niedrigstufige Abstraktionsebene der determinates. Dagegen wird die subjektive mentale Ebene („Ich erwarte ein emotionales Bild" o. ä.) nicht allein durch detaillierte Form- und Farbstimuli instanziiert. Dies kann auch trivialerweise für die umgekehrte Richtung festgehalten werden: Neuronale Regionen, die basale Funktionen prozessieren, umfassen keine höherstufigen Funktionen.

Diese vertikale Ebene ist jedoch für das Prinzip der Proportionalität nicht allein entscheidend. Hierfür muss die horizontale Ebene zwischen zwei prozessierten Funktionen („emotionale Aufmerksamkeit" und „emotionales Bildsehen") bzw. deren subvenienten neuronalen Basen („antizipatorisches Netzwerk" und „emotional-perzeptives Netzwerk") in einer zeitlichen bzw. kausalen Abfolge betrachtet werden. Nach dem Prinzip der Proportionalität würde dann gelten, dass die Aktivität des antizipatorischen Netzwerks zwar notwendig, aber nicht hinreichend ist für die Instanziierung des subjektiven Zustands des emotionalen Bildsehens „Ich sehe ein positives Bild besonders aufmerksam". Beide Ebenen, die vertikale wie die horizontale, sind, wie auch die höherstufigen und niedrigstufigen Eigenschaften, hinsichtlich der Kausalität miteinander verwoben. Nach dem Prinzip der Proportionalität kann nun der subjektive Zustand der emotionalen Aufmerksamkeit die abstrakte Minimalbedingung für diese variablen Hirnzustände darstellen, unter die unterschiedliche Amplituden oder Dauer genauso wie im Detail unterschiedliche Vernetzungen subsumiert werden können.[167] Doch handelt es sich nach Schröder hierbei nicht nur

[167] Das Prinzip der Eigenschaftsproportionalität berücksichtigt neben dem erforderlichen Detailumfang auch die höchste abstrakte Eigenschaft. Dabei soll als sparsamste Erklärung diejenige Eigenschaft gelten, die auch durch die Wirkung von unterschiedlichen Ereignissen instanziiert wird. Dieser Eigenschaft soll als Minimalbedingung die proportionale kausale Kraft zugeschrieben werden (vgl. Kap. 4.6). Übertragen auf die vorliegenden Studien kann gesagt werden, dass die subjektiven Erlebniszustände (emotionale Aufmerksamkeit), emotionales/neutrales Bildsehen oder (nicht-) erwartete Bildbewertung a) einmal auf verschiedenen Probandengehirnen supervenieren und b) zusätzlich die Amplitude oder die Dauer innerhalb eines einzelnen Probandengehirns für diese subjektiven Zustände variiert. Damit ergibt sich eine Situation, die mit dem Chauvinismus-Dilemma aus der Funktionalismus-Diskussion vergleichbar ist. Dort ergibt sich das Problem, dass eine zu enge Bindung der Rollen-Eigenschaft an die Realisierer-Eigenschaft den Kreis der Wesen, die z. B. Schmerzen haben können, zu stark

um eine epistemische, sondern auch um eine ontologische Trennung. Dies macht er dadurch deutlich, dass auf den Ebenen unterschiedliche Ereignisse instanziiert werden (vgl. Kap. 4.7.2). Insofern richtet sich das Prinzip der Proportionalität auch auf der ontologischen Ebene gegen das Ausschlussargument.

In der Funktionszuschreibung neuronaler Regionen oder sogar einzelner Neuronen ergibt sich aus deren vielfacher Einsetzbarkeit in unterschiedlichen Netzwerken auch eine andere Perspektive für die Anforderung an Adäquatheit. Ein Beispiel gibt die neuronale Netzwerkstruktur im visuellen System. Sie ist durch parallele Prozessierung gekennzeichnet, Konvergenzzentren scheinen dagegen zu fehlen (vgl. Singer 2002c, 149f.). Auf allen Verarbeitungsstufen reagieren die Neuronen auf mehr als nur ein Merkmal:

> Weil einzelne Neuronen relativ unselektiv durch ein breites Spektrum verschiedener Merkmale erregt werden, überlappen die neuronalen Populationen beträchtlich, die durch verschiedene Merkmale aktiviert werden. Eine bestimmte Zelle kann zu verschiedenen Zeitpunkten Mitglied unterschiedlicher Populationen werden und dadurch an der Repräsentation sehr vieler verschiedener Merkmale partizipieren. (ebd., 157)

Man könnte nun sagen, dass die Populationen bzw. Ensemblebildungen proportional zu einer externen Ursache oder Wirkung sind. Hier stellt sich erneut die Frage der Grenzziehung (vgl. Kap. 6.4): Auf welcher Ebene ist Proportionalität gewährleistet, wie umfassend darf bzw. muss ein Netzwerk sein, inwiefern muss differenziert werden? Ziel des Prinzips der Proportionalität ist nicht zu zeigen, dass mentale Zustände neuronale Basen haben und auf diesen supervenieren, sondern eine ebenenspezifische proportionale Kausalität zu sichern. Während auf der neuronalen Ebene gezeigt werden kann, dass die Amplitude/Aktivität in neuronalen Detektoren zu weiterer Aktivität in Ganglienzellen etc. führt, ist auf der mentalen subjektiven Ebene die Frage, ob die Aktivität der neuronalen Detektoren oder die Aktivität einer entsprechenden Population zum Zeitpunkt $t_1$ adäquate Ursache für z. B. die Handlung „Knopfdruck" zum Zeitpunkt $t_2$ sein kann oder ob nur der subjektive Gedanke „Das Bild ist positiv, ich drücke den Knopf" als adäquat gelten kann. Hierbei handelt es sich um zwei verschiedene Erklärungsbereiche, die sich jedoch beide auf dasselbe Ereignis auf der ontologischen Ebene beziehen. Sobald eine Erklärungsebene wegfällt, wird der adäquate Bezug zwischen der epistemischen und der ontologischen Ebene aufgegeben.

Empirisch kann das Prinzip der Proportionalität auch dahingehend interpretiert werden, dass es für die neuronale Aktivität einen bestimmten „minimalen" Grenzwert geben muss, ab welchem der subjektive Zustand der emotionalen Aufmerksamkeit instanziiert wird. Ist der er-

---

begrenzt, während eine zu lose Bindung die Rollen-Eigenschaft zu vage und kausal wirkungslos erscheinen lässt (vgl. Kap. 2.2).

reicht, wird auf der ontologischen Ebene ein Ereignis instanziiert, auf welches sich entsprechend beide Eigenschaftsbereiche beziehen. Allerdings stellt sich die Frage, ob die neuronale Aktivität auf der Detailebene ohne die abstrakte Ebene – die determinables – geeignete Signale empfangen kann bzw. passend aktiv wird. Schließlich muss der Inhalt bzw. müssen die Eigenschaften mit einem hohen Abstraktionsgrad verarbeitet werden (vgl. Kap. 4.3.3).

***Quintessenz Vergleich Proportionalität***

- *empirische Bestätigung zweier unterschiedlicher Abstraktionsebenen mit spezifischer Aussagekraft*
- *Ablehnung ontologisches Primat aufgrund epistemischer Pluralität*
- *Ausdifferenzierung des Abstraktionsgrades:*
  *niedrigstufig-neuronal, höherstufig-neuronal,*
  *niedrigstufig-mental-subjektiv, höherstufig-mental-subjektiv*
- *Unterscheidung zwischen Erklärungsebenen (vertikal) und Erklärungstypen (mental versus neuronal)*
- *adäquate Ursache: zeitlich vorhergehend und ebeneneinheitlich*
- *mentale Eigenschaften als Bestandteil von nicht physisch beschreibbaren Generalisierungen*
- *Sicherstellen der Zielrepräsentation: Erfassung der Variabilität physischer Ereignisse durch mentale Repräsentationen*

- *Prinzip der Proportionalität für mentale Kausalität: kausale Relation zwischen zwei Korrelationen*
- *Aktivität neuronaler Populationen proportional zu äußeren Ereignissen*
- *Kausale Verwebung: neuronale Aktivität Basis für Subjektivität; Instanziierung bewusster subjektiver Ereignisse kausal notwendig für Instanziierung weiterer Ereignisse; Abhängigkeit niedrigstufiger Eigenschaften von höherstufigen Eigenschaften*
- *Adäquatheit der Ereignisinstanziierung: Lediglich empirisch-epistemische Differenzierung bei ontologischer Verschränkung*

Für die Prüfung der kausalen Adäquatheit muss letztlich diese „Adäquatheit der Instanziierung" berücksichtigt werden. Wenn nun einer Bildbetrachtung die Bildbewertung und schließlich die Handlung „Knopfdruck" folgen, so müssen Neuronenensemble diesen Zuständen vertikal zugrunde liegen. Ein neuronales Ensemble zum Zeitpunkt $t_1$ kann dabei für die bewusste subjektive Bildbewertung oder den bewussten subjektiven Knopfdruck $t_2$ nur adäquat sein, wenn es selbst Basis des vorhergehenden bewussten subjektiven Zustands ist. In diesem Sinne kann nur die Instanziierung des bewussten subjektiven Zustands kausal adäquat für die nachfolgenden bewussten subjektiven Zustände sein.

Das Ausschlussargument scheint dem tatsächlichen Hirnereignis nicht gerecht zu werden, da weder die mentalen subjektiven noch die neuronalen Eigenschaften irrelevant sind. Die neu-

ronale Aktivität ist notwendig, um, basierend auf der Population, das subjektive bewusste Ereignis zu instanziieren. Die Instanziierung des subjektiven bewussten Ereignisses scheint dagegen erst hinreichend für die Instanziierung weiterer bewusster subjektiver Ereignisse zu sein.[168] Das Prinzip der Proportionalität beschreibt hinsichtlich der kausalen Adäquatheit zwei Erklärungsbereiche, die a) empirisch scheinbar vollständig zu trennen sind, aber b) beide aufgrund ihrer Verschränkung auf eine gemeinsame ontologische Relation verweisen. Würde einer der beiden Eigenschaftsbereiche fehlen bzw. wäre nicht erfüllbar, würde eine andere ontologische Relation beschrieben werden.

[168] Natürlich ist neben der Folge zweier bewusster Ereignisse aufeinander auch die Folge unbewusstes Ereignis und bewusstes Ereignis möglich. Hierbei muss allerdings untersucht werden, welche Eigenschaften ein unbewusstes Ereignis haben muss, damit ihm ein bewusstes folgt. Schließlich folgt nicht jedem unbewussten Hirnereignis ein bewusstes, sondern nur in besonderen Fällen.

## Kapitel 7

## Entwicklung der neurophilosophischen Hypothese

Die Ergebnisse aus dem Vergleich der Kriterien für mentale Kausalität mit den neurowissenschaftlichen Paradigmen zur emotionalen Aufmerksamkeit und zur emotionalen Beurteilung verweisen, neben einigen positiven Resultaten möglicher Übereinstimmung und damit Unterstützung der Position, auf bestehende Schwierigkeiten der mentalen Kausalität – jenseits theoretischer Differenzen, allein auf empirischer Ebene. Dies liegt möglicherweise daran, dass die Kriterien und die Voraussetzungen zu diesen Schwierigkeiten führen, da sie keine Auflösung des Problems zulassen (vgl. Kap. 3; Kap. 4). Das Problem der mentalen Kausalität wäre somit hausgemacht, wir befänden uns in einem argumentativen Zirkel.

In einer kausalen Erklärung sollen die zu erklärenden Ereignisse als Wirkung bestimmter vorhergehender Ursachen folgen. Die Geschlossenheit der physischen Welt lässt mentale Eigenschaften als Ereignisursachen scheinbar nicht zu. Auf der anderen Seite funktionieren psychologische Erklärungen für Ereignisse im Alltag sehr wohl. Sätze wie „Er sprang in den Tod, da er finanzielle Probleme hatte" werden als Erklärungen von Ereignissen akzeptiert, allerdings sind sie unvollständige kausale Erklärungen und entbehren eines strikten Gesetzescharakters. Andererseits hat die Anwendung der Kriterien für mentale Kausalität auf die neurowissenschaftlichen Paradigmen verdeutlicht, wie wichtig die Regularitäten auf der psychologischen, subjektiven Ebene für die neurowissenschaftliche Forschung sind (vgl. Kap. 6.1).

Zudem muss eingeräumt werden, dass der Begriff der Kausalität selbst problematisch ist. In der philosophischen Tradition hat vor allem David Humes Entwicklung der Regularitätstheorie Entscheidendes zur Analyse des Kausalitätsbegriffs beigetragen.[169] Der Empirist Hume hat gezeigt, dass Kausalurteile über die Erfahrung hinausgehen, da sie induktiv, also als Einzelbeobachtung verallgemeinert geschlossen werden. Für die Vorstellung, eine Ursache bringe mit einer Kraft eine Wirkung hervor, kann in der Erfahrung nichts Entsprechendes gefunden werden. In der Natur können kausale Prozesse lediglich als regelmäßige zeitliche Abfolge gleichartiger Ereignisse beobachtet werden, eben als Regularitäten:

> In Übereinstimmung mit dieser Erfahrung mögen wir also eine Ursache definieren als: einen Gegenstand, dem ein anderer folgt, wobei allen Gegenständen, die dem ersten gleichartig sind, Gegenstände folgen, die dem zweiten gleichartig sind. (Hume 1973, 92)

[169] Hume hat damit Kant aus seinem sog. dogmatischen Schlummer geweckt. Für diesen gilt allerdings das Kausalprinzip, insofern es die Voraussetzung von Erfahrung bildet. Die Notwendigkeit der Verknüpfung von Ursache und Wirkung versteht er als strenge regelhafte Allgemeinheit und nicht lediglich als Gewöhnung an gleichartige Eindrücke (vgl. Kant 1998, 48f.). Letztendlich beruft sich Kant damit auch auf Leibniz, der das Kausalprinzip mit dem Satz vom zureichenden Grunde erst aufgestellt hat (vgl. Titze 1992, 13).

Es scheint zwar, dass mentale und nicht-mentale Ereignisse in allen Raum-Zeit-Dimensionen Ursachen haben oder Ursachen sind, dennoch sind kausale Relationen nicht einsehbar. Allein aus der Beobachtung zweier aufeinander folgender Ereignisse ergibt sich keine Notwendigkeit. Als Teil der Struktur der Welt sind kausale Relationen metaphysisch, aber nicht epistemisch begründbar. Das Kausalitätsprinzip, wonach jedes Ereignis eine Ursache hat, wird von Hume lediglich als Regelmäßigkeit zwischen Ursachen und Wirkungen in der Erfahrungswelt aufgefasst (vgl. Kutschera 2003, 73).

Die modernen Varianten der Regularitätstheorie[170] setzen sich im Prinzip mit denselben Fragen wie Hume auseinander. Ein Hauptproblem ist dabei die Unterscheidung zwischen kausalen und zufälligen Regularitäten. In der zeitgenössischen Debatte ist nach wie vor strittig, wie der Begriff der Kausalität zu analysieren ist. Was sind beispielsweise Träger von Kausalbeziehungen, welchen ontologischen Status haben sie? Wie lassen sich genuine Ursachen bestimmen? Welche Rolle spielt die Zeit für die Ursache-Wirkung-Relation (vgl. z. B. Lewis 1986, Mackie 1976, Salmon 1984, Brand 1976)? Unter Zuhilfenahme von Bedingungsanalysen kann ermittelt werden, inwiefern eine Ursache notwendige oder hinreichende Bedingung für die Wirkung ist. Eine übliche Formulierung lautet: „Ein Ereignis q ist Ursache eines Ereignisses p genau dann, wenn es eine kausale Erklärung von p mit q gibt." (Kutschera 1982, 95) Immer, wenn die Ursache auftritt, soll auch die Wirkung eintreten, wobei eine Ursache unter gegebenen Umständen hinreichend für das Eintreten der Wirkung ist. Da also ein Set von Ursachen die Vorbedingung für die Wirkung stellt, muss zudem ermittelt werden, welche Ursachen notwendig und welche hinreichend für die Wirkung sind.[171] Eine Schwierigkeit ist dabei, alle relevanten Bedingungen aufzuzeigen, die als Ursache eines Ereignisses in Frage kommen. Für einen Hausbrand wäre genügend Sauerstoff in der Luft eine notwendige Bedingung; mögliche hinreichende Bedingungen, die den Brand tatsächlich auslösen, sind ein Blitzschlag oder ein Kurzschluss (vgl. Stegmüller 1983, 508).

In der zeitgenössischen Diskussion um mentale Verursachung hat sich gezeigt, dass sämtliche Beiträge einen Kausalitätsbegriff voraussetzen, ohne diesen zu präzisieren. Implizit wird der oben genannte Kausalitätsbegriff verwendet, wonach jeder Ursache eine Wirkung folgt. Auch der Identitätstheorie, funktionalistischen Positionen, dem Anomalen Monismus oder der Superve-

[170] Daneben gibt es natürlich noch alternative Definitionen des Kausalitätsbegriffs beispielsweise aus der Handlungstheorie (vgl. Gasking 1976) oder der Wahrscheinlichkeitstheorie (vgl. Salmon 1984). Mit Verweis auf mikrophysische Prozesse in der Quantenphysik hat sich in der Wissenschaft zudem die Einsicht durchgesetzt, dass Kausalität entgegen dem klassischen Determinismus statistischen Prozessen und Wahrscheinlichkeiten unterliegt (vgl. z. B. Heisenberg 1970; Kutschera 2003, 77).

[171] Ein Beispiel für die Bedingungsanalyse liefert Mackie mit seinem Begriff der INUS-Bedingung: „an *i*nsufficient but *n*onredundant part of an *u*nnecessary but *s*ufficient condition" (vgl. Mackie 1976). Von solchen extensionalen Bedingungsanalysen sind wiederum intensionale zu unterscheiden. Sie legen einen starken Notwendigkeitsbegriff zugrunde, der z.T. auf intensionale Gesetze zurückgeführt wird (vgl. z. B. Lewis 1986).

nienztheorie liegt diese Form der Kausalität zugrunde.[172] Die Standarddefinitionen von Kausalität scheinen nur bedingt im Zusammenhang des Gehirn-Bewusstsein-Problems bzw. bei neuronalen Prozessen anwendbar zu sein. So versagt die Regelmäßigkeitsanalyse, da nach ihr eine Unterscheidung zwischen Kausalbeziehung und Korrelation mentaler und physischer Ereignisse nicht möglich ist:

> [Es] lässt sich nicht unterscheiden zwischen dem Fall, dass ein psychisches Ereignis P, das regelmäßig ein physisches Ereignis $M_p$ begleitet, ein bestimmtes physisches Ereignis $M_q$, das regelmäßig auf $M_p$ folgt, verursacht, und dem Fall, dass P, als Epiphänomen von $M_p$, $M_q$ lediglich regelmäßig vorausgeht, ohne an seiner Verursachung beteiligt zu sein. (Birnbacher 1990, 70)

Kutschera verweist deshalb noch auf eine zweite, eine modallogische Erklärung der Kausalrelation, deren Rahmen alternative Weltgeschichten bilden. Zur tatsächlichen Weltgeschichte *w* in ihrem detaillierten zeitlichen Verlauf wird die Möglichkeit von Alternativen zu einem Zeitpunkt *t* gedacht, die bis zu diesem Zeitpunkt deckungsgleich sind, sich von *t* jedoch grundlegend unterscheiden. Der Zeitpunkt *t* ist dabei ein Verzweigungspunkt, für den die Zukunft dieser Welt nicht eindeutig feststeht. Voraussetzung ist somit, dass es Zufallsereignisse gibt und kein Determinismus herrscht, sondern eine zeitabhängige Notwendigkeit, wonach „in der Welt *w* im Zeitpunkt *t* der Sachverhalt *A* notwendigerweise besteht, wenn er in allen Welten besteht, die mit *w* bis hin zu *t* übereinstimmen" (Kutschera 2003, 72f.). Was aber notwendig ist, hängt von dem Zustand der Welt zu einem gegebenen Zeitpunkt ab und von den alternativen Welten, die mit der Welt bis zu dem bestimmten Zeitpunkt übereinstimmen. Bemerkenswert an der modallogischen Kausalrelation ist im Vergleich zur regularitätstheoretischen Konzeption, dass zum einen Zufallsereignisse oder freie Handlungen zugelassen werden und es zum anderen keine Kausalketten gibt, da Ursachen immer Erstursachen sein müssen (vgl. ebd.).

Die modallogische Kausalrelation als Variante der Ereigniskausalität verbindet Kutschera zudem mit der Variante der Agenskausalität, bei der anstelle einer Ursache ein Urheber ein Ereignis bewirkt. Grob skizziert, geht eine Handlung immer aus einer Situation mit mindestens zwei Alternativen hervor. Im Modell der alternativen Weltgeschichten hat eine Menge von Agenten in jeder möglichen Welt zu jedem Zeitpunkt eine Menge von Alternativen. Durch die Wahl ihrer Alternativen bestimmen alle Agenten zusammen, wie es nach einem bestimmten Zeitpunkt *t* im Detail weitergeht. Die Handlung eines Agenten ist eine echte Alternative, die er realisiert. Entscheidend ist dabei, dass Personen keine Handlungen, sondern mit ihren Handlungen etwas, ein Ereignis, bewirken (vgl. ebd., 74f.). Im Ergebnis ist damit z. B. das Armheben als Handlung ein Ereignis, das keine Ursache hat. Die Suche nach physiologischen Ursachen für diese Handlung ist

[172] Davidson bildet eine Ausnahme, da er sich in verschiedenen Aufsätzen mit kausalen Relationen bzw. der logischen Form singulärer Kausalaussagen auseinandersetzt (vgl. z. B. Davidson 1998, 155-177).

nach Kutschera vergeblich, da lediglich begleitende physiologische Vorgänge – Muskelaktivität, elektrische Nervenimpulse, neuronale Aktivität etc. – gefunden werden können. Aber auch Absichten kommen als Ursachen für Handlung nicht in Frage, da sie alleine nichts bewegen. Im modalen Sinn haben Handlungen keine Ursachen, sondern sind selbst Erstursachen. Es sind keine Ursachen, sondern Gründe, die Handlungen motivieren (vgl. ebd., 79).

Weitere Probleme ergeben sich aus Überlegungen, inwiefern physische Ereignisse notwendige und hinreichende Bedingungen für mentale Ereignisse sind. Welches sind hier die entscheidenden Kriterien, wonach physische Ereignisse mentale realisieren? Welche Rolle spielt etwa die Anzahl oder die Dauer physischer Ereignisse? Henrik Walter bemerkt zu Recht, dass verschiedene Modelle für verschiedene mentale Phänomene, die sich in mehreren Dimensionen unterscheiden, entwickelt werden müssen, anstatt sich auf ein theoretisch umfassendes Konzept zu konzentrieren:

> Warum sollten mentale Zustände oder Prozesse wie Denken, Fühlen oder Entscheiden genau die gleiche Relation zu physischen Zuständen haben? Natürlich wäre es eleganter, mit einer einzigen Theorie alle mentalen Phänomene zugleich zu erfassen. Mir ist allerdings kein einziges Argument bekannt, das überzeugend dafür spricht, dass das so sein muss. Tatsächlich wird es meist gar nicht für nötig befunden, für diese Annahme zu argumentieren! Sie wird einfach vorausgesetzt. (Walter 1999, 166; vgl. auch Kutschera 2003, 74)

Bisher ist die Suche nach einem solchen Konzept gescheitert. Vielleicht bedarf es also auch verschiedener Modelle für den Bereich der mentalen Kausalität. Thomasson verweist in Anlehnung an Aristoteles darauf, dass es verschiedene Relationen gibt, die gleichsam undifferenziert als Kausalrelation bezeichnet werden:

> Different sorts of explanation may appeal to different underlying relations. Thus there may be more than one explanation of an event without these explanations being in competition, and there may be more than one such relation involving the event without its being causally overdetermined. Indeed far from holding that explanations involving different types of cause compete, Aristotle held that a complete explanation of an event is one which appeals to all of these types of cause. (Thomasson 1998, 183)

Von dieser Überlegung ausgehend, werden zwei Erklärungstypen abgeleitet, die auf einer unterschiedlichen Ebene angesiedelt für eine Kausalrelation gemeinsam eine vollständige Erklärung liefern.[173] Die aristotelische Vier-Ursachenlehre kann als Schema dienen, einen neurophilosophi-

[173] Thomasson unterscheidet zwischen Erklärungen der Kausalität und Erklärungen der Determination. Kausale Erklärungen werden den Makro-Leveln zugeordnet, während Makro-Eigenschaften durch die Mikro-Eigenschaften determiniert sind. Beispielsweise ist die Tatsache, dass Plätzchen hart sind, damit zu erklären, dass sie 1. zu lange im Ofen waren und 2. ihre Moleküle eng gebunden sind. Beide Erklärungen widersprechen

schen Begriff für mentale Kausalität zu entwickeln, möglicherweise auf einer rein physikalischen Ebene.

Im Folgenden wird deshalb auf das aristotelische Kausalmodell zurückgegriffen. Es bestimmte lange Zeit in der westlichen Kultur das Bild der Kausalität. In Abgrenzung zu seinen Vorgängern, beispielsweise Anaxagoras, den Atomisten sowie seinem Lehrer Platon, versuchte Aristoteles den Begriff der Verursachung zu präzisieren. Er entwickelte ein Einteilungsschema, wonach sich in der Natur oder in der Kunst eine Wirkung auf vier Ursachen zurückführen lässt. Als solche werden bestimmt: eine materiale Ursache, auf die als passive Substanz die übrigen Ursachen einwirken; daneben die Formursache, welche das Wesen oder die Idee des betreffenden Dinges ausmacht; ferner die Wirkursache, die als bewegende Kraft und äußerer Antrieb auf eine Sache einwirkt; schließlich die Zweckursache, die das Ziel bestimmt, wohin die Dinge streben (vgl. Gigon 1961, 63f.).[174] Seit der Renaissance gibt es nun eine Beschränkung auf die Wirkursache, die causa efficiens, die anderen drei Ursachen sind aus der wissenschaftlichen Forschung verschwunden. Das liegt daran, dass sie begrifflich leichter zu fassen und auch zu formalisieren ist. Zudem gibt es für die Wirkursache empirische Entsprechungen; kausale Vorgänge sind empirisch messbar und in der Erfahrung überprüfbar. Kausalgesetze, wie jene der Newtonschen Mechanik, fordern, dass die Ursachen ihrer Wirkung zeitlich vorausgehen und über Druck- und Stoßbewegungen die nachfolgenden Ereignisse beeinflussen. Da die Wirkursache experimentell erfassbar ist, passt sie zudem als einzige in ein instrumentales Wissenschaftskonzept, das vor allem die Beherrschung der Natur ins Auge fasst (vgl. Bunge 1987, 35f.). Der Rekurs auf Aristoteles' Ursachenlehre soll die Perspektive erweitern und Anknüpfungspunkte für mentale Kausalität prüfen.

### 7.1 Die aristotelische Ursachenlehre

Eine genaue Bestimmung der vier Ursachen fällt insofern schwer, als Aristoteles diese sowohl in seiner Physik (Aristoteles Phy. B3)[175] als auch in der Metaphysik (Aristoteles Met. Δ 2)[176] einführt und verwendet, ohne ihre Herleitung anzugeben. „We do not know how Aristotle arrived at the doctrine of the four causes; where we find the doctrine in him, we find it not argued for but

sich nicht, sondern ergänzen einander. Thomasson würde allerdings lediglich Erklärung 1 als Kausalerklärung anerkennen (vgl. Thomasson 1998, 183f.).

[174] Die Scholastik hat diese vier Ursachen als causa materialis, causa formalis, causa efficiens und causa finalis definiert (vgl. Weiss 1942, 48). Im Folgenden werden diese Bezeichnungen übernommen.

[175] Im Folgenden zitiert nach Aristoteles: Physik. 1. Halbbd., Bücher I-IV, Hamburg 1987; 2. Halbbd. Bücher V-VIII, Hamburg 1988.

[176] Im Folgenden zitiert nach Aristoteles: Metaphysik. 1. Halbbd., Bücher I-VI, 1. Aufl., Hamburg 1978; 2. Halbbd., Bücher VII-XIV, Hamburg 1980.

presented as self-evident." (Ross 1949, 37). Wieland schlägt vor, sich von der Vorstellung zu lösen, es handele sich bei der Ursachenlehre um eine hintergründige Theorie von metaphysischen Grundprinzipien. Das aristotelische Konzept ist vielmehr das Ergebnis einer Analyse des Sprachgebrauchs. Die mehrfache Bedeutung des Begriffs Ursache wird untersucht und in verschiedene Gesichtspunkte eingeteilt:

> Streng genommen handelt es sich hier also nicht um vier Ursachen, sondern um den vierfachen Sinn, in dem wir von Ursache sprechen. Die formale Einheit dieser unterschiedlichen Bedeutungen wird auch hier durch ein funktionales Element hergestellt. (Wieland 1962, 262)

Die vier Ursachen bieten somit ein Schema, anhand dessen über verschiedene Untergruppierungen die Warum-Frage beantwortet werden soll. Auf diese Weise sollen alle Möglichkeiten erfasst werden, nach dem Warum bzw. einer Ursache zu fragen.

### 7.1.1 Hylëmorphismus als ontologischer Hintergrund

Aristoteles' Philosophie zeichnet sich durch das Induktionsprinzip aus. Er hat in Abgrenzung zu Platon Respekt vor den Erscheinungen, sie sind evident, und darüber hinaus ist in ihnen ihr Wesen als Möglichkeit und Anlage enthalten. Zwischen den wahrnehmbaren Erscheinungen und der unsichtbaren Wirklichkeit kann es keinen Widerspruch geben, mit seinen Sinnesorganen ist der Mensch grundsätzlich in der Lage, über die Erscheinungen den inneren Bau der Natur zu erkennen (vgl. Gigon 1961, 56). Dies kann als Hinweis darauf gewertet werden, dass Aristoteles im Gegensatz zu Hume mit einem realistischen Ursachenbegriff arbeitet. Das Kausalitätsprinzip ist nicht bloß ein psychologisches Phänomen, das aufgrund seiner Regelmäßigkeit akzeptiert werden muss, sondern steht fest auf einem metaphysischen Boden. Allerdings muss in diesem Zusammenhang bereits darauf hingewiesen werden, dass sich Aristoteles' Ursachen zunächst auf Dinge beziehen und erst in zweiter Linie auf Prozesse. Die Dinge werden auf ihre Ursachen hin untersucht, nach Wirkungen von Ursachen im neuzeitlichen Sinne wird dagegen nicht gefragt. Der moderne Wirkungsbegriff fehlt der aristotelischen Physik. So ist die Bewegungsursache kein Vorgang eines physischen Druck- und Stoßprozesses, sondern ein Ding, das die Bewegung auslöst, ebenso wie das Resultat dieser Bewegung ein Ding ist (vgl. Wieland 1962, 266).[177]

[177] Ebenso ist beispielsweise das Telos oder die causa finalis nicht mit neuzeitlichen Telosbegriffen zu messen. Das Telos ist nicht eine zielstrebig erreichte Wirkung, sondern selbst Ursache, allerdings derart, dass eine außerhalb von ihr liegende Wirkung nicht erfragt werden kann. Wieland macht zu Recht darauf aufmerksam, dass der aristotelische Ursachebegriff sich in diesen Punkten von dem heutigen wissenschaftlichen unterscheidet, weist aber auch darauf hin, dass das Ursachenschema sich nicht ausschließlich auf Dinge bezieht, sondern gelegentlich auch auf bloße Vorgänge angewendet wird (vgl. Wieland 1962, 266).

Das Induktionsprinzip verweist auf einen weiteren Grundsatz, nämlich die Zielrichtung der aristotelischen Philosophie. Es sollen die Ursachen für die Erscheinungen erforscht werden, die Grundprinzipien des Seins. Die ontologische Dimension der Ursachenlehre darf deshalb in ihrem Ausgangspunkt nicht außer Acht gelassen werden. Dabei überschneidet sich die Frage nach der Arché, dem Ursprung, mit der Frage nach der Aitia, der Ursache. Eine scharfe Trennung beider Begriffe ist nicht möglich. So erklärt Aristoteles verschiedene Arten von Prinzipien (Archai), um anschließend fortzufahren: „In ebensoviel Bedeutungen spricht man von Ursachen, denn alle Ursachen sind Prinzipien." (Aristoteles Met. Δ 5, 1013a 16)[178] Die Natur und die Naturdinge sind Arché (Ursprung) mit dem Charakter einer Aitia (Grund). Dinge, die von Natur aus sind, haben ihren Ursprung und ihren Grund in sich selbst. Beispiele sind eine Pflanze, die wächst, sich entfaltet, blüht und schließlich verwelkt oder ein Tier, das wächst und in seiner Entwicklung von seinen Trieben geleitet wird oder ein Stein, der über die Zeit hin verwittert und seine Gestalt verändert.

In all diesen Lebewesen und Naturdingen ist der Ursprung ihrer Wandlung und Bewegung in ihnen selbst enthalten, sie werden nicht allein von außen her verändert. Die Arché ist ihnen gewissermaßen eingewachsen, sie tragen das Prinzip der Wandlung von Natur aus in sich selbst. „Die Von-Natur-Seienden zeigen sich sämtlich als solche, die in sich selbst den Ausgang haben für das In-Bewegung-Geraten und für das Zum-Stillstand-Kommen." (Aristoteles Phy. B 1, 192b 13-14) An dieser Stelle wird deutlich, dass die Frage nach dem Ursprung der Natur und des Seins auch die Frage nach dem Woher der Bewegung ist, denn Werden ist für Aristoteles mit Bewegung verbunden. Alle seienden Dinge, die uns umgeben, sind in Bewegung. Interessant ist nun, dass einige Dinge die Arché und Aitia für ihre Bewegung von Natur aus in sich selbst haben, während insbesondere künstliche Dinge von äußeren Ursachen bewegt bzw. verändert werden (vgl. Weiss 1942, 45). Offenbar reicht die Arché nicht aus, um zu erklären, warum die Dinge bewegt werden, die Ursprünge zeigen lediglich, woher etwas ist bzw. wird. Es handelt sich bei ihnen gewissermaßen um die „formalen Bedingungen der Möglichkeit des Werdenden als solchen" (ebd., 47). Da sie aber generell vorhanden sind, muss etwas Zusätzliches die Veränderung bewirken. Deshalb wird das Phänomen des Werdens in der Natur auf seine Arché im Sinne der Aitia hinterfragt.

Aber es wird hier noch in einem weiteren Punkt deutlich, dass sich die Bereiche der Arché und Aitia überschneiden. So sind zwei der vier Ursachentypen bereits als Strukturmerkmale der

[178] Der Zusammenhang von Arché und Aitia ist sehr eng, sie werden vielfach gleichbedeutend gebraucht. Eine Unterscheidung kann u. a. getroffen werden, da die Arché im Gegensatz zur Aitia die unmittelbare Bedeutung des Ersten hat, des Anfangs oder des Äußersten. So sind alle Aitia immer schon Arché, aber nicht umgekehrt. Von der Richtung, in der die Frage nach der Arché gestellt wird, hängt ab, inwiefern der Begriff der Aitia zu bestimmen ist (vgl. Weiss 1942, 14). Eine detaillierte Auseinandersetzung mit dieser Problematik kann an dieser Stelle nicht erfolgen. Die Arbeit wird sich im folgenden Verlauf stärker auf das eigentliche Interesse, die Formen der Aitia, konzentrieren.

Arché vorhanden, die Hylë (Materie) und der Eidos bzw. die Morphë (Form).[179] Alles Gewordene ist ein Zusammengesetztes aus Stoff und Form, beide Prinzipien leitet Aristoteles abermals zunächst aus dem Sprachgebrauch und den Denkgewohnheiten ab. Die Menschen reden davon, dass Etwas aus Etwas entsteht, beispielsweise ein Schmetterling aus einer Raupe oder eine Statue aus Erz. In allen Bereichen der Physis zeigt sich das Zusammenspiel der beiden Prinzipien, sie stehen sich als korrelative Begriffe gegenüber. Beide Prinzipien sind in Aristoteles' Stoff-Form-Metaphysik, dem sogenannten Hylëmorphismus, untrennbar miteinander verwoben. Die Materie ist dabei zunächst ungestaltet vorhanden und bildet die Basis alles Seienden. Der Stoff selbst ist nicht fähig, sich aus eigenem Antrieb Form zu geben oder zu verändern, dafür ist er durchgängig auch in den gestalteten Dingen noch vorhanden. Die Form macht nun das Wesen des Stoffes aus, sie verleiht dem Seienden seine Gestalt und Bestimmung. Dabei ist zu beachten, dass die Natur eines Dinges nicht darin liegt, ungestalteter Stoff zu sein, sondern Gestalt und damit auch sein Wesen zu erhalten. Hinsichtlich der Statue ist das Erz zunächst ungestaltet, erst in der Bearbeitung kann es Form annehmen. Alles Werden ist somit Formempfangen, die Form bestimmt den Prozess des Werdens (vgl. Bröcker 1935, 53ff.).

### 7.1.2 Vier Ursachen als einheitliches Phänomen

Mit der Untersuchung der vier Ursachen ergänzt Aristoteles nicht einfach das Form- und das Stoffprinzip um zwei weitere Begriffe. Wie bereits gezeigt wurde, macht diese Vierheit einen eigenständigen Sinn. Zudem erhalten Materie und Form als Gründe eine andere Gewichtung und Bedeutung. Dass die Prinzipien mit den Ursachen verschränkt sind, liegt an der einfachen sprachlichen Tatsache, dass die Frage nach dem Woher mit der Frage nach dem Warum in Beziehung steht (vgl. Wieland 1962, 179). Aristoteles nennt die vier Ursachen sowohl in der Physik als auch in der Metaphysik:

> Indem es nun vier Ursachen sind, so ist es Aufgabe des Natur-Forschers, bezüglich aller sich ein Wissen anzueignen, und wenn er die Rückleitung auf alle (vier) vollzieht, dann wird er die Weshalb-Frage auf naturbezogene Weise beantworten: Stoff, Form, das in Gang Setzende, das Weswegen. (Aristoteles Phy. B 7, 198a 13)

[179] Als drittes Strukturmerkmal wird die privatio genannt, das Prinzip der Beraubung. Es zielt darauf ab, dass der Stoff in seiner Veränderung Form verliert, um eine neue zu erhalten. Form und Beraubung sind somit verschiedene Gesichtspunkte des Stoffs. Werden ergibt sich somit nicht allein aus bestehenden Strukturen, sondern auch aus der Richtung, in die die Veränderung zielt. Aus einem Woher, der Beraubung, entwickelt sich ein Gegenstand zu etwas hin, das er noch nicht war. Hier zeigen sich bereits die Anknüpfungspunkte der aristotelischen Teleologie (vgl. Bröcker 1935, 59).

> [...] man aber von den Ursachen in vierfacher Bedeutung spricht, einmal nämlich als Wesen und als Was-es-ist-dies-zu-sein (denn das Weshalb wird zuletzt auf den Begriff zurückgeführt, Ursache und Prinzip aber sind das erste Weshalb), zweitens als Stoff und Substrat, drittens als das, woher der Anfang der Bewegung kommt, viertens – im Gegensatz zur dritten Bedeutung – als das Weswegen und das Gute (denn dies ist das Ziel aller Entstehung und aller Bewegung) [...]. (Aristoteles Met. A 3, 983a, b)

Was verbirgt sich nun hinter diesen vier Bedeutungen von Ursache, die Aristoteles nennt? Die beiden ersten sind bereits als Archai bekannt, bei ihnen handelt es sich um die so genannten inneren Ursachen, während die beiden letzteren die äußeren Ursachen bilden (vgl. Bröcker 1935, 53):

> 1. causa formalis: Mit Was-es-ist-dies-zu-sein ist die Formursache gemeint. Sie verleiht der rohen Materie Gestalt und bestimmt das Wesen der Dinge. Der Grund des Gutseins eines Nahrungsmittels, z. B. eines Schokoladenriegels, ist das Gutsein überhaupt, das Wesen „Güte". Aber auch dem Schokoladenriegel selbst liegt durch die Formursache bestimmt sein Wesen zugrunde, nämlich ein Schokoladenriegel zu sein (vgl. Aristoteles Met. Δ 2, 1013b 25).[180]
>
> 2. causa materialis: Mit Stoff und Substrat ist die Stoff- oder Materialursache gemeint, derjenige Teil eines Dinges, von dem aus innewohnend etwas entsteht. Als Beispiele werden das Erz als Ursache für die Statue und das Silber als Grund für eine silberne Schale genannt (vgl. ebd.).
>
> 3. causa efficiens: Der Anfang der Bewegung ist die Antriebsursache. Der Grund eines Dinges ist außerhalb zu suchen, es erhält einen Anstoß von etwas anderem. So ist der Vater Urheber des gezeugten Kindes, der Arzt Urheber des geheilten Kranken und der Schokoladenfabrikant Urheber des Schokoladenriegels. Alles Bewirkende ist Ursache des Bewirkten, sowie jede Bewegung Ursache des Bewegten ist (vgl. ebd.).
>
> 4. causa finalis: Das Weswegen und das Gute schließlich beschreibt die Ursache als Ziel oder Zweck. Das Ende und die Richtung des Werdensprozesses sind hier im Blick. Das Ziel wird als Ursache verstanden; es liefert den Grund, um dessentwillen überhaupt etwas geschieht, entsteht oder sich verändert. Ein Spaziergang wird um der Gesundheit willen gemacht, ebenso wird ein Schokoladenriegel produziert, damit man ihn essen kann. Das Gute ist dabei stets mit inbegriffen, da als Ziel das Beste bzw. Gute angestrebt wird, egal ob es sich nun um das tatsächlich oder nur scheinbar Gute handelt (vgl. ders., 1013b 26f.).

In welchem Verhältnis stehen diese vier Ursachen nun zueinander? Inwiefern sind sie als einheitliches Phänomen zu verstehen? Aristoteles macht deutlich, dass die vier Ursachen gemeinsame Gründe für eine Sache sind. „Da aber das Grundsein in mehrfacher Bedeutung gemeint ist, ergibt es sich auch, dass für ein und dasselbe Mehreres Grund ist." (Aristoteles, Phys. B 3, 195a 4f.) Ein Beispiel für das Zusammenwirken zweier Gründe ist die Statue, die ihre Ursache sowohl im Erz

[180] Aristoteles verwendet mitunter zwei Begriffe für die Form, Morphë und Eidos. Während Morphë vor allem auf die Formgestalt eines Dinges abzielt, nimmt Aristoteles mit der Wendung Eidos Bezug auf Platons Ideenbegriff. Insofern steuert die Form das Geschehen und Vorbild und Grund für die Erscheinungen. Die Form ist aber bei Aristoteles immanent im Körper vorhanden und konkret realisiert, und nicht einem Allgemeinbegriff entlehnt (vgl. Aristoteles Met. Δ 2, 1013b 25).

findet, aus dem sie gemacht ist, als auch in der Bildhauerei bzw. der schaffenden Hand des Bildhauers. Hier wirken die causa materialis und die causa efficiens zusammen (vgl. ebd., 195a 6). Ein weiteres Beispiel soll die Beziehung zweier Ursachen zeigen, die untereinander ein gegenseitiger Grund sind. So sind sportliche Aktivität und Fitnesstraining Grund für eine gesunde Verfassung des Körpers. Auf der anderen Seite ist das körperliche Wohlbefinden der Grund für die unternommene Anstrengung. Derselbe Sachverhalt, aus unterschiedlicher Richtung betrachtet, zeigt die Ursächlichkeit von causa efficiens und causa finalis (vgl. ebd., 195a 8-11). Die gegenseitige Verursachung beider Gründe ist nicht zufällig, sondern liegt in ihrem Wesen als Ursache. Die Ursächlichkeit der causa finalis beruht auf der Ursächlichkeit der causa efficiens und umgekehrt:[181]

> Es ist das Wesen des Telos als Aition, dass es selbst Aition wird für das Erste Bewegende, d. h. eben Grund für sein Bewegen [...]. Nur dadurch, dass das Telos den Grund abgibt für das In-Bewegung-Setzen, ist es Grund für das, was durch die Bewegung ins Sein kommt. (Weiss 1942, 59)

Die vier verschiedenen Ursachen sind also auch strukturell miteinander verknüpft. Ein übliches Beispiel, das Zusammenspiel aller vier Ursachen hinsichtlich eines Werdens aufzuzeigen, ist der Bau eines Hauses.[182] Auch hier stehen causa efficiens und causa finalis in ursächlichem Zusammenhang. Denn Aristoteles unterscheidet zu Beginn des Hausbaus zwischen zwei unterschiedlichen Bewegungen, der des Denkens und der des Bewirkens. „Der Teil nämlich, der vom Anfang und der Form ausgeht, heißt Denken, und der Teil, der von dem ausgeht, was den Endpunkt des Denkens darstellt, heißt Bewirkung." (Aristoteles Met. Z 7, 1032b 16) Der Bauherr oder Architekt muss demnach einen Bauplan haben. Da er die Richtung für den weiteren Verlauf bestimmt, muss er ein Bild oder Konzept vor sich haben, dieses wäre das Eidos, das Wesen des Hauses. Hier wirken causa formalis und causa finalis als Gründe für den Bau zusammen, da die Form des Hauses Ziel der Herstellung ist. Von dieser Zielvorstellung ausgehend, wird der Bauherr zurückschließen und überlegen, welche Schritte benötigt werden, um das Haus zu realisieren, bis er zu den Bedingungen gelangt, die er selbst verwirklichen kann. Von hier aus geht die Bewegung wieder in die andere Richtung, nacheinander werden die Bedingungen verwirklicht, um das Ziel zu erreichen. Der Bauherr repräsentiert dabei bereits in der Planungsphase und auch im Handwerk selbst die causa efficiens in Verbindung mit der causa materialis, die auf das Ziel hinwirkt. Denn

[181] Anhand des Zusammenwirkens zwischen causa efficiens und causa finalis zeigt sich im Kleinen Aristoteles' teleologische Weltkonzeption (vgl. Wieland 1962, 260f.). Die Problematik dieser Konzeption soll uns an dieser Stelle nicht interessieren, wir beschränken uns auf die Analyse der vier Ursachen sozusagen auf der Mikroebene.
[182] Das Beispiel bezieht sich zunächst auf die vierfache Ursächlichkeit bezüglich des künstlichen Weltbereichs, der Technë, und wird bei Aristoteles daraufhin auch auf die Natur, die Physis, angewendet. Dabei zeigt er, das sein Ursachebegriff für beide Bereiche gilt (vgl. Wieland 1962, 269). Für diese Arbeit ist interessant, inwiefern sich Aristoteles' Schema auf das Gehirn übertragen lässt.

um letztendlich wirksam werden zu können, benötigt der Bauherr Materialien, z. B. Steine, Mörtel und Beton. Von dem Stoff ausgehend, wird nun das Haus in umgekehrter Richtung zum Ziel hin erbaut (vgl. Aristoteles Met. Z 7, 1032b-1033a).

Das Beispiel des Hausbaus macht deutlich, dass die vier Ursachen nicht nur Grund für die Entstehung eines Dinges sein können, sondern müssen. Anfang und Ziel bedingen sich wechselseitig. Ausgehend von ihnen sind, strukturell miteinander verbunden, alle vier Ursachen am Werden eines Dinges beteiligt (vgl. Wieland 1962, 270). Zudem zeigt das Beispiel, wie sich die Bedeutung von Hylë und Eidos im Werdensprozess gewandelt hat: Sie sind zunächst gleichbedeutend zwei von vier Ursachen. Allerdings lassen sich die vier Ursachen wiederum auf diese beiden bekannten reduzieren; genauer gesagt, causa formalis, causa efficiens und causa finalis fallen im Eidos zusammen. Es wurde bereits gezeigt, dass das Was-es-ist-dies-zu-sein und das Weswegen beim Verursachungsprozess strukturell zusammenwirken, da das Erreichen des Wesens eines Dinges stets Zielrichtung ist. Ebenso fällt der Ausgang der Bewegung unter das Eidos, da es auf seine Verwirklichung hinzielt. „Das aber, was erster Ausgang der Bewegung ist, ist mit jenen beiden dem Eidos nach ein und dasselbe." (Aristoteles Phy. B 7, 198a 24ff.) Wenn diese drei genannten Ursachen in einer aufgehen, stehen sie nur noch der Materialursache, der Hylë, gegenüber. Das Eidos muss nun nicht mehr bloß als ein Vorhandensein gegenüber einem Mangel verstanden werden, sondern in seiner dreifachen Funktion im Werdensprozess.

Die Hylë ist für den Werdensprozess ebenso unentbehrlich, obwohl sie eine untergeordnete Rolle spielt. So ist sie als Ursache zwar notwendig, aber nur, als von vornherein bestimmt ist, was sie notwendig macht. Holz und Steine sind für den Hausbau notwendige Stoffe; dennoch geht aus ihrem Vorhandensein nicht zwingend ein Haus hervor. Aus Holz ließe sich ebenso gut ein Schreibtisch zimmern und die Steine können im Straßenbau verwendet werden. Die Stoffursache ist, wenn man so will, zwar notwendig, aber nicht hinreichend für das Entstehen eines Hauses. Sie ist untergeordnet, da das Eidos in Verbindung mit Telos und Arché der Bewegung die Auswahl der Hylë diktiert und diese erst hinsichtlich jener Ursache wird. Die Hylë „[...] ist Grund in der Weise, dass das Telos auf sie angewiesen ist, sie aufsucht und sie gebraucht, indem es sie umformt, um aus ihr heraus selbst zu werden" (Weiss 1942, 76).

### 7.1.3 Möglichkeiten für mentale Kausalität

Halten wir fest: Bei Aristoteles' Vier-Ursachenlehre handelt es sich um ein Untersuchungskonzept, mit dem durch Einteilung in verschiedene Gesichtspunkte umfassend die möglichen Warum-Fragen hinsichtlich eines Dinges oder Sachverhalts aufgestellt und beantwortet werden sollen. Denn letztendlich beziehen sich alle vier Ursachen bzw. ist in allen vier Ursachentypen der

Ursprung, die Arché erkennbar (vgl. Gigon 1961). Erst die Verschränkung der Woher-Frage mit der Weshalb-Frage, also die Verschränkung von Arché und Aitia, macht die ontologische Dimension der Vier-Ursachenlehre deutlich. Die Phänomene geben uns im Induktionsverfahren Aufschluss über die Beschaffenheit der Welt. Wir müssen also alle Phänomene zunächst ernst nehmen, um wissenschaftlich forschen zu können. Nach Aristoteles erhalten wir über die Erklärung von Etwas Aufschluss über seinen Ursprung. Dahinter verbirgt sich die Frage, inwiefern epistemische Tatsachen ontologischen Tatsachen als Ausgangspunkt von Forschung vorzuziehen sind.

Außerdem ist für die aristotelische Ursachenlehre die Richtung bzw. das Objekt der Ursachenforschung spezifisch. Wie erwähnt, nimmt Aristoteles in erster Linie Dinge in den Blick, er bezieht sich weniger auf Prozesse. Die Richtung der Fragestellung ist dabei interessant, da Aristoteles sich weniger an der Wirkung von Etwas auf etwas Anderes konzentriert, sondern vielmehr rückwirkend die Ursachen von Etwas hinterfragt. Anders formuliert, lautet die Frage nicht: „Was ist die Wirkung Y von Ding X?“, sondern: „Was ist die Ursache X für Y?“ Diese Form der Fragestellung ermöglicht es, einen mehrfachen Ursachebegriff zuzulassen, anstatt die Frage von vornherein auf einen einzelnen Prozess zu beschränken.

Für die Frage der mentalen Kausalität ist von Interesse, welche Merkmale sich für einen Ursachebegriff über diese Voraussetzungen hinaus aus den vier Ursachen ableiten lassen. Da ist zunächst festzuhalten, dass der Ursachebegriff vierfach bestimmbar ist. Aristoteles nennt vier Ursachen und teilt ihnen einen gleichwertigen Spielraum zu. Sie zeigen zudem weitere Kriterien für den Ursachebegriff: Kausalität ist wesenhaft, dynamisch, material und zielgerichtet. Sind dies auch vier mögliche Attribute einer mentalen Kausalität? Jede dieser vier Ursachen ist notwendig für den Werdensprozess eines Dinges und kann aus diesem nicht ausgeschlossen werden. Zugleich muss beachtet werden, dass keine von ihnen allein notwendige und hinreichende Ursache sein kann. Somit stehen die vier Ursachen nicht in einem Konkurrenzverhältnis zueinander, sondern ergänzen sich. Das Prinzip des kausalen Ausschlusses gilt bei Aristoteles nicht für die vier Ursachen. Ferner unterteilt Aristoteles zwischen Naturdingen, die ihren Ursprung und Grund in sich selbst haben und Dingen oder Prozessen, die von außen bewegt werden.[183] Es gibt also eine Art Polarität zwischen inneren und äußeren Ursachen, wobei allerdings offen bleibt,

[183] Es verwundert auf den ersten Blick, warum Naturdinge nur innere Gründe haben können, obwohl doch alle vier Ursachen beim Werdensprozess eine Rolle spielen sollen. Dieser vermeintliche Widerspruch lässt sich über den Eidosbegriff ausräumen. Denn wie gezeigt wurde, erhält dieser durch das Vier-Ursachen-Schema eine erweiterte Bedeutung: Er subsumiert die Ursachen der Form, Bewegung und des Zwecks. Die äußeren Ursachen sind somit Bestandteil der inneren. Bei der Wandlung von der Raupe zum Schmetterling beispielsweise muss die Phase der Verpuppung und der Metamorphose als Bewegung auf ein Telos hin betrachtet werden. Die Unterscheidung zwischen äußeren und inneren Ursachen kommt dagegen viel stärker bei Artefakten zum Tragen. Das Beispiel des Bildhauers zeigt deutlich, dass hier von außen Stoff und Form bearbeitet und auf ein Ziel hin verändert werden.

welche Konsequenzen sich aus dieser Unterteilung ergeben. Sie scheint vielmehr aus dem Hylëmorphismus heraus zu resultieren, wonach Stoff und Form als Ursprungsprinzipien die Rolle innerer Ursachen zufällt. Dabei ist entscheidend, dass die Form das Wesen des Stoffes ausmacht, dieser ist alleine antriebslos. Die äußeren Ursachen causa efficiens und causa finalis machen wiederum auf einen möglichen relationalen Charakter von Kausalereignissen aufmerksam. Gegenstände und Prozesse stehen im Kontext ihrer Umwelt und werden auch von äußeren Bedingungen beeinflusst.

Übertragen auf das Problem der mentalen Kausalität kann damit festgehalten werden, dass innere Ursachen allein womöglich nicht ausreichen, um mentale Eigenschaften hervorzurufen. Eine weitere Konsequenz, die sich aus der aristotelischen Vier-Ursache-Lehre ergibt, ist jedoch die strukturelle Verknüpfung der verschiedenen Ursachen. Sie bedingen sich gegenseitig, sind also aufeinander angewiesen. Weiterhin bleibt die Dichotomie von Hylë und Morphë bestehen, beide Begriffe haben nun eine veränderte Bedeutung. Für die Hylë ist dabei entscheidend, dass sie nicht bloß formloser Stoff ist, sondern erst in ihrer Gerichtetheit auf ein Ziel hin zu einer wirklichen Ursache wird. Holz ist einfach Holz, aber erst als Material für einen Tisch und nicht einen Stuhl auch Ursache für ersteren. Ohne causa finalis also keine causa materialis. Ähnliches gilt für die Bewegungsursache, denn sie ist auf den Stoff angewiesen, um ihr Ziel zu erreichen, und kann erst durch den Stoff selbst tätig werden. Für sie gilt somit: Ohne causa materialis (und causa finalis) keine causa efficiens. Die Zielursache wiederum erfährt erst durch die Formursache ihre nähere Bestimmung. Telos und Morphë arbeiten Hand in Hand, es gilt also: Ohne causa formalis keine causa finalis. Als Beispiel kann hier abermals die Entstehung eines Schmetterlings herhalten, denn seine Form mit Flügeln und Saugrüssel gibt dem Ziel erst den Inhalt bzw. Sinn. Schließlich gilt auch die Umkehrung, dass das Telos die Morphë erst als Ziel setzt, und darüber hinaus die Form erst in ihrer Realisierung durch Stoff- und Bewegungsursache selbst tatsächliche Ursache sein kann. So ließen sich sämtliche Zweifachverbindungen der vier Ursachen durchspielen. Entscheidend ist jedoch, dass sie alle vier in unterschiedlicher Variation stark oder schwach miteinander und übereinander verknüpft sind. Am deutlichsten tritt das anhand des Eidos zutage, der gleich drei Ursachen in sich fasst. Indem Aristoteles causa efficiens, causa finalis und causa formalis in einem Begriff aufgehen lässt, macht er ihre strukturelle Verbundenheit untereinander besonders stark.

Schließlich muss nochmals der prozesshafte Charakter des Kausalitätsmodells betont werden, auch wenn Aristoteles sich auf Dinge und nicht auf Prozesse selbst bezieht. Alles Seiende ist in Bewegung und alles Werden ist Formempfangen. Diese Idee, dass im Rahmen eines Hylëmorphismus der Stoff erst noch eine Form erhalten muss, kann auch auf die Situation der mentalen Kausalität übertragen werden, wo ebenfalls eine Dichotomie zwischen physischen und mentalen oder intrinsischen wie extrinsischen Eigenschaften besteht. Auch die neuronalen Struk-

turen müssen sich bei einem Menschen erst ausbilden und das Gehirn seine individuelle Form erhalten. Hinsichtlich des von Aristoteles beschriebenen kausalen Prozesses selbst lässt sich festhalten, dass es zumindest eine zeitliche Rangordnung gibt. So geht die Denkbewegung mit Zielsetzung und Formbestimmung dem Schaffensprozess voraus, dieser erfolgt als Gegenbewegung der ersteren nach. Kausale Ereignisse kann man sich unter diesem Gesichtspunkt als Rückkopplungsprozess vorstellen, bei dem das Ziel den Grad der Aktion vorgibt.

In Tabelle 3 werden die physischen und mentalen Eigenschaften den vier Ursachen gegenübergestellt, um zu ermitteln, ob sich erstere und letztere einander zuordnen lassen. Inwiefern ist das Vier-Ursachen-Schema überhaupt auf mentale Kausalität übertragbar? Die Tabelle zeigt auf der vertikalen Achse die mentalen und physischen Eigenschaften und auf der horizontalen Achse die vier aristotelischen Ursachen, letztere in ihrer Aufteilung als Hylë und Eidos sowie innere und äußere Ursachen. Das „X" steht jeweils für die Zuordnung von Ursache und Eigenschaft. Eine mögliche Interpretation der Tabelle wäre demnach, dass 1.) mentale Eigenschaften und äußere Ursachen sowie physische Eigenschaften und innere Ursachen zusammenfallen, 2.) physische Eigenschaften in den Bereich der Hylë, mentale Eigenschaften aber in den Bereich des Eidos fallen, 3.) causa formalis und causa efficiens eine Art Schnittmenge zwischen mentalen und physischen Eigenschaften bilden,[184] 4.) causa materialis mit den physischen und causa finalis mit den mentalen Eigenschaften kompatibel sind.

| | **Hylë** | **Eidos** | | |
|---|---|---|---|---|
| | innere Ursachen | | äußere Ursachen | |
| | **causa materialis** | **causa formalis** | **causa efficiens** | **causa finalis** |
| **physische Eigenschaften (intrinsisch)** | X | X | X | |
| **mentale Eigenschaften (extrinsisch)** | | X | X | X |

Tabelle 3: Es erfolgt eine Zuordnung von Ursachen und Eigenschaften. Nach der in der analytischen Diskussion üblichen Dichotomie wäre ein mögliches Ergebnis aus der Tabelle z. B. die Zuordnung der extrinsischen mentalen Eigenschaften mit weitem Inhalt zu den äußeren Ursachen sowie die Zuordnung der intrinsischen physischen Eigenschaften mit engem Inhalt zu den inneren Ursachen. Ferner scheint die causa materialis lediglich mit den physischen und die causa finalis lediglich mit den mentalen Eigenschaften kompatibel zu sein.

[184] Hier fällt eine eindeutige Zuordnung schwer. Es ist möglich zu argumentieren, dass die mentale Eigenschaft „Schmerz" erst das Wesen der C-Fasern ausmacht; oder aber „Schmerz" gehört nicht wesenhaft zu den C-Fasern dazu, sie sind für sich causa materialis und finalis. Hinsichtlich der causa efficiens lautet die entscheidende Frage, ob mentale Eigenschaften äußere Wirkursache sein können oder ob causa efficiens nur als physische Kausalität möglich ist.

Nach dem aristotelischen Modell müsste demnach – unabhängig von der kausalen Kraft mentaler Eigenschaften – die kausale Kraft der physischen Eigenschaften als causa efficiens und causa materialis gewertet werden, während den mentalen Eigenschaften eine spezifisch finale Kausalität zufiele, evtl. ergänzt durch die causa formalis. Allerdings muss der hypothetische Charakter eines solchen Modells betont werden, da die Ergebnisse stark von der Interpretation der Tabelle abhängen. Ironischerweise verweist der moderne Kausalitätsbegriff hauptsächlich auf die causa efficiens, die causa materialis wird vom Physikalismus als materialistische Position allenfalls implizit als Ursache dargestellt. Aus dem Schema ist jedoch ersichtlich, dass für die kausale Kraft physischer Eigenschaften neben der causa efficiens auch die causa materialis und die causa formalis relevant sind. Schließlich implizieren physische Eigenschaften einen Materiebegriff, der aber bei Aristoteles nicht von dem Eidos zu trennen ist – alles Werden ist Formempfangen! Dieser einheitliche Gedanke ist eine gute Ergänzung zu dem Vergleich zwischen den aufgestellten Kriterien für mentale Kausalität und den empirischen Paradigmen; es wurde gezeigt, dass die herkömmliche Dichotomie zwischen mentalen und physischen Eigenschaften in verschiedener Hinsicht aufgehoben werden muss (vgl. Kap. 6.4; Kap. 6.4.2; Kap. 6.5; Kap. 6.6). Auf der Basis von causa efficiens und causa materialis lässt sich allenfalls die physisch-physische Kausalität gut erklären. Bei den Relationen der mental-mentalen, physisch-mentalen sowie mental-physischen Kausalität wird aber bereits deutlich, dass nach aristotelischem Schema die Relation durch die causa formalis unbedingt ergänzt werden muss. Mentale Eigenschaften machen einen wesentlichen Teil eines Ereignisses oder Zustands aus. Die physischen Eigenschaften müssen also eine spezifische „Form“ haben, um beispielsweise tokenidentisch zu sein oder als Supervenienzrelation andere mentale Eigenschaften zu verursachen.

Sicherlich ist es möglich, mentale Kausalität allein aufgrund der drei aristotelischen Ursachen causa materialis, causa formalis und causa efficiens zu erklären. Voraussetzung wäre dabei weiterhin, die mentalen Eigenschaften auf physische zurückzuführen. Allerdings ginge dabei der spezifische Charakter mentaler Ereignisse als normativ, repräsentational bzw. intentional verloren. Die vierte aristotelische Ursache, die causa finalis, scheint dagegen diesem Aspekt mentaler Ereignisse Rechnung zu tragen. Im Folgenden soll deshalb untersucht werden, welchen Beitrag das Konzept einer finalen Kausalität hinsichtlich der Frage der mentalen Kausalität leisten kann.

### 7.2 Das Modell der finalen Kausalität

Inwiefern kann nun eine causa finalis den extrinsischen Charakter mentaler Eigenschaften und ihre Umweltbezogenheit besser konzeptuell aufgreifen als die causa efficiens? Aristoteles beschreibt die causa finalis als ein Streben eines Dinges oder Zustands hin zu seiner Verwirklichung. Die Ursache des Ziels stellt die Frage, weswegen etwas geschieht: Ziel eines Spaziergangs ist die Gesundheit, Ziel des Bauens ein Haus usw.: „[A]lles dies ist ja um des Ziels willen da, der Unterschied untereinander besteht nur darin, dass es sich teils um Tätigkeiten, teils um Werkzeuge handelt." (Aristoteles Phy. B 3, 195a.) Gerade das Beispiel des Hausbaus hat gezeigt, dass das Ziel die Richtung angibt, wohin die anderen Ursachen steuern. Nach der Planung des Hauses wirken sich Bewegungsursache und Materialursache auf dessen Verwirklichung hin (s. o.). Die Kriterien, die aus dem aristotelischen Ursachebegriff entwickelt worden sind, haben vor allem die strukturelle Verknüpfung der vier Ursachen miteinander sowie deren Prozesscharakter, der sich aus einer zeitlichen Rangordnung ergibt, betont:

> Das Ziel wird vom Denken gesetzt, das nunmehr auf die Bedingungen zurückschließt, bis es zu einer Bedingung gelangt, die der Herstellende selbst verwirklichen kann. [...] Von hier aus geht die Bewegung wieder in die andere Richtung; die Bedingungen werden nacheinander verwirklicht, bis das Ziel erreicht ist. Die beiden Bewegungen des Denkens (νοειν) und des Tuns (ποειν) sind also zueinander gegenläufig. (Wieland 1962, 270)

Causa formalis und causa finalis bilden zusammen das zu erreichende Ziel, die causa materialis liefert der causa efficiens den Stoff, der auf das Ziel hin verwirklicht werden soll, und die causa efficiens schließlich bildet den Gegenpol zur causa finalis, zu der hin die Bewegung erfolgt. Übertragen auf mentale Kausalität, muss also auch hier das Ziel der Realisierung vorangehen, woraufhin sich die Bedingungen zum Ziel hin entwickeln. „Um von Notwendigkeit in der Natur sachgerecht sprechen zu können, muss man den Zweck selbst voraussetzen und dann fragen, was für ihn notwendig ist." (ebd., 264)

Die strukturelle Verknüpfung, der Prozesscharakter sowie die Verschränkung von Ziel- und Bewegungsursache sind dabei wesentliche Merkmale einer finalen Kausalität für mentale Phänomene. Der erweiterte Kausalitätsbegriff der finalen Kausalität liefert andere Erklärungsmöglichkeiten als die „lineare Kausalrelation" von Ursache und Wirkung. Dennoch wird das Prinzip der causa efficiens in das Modell miteinbezogen und der causa finalis gegenübergestellt. Für mentale Eigenschaften soll damit deutlich werden, dass sie nach wie vor eine physische Basis haben und durch physische Eigenschaften realisiert werden. Diese sind notwendig, jedoch nicht hinreichend für mentale Eigenschaften.

Mit welcher Notwendigkeit also realisieren physische Eigenschaften mentale Eigenschaften? Ein übliches Beispiel für die Realisierungsrelation sind funktionale Eigenschaften, die als Eigenschaften mit einer kausalen Rolle innerhalb eines Systems von zugrunde liegenden physischen Eigenschaften implementiert werden. Mentale Eigenschaften werden indirekt über die Instanziierung physischer Eigenschaften instanziiert (vgl. Kap. 2.2; Raymont 2003, 238). Levine zieht die Realisierungsrelation einer nomologischen bzw. Kausalrelation zwischen mentalen und physischen Eigenschaften vor. Im ersten Fall bleibt eine signifikante ontologische Unabhängigkeit zwischen der Instanziierung beider Eigenschaften bestehen. Das Auftreten der Ursache führt nicht zwangsläufig zum Auftreten einer Wirkung, das Bewirken wird vielmehr als zusätzliches Merkmal über die Instanziierung hinaus verstanden. Realisierung beschreibt dagegen eine wesentlich engere Beziehung. Die Instanziierung der einen Eigenschaft wird durch die Instanziierung der anderen erreicht. Mit der Instanziierung der realisierenden Eigenschaft wird zugleich die realisierte Eigenschaft instanziiert. Beide Eigenschaften sind damit notwendig viel stärker miteinander verbunden als in einer Kausalrelation (Levine 2001, 14).

Auch in der Supervenienztheorie wird die Relation zwischen mentalen und physischen Eigenschaften als eine Realisierung beschrieben (vgl. Kap. 2.4.1; Kap. 2.4.2). Es handelt sich dabei um eine asymmetrische Beziehung, da die mentalen Eigenschaften von den physischen Eigenschaften abhängen (vgl. Kim 1993a, 354). Dennoch wird auch hier die enge, gleichzeitige Instanziierung beider Eigenschaften betont, so dass beide sich aufeinander beziehen und die Relation eben als Realisierungs- und nicht als Kausalrelation aufgefasst werden sollte (vgl. ebd.; Raymont 2003, 226).

Die Relation zwischen den einzelnen funktionalen Zuständen und dem System, in dem diese operieren, unterscheidet Shoemaker weiter in eine „Kernrealisierung“ und eine „Gesamtrealisierung“ (vgl. Shoemaker 1981, zitiert nach: Levine 2001, 34). Die Kernrealisierung entspricht prinzipiell dem Realisierungsmodell, wonach eine niedrigstufige Eigenschaft eine höherstufige Eigenschaft realisiert. Beispielsweise ist für Menschen der Gehirnzustand B notwendig und hinreichend für die höherstufige Eigenschaft „Schmerz haben“. Die Gesamtrealisierung erweitert den Rahmen und bezieht sich auf ein externalistisches Szenario (vgl. Kap. 6.3.2). Es wird für ein Subjekt bzw. sein Gehirn eine innere symbolische Struktur S, realisiert in einer neurologischen Struktur, angenommen sowie eine „belief box“ B. Das Subjekt befindet sich in einer Glaubensrelation, der Kernrealisierung, wenn S sich in B befindet. Vor dem Hintergrund externalistischer Annahmen ist der Zustand S in der belief box B jedoch nicht hinreichend für z. B. die Annahme „Wasser ist nass“. Hierfür ist zusätzlich notwendig, dass sich das Subjekt im entsprechenden Kontext befindet: „To get that belief, we must also ensure that the subject is embedded in the right context – which, presumably, includes being in a world with water in it.“ (Levine 2001, 34) Die Gesamtrealisierung umfasst beides, die Kernrelation, wonach S in B ist, sowie zusätzliche

physische Konditionen, die gewährleisten, dass sich S in einem passenden Kontext ereignet. Wenn davon gesprochen wird, dass Glaubensannahmen als Gehirnzustände realisiert werden, sind damit üblicherweise Kernrealisierungen gemeint. Genau genommen sind es aber nur die Gesamtrealisierungen, die vollständig für jene Glaubensannahmen hinreichend sind.

Sowohl die Gesamt- als auch die Kernrealisierung sind für die Frage, inwiefern mentale Zustände menschliches Verhalten verursachen, von Bedeutung. Schließlich geben sie Auskunft darüber, wie Glaubensannahmen, Wünsche oder Intentionen realisiert werden. Der Glaube, wonach das Trinken von Wasser Durst löscht, umfasst mehr als einen bestimmten Gehirnzustand. Die Gesamtrealisierung dieser Glaubensannahme umfasst einen bestimmten Gehirnzustand, der sich wiederum in nomologischer Relation mit weiteren internalen Zuständen, mit Wasser innerhalb einer globalen Umwelt sowie letztlich einem konkreten Glas Wasser befindet. Levine betont allerdings, dass die Gesamtrealisierung die Kernrealisierung einschließt und der Kernrealisierer als zentrale Eigenschaft in seiner Relation zu den übrigen Eigenschaften bzw. Zuständen entscheidend für die Realisierung ist – wenn auch unter den Konditionen der Gesamtrealisierung (vgl. ebd., 37).

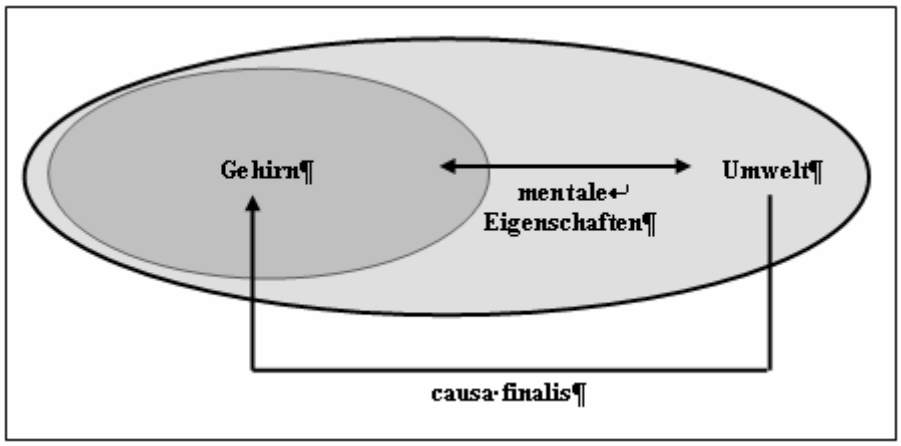
Der Vergleich zwischen den Kriterien für mentale Kausalität und den neurowissenschaftlichen Paradigmen verweist auf einer empirischen Ebene ganz konkret auf die Interaktion zwischen subjektivem phänomenalen Erleben, neuronaler Aktivität und Umweltereignissen außerhalb des Gehirns bzw. der Probanden. Dies ging aus der Diskussion der neurowissenschaftlichen Paradigmen u. a. mit den Kriterien Inhalt (vgl. Kap. 6.3), Intentionalität (vgl. Kap. 6.5) oder Adäquatheit (vgl. Kap. 6.6) hervor. Neben der Kausalrelation zwischen physischen und mentalen Eigenschaften muss also auch die Relation zwischen den Eigenschaften und dem Umweltereignis berücksichtigt werden. Weder ein Umweltereignis noch eine physische Eigenschaft reicht allein aus, um mentale Zustände nicht nur notwendig, sondern auch hinreichend zu verursachen. Erst die Konjunktion von Umweltereignis sowie physischer und mentaler Eigenschaft reichen als hinreichende Bedingung für die Verursachung mentaler Zustände aus. Finale Kausalität soll diese zweifache Relation beschreiben:

1. Mentale Eigenschaften setzen physischen Eigenschaften das Ziel der Instanziierung.
2. Physische Eigenschaften realisieren mentale Eigenschaften, um das Ziel der Instanziierung zu erreichen.

Der entscheidende Unterschied zu den oben genannten Positionen besteht im Modell der finalen Kausalität darin, dass sich mentale und physische Eigenschaften wechselseitig realisieren, was sich in der üblichen Realisierungsrelation über die Tatsache der gleichzeitigen Instanziierung bei-

der Eigenschaften andeutet. Zudem hat Levine betont, dass dadurch beide Eigenschaften stärker miteinander verbunden sind als in einer nomologischen Relation (vgl. Levine 2001, 14; s. o.).

Es bedarf zusätzlich einer Zielursache, um mentale Prozesse auszulösen. Anders formuliert: Durch Reize aus der Umwelt wird neuronale Aktivität ausgelöst, die Anlass erhalten, bestimmte Gedanken oder Wünsche zu realisieren. Umwelt und physische Eigenschaften sind strukturell miteinander über die mentalen Eigenschaften verknüpft. Ohne die subjektiv mentale Ebene könnte keine Interaktion zwischen Gehirn und Umwelt stattfinden, zumindest nicht so reich und komplex, wie dies geschieht. Als Ziel der Realisierung soll Interaktion mit Umwelt ermöglicht werden. Voraussetzung finaler Kausalität ist, dass Wünsche, Absichten, propositionale Einstellungen etc. als Kennzeichen des Mentalen sich nicht aus dem Nichts entwickeln, sondern in der Beziehung zwischen Subjekt und Umwelt ihren Ursprung haben. Die erste der beiden oben genannten Prämissen bringt diese Relation zum Ausdruck. Die Absicht, den Experimentalfilm „Inland Empire" zu sehen, kann sich erst entwickeln, nachdem ein Plakat des Films oder eine TV-Werbung, also ein Umweltereignis, darauf aufmerksam gemacht hat. Ebenso veranlasst die schwärmerische Buchbeschreibung eines Freundes erst zur Entscheidung, dieses Buch zu kaufen. Die zweite Prämisse beschreibt die in der philosophischen Diskussion übliche Realisierung mentaler Eigenschaften, allerdings erweitert und zielorientiert. Insofern erfüllt das Modell der mentalen Kausalität als finale Kausalität verschiedene Kriterien des aristotelischen Ursachebegriffs. Zum einen sind causa finalis und causa efficiens strukturell miteinander verknüpft, zum anderen geht die causa finalis der causa efficiens zeitlich voran (vgl. Kap. 7.1.2; Kap. 7.1.3).




Grafik 19: In der Grafik wird vereinfacht das Modell der finalen Kausalität dargestellt. Das Gehirn ist durch die Umwelt prädisponiert. Auf der einen Seite erkennt es Umweltreize, auf der anderen Seite muss es auf unbekannte Umweltreize reagieren. Ziel ist es dabei, den Informationsaustausch zwischen Umwelt und Person/Gehirn aufrechtzuerhalten. Insofern gibt die Umwelt den Kontext vor, auf den das Gehirn „final" reagiert. Die mentalen Eigenschaften werden von physischen Eigenschaften und Umweltereignis verursacht.

Auch wenn Umwelteinflüsse relevant für mentale Eigenschaften sind, darf das Prinzip der finalen Kausalität nicht als direktes Reiz-Reaktions-Schema verstanden werden. Vielmehr sind die Umweltreize als Auslöser und Ursache einer Kette physischer bzw. neuronaler Kausalprozesse zu

verstehen, die zeitlich versetzt und unabhängig erfolgen können. Weder Kinobesuch noch Buchkauf erfolgen direkt, der Entschluss kann geraume Zeit später getroffen werden. Mit seiner Unterscheidung zwischen „triggering causes“ und „structuring causes“ macht auch Dretske (1999) darauf aufmerksam, dass zumindest einige Ereignisse zwei zeitlich versetzten Ursachen unterliegen. Während ein triggering cause ein direkter Auslöser – im Sinne einer Wirkursache – für ein Ereignis ist, liegt der structuring cause zeitlich weiter zurück.[185] Letzterer liefert aber eine dauerhafte Vorbedingung, die durch ein späteres Ereignis ausgelöst wird. Übertragen auf finale mentale Kausalität, wären die structuring causes bereits gelernte Informationen über die Welt, die die Menschen in Form von propositionalen Einstellungen oder anderen mentalen Zuständen in sich tragen. Diese werden durch die triggering causes, beispielsweise ein Kinoplakat, aktiviert. Eine Entscheidung unterliegt zudem einer Reihe innerer physischer Abwägungsprozesse gemäß der causa efficiens, die womöglich teilweise unbewusst verlaufen. Die philosophischen Modelle zur mentalen Kausalität, wie die Supervenienzrelation, bleiben innerhalb des Modells einer finalen Kausalität nur eingeschränkt bestehen. Sie werden vielmehr modifiziert. Physische Eigenschaften realisieren zwar mentale Eigenschaften, sind aber von letzteren abhängig. Denn Aufgrund der Zielrichtung mentaler Eigenschaften werden die physischen organisiert. Für die finale mentale Kausalität ist aber wichtig, dass der Austausch zwischen Person und Umwelt sich über mentale Eigenschaften vollzieht.[186] Stegmüller verweist im Zusammenhang mit der Frage der notwendigen und hinreichenden Bedingungen darauf, dass menschliche Handlung zielgerichtet erfolgt.

> Erscheint ein Ereignis E als erstrebenswertes Ziel, welches den Effekt einer menschlichen Tätigkeit bilden soll, so wird das Wort „Ursache“ approximativ im Sinn einer hinreichenden Bedingung von E zu verstehen sein, da der Handelnde alle nicht bereits „durch die Natur realisierten“ notwendigen Bedingungen von E in ihrer Totalität verwirklichen muss, um eine Situation zu erzeugen, die E zur Folge hat. (Stegmüller 1983, 508)

Das Kausalitätsmodell, das die Bedingungen menschlichen Handelns beschreiben soll, kann auch für die Bedingungen mentaler Kausalität gelten. Einen Unterschied gibt es allerdings, da nicht ein

[185] Grundsätzlich besteht eine Nähe des Modells der finalen Kausalität zu informationstheoretischen (vgl. Dretske 1981) oder teleologischen Ansätzen. Allerdings ist deren grundsätzliche Idee, dass es Wunsch- oder Glaubensannahmen-produzierende Systeme mit bestimmten biologischen Funktionen gibt. Die Produktion jener Wünsche oder Annahmen hat eine bestimmte Wirkung zum Ziel, die einen evolutionären Vorteil sichern soll (vgl. Milikan 1984; Papineau 1993). Allerdings liegt hierin ein wesentlicher Unterschied zum Modell der finalen Kausalität, da bei diesem nicht zwingend ein evolutionstheoretischer Gedanke zugrunde liegt. Mentale Ereignisse haben viele verschiedene mögliche Wirkungen. Eine feste Bestimmung eines mentalen Inhalts auf eine einzige erscheint somit problematisch. Insofern ist es auch fragwürdig, ob die Evolution konkrete mentale Inhalte aufgrund ihrer Effekte hin selektiert. Das Modell der finalen Kausalität betont dagegen eine höhere Flexibilität für aktuelle Prozessierung mentaler Ereignisse.

[186] Natürlich ist auch reziproke Interaktion zwischen Person und Umwelt auf rein physischer Basis möglich, beispielsweise werden visuelle Reize über die Augen und akustische Signale über die Ohren an das zentrale Nervensystem übermittelt. Dies sind aber Beispiele einer physisch-physischen Kausalität im Modell der finalen Kausalität. Hier geht es aber vor allem um mental-physische Kausalität innerhalb der mentalen Eigenschaften, die auch eine physische Basis haben, als mentale Eigenschaften erkannt und verarbeitet werden.

Handelnder die bereits naturgegebenen notwendigen Bedingungen durch eine hinreichende ergänzt, sondern die im Gehirn vorhandenen physischen Eigenschaften als notwendige Bedingungen durch eine weitere Eigenschaft ausgelöst werden, um der Zielsituation zu entsprechen. Diese Eigenschaft muss relational sein, da sie den Bezug der Person zur Umwelt sichert; also kann es sich nur um eine mentale Eigenschaft handeln. Das Modell der Multirealisierung bzw. Supervenienz mentaler Eigenschaften wird insofern umgekehrt, als nicht allein die physischen Eigenschaften untereinander kausale Wirkung entfalten, sondern zuvor die mentalen Eigenschaften den Kontext angeben, der zwischen Person und Umwelt besteht. Zielgerichtet werden die Neuronenensembles aktiviert, die notwendig sind, eine Kinokarte zu kaufen und nicht eine Tasse Kaffee zu kochen. „[W]e can say that mind is a psychophysical process that encodes information, developing over time." (Velmans 2000, 250)[187] Das Kinoplakat, die TV-Werbung oder die Buchvorstellung werden über mentale Eigenschaften aus der Umwelt als Informationen auf dem Makrolevel an die Person weitergegeben und wirken somit auf die bereits vorhandenen physischen bzw. mentalen Eigenschaften der Person ein. Das Modell einer finalen Kausalität zeigt also, wo mental-physische Kausalität möglich ist: In der Schnittstelle zwischen Umwelt und Person.

### 7.2.1 Finale Kausalität als Ereignisontologie

Während die strukturelle Verknüpfung zwischen mentalen und physischen Eigenschaften, Gehirnereignissen und Umweltereignissen die kausale Relation auf einer Detailebene beschreibt, soll das Zustandekommen einer solchen Relation im Modell der finalen Kausalität als Instanziierung eines „Gesamtereignisses" bzw. „mentalen Ereignisses" beschrieben und damit auf eine jene Detailebene umfassende höhere Ebene geführt werden. Dabei wird die epistemische wie ontologische Bestimmung als „mentales Ereignis" gegenüber einer Bestimmung als „mentaler Zustand" favorisiert, da das Gehirn sich als biologisches Organ durch seinen prozesshaften, dynamischen Charakter auszeichnet (vgl. Northoff 2004, 175 ff.), was auch innerhalb der vorliegenden neurowissenschaftlichen Studien aufgrund der spezifischen Funktionen wie der Prozessierung oder Modulation der emotionalen Dimensionen oder der emotionalen Aufmerksamkeit deutlich wurde (vgl. Bermpohl et al. 2006; Grimm et al. 2006). Die dabei gemessenen Werte sind ein jeweiliger Ausschnitt aus einem Prozess, dem das Gehirn unterliegt. Insofern kann auch eine Zustandsbestimmung des Gehirns zu einem ausgewählten Zeitpunkt t erfolgen, um die Extreme eines ausgewählten Prozesses zu bestimmen. Im Grunde wird aber die Bestimmung von Gehirnereig-

[187] Als Vertreter des so genannten „dynamic approach" unterscheidet Velmans „mind" weiter durch „conscious experience" und „brain". Während die Gehirnaktivität eine beständige Begleiterscheinung ist, wechseln bewusstes und unbewusstes Erleben einander ab. Bewusstes Erleben ermöglicht fokussiertes und genaueres Reagieren auf Umweltereignisse (vgl. Velmans 2000, 255).

nissen dem Organ eher gerecht, da für die Prozessierung eines Umweltgeschehens verschiedene Areale und Netzwerke im Gehirn in unterschiedlicher zeitlicher Folge notwendig sind – auch wenn diese Zeiträume zugegebenermaßen mitunter äußerst kurz sind und den Bruchteil einer Sekunde einnehmen. Genauso wie dem Umweltgeschehen ein kontinuierlicher zeitlicher Fluss unterliegt, der verschiedene längere oder kürzere Ereignisse umfasst, lassen sich auch die Gehirnereignisse zeitlich bestimmen.

Dabei wurde bereits auf Schwierigkeiten bei der Bestimmung des Ereignisbe-griffs durch Davidson oder Kim hingewiesen: Einmal wird das Ereignis als konkretes Partikular an einem speziellen raum-zeitlichen Ort aufgefasst, dann wieder als Instanziierung einer Eigenschaft zu einer bestimmten Zeit (vgl. Kap. 3.1.1). Epistemisch muss ein gewisses Einverständnis über die Eingrenzung eines Ereignisses bei den verschiedenen Gesprächspartnern o. ä. herrschen (vgl. Davidson 1986a, 284; 1990; vgl. Kap. 2.3.1). Die Identität der Ereignisse lässt sich ferner aufgrund ihrer „kausalen Rolle" im kausalen Netzwerk der Ereignisse bestimmen. Kennzeichen der Ereignisse ist, dass sie Ursache und Wirkung haben und auch selbst sind (vgl. Davidson 1998, 29ff.). Als Beispiele für eben solche unwiederholbare, datierte Ereignisse nennt Davidson z. B. „Der Tod von Walter Scott" = „Der Tod des Verfassers von Waverly", „Die Ermordung des Erzherzogs Ferdinand" = „Das Ereignis, das den Ersten Weltkrieg auslöste", „Der Ausbruch des Vesuvs im Jahre 79 n. Chr." = „Die Ursache der Zerstörung von Pompeji" (vgl. Davidson 1970, 294; vgl. Kap. 2.3.1). Die gewählten Beispiele formulieren für Davidson bereits Identitätsbehauptungen. Solche Ereignisse kennzeichnen auch im Rahmen des Modells der finalen Kausalität die Umweltereignisse. Auch wenn sie zeitlich eingegrenzt oder relativ genau bestimmt werden können, wird anhand der vorliegenden Beispiele deutlich, dass die Ereignisse unterschiedliche zeitliche Dimensionen umfassen.[188] Es ist wichtig, nochmals zu betonen, dass die Bestimmung der Ereignisse nach Davidson eine epistemische ist, während die Ereignisse auf der ontologischen Ebene beschreibungsunabhängig bleiben (vgl. 1998; Kap. 2.3; Kap. 3.1.1).

Kim betont den zeitlichen Aspekt der Konstituierung von Ereignissen (vgl. Kim 1976, 160f.; Kap. 3.1.1). Sie haben eine Dauer innerhalb einer bestimmten Zeitspanne. Ereignisse vollziehen sich an einem Träger, instanziieren aber in jener konkreten Zeitspanne selbst Eigenschaften. Aufgrund dieses letzten konstitutiven Elements gilt Kims Ereignisbegriff als sehr „feinkörnig".[189] Denn scheinbar wird durch jede Beschreibung, die neue Umstände berücksichtigt, das

[188] Die zeitliche Dimension der Umweltereignisse unterscheidet sich insofern von der zeitlichen Dimension der Gehirnereignisse, als erstere als aktuelles Ereignis einen spezifischen zeitlichen Prozess umfassen, danach aber als historische Ereignisse lediglich einen konkreten Referenzpunkt angeben, während letztere stets einen aktuellen Prozess umfassen, wenn auch in unterschiedlichen zeitlichen Dimensionen. So ist das Gehirnereignis einer Person, die den Tod von Walter Scott oder die Ermordung des Erzherzogs Ferdinand miterlebt, trivialerweise ein anderes, als das Gehirnereignis einer Person, die diese Ereignisse aus der historischen Distanz nachvollzieht.

[189] Beispielsweise favorisiert Raymont (2003) eher einen „grobkörnigen" Ereignisbegriff à la Davidson. Ihm zufolge haben nur Ereignisse kausale Kraft; Eigenschaften sind lediglich relevant in Bezug auf ihre Träger, die Ereignisse. Mentale wie physische Eigenschaften können kausale Erklärungen liefern, aber selbst nicht verursa-

Ereignis ein anderes, auch wenn es sich zur selben Zeit am selben Träger vollzieht. Dies ist tatsächlich ein Nachteil dieses Ereignisbegriffs und man kann fragen, ob nicht zwischen verschiedenen Beschreibungen, wonach am Computer zur Zeit t „getippt", „ein Text getippt" oder „ein Text sehr schnell getippt" wird, sowie dem eigentlichen Ereignis auf der ontologischen Ebene unterschieden werden muss. Andererseits betont dieser Ereignisbegriff selbst minimale Veränderung. Dieser Aspekt scheint insbesondere hinsichtlich einer mentalen Kausalität für das Gehirn vorteilhaft, da auch hierbei auf der neuronalen Ebene sehr feinkörnig unterschieden werden kann. Es ist eben die Frage, ob dadurch, dass dem subjektiven Erleben eines bestimmten emotionalen Bildes bei Person A teilweise andere Neuronen aktiv sind als bei Person B, auch direkt ein anderes Ereignis konstituiert wird. Beide Personen A und B können sich auf ein subjektives Erlebnis in der Beschreibung einigen, ohne die Unterschiede ihres jeweiligen Erlebens zu berücksichtigen. Ein solches Ereignis, bei dem die Konstituenten variabel sind, gilt üblicherweise als multirealisierbar.

Eine Identitätsbestimmung kann erfolgen, weil Ereignisse Elemente von Kausalbeziehungen sind, und in unterschiedlichen Abläufen unterschiedliche Ursachen und Wirkungen haben können (vgl. Kim 1976, 169f.; vgl. Kap. 3.1.1). Hierbei ist es möglich, dass wir aufgrund epistemischer Einschränkungen die detaillierten Veränderungen auf einer Mikroebene nicht einsehen können und somit zu falschen Schlussfolgerungen gelangen. Solange jedoch das fMRT von Person A ein auch nur minimal abweichendes funktionelles Bild aufweist als das fMRT von Person B, ist es prinzipiell möglich, dass es sich hierbei um zwei unterschiedliche Ereignisse $E_1(A)$ und $E_1(B)$ handelt. Damit stellt sich das Problem, dass es für den mentalen Bereich bzw. für eine mentale Kausalität im Gehirn streng genommen keine wiederholte Konstituierung ein und desselben Ereignisses geben kann, sondern die Gehirnaktivität eine fortlaufende Prozessierung teilweise ähnlicher Einzelereignisse ist. Dies muss nicht weiter erschrecken, denn Ziel einer theoretischen Beschreibung der mentalen Ereignisse muss nicht der Durchschnitt sein, sondern eine „ontologische Realität". In ihrer Beschreibung auf der epistemischen wie auch empirischen Ebene scheint es aber zumindest Schnittmengen zu geben, die eine Identifizierung oder auch Typisierung der verschiedenen Einzelereignisse zulassen.

Während Davidson und Kim den Ereignisbegriff, wie dargestellt, als ontologische Kategorie für verschiedene Ereignisse in der Umwelt oder auch in einer Person – also für sämtliche Ereignisse in der Welt – beschreiben, fokussiert Northoff (2004) den Ereignisbegriff auf die Relation zum Gehirn: Die „Ereignis-Kodierung" ist danach ein wesentliches empirisches Merkmal des Gehirns. Ontologisch unterscheidet er ein „eingebettetes Gehirn" von einem „isolierten Gehirn", wie es üblicherweise in der Diskussion betrachtet wird. Letzteres bedeutet, dass das Gehirn

chen: „[…] it remains true that within a model of coarse events, properties (mental or physical) enjoy explanatory relevance, but not real efficacy" (ebd., 235).

getrennt von seiner Umwelt oder auch von seinem Körper betrachtet wird, woraus die unüberbrückbare Dichotomie zwischen mental und physisch sowie Erste-Person- und Dritte-Person-Perspektive resultiert. Die Hypothese der Einbettung führt dagegen zu einer neuen empirischen, epistemischen und ontologischen Bestimmung des Gehirns als ein „eingebettetes Gehirn".

Auf der empirischen Ebene wird das „physische Gehirn" durch ein „dynamisches Gehirn" ersetzt, welches sich durch dynamische Zustände auszeichnet. Die dynamischen Zustände werden durch „dynamische Gesetze" und nicht durch klassische „physikalische Gesetze" beschrieben. Zwar sind neuronale Zustände als physische Zustände notwendig, aber nicht hinreichend für dynamische Zustände, während diese als notwendige und hinreichende Bedingung für mentale Zustände angesehen werden können. Wenn sich dynamische Zustände auf dieselben Umweltereignisse beziehen wie mentale Zustände, können Gehirnzustände insgesamt nicht länger getrennt von mentalen Zuständen gedacht werden. Diese empirische Trennung neuronaler und mentaler Zustände wird durch die Bestimmung der dynamischen Zustände aufgehoben. Diese Bestimmung des Gehirns über seine dynamischen Zustände kann als eine biologische Bestimmung mit dynamischen Gesetzen als biologischen Gesetzen aufgefasst werden (vgl. Northoff 2004, 338). Parallel dazu wird das Gehirn nicht allein über physische Eigenschaften und eine physikalische Ontologie definiert, sondern über seine „dynamischen Konfigurationen" in einer „Ontologie der Einbettung" bestimmt; deshalb ist eine ontologische Ablösung des Geistes vom Gehirn nicht mehr möglich. Der Geist ist abhängig vom Gehirn, da das Gehirn nicht isoliert, sondern eingebettet ist (vgl. ebd., 340). Die dynamische Konfiguration beschreibt eine intrinsische Beziehung zwischen dem Gehirn, dem Körper und der Umwelt (ebd., 341).

Die Ereignis-Codierung beschreibt auf der empirischen Ebene den neuronalen Code des Gehirns und zeigt sich in der Prozessierung sowie der Organisation der neuronalen Zustände in ihrer Ausrichtung auf beobachtbare und zu erzielende Ereignisse innerhalb der Umwelt. Der Begriff der Ereignis-Kodierung kann von einer bloßen Stimulus-Kodierung unterschieden werden; letztere ist lediglich durch die Prozessierung und Organisation von Stimuli bestimmt. Organe wie das Herz, die Niere oder Muskeln sowie Computer gelten für Northoff als Beispiele für eine solche Stimulus-Kodierung. Ereignis-Codierung betrachtet er im Gegensatz dazu als eine konstitutive empirische Fähigkeit des Gehirns. Diese setzt eine intrinsische Integration des Gehirns mit dem Körper und der Umwelt voraus, andererseits wäre eine Organisation neuronaler Zustände in Orientierung auf Umweltereignisse nicht möglich (vgl. ebd., 341f.). In Ergänzung zu empirischen Konsequenzen der Ereigniskodierung beschreibt Northoff auch die epistemischen und die ontologischen Konsequenzen. Epistemisch ist vor allem die „autoepistemische Limitation" des Gehirns relevant, die in enger Verbindung zur Ereignis-Kodierung steht: Das Gehirn ist

epistemisch nicht fähig, seine eigenen Zustände als solche zu erkennen.[190] Ontologisch wird abermals auf die intrinsische Integration von Gehirn, Körper und Umwelt verwiesen, die sich auf der ontologischen Relation der Einbettung sowie den dynamischen Konfigurationen ergibt (vgl. ebd., 342).

Für das Modell der finalen Kausalität sind insbesondere Northoffs Ausführungen zu den empirischen Konsequenzen von Interesse. Sie bestätigen beispielsweise die Sichtweise, die sich aus dem Vergleich der Kriterien für mentale Kausalität mit den neurowissenschaftlichen Paradigmen ergeben hat, wonach der Rahmen weiter und über die Dichotomie zwischen mentalen und physischen Eigenschaften hinaus gefasst werden muss. In der Diskussion des Vergleichs wurde wiederholt auf die Relevanz der „extramentalen Bildstimuli" z. B. für die Kriterien der Regularität (vgl. Kap. 6.1), der Multirealisierbarkeit (vgl. Kap. 6.2) oder der extrinsischen Relation (vgl. Kap. 6.4) hingewiesen. Northoff betont ebenfalls u. a. über die Definition der Begriffe „Ereignis-Codierung" oder „dynamische Konfiguration" die strukturelle Verknüpfung des Gehirns mit der Umwelt sowie den Prozesscharakter für Gehirnzustände, sowohl auf der empirischen als auch auf der ontologischen Ebene. Die Hypothese der Einbettung führt konsequenterweise auch zu einer Neubestimmung der Begrifflichkeiten und des Verhältnisses zwischen mentalen und physischen Eigenschaften. Ähnliches wird auch im Modell der finalen Kausalität angestrebt. Dort wird auf der ontologischen Ebene über die Beschreibung einer wechselseitigen Realisierungsrelation mentaler wie physischer Eigenschaften versucht, die starre Dichotomie aufzubrechen. Während Northoff auf der empirischen Ebene mit dem Begriff der „Ereignis-Kodierung" die Prozessierung sowie die Organisation der neuronalen Zustände in ihrer Ausrichtung auf Ereignisse innerhalb der Umwelt beschreibt, wird dieser Prozess im Modell der finalen Kausalität in der Verbindung zwischen mentalen und physischen Eigenschaften als Eigenschaften neuronaler Aktivität beschrieben. Unter dem aristotelischen Stichwort des „Hylëmorphismus" wird zwischen material- und formgebenden Anforderungen an neuronale Prozessierung in Orientierung auf Umweltereignisse unterschieden (vgl. Kap. 7.1.1; Kap. 7.1.3). Mit dieser Differenzierung unterscheidet sich das Modell der finalen Kausalität von Northoffs Position. Ein weiterer Unterschied kann – trotz der gemeinsamen Betonung des dynamischen Prozesses – in der Verwendung des Ereignisbegriffs gesehen werden. Schließlich wird bei der finalen Kausalität das Ereignis nicht nur als extramentales Ereignis beschrieben, sondern betont in das Subjekt hineinverlegt: einmal als Gehirnereignis und ein weiteres Mal als umfassendes mentales Ereignis, das auf einer überge-

[190] Die autoepistemische Limitation mag, evolutionär gesehen, ein Grund dafür sein, dass das Gehirn über den „Umweg" der mentalen Zustände indirekte Informationen über seine Gehirnzustände erhält. Außerdem mag sie zur besseren Unterscheidung zwischen der Erste-Person- und der Dritte-Person-Perspektive bzw. der besseren Orientierung in der Umwelt dienen (vgl. Northoff 2004, 344f.). Hinsichtlich finaler Kausalität liefert die autoepistemische Limitation möglicherweise einen Grund für die Differenzierung zwischen der epistemischen Ebene mit mentalen und physischen Eigenschaften sowie Gehirn- und Umweltereignissen einerseits, und der ontologischen Ebene mit den mentalen Ereignissen an sich andererseits.

ordneten Ebene die Dichotomie zwischen Gehirnereignis und Umweltereignis aufhebt bzw. umfasst.

In einer Variante einer „Ereignis-Ontologie" bzw. „Prozess-Ontologie" beschreibt Manzotti den Betrachter als wesentliches konstitutives Element eines Ereignisses. Er argumentiert von einem radikalen externalistischen Standpunkt aus. Demnach weisen mentale Phänomene über die Grenzen des Körpers hinaus, weshalb die traditionelle Sichtweise, wonach zwischen den Dingen in der Welt und der Erfahrung solcher Dinge unterschieden wird, nicht aufrechterhalten werden kann. Die Existenz von Objekten oder Ereignissen setzt für Manzotti die Interaktion mit etwas anderem, einem Beobachter, voraus. Deshalb schlägt er eine Neudefinierung einer prozessbasierten Ontologie vor (vgl. Manzotti 2006, 200). Ausgangspunkt ist die Rationale, wonach nur etwas existiert, was Wirkung verursachen kann. Verschiedene Beispiele, wie der Regenbogen, generell Muster, Gesichter, Landschaften, Musik oder Farben werden als Phänomene diskutiert (vgl. ebd., 201-207). Sie sollen den Wechsel von der reinen Abstraktion zu einer tatsächlichen Entität verdeutlichen, die auf dem Prozess zwischen physischen Systemen basiert. All diese Objekte bzw. Phänomene existieren, da sie für physische Prozesse ursächlich sein können. Hierfür benötigen sie allerdings die Interaktion mit einem weiteren geeigneten physischen System – beispielsweise dem menschlichen Gehirn (vgl. ebd., 207). Die Trennung zwischen Subjekt und Objekt scheint für die Phänomene aufgehoben, da die Objekte nur durch das Subjekt als Beobachter existieren. Ohne einen Beobachter hätten die verschiedenen Lichtwellen keinen Gesamteffekt als Regenbogen, würden Muster nicht erkannt, Konstellationen oder Landschaften nicht nachvollzogen oder akustische Signale nicht als Musik zusammengeführt und erkannt. Insofern sind die Phänomene als Ursachen nicht isoliert von ihrer Wirkung im menschlichen Gehirn. In dem Prozess selbst ist dabei die Trennung zwischen dem wahrgenommenen Objekt und der Wahrnehmung aufgehoben. Beide beschreiben lediglich zwei Aspekte desselben Ereignisses.[191] Der physische Prozess zwischen der externalen Welt und dem Gehirn wird als Einheit bzw. als einheitlicher Bezugsrahmen für die mentale wie physische Realität aufgefasst. Der kausale Prozess konstituiert die externalen Objekte, wie den internalen Inhalt des Geistes. Dabei wird betont, dass das Gehirn nicht hinreichend ist, um mentale Ereignisse zu konstituieren (vgl. ebd., 210f.).

Manzotti bezieht sich in seinen Überlegungen auf einen radikalen externalistischen Standpunkt von Honderich (vgl. Honderich 1998, 2004), wonach dieser jegliche Position zurückweist,

[191] Manzotti begründet den erweiterten Geistbegriff durch seine „Prozess-Ontologie": Subjekte und ihre Umwelt konstituieren sich durch Prozesse, einige werden von beiden geteilt. Die Prozesse formen sowohl die erfahrene Welt, als auch die Menschen als Subjekte von Erfahrung. Körper und Geist reagieren kontextsensibel mit denselben Prozessen auf identische externale Systeme. Bewusstsein, Existenz und Relationalität sind für Manzotti untrennbar. Sie beschreiben lediglich auf verschiedene Weise denselben Prozess (vgl. Manzotti 2006, 208). Manzottis Ansatz ist allerdings insofern radikaler als das Modell der finalen Kausalität, als er anscheinend die Existenz von Umweltereignissen generell von einem Beobachter abhängig macht. Im Unterschied dazu verlagert sich der Existenzbegriff bei der finalen Kausalität von Objekten hin zu mentalen Ereignissen. Die Objekte existieren auch unabhängig von ihrem Beobachter, werden allerdings anders bzw. überhaupt nicht interpretiert.

die Geist bzw. Bewusstsein nur als internales Ergebnis alleiniger Hirnaktivität auffasst. Im Gegensatz dazu wird bewusste Wahrnehmung des einzelnen Subjekts als notwendige Voraussetzung für die Existenz von Dingen in der Welt verstanden: „The particular subject is a necessary condition of the state of affairs. It is so because it is in some manner a part of it. But that is not all. In the absence of the subject, there would not exist anything of the world whose existence is what perceptual consciousness consists in." (Honderich 2004) So weit muss das Modell der finalen Kausalität nicht geführt werden, denn im Kern interessieren zunächst nicht Existenzaussagen der Dinge in der Welt, sondern die Konstituierung mentaler Ereignisse. In diesem Sinne kann die Aussage Honderichs zunächst interpretiert werden, dass sich für eine Person bzw. ein Gehirn durch Wahrnehmung die Welt erst konstituiert. In dieser Wahrnehmung, dem mentalen Ereignis, fallen Welt, Gehirn und Person zusammen. In einer radikalisierten, vor-epistemischen Variante könnte davon gesprochen werden, dass sich die Grenzen zwischen Umwelt und Person aufheben und lediglich ein mentales Ereignis geschieht.

Auch wenn Manzottis Modell des „erweiterten Geistes" mitunter etwas skizzenhaft und ontologisch vage begründet erscheint, lassen sich für das Modell der finalen Kausalität wertvolle Impulse ableiten. Insbesondere die Betonung der Relation zwischen Umwelt und Subjekt als Prozess, der als physisch interpretiert wird und dem ein eigener ontologischer Status zufällt, erscheint vergleichbar mit dem Status des mentalen Ereignisses: In beiden Fällen wird der dynamische Charakter einer Relation zwischen Gehirn und Umwelt betont.

### 7.2.2 Finale Kausalität und Regularität

Im vorhergehenden Kapitel wurde das Modell der finalen Kausalität bereits grob beschrieben, wobei auch verschiedene Aussagen hinsichtlich der Relation gemacht wurden. Es handelt sich einmal um eine Relation zwischen Gehirn und Umwelt und zusätzlich um eine Relation aus mentalen und physischen Eigenschaften. Diese Relation soll nun präzisiert und hinsichtlich ihrer Regularität diskutiert werden.

Mit dem Begriff der Regularität wurde dieses Kriterium für mentale Kausalität insofern recht weit gefasst, als nicht definitiv bestimmt werden konnte, ob mentale Kausalität ausschließlich strikten Gesetzen unterliegen muss oder ceteris-paribus-Gesetzen unterliegen kann (vgl. Kap. 3.5; Kap. 4.1). In der Diskussion des Kriteriums vor dem Hintergrund der neurowissenschaftlichen Paradigmen zeigte sich, dass mentale Kausalität drei Gesetzesarten impliziert: neuronale und psychische Gesetze sowie Korrelationsgesetze. Die Basisannahme lautet, dass die mentale Kausalität eine nomologische, also eine gesetzesartige Relation sein muss (vgl. Birnbacher 2006, 116; Levine 2001, 13). Während der Blick üblicherweise lediglich auf der Relation zwischen

mentalen und physischen Eigenschaften liegt, kann nun der Rahmen aufgrund des Vergleichs zwischen dem Kriterium und dem Paradigma erweitert werden (vgl. Kap. 6.1.1; Kap. 6.1.2). Tatsächlich kann eine Relation zwischen drei Komponenten beobachtet werden: zwischen den Bildsignalen und -stimuli, der subjektiven Perspektive sowie den fMRT-Daten bzw. der neuronalen Aktivität. Eine horizontale Kausalität allein zwischen physischen Eigenschaften wird hierdurch in Frage gestellt, da offensichtlich externe Stimuli von Bedeutung sind. Somit erfolgt zumindest eine Art Ebenenwechsel zwischen den Relata. Die Bildsignale sind dabei als die Zielursache zu verstehen, auf die hin neuronale Aktivität, physische und mentale Eigenschaften ausgerichtet werden. Die neuronale Aktivität, die physischen wie mentalen Eigenschaften wiederum wären im aristotelischen Sinne als causa efficiens, causa materialis und causa formalis aufzufassen (vgl. Kap. 7.1.3). Das Telos bestimmt den Grund für die Bewegung der Einzelursachen auf das Ziel hin. Für die mentale Kausalität als finale Kausalität wäre das Ziel ein Gedanke oder eine Repräsentation eines Bildsignals. Lediglich das Vorhandensein von Gehirn bzw. Neuronen (causa materialis) reicht jedoch nicht aus, dies zu leisten. Sie müssen zudem aktiviert werden (causa efficiens) und in einem Netzwerk mit einer spezifischen Funktion zusammengeschlossen sein (causa formalis).[192]

Es muss genauer untersucht werden, in welcher Relation die Einzelkausalitäten zueinander stehen. Dabei kann von der Annahme, dass sich psychische und physische Eigenschaften nicht verursachen können, Abstand genommen werden (vgl. Kap. 2.3.2; Kap. 3.5.1). Das Modell der finalen Kausalität weist eine Anomalie des Mentalen zurück und geht stattdessen von einer idealen Psychologie wie auch Physik aus. Natürlich unterliegt der mentale Bereich Rationalitätsprinzipien und gegebenenfalls ceteris paribus Gesetzen. Es wurde jedoch darauf hingewiesen, dass dieser Gesetzestypus prinzipiell auch für den physischen Bereich gilt bzw. Basis für eine Idealisierung hin zu so genannten strikten Gesetzen ist (vgl. Menzies 2003, 215). Warum sollte dies nicht auch für den mentalen Bereich gelten? Psychologische Generalisierungen sind zudem methodisch für die neurowissenschaftliche Forschung als Kompass zur Bestimmung der neuronalen Aktivität wichtig (vgl. Kap. 3.6.2). Daraus folgt nicht automatisch die Reduktion des mentalen auf den physischen Bereich – dagegen spricht das Prinzip der Multirealisierbarkeit (vgl. Kap. 4.2.1). So erscheint es durchaus plausibel, für einzelne konkrete mentale Ereignisse Korrelationsgesetze für die entsprechenden mentalen und physischen Eigenschaften anzunehmen, selbst wenn diese nicht extensionsgleich sind.[193]

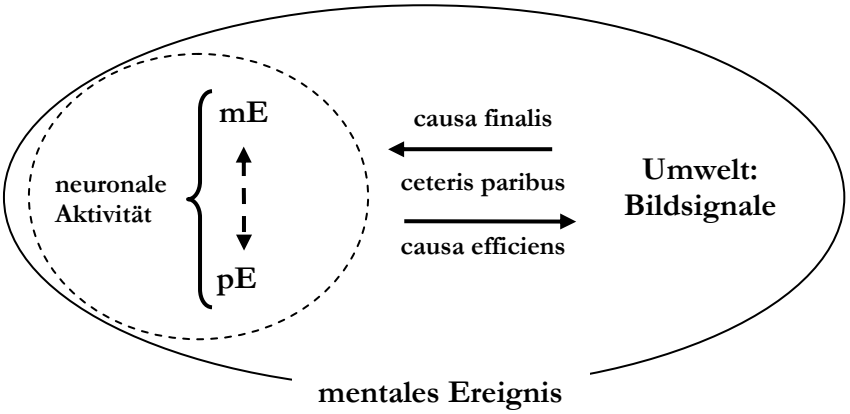
[192] Eine eingehendere Beschreibung dieser Funktion und der Differenzierung der finalen Kausalität erfolgt im Kapitel 7.2.5. An dieser Stelle soll im Schwerpunkt die Regularität betrachtet werden.

[193] Zur Erinnerung: der Begriff der Korrelation ist u. a. problematisch, da ein hoher T-Wert nicht gleich Identität ist. Dies legen beispielsweise die negative Korrelation, die konverse Korrelation oder die reziproke Modulation nahe (vgl. Kap. 5.1.3; Kap. 5.2.3). Es kann nicht definitiv bestimmt werden, dass Inhalt und neuronale Aktivität extensionsgleich sind (vgl. Kap. 6.3.3).

Es ist nicht zwingend erforderlich, dass mentale und physische Eigenschaften sich wechselseitig verursachen. Da sie innerhalb der finalen Kausalität zeitgleich erscheinen und nicht eine Ursache einer Wirkung zeitlich folgt, wäre es ebenso denkbar, dass es sich hierbei um eine Realisierungsrelation handelt, wobei sich mentale und physische Eigenschaften wechselseitig bedingen bzw. wechselseitig realisieren. Danach würde allein die Instanziierung der physischen Eigenschaf als causa materialis die Instanziierung der mentalen Eigenschaft als causa formalis begründen und umgekehrt (vgl. Levine 2001, 13; Kap. 2.2). Hierin liegt der entscheidende Unterschied zur herkömmlichen Idee der Realisierung beim Funktionalismus oder zur Relation der Supervenienz, die asymmetrisch ist: Mentale und physische Eigenschaften sind gleichberechtigt und realisieren sich wechselseitig in der Bewegung auf ein Ziel, einen spezifischen Gehirnzustand hin.

Dieses Modell eignet sich auch für die neurowissenschaftliche Beschreibung der Modulation oder Prozessierung mentaler Ereignisse. Dies wird auch dem Umstand gerecht, dass biologische Prozesse nicht zwangsläufig strikten kausalen Gesetzen unterliegen. Die Realisierung eignet sich gut für die Relation zwischen mentalen und physischen Eigenschaften, weil hierbei die starke wechselseitige ontologische Abhängigkeit deutlich wird (vgl. Marras 2003, 260). Sie kann dabei als Tokenrelation verstanden werden. Beide Eigenschaften können dabei einerseits als Eigenschaften der Neuronen aufgefasst und auf diese zurückgeführt werden. Umgekehrt lässt sich sagen, das Neuronen physische und mentale Eigenschaften haben. Die physischen Eigenschaften eines Neuronenensembles sind aber nur deshalb genau diejenigen physischen Eigenschaften, die für eine spezifische Funktion notwendig sind, da es zugleich bestimmte mentale Eigenschaften hat. Deshalb wird hier von wechselseitiger Abhängigkeit gesprochen. Jene mentalen Eigenschaften sind andererseits auch als Eigenschaften der Zielursache, z. B. eines Stimulus, aufzufassen. Als solche prägen sie die physischen Eigenschaften eines Neuronenensembles und dieses selbst mit und richten es als causa formalis auf die Zielursache hin aus. Die mentalen Eigenschaften sind zentrale Schaltstelle bzw. Filter zwischen Umweltereignis und Gehirn.

Die Relation zwischen Umwelt und Gehirn dagegen ist als kausale, nomologische Relation zu verstehen. Beide sind ontologisch eigenständig, auch wenn im Rahmen der mentalen Kausalität eine Verknüpfung entsteht. Dabei geht die (Ziel-) Ursache der Wirkung voran; neuronale Aktivität im Gehirn wird durch extramentale Stimuli bzw. Umweltereignisse verursacht (vgl. Kap. 7.2.5).




Grafik 20: Neuronale Aktivität verfügt über mentale und physische Eigenschaften (mE, pE). Diese werden als Realisierungsrelation instanziiert und bedingen sich als causa formalis und causa materialis wechselseitig. Die Realisierungsrelation ist Voraussetzung für die kausale Relation zwischen Umwelt und Gehirn. Diese Relation beschreibt über ceteris-paribus-Gesetze als causa finalis und causa efficiens die eigentliche Kausalrelation und damit auch das mentale Ereignis.

Während wir im Kernmodell der finalen Kausalität die Relation zwischen mentalen und physischen Eigenschaften anstelle einer vertikalen Kausalität als eine Realisierungsrelation beschreiben, ist die Relation zwischen Umwelt, subjektivem Erleben und Gehirn die einer horizontalen Kausalität. Diese Relation ist die eigentlich entscheidende, da aus ihr das eigentliche mentale Ereignis hervorgeht. Das eigentliche mentale Ereignis – ein Gedanke, eine Absicht, eine Wahrnehmung oder eine Empfindung – droht in der subtilen Diskussion um Tokenrelationen, Supervenienzen oder Ausschlussargumente aus dem Blickfeld zu geraten. Dabei umfasst das mentale Ereignis als ein raum-zeitliches Ereignis (vgl. Kap. 3.1.1) sowohl die Relation zwischen mentalen und physischen Eigenschaften als auch die Relation zwischen Umwelt, subjektivem Erleben und Gehirn. Es führt diese verschiedenen Bereiche zusammen; diese Bereiche sind gewissermaßen Eigenschaften des Ereignisses. Insofern wird die Tokenrelation um die externale Komponente erweitert. Nur durch ein mentales Ereignis ist es möglich, Umwelt, subjektives Erleben und Gehirn miteinander zu verknüpfen und nur durch ein mentales Ereignis werden im Kern mentale wie physische Eigenschaften realisiert. Dies wird auch dem Umstand gerecht, dass Neuronen eben nicht alleinige Ursache für Wirkung in einem anderen Netzwerk oder eines anderen Netzwerks sind. Daneben sind andere Netzwerke sowie externe Stimuli entscheidend (vgl. Kap. 6.1; Kap. 6.3; Kap. 6.5). Ein neuronaler Zustand zu einem bestimmten Zeitpunkt $N_{t1}$ ist somit Kennzeichen für ein Ereignis der finalen Kausalität, wobei Gehirn und Umwelt in einer Relation zueinander stehen. Das mentale Ereignis umfasst letztlich das gesamte Gehirn mit einigen besonders aktiven Hirnregionen und ein entsprechendes Umweltereignis.

Insofern kann für die mentale Kausalität als finale Kausalität auf Davidsons Ereignisontologie verwiesen werden. Ereignisse bilden das ontologische Netz, sind dabei „vor-epistemisch“ und bedürfen im Grunde keiner weiteren Unterteilung. Mentale Ereignisse umfassen das mentale

Phänomen, den Gedanken, den Wunsch, die Wahrnehmung oder die Absicht selbst und ohne eine weitere differenzierende ontologische Priorität. Jenseits von Davidson sind sie aber nicht nur sowohl physisch als auch psychologisch beschreibbar, sondern sie besitzen physische wie mentale Eigenschaften, Umwelt- wie Gehirneigenschaften. Hier kommt auch Kims Version einer Ereignisontologie zum Zug, da jedes (mentale) Ereignis im Grunde sehr feinkörnig zu bestimmen wäre (vgl. Kap. 3.1.1). Dadurch, dass das Ereignis Umwelt und Gehirn umfasst, ist nun auch verständlich, inwiefern mentale Eigenschaften als Eigenschaften beider Bereiche gelten können. Weil die Ereignisse sehr spezifisch sind, müssen auch die sie betreffenden Gesetze sehr individuell zu bestimmen sein. Dabei handelt es sich im Grunde um ceteris-paribus-Gesetze, denen aufgrund der üblichen Idealisierung durchaus stabiler nomologischer Charakter zukommt (vgl. Menzies 2003, 215; Kap. 3.5).[194]

### 7.2.3 Finale Kausalität und Multikausalität

Aufbauend auf der Relation und der Nomologozität soll nun untersucht werden, inwiefern das Modell der finalen Kausalität mit dem Kriterium der Multikausalität kompatibel ist. Für das Kernmodell der finalen Kausalität wurde Relation als wechselseitige Realisierung beschrieben. Nun stellt sich die Frage, inwiefern diese Kernrelation, aber auch die Relation auf der Ereignisebene (kausal) multirealisierbar ist. Die Anwendung des Kriteriums der Multkausalität auf die empirischen Paradigmen hat für die psychologische wie auch die neuronale Ebene zu ähnlichen Ergebnissen geführt: Es gibt Indizien, die für „multirealisierte Kausalität" bzw. „inverse multirealisierte Kausalität" sprechen, und Indizien, die dagegen sprechen. In der Summe kann zwar im Sinne eines Rollenfunktionalismus pro Multirealisierbarkeit und multirealisierte Kausalität argumentiert werden. Dennoch bleibt es problematisch, dass einerseits die psychologischen Begriffe im Fall der emotionalen Dimensionen sehr grob und vage sind, und dass andererseits einzelne Hirnregionen funktionsspezifisch stabil erscheinen (vgl. Kap. 6.2). Was kann das Modell der finalen Kausalität in diesem Zusammenhang leisten? Zur Klärung soll nochmals auf das Prinzip der Multirealisierbarkeit mentaler Kausalität verwiesen werden:

> Für mentale Kausalität gilt, dass Eigenschaft F in einem System A als Ursache Eigenschaft G oder Eigenschaft H oder Eigenschaft I usw. haben kann, wobei die Eigenschaften G, H, I usw. für F hinreichend sein müssen, auch wenn sie selbst durch eine bestimmte Zahl variabler Eigenschaften realisiert sind. (vgl. Kap. 4.2.2)

[194] Menzies stellt dar, dass prinzipiell auch so genannte strikte wissenschaftliche Gesetze Generalisierungen und Idealisierungen von Einzelgesetzen sind. Es ist gängige wissenschaftliche Praxis, aufgrund empirischer Beobachtungen allgemeingültige Gesetze aufzustellen. Im Grunde, so Menzies, hätten aber auch sie letztlich nur den Status von ceteris-paribus-Gesetzen (vgl. Menzies 2003, 215; Kap. 3.5).

Das Prinzip bezieht sich auf die verschiedenen Fälle multirealisierter mentaler Kausalität, denn

1) auf der psychologischen Ebene kann der Gedanke $G_2$ durch verschiedene Gedanken $G_1$, $G_3$ oder $G_4$ bei einer Person S verursacht werden,

2) unter Supervenienzbedingungen und der Annahme, dass die Gedanken $G_1$ und $G_2$ auf physischen Eigenschaften $P_1$ und $P_2$ supervenieren, wird $G_2$ entweder

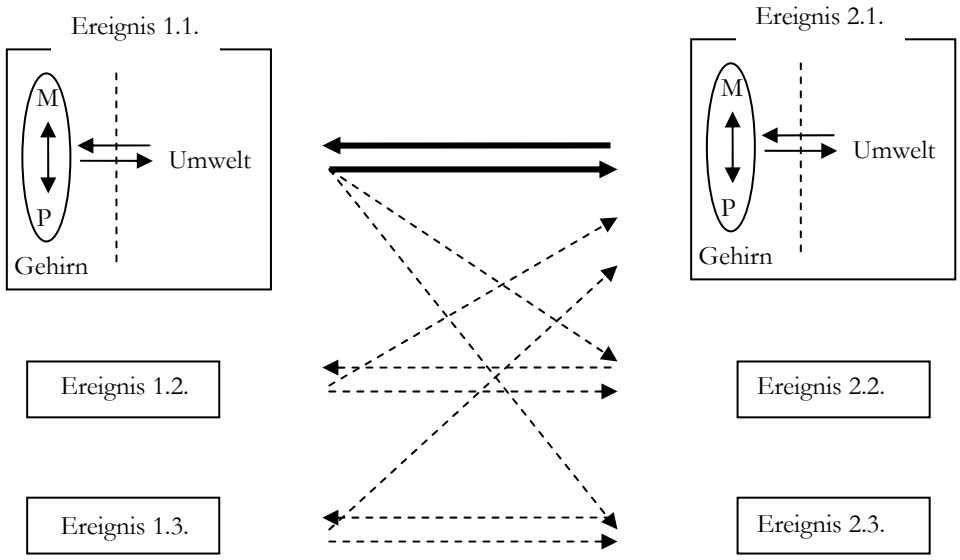
   a) direkt von $P_1$ verursacht, wobei $P_1$ verschiedene physische Eigenschaften sein können, da $G_1$ multirealisierbar ist, oder
   b) $G_2$ wird vertikal von $P_2$ klassisch multirealisierbar verursacht oder
   c) $G_2$ wird indirekt von $P_1$ über $P_2$ verursacht bzw. $P_1$ verursacht $P_2$, wobei auch hier Multirealisierbarkeit gilt, da $G_1$ multirealisierbar ist und somit auch $P_2$ durch verschieden physische Eigenschaften verursacht werden kann (vgl. ebd.).

Die Anwendung des Kriteriums auf die Paradigmen hat gezeigt, dass ein duales, asymmetrisches Modell durch ein „tridirektionales" Modell ersetzt werden kann. Insofern müssten die Fälle 2a), 2b) und 2c) prinzipiell modifiziert werden. Der Fall 1) lässt sich dagegen nach wie vor auf der Ereignisebene, der Ebene der Gedanken etc. darstellen: Ein mentales Ereignis $E_1$ zum Zeitpunkt $t_1$ verursacht ein weiteres mentales Ereignis $E_2$ zum Zeitpunkt $t_2$. Letzteres kann aber auch durch die mentalen Ereignisse $E_3$ oder $E_4$ zum Zeitpunkt $t_1$ verursacht werden: Das Ereignis $E_2$ ist als Wirkung multikausal. Das Beispiel des Schachspielers, den mehrere alternative Gedanken „Läufer durch Dame in Gefahr", „Es ist immer besser, den Läufer proaktiv zu schützen" oder „Eventuell wird der Gegner versuchen, meinen Läufer zu attackieren" zum Gedanken „Läufer durch Springer schützen" bewegen, kann an dieser Stelle aufgegriffen werden (vgl. ebd.). Im Unterschied zur gängigen Multirealisierbarkeit ist die Multikausalität oder Multieffektivität[195] keine vertikale, sondern eine horizontale Relation. Und im Unterschied zu einem unidirektionalen Kausalitätsmodell kann unter dem Modell der finalen Kausalität die wechselseitige Kausalrelation der mentalen Ereignisse $E_1$ zum Zeitpunkt $t_1$ bzw. $E_2$, $E_3$ oder $E_4$ zum Zeitpunkt $t_2$ betont werden. Einerseits verursachen die mentalen Ereignisse zum Zeitpunkt $t_1$ das mentale Ereignis zum Zeitpunkt $t_2$. Andererseits verursachen jene mit Notwendigkeit dieses Ereignis und kein anderes. Es gilt dabei, die Bewegung vom Ziel her zu denken.

Aus dem Vergleich zwischen dem Kriterium der Multikausalität und den Paradigmen hat sich allerdings auch das Dilemma ergeben, dass nicht nur verschiedene Ursachen alternativ eine Wirkung erzielen können, sondern auch eine Ursache verschiedene Wirkungen erzielen kann. Hier droht das Prinzip der Multikausalität in völlige Willkür zu münden. Diesem Dilemma setzt das Modell der finalen Kausalität eben diese gerade beschriebene Notwendigkeit der konkreten Rela-

[195] Streng genommen müsste man davon sprechen, dass das Ereignis $E_2$ „multieffectus" ist, also vielfach verursachbar.

tion zwischen zwei mentalen Ereignissen entgegen. Zugegeben, ein möglicher drohender Chauvinismus wird nur durch einen konsequenten Individualismus vermieden (vgl. Kap. 2.2). Dabei lässt sich andererseits konstruieren, dass es auf der Ereignisebene Ähnlichkeiten über die Individuen hinaus gibt; dieses gilt sogar für die Gedanken eines einzigen Individuums. Schließlich wäre es die Frage, ob ein Individuum denselben Gedanken zu zwei verschiedenen Zeitpunkten haben kann oder nicht. Auf der Detailebene müsste diese Frage verneint werden, auf der Ereignisebene rettet möglicherweise die Un- schärfe der Gedanken bzw. der mentalen Ereignisse. Kann beispielsweise ein Mensch das Wort „Bus" in seinem Leben zweimal genau identisch denken?



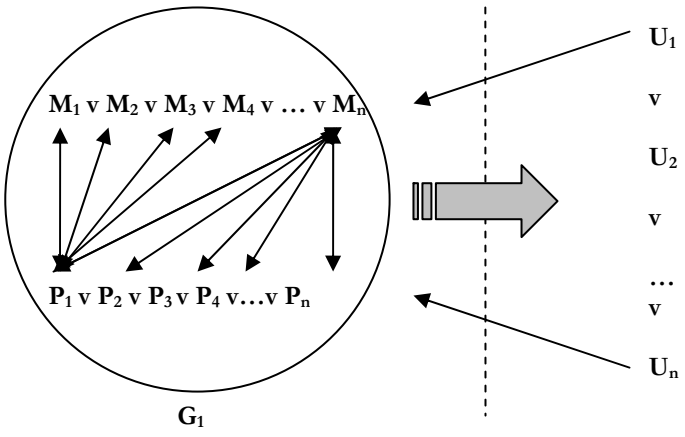

Grafik 21: In der Grafik wird das Modell der finalen Kausalität in Verbindung mit Multikausalität dargestellt. Die (mentalen) Ereignisse repräsentieren in dem Modell Gedanken, Wünsche etc. Die Ereignisse 1.1., 1.2. und 1.3. sind potenzielle Ursachen für das Ereignis 2.1. im Sinne der Bewegungsursache. Die Ereignisse 2.1., 2.2. und 2.3. sind potenzielle „finale" Ursachen für das Ereignis 1.1. Tatsächlich bilden Ereignis 1.1. und Ereignis 2.1. eine kausale Relation aufgrund ihrer konkreten Gehirn- und Umwelteigenschaften zu einem bestimmten Zeitpunkt t. Dabei wird die Gehirnaktivität innerhalb des Ereignisses 1.1. in ihrer Bewegung auf die Zielursache Ereignis 2.1. hin ausgerichtet. M und P symbolisieren die mentalen und physischen Eigenschaften der neuronalen Gehirnaktivität.

Der Einwand lautet hier, dass sich die Konditionen – Umwelt, Biorhythmus etc. – fortwährend verändern, weshalb dies streng genommen eben nicht möglich ist. Trotzdem denken wir oft „Bus" und haben nicht den Eindruck, dass es einen Unterschied macht.

Dieses Dilemma kann zusätzlich auf der Detailebene aufgelöst werden. Schließlich werden die mentalen Ereignisse auf dieser Ebene konstituiert: durch die Relation zwischen Umwelteigenschaften und Gehirneigenschaften, wobei letztere sich aus der wechselseitigen Relation der physischen und mentalen Eigenschaften konstituieren (vgl. Kap. 7.2.1). Auf dieser Ebene kann von einer wechselseitigen Multirealisierbarkeit der mentalen wie physischen Eigenschaften gesprochen werden, da einerseits eine mentale Eigenschaft durch verschiedene physische Eigenschaften realisiert wird, andererseits eine physische Eigenschaft durch verschiedene mentale Eigenschaften realisiert werden kann. Der erste Fall der Realisierung ist mit der konventionellen Multirealisie-

rung vergleichbar (vgl. Kap. 4.2.1), auch wenn die mentale Eigenschaft nicht mit einem Gedanken und die physische Eigenschaft mit einem Neuron o. ä. gleichzusetzen ist. Wie ist aber die umgekehrte Variante der Multirealisierung und darüber hinaus die wechselseitige Multirealisierung zu begründen? Die Multirealisierung der physischen durch die mentalen Eigenschaften lässt sich in der Überlegung fassen, dass die Neuronen mit ihren physischen Eigenschaften zunächst einem unbeschriebenen Blatt ähneln. Erst durch die konkrete Einbettung in das gesamte neuronale System, gegebenenfalls die Verortung in einem bestimmten Hirnareal, die Vernetzung mit anderen Neuronen zu einem Ensemble bzw. einer Neuronenpopulation wird ein Neuron Teil einer (oder mehrerer) spezifischen Funktion(en) und erhält somit konkrete physische Eigenschaften. Diese sind wiederum abhängig von den verschiedenen mentalen Eigenschaften respektive den Umwelteigenschaften (vgl. Kap. 7.2.2).

Schließlich bilden sich die Neuronenpopulationen und die neuronalen Vernetzungen durch die Interaktion mit Umweltereignissen aus. Stichwort ist hier die „kausale Realisierung" als neuronale Verknüpfung durch externale Stimuli (vgl. Kap. 6.2.3). Es erscheint attraktiv, den aristotelischen Begriff des Hylëmorphismus (vgl. Kap. 7.1.1) auf die mentalen wie physischen Eigenschaften der Neuronen anzuwenden, da diese im konkreten Fall als Einheit zu denken sind. Dass Stoff und Form sich wechselseitig bedingen, gilt für das einzelne Neuron wie auch zusammenhängende Neuronenpopulationen.




Grafik 22: In der Grafik wird das Modell der finalen Kausalität auf der Detailebene der mentalen Ereignisse dargestellt. Innerhalb des Gehirns im Gesamtzustand $G_1$ realisieren sich mentale und physische Eigenschaften wechselseitig. Unterschiedliche physische Eigenschaften realisieren eine mentale Eigenschaft, unterschiedliche mentale Eigenschaften realisieren eine physische Eigenschaft. Dabei instanziieren mentale wie physische Eigenschaften als Einheit den neuronalen Gesamtzustand des Gehirns $G_1$. Dieser wird durch ein Umweltereignis, die Zielursache, final verursacht. Diese Relation bildet den Kern der finalen Kausalität. Auch der Gehirnzustand G1 ist zunächst „multikausal" verursachbar, in der konkreten Relation zu einem bestimmten Zeitpunkt jedoch an genau ein Umweltereignis gebunden. Die Kernfrage für die Multikausalität mentaler Ereignisse lautet vor diesem Hintergrund, wie die Flexibilität des Gehirns zu begründen ist.

Aus der Perspektive der mentalen Ereignisse ist in der Relation zwischen Gehirneigenschaften und Umwelteigenschaften die eigentliche finale Kausalität anzusiedeln. Hier geht die Umwelteigenschaft bzw. das Umweltereignis der Gehirneigenschaft voraus, wobei die Gehirnaktivität sich auf Umweltereignisse als Ziel hin ausrichtet (vgl. Kap. 7.2.1). Auch für diesen Fall kann von einer multikausalen Relation gesprochen werden. Schließlich sind es verschiedene Umweltereignisse U, die zu einem bestimmten Zeitpunkt t einen Gehirnzustand G hervorrufen oder bewirken können. Tatsächlich scheint dabei jedem Umweltereignis U in der konkreten Umsetzung genau ein Gehirnzustand G zu genau einem Zeitpunkt t innerhalb einer Person P zuzukommen. Dafür spricht die funktionsspezifische Stabilität der Hirnareale für konkrete Zustände, auch wenn solche Korrelationen neurowissenschaftlich derzeit nur vereinzelt bestimmbar sind. In der Ausgangslage sind jene Gehirnzustände jedoch multikausal von den Umweltereignissen abhängig. Hierfür spricht schließlich auch die grundsätzliche Variabilität des Gehirns, die auch auf individuelle sensorische Erfahrungen hinweist.

Dabei stellt sich die Flexibilität des Gehirns schließlich als ein Kernproblem für Multirealisierbarkeit genauso wie die Festlegung spezifischer Funktionen heraus. Wie ist es möglich, dass das Gehirn auf lediglich ähnliche bzw. vom Standard abweichende Umweltereignisse reagieren kann? Hierin liegt ggf. der Vorteil der finalen Kausalität gegenüber direkten Kausalmodellen, die lediglich auf die Identifizierung physischer Eigenschaften als die eigentlich entscheidenden inhaltlichen wie kausalen Eigenschaften abzielen. Denn innerhalb eines solchen starren unilateralen Kausalmodells müsste für jedes Ereignis genau eine physische Eigenschaft definiert werden. Hier stellt sich aber die Frage, wie das Gehirn unter diesen Voraussetzungen auf zukünftige Ereignisse reagieren kann. Unter dem Modell der finalen Kausalität wird diese einseitige Belastung von den physischen Eigenschaften weggenommen und auf physische, mentale sowie Gehirn- und Umwelteigenschaften verteilt: Das mentale Ereignis konstituiert sich über das Zusammenwirken aller genannten Eigenschaften. Rein physiologisch kann hier möglicherweise zwischen einem Kernereignis, das insbesondere die Gehirneigenschaften bzw. mentale und physische Eigenschaften prägt, und späteren variablen Ereignissen unterschieden werden. Während das Kernereignis in jungen Jahren von einem Menschen erlernt wird, kann dieses als Prototyp zu späteren Zeitpunkten hin aktiviert werden, ausgelöst durch eine adäquate Zielursache. Diese muss hinreichend sein, die entsprechenden Gehirneigenschaften zu aktivieren. Letztere wiederum sind flexibel genug, auf verschiedene Ursachen hin zu reagieren – und somit in einem tatsächlichen Sinn „multieffektiv“. Dies scheint insbesondere dadurch gegeben, dass die Umwelteigenschaft immer auch Teil des gesamten mentalen Ereignisses ist.

Sowohl die Umwelteigenschaft als auch die Gehirneigenschaft, respektive die mentalen und physischen Eigenschaften sind notwendig und erst zusammen hinreichend zur Konstituierung des mentalen Ereignisses. Da sie multirealisierbar bzw. multikausal bzw. multieffektiv sind, si-

chern sie die notwendige Flexibilität, auf neue Ereignisse hin aktiv zu werden, die sich aus den permanent wechselnden Konditionen im Fluss der Zeit ergeben (vgl. Kap. 7.2.6). Neuronale Aktivität ist tatsächlich immer in einen größeren Gesamtzusammenhang einzuordnen, aus dem wiederum individuelle Unterschiede hervorgehen.

### 7.2.4 Finale Kausalität und Inhalt

Die Frage, wie Inhalt kausal relevant bzw. wirksam sein kann, ist letztlich die Kernfrage der mentalen Kausalität. Diese Frage bleibt deshalb so drückend, da propositionale Inhalte, Gedanken oder Wünsche als etwas Immaterielles, etwas Abstraktes erscheinen und damit nicht direkt auf physische oder neuronale Prozesse zu beziehen sind. Umgekehrt scheinen neuronale Prozesse per se nicht empfindsam für Inhalt zu sein (vgl. Kap. 4.3). Die Spannung aus der Relation zwischen Gehirn und Inhalt ergibt sich somit aus der Frage, wie Inhalte im Gehirn und darüber hinaus in der physischen Welt etwas auslösen können sowie aus der Frage, wie das Gehirn selbst Inhalte produzieren oder realisieren kann.

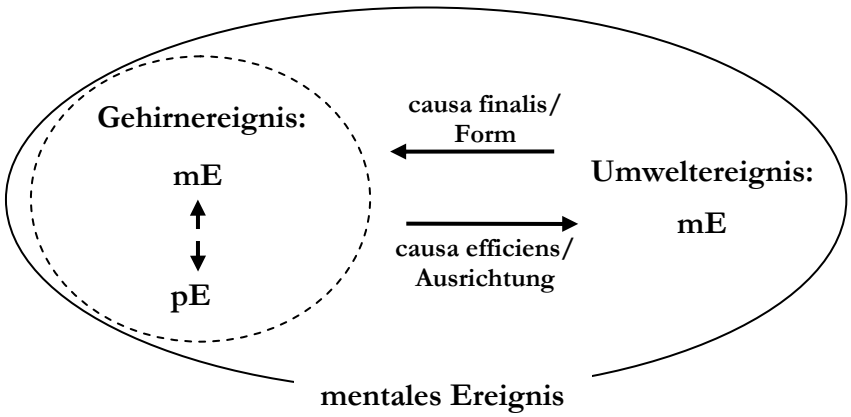
Wie schwierig eine Beantwortung dieser Frage ist, zeigen die unterschiedlichen Bemühungen, Inhalt zu naturalisieren. Sie alle bieten kluge Ansätze an, die jedoch nicht kritiklos bleiben und auch nicht vollständig erklären können (vgl. Kap. 4.3; Kap. 6.3). Hinsichtlich der repräsentationalen Theorie des Geistes lauten einige Kritikpunkte allein auf theoretischer Ebene, dass z. B. Fehlrepräsentationen möglich seien oder dass die genaue Verursachung und Eingrenzung eines Begriffs offen bleibt (vgl. Kap. 4.3.3). In Ergänzung hierzu wurden im Vergleich mit den neurowissenschaftlichen Paradigmen auch Schwierigkeiten auf der empirischen Ebene deutlich. Insbesondere die Unterscheidung zwischen weitem und engem Inhalt wurde hierbei kritisiert. Ein einfaches Kausalitätsmodell scheint nicht auszureichen, um mentale Kausalität zu erklären, weshalb eine Modifizierung des Begriffs für mentale Kausalität eingefordert wurde.

Bevor an dieser Stelle erörtert werden kann, was das Modell der finalen Kausalität in diesem Zusammenhang leistet, soll noch einmal deutlich die übliche Unterscheidung zwischen Inhalt als etwas Abstraktem und neuronaler Aktivität als etwas Materiellem betont werden. Diese Differenz ist der Hintergrund, weshalb Davidson zwischen einem anomalen mentalen Bereich und einem nomologischen physischen Bereich unterscheidet. Die Differenz ist auch der Hintergrund, weshalb Funktionalisten zwischen funktionaler Rolle und Realisierer oder funktionalen Eigenschaften erster und zweiter Ordnung unterscheiden, und weshalb in der Supervenienztheorie zwischen supervenienten Eigenschaften und subvenienten Basen unterschieden wird (vgl. Kap. 2.2.2; Kap. 2.4.2). Damit erkennen sie alle die Differenz an und räumen Inhalt einen eigenen nicht-

physischen Eigenschaftsbereich ein, der damit allerdings per se nicht weiter naturalisiert werden kann.

Tatsächlich zeigen auch die Argumente gegen Identität oder Reduktion, dass sich die abstrakten, psychologischen Inhalte mentaler Zustände von den konkreten, physischen Eigenschaften unterscheiden. Deshalb ist es auch nicht möglich, ohne weiteres etwaige psychologische Gesetze auf physische Gesetze zu reduzieren. Als Beispiele für rein abstrakte Inhalte können die in der Philosophie üblichen Beispiele der reinen Mathematik und Logik angeführt werden, die sich in ihrer Systematik einer Reduzierung auf neuronale physische Prozesse entziehen. Daran anknüpfend, muss generell gefragt werden, ob die psychologischen Gedankenwege generell anderen Gesetzmäßigkeiten unterliegen als die neurologischen Gedankenwege (vgl. Kap. 7.2.2). Und doch muss beides miteinander zusammenhängen.

Im Vergleich mit den empirischen Paradigmen wurde deutlich, dass weder der Ausschließlichkeitsanspruch radikaler Internalisten wie Externalisten noch das Konstrukt von weitem und engem Inhalt durchzuhalten ist. Dagegen ist eher eine Differenzierung zwischen dem äußeren Inhalt einer Sprachgemeinschaft und dem inneren Inhalt bzw. der subjektiven Konnotation des Individuums möglich (vgl. Kap. 6.3.2). Zur weiteren Bestimmung von Inhalt soll nochmals auf die Differenzierung verschiedener Inhaltstypen verwiesen werden (vgl. Kap. 6.3.3). Es wurde im Rahmen des Vergleichs zwischen a) aktuellem Inhalt, b) erlerntem Inhalt und c) assoziativem Inhalt unterschieden. Der aktuelle Inhalt a) bezog sich dabei vor allem auf aktuelle Wahrnehmungen, z. B. durch die IAPS-Bilder des Versuchs.




Grafik 23: Für den Inhaltstyp a) aktueller Inhalt wird die Relation zwischen Gehirnereignis und Umweltereignis beschrieben. Entscheidend sind einerseits die Passung zwischen der mentalen Eigenschaft des Gehirnereignisses und der mentalen Eigenschaft des Umweltereignisses sowie andererseits die Kernbewegung zwischen mentaler und physischer Eigenschaft des Gehirnereignisses. Die mentale Eigenschaft gibt die Form vor, wonach sich die physische Eigenschaft hin ausrichtet.

Für das Modell der finalen Kausalität wird dabei die Relation zwischen Gehirnereignissen und Umweltereignissen beschrieben. Es sind die abstrakten Eigenschaften der Umweltereignisse, wel-

che final das Gehirnereignis verursachen. Dabei handelt es sich, wenn wir von Inhalt sprechen, zunächst um eine abstrakte Eigenschaft des Umweltereignisses, also um eine abstrakte mentale Eigenschaft. Die Gehirnaktivität verursacht das Gehirnereignis in einer Gegenbewegung zum Umweltereignis als causa efficiens. Die Kernbewegung findet hierbei in der wechselseitigen Realisierung von mentalen und physischen Eigenschaften statt: Die mentalen Eigenschaften korrespondieren mit den mentalen Eigenschaften des Umweltereignisses, sie geben die Form vor, nach der sich die physischen Eigenschaften bzw. die neuronale Aktivität insgesamt hin ausrichten. Die mentalen Eigenschaften des Gehirnereignisses und des Umweltereignisses sind dabei nicht notwendig identisch, sondern kompatibel. Die Schnittmenge muss groß genug sein, dass das Gehirnereignis im Ergebnis zum Umweltereignis passt, aber zugleich grob genug, dass ein Gehirnereignis variabel zu einem abweichenden Umweltereignis entsteht. Vor diesem Hintergrund scheint auch die Möglichkeit von Fehlrepräsentationen plausibel zu sein (vgl. Kap. 4.4.1; Kap. 4.4.3).

Diese Forderung der Variabilität knüpft an das Problem der kontinuierlichen Veränderung und b) erlernten Inhalt an. Für diesen Inhaltstypen wurde festgestellt, dass er wesentlich external sein muss, da Lernen als eine neuronale Prägung des Gehirns durch äußere Ereignisse aufgefasst werden kann (vgl. Kap. 6.3.3). In einer aktuellen, erstmaligen Lernsituation wird das Gehirn mit vollkommen neuen Stimuli versorgt. Das bedeutet, dass das Gehirnereignis noch keine mentalen Eigenschaften hat, die mit den mentalen Eigenschaften des Umweltereignisses kompatibel sein können. Vielmehr müssen die mentalen Eigenschaften ausschlaggebend für die Realisierung der physischen Eigenschaften des Gehirnereignisses sein.

Hier gelangen wir an einen kritischen Punkt: während für konkrete Gegenstände noch sehr einfach gefolgert werden kann, dass rein physische Eigenschaften als visuelle Stimuli das Gehirnereignis prägen, gelangen wir hinsichtlich abstrakter Ereignisse in Schwierigkeiten. Dies ist gewissermaßen eine Übertragungslücke in Ergänzung zur Erklärungslücke (vgl. Kap. 3.2.4). Churchland nimmt für diesen Fall an, dass es im Gehirn nicht nur einen physischen Speicher, sondern auch einen Speicher für soziale Ereignisse gibt (vgl. Churchland 1997, 150). Das ist sehr wohl möglich, erklärt jedoch nicht, wie dieser Speicher angelegt werden kann. Sicherlich ist es möglich, dass wir an dieser Stelle an menschlichen epistemischen Grenzen angelangt sind und wir keine Aussage zur Übertragung machen können. In der Folge kann man entweder annehmen, dass abstrakte Inhalte doch irgendwie physisch sind, gegebenenfalls nur anders, oder annehmen, dass sie eben grundsätzlich verschieden sind von konkreten, physischen Inhalten. Im ersteren Fall müssen die Begriffe „Inhalt“ und „physisch“ modifiziert werden, im letzteren muss erklärt werden, wie eine Übertragung zwischen zwei verschiedenen Bereichen möglich ist. Insofern

scheinen beide Alternativen wenig attraktiv.[196] Das Modell der finalen Kausalität bietet hier vor allem auf der Ereignisebene einen Perspektivwechsel an. Danach prozessiert das Gehirn Ereignisse, sowohl physische als auch mentale oder abstrakte. Diese Differenz ist lediglich eine epistemische, die zwar wir Menschen treffen, die jedoch „das Gehirn" selbst nicht kennt. Konkrete Ereignisse können direkter, einfacher oder schneller erlernt werden. Abstrakte Ereignisse sind schwieriger zu prozessieren, da verschiedene konkrete Ereignisse plus X (z. B. eine Situation, eine Metapher = das Abstrakte) zusammengefasst werden. Dies würde erklären, warum Kinder eben viele Situationen erst nach und nach interpretieren lernen und auch Erwachsene in der Einschätzung abstrakter Ereignisse eine höhere Fehlerquote aufweisen, als wenn es darum geht, einen Hund zu erkennen.

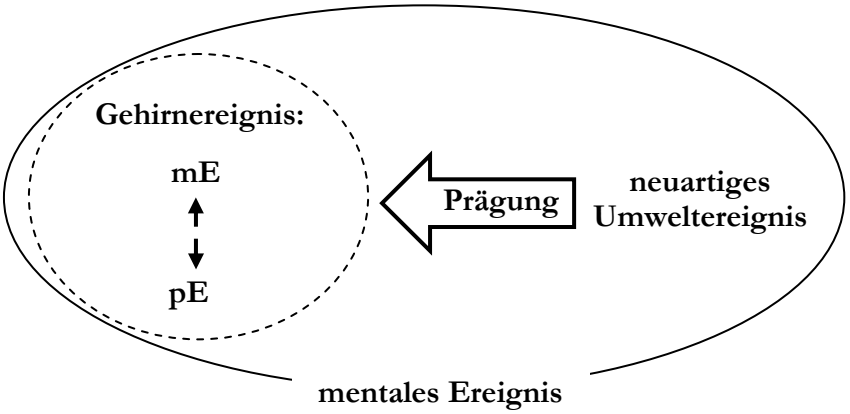
Für das Gehirn ist dabei jedoch entscheidend, dass ein einfaches oder eben ein sehr komplexes Ereignis prozessiert wird. Ob diese Prozessierung einer syntaktischen Struktur entspricht, darüber muss ein Modell der finalen Kausalität zunächst keine Aussage treffen. Wichtig ist jedoch, dass tatsächlich eine (syntaktische) Struktur für die Prozessierung von semantischen Inhalten vorhanden sein muss. Wenn wir dagegen ausschließlich die Prozessierung derjenigen Stimuli, die uns im Kern als physisch erscheinen, akzeptieren, trauen wir dem Gehirn weniger zu, als es zu leisten vermag. Eine Prozessierung, die lediglich auf der Ebene einfacher physischer Reize verweilt, bleibt unter den epistemischen Möglichkeiten des Gehirns und wird dessen Leistungsfähigkeit nicht gerecht. Das Gehirn ist nicht nur fähig, konkrete Gegenstände, sondern wesentlich auch konkrete Situationen, sogar hypothetische Situationen oder Irreales zu prozessieren.

Dass unter finaler Kausalität auch ein aktiver, konstruktiver Prozess beschrieben wird, kann als weiterer Vorteil dieses Modells gewertet werden. Die neuronale Aktivität repräsentiert nicht nur passiv vorhandene „reale" Umweltereignisse, sondern ist zusammen mit diesen Ereignissen konstitutiv. Wahrnehmung ist aktives Interpretieren (vgl. Kap. 6.3.3). Wenn die Bewegung der Gehirnaktivität aktiv auf einen Stimulus hin verläuft, können neue Inhalte mit bereits bestehenden Inhalten abgeglichen und verarbeitet werden. Denn ein Bildsignal löst zunächst im Gehirn neuronale Aktivität aus, woraufhin bestehende Neuronenpopulationen aktiviert werden, die geeignet sind, den Stimulus zu prozessieren. Dabei müssen grundsätzliche Eigenschaften oder Parameter z. B. eines Stuhls erkannt werden, weshalb zunächst eine Schnittmenge zwischen dem Umweltereignis und dem Gehirnereignis bzw. dem neuronalen Prototypen „Stuhl" prozessiert werden muss.[197] Zusätzlich müssen aber die Informationen, die über diesen Prototypen hinausgehen – Farb-, Formvariationen, aktuelle Situation – prozessiert werden. Hierin liegt eine kombi-

[196] Eine weitere Alternative ist natürlich, zu erklären, dass es abstrakte Inhalte eigentlich nicht gibt und sie lediglich eine Täuschung sind. Hier lässt sich allerdings einwenden, dass abstrakte Inhalte offensichtlich in vielen Gehirnen verarbeitet werden und damit subjektiv und objektiv real sind. Insofern scheint die Strategie, abstrakte Inhalte zu ignorieren, nicht dienlich.

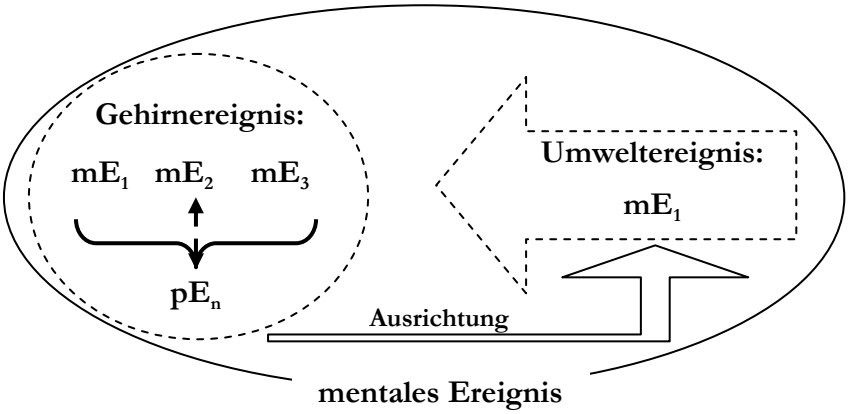
[197] Der Prototyp „Baum" ist für Menschen, die in einem Laubwald aufgewachsen sind, ein anderer als für Menschen, die in einem Nadelwald aufgewachsen sind.

natorische und zugleich kreative Leistung des Gehirns (vgl. Kap. 4.3.1). Während in einem vorbewussten Prozess die finale Kausalität zwischen Umweltereignis und Gehirnereignis verläuft, stellt das mentale Ereignis den zugehörigen bewussten Prozess dar. Das mentale Ereignis ist dabei die konstruktive Verbindung aus Gehirneigenschaft und Umwelteigenschaft, es ist unsere einzige Verbindung zur Welt. Es konstituiert unsere Welt.




Grafik 24: Für den Inhaltstyp b) erlernter Inhalt ist die Phase der Prägung entscheidend. In einer erstmaligen Lernsituation wird das Gehirn mit neuen Stimuli konfrontiert. Der Aufbau z. B. entsprechender neuronaler Vernetzung ist nicht nur passiv, sondern auch konstitutiv. Durch die entwickelten Prototypen werden in späterer Konfrontation mit ähnlichen Umweltereignissen kreativ und kombinatorisch diese generiert. Zu beachten ist weiterhin, dass die Unterscheidung zwischen physischen und mentalen Eigenschaften der Gehirn- und Umweltereignisse eine epistemische ist.

Hinsichtlich des assoziativen Inhalts wurde bereits festgestellt, dass insbesondere an ihm internale Eigenschaften festzumachen sind (vgl. Kap. 6.3.3). Ihm kann komplementär die logische Gedankenfolge an die Seite gestellt werden. Auch innere Gedanken können in Ergänzung zur Wahrnehmung zu a) aktuellem Inhalt gezählt werden. Natürlich greifen Erinnerungen auf Wahrnehmungen und Erfahrungen zurück, die zu einer vorhergehenden Zeit gemacht wurden. Auch diese Inhalte sind external und abhängig von äußeren Ereignissen. Andererseits können assoziative Inhalte und logisches Nachdenken als „innere" Gedanken auch unabhängig von einer direkten Einflussnahme durch äußere Stimuli betrachtet werden. Im Zusammenhang mit dem Versuch wurden diese Inhalte durch ein Bild ausgelöst, allerdings äußerst individuell, was ihren informativen Gehalt angeht und insofern keine notwendige Wirkung des Stimulus. Notwendig allein deshalb, weil gerade dieser eine Proband die assoziativen Gedanken hat.




Grafik 25: Der Inhaltstyp c) assoziativer Inhalt folgt unabhängig von externalen Ereignissen einer „inneren Logik". Diese kann jedoch auf eine Prägung zu einem früheren Zeitpunkt beruhen oder durch ein aktuelles Umweltereignis initiiert sein. Dadurch wird schließlich die Ausrichtung der neuronalen Aktivität bestimmt. Unter holistischen Gesichtspunkten und unter Berücksichtigung der Variabilität physischer Eigenschaften für repräsentationale Eigenschaften wird die Realisierung der mentalen Eigenschaften $mE_1$, $mE_2$ und $mE_3$ durch die Variable $pE_n$ dargestellt.

Insbesondere die assoziative Gedankenfolge verweist darauf, dass mentale und physische Eigenschaften sich zwar in formaler und materieller Hinsicht ergänzen, jedoch nicht aufeinander reduzierbar sind. Finale Kausalität ist deshalb nicht nur ein monokausales Modell, sondern berücksichtigt wechselseitige Modulation auf der neuronalen Ebene sowie das reziproke Verhältnis zwischen mentalen und physischen Eigenschaften. Schließlich sind es die abstrakten Eigenschaften des Inhalts, die eine solche Ereignisfolge auslösen, die zunächst nur aus der subjektiven Perspektive wahrgenommen werden kann. Eine mentale Assoziation bewirkt weitere mentale Assoziationen. Wer „Hund" denkt, denkt auch „Knochen" oder „Hundehütte". Entscheidend ist, dass sich die entsprechende neuronale Aktivität dabei eben nach der Form des mentalen Ereignisses hin ausrichten muss.

Diesen Umstand greift auch die Bestimmung inhaltlicher Dispositionen auf, wonach diese als zunächst „innerer Inhalt" aufgrund extramentaler Ereignisse ausgelöst und bewusst werden (vgl. Kap. 6.3.3). Der umgekehrte Weg, wonach allein die physische Eigenschaft die Assoziationsfolge bestimmt, scheint dagegen weniger sinnvoll, da so keine annähernd sinnvolle Assoziation möglich scheint.[198] Denken wird weiterhin als kausales Fortschreiten verschiedener geistiger Zustände in einem semantischen Sinn aufgefasst. Konzepte und Begriffe konstituieren Gedanken und deren (neuronale) Beziehungen. Eine innere, assoziative oder auch logische Gedankenfolge scheint eigenen Maßstäben zu folgen, einer eigenen inneren Logik, die nicht auf rein physische Mechanismen zu reduzieren ist. Hier könnten erneut die Mathematik oder die ideale Logik als extreme Beispiele für Eigenständigkeit innerer Gedanken, unabhängig von externalen Ereignissen

[198] Hier lässt sich einwenden, dass einige Assoziationen tatsächlich nicht sinnvoll sind. Deshalb kann jedoch nicht geschlossen werden, dass in solchen Fällen nur die physischen Eigenschaften der neuronalen Aktivität die Folge initiieren. Ein mögliches Beispiel für ein mögliches Dominieren der physischen Eigenschaften bei einer Assoziationsfolge scheinen surreale Träume zu sein.

aufgeführt werden. Eingeschränkt mag dies auch für ästhetische Traumwelten, Fiktionen oder Illusionen zutreffen. Der Bereich der logischen Möglichkeit bzw. des denkbar Möglichen ist per se weiter als der natürliche Bereich. Allerdings schöpfen diese Welten aus dem Reichtum der empirischen Erfahrung, weshalb sie auch external sein müssen, selbst wenn sie nicht allein externalen Wahrheitsbedingungen unterliegen. Trotzdem ist es auch dieser letzte Bereich des assoziativen bzw. inneren Inhalts, der die Naturalisierung von Inhalt verhindert.

Diese Spannung zwischen abstraktem, mentalem Inhalt und konkreten, physischen Neuronen kann auch unter dem Modell der finalen Kausalität nur bedingt aufgelöst werden. Für das Modell der finalen Kausalität ist der Begriff der Ereigniskausalität, wie dargestellt wurde, von zentraler Bedeutung. Er ermöglicht eine holistische Perspektive, die dem Zusammenspiel sämtlicher Hirnregionen in ihrer internalen und externalen Vernetzung Rechnung trägt.

### 7.2.5 Finale Kausalität und Relationalität

Der Vergleich mit den neurowissenschaftlichen Paradigmen hat das philosophische Problem bestätigt, wonach nicht eindeutig zwischen intrinsischen und extrinsischen Eigenschaften unterschieden werden kann (vgl. Kap. 4.4; Kap. 6.4). Zwar scheint die subjektive Ebene vornehmlich extrinsisch zu sein, da die Gedanken fast zwangsläufig relational zur Außenwelt sind. Andererseits wurde deutlich, dass insbesondere die emotionalen Dimensionen individuell geprägt sind und teilweise unabhängig von sensorischen Qualitäten erscheinen. Deshalb ließe sich durchaus fragen, ob Valenz, Intensität oder Wiedererkennen nicht einen intrinsischen Kern haben. Zusätzlich wurde argumentiert, dass per definitionem intrinsische Eigenschaften sich ausschließlich auf ihr Bezugsobjekt beziehen (vgl. Lewis 1986), weshalb das subjektive Erleben im weiten Sinne als intrinsisch gedeutet werden kann. Schließlich sind das direkte Bezugsobjekt die Bilder bzw. äußere Stimuli. Allerdings würde sich bei dieser weiten Interpretation der Schwerpunkt des Ereignisses vom Innenleben der Person weg hin zu einer Schnittstelle zwischen Person und Umwelt verlagern.

Wesentlich kritischer muss die Frage der Realisierung aufgrund des Vergleichs mit den neurowissenschaftlichen Paradigmen hinsichtlich der Einordnung auf der neuronalen Ebene erörtert werden. Schließlich ist insbesondere auf dieser Ebene eine eindeutige Einordnung als intrinsisch oder extrinsisch nicht möglich. Wesentliches Argument ist hier, dass eine Gleichsetzung neuronaler Aktivität mit „intrinsisch“ nicht zutreffend ist, da Neuronen untereinander relational sind, aber auch einen Bezug zu bzw. eine Abhängigkeit von äußeren Stimuli haben und somit letztlich als relational interpretiert werden können. So kann sehr allgemein ausgesagt werden, dass alles in Relation zu etwas anderem steht, Eigenschaften zu Eigenschaften und Neuronen zu Neuronen.

Die Amplitude der neuronalen Aktivität einer Hirnregion ist messbar und selbstverständlich in Relation zur neuronalen Aktivität einer weiteren Hirnregion zu setzen. Gefordert wurde deshalb eine Binnendifferenzierung, um den intrinsischen und extrinsischen Eigenschaften der neuronalen Aktivität gerecht zu werden (vgl. Kap. 6.4.3).

Diese Binnendifferenzierung ist gut mit dem Modell der finalen Kausalität vereinbar. Wie bereits erörtert, wird bei diesem Modell zwischen verschiedenen Ebenen, der Ereignisebene sowie der Relation zwischen Gehirneigenschaft und Umwelt sowie weiter zwischen mentalen und physischen Eigenschaften unterschieden. Eine Differenzierung zwischen Neuronen, Regionen oder neuronalen Netzwerken lässt sich auf physische Eigenschaften oder aber auf Gehirneigenschaften beziehen. Allerdings sollte nicht verkürzend die physische Eigenschaft mit Neuronen und die Gehirneigenschaft mit neuronalen Netzwerken identifiziert werden. Vielmehr kann das Modell der finalen Kausalität jeweils auf die entsprechende Ebene der einzelnen Neuronen, Regionen oder Netzwerke bezogen werden. Wobei hier eine Einschränkung gemacht werden muss, da einzelne Neuronen nach dem Modell der finalen Kausalität sich nicht allein, sondern tatsächlich im Netzwerk gedacht, in einer kausalen Relation mit Umwelteigenschaften befinden.

Die funktionelle Spezifizierung neuronaler Netzwerke muss dabei von zwei Seiten betrachtet werden: Zum einen kann sie für Neuronen ähnlich wie Masse, Gewicht oder Form als intrinsische Eigenschaft interpretiert werden, zum anderen verweist sie auf den extramentalen Bereich. Insofern sind den einzelnen Neuronen als causa materialis die intrinsischen Eigenschaften zuzuordnen, während auf der Ebene der Gehirneigenschaften bereits causa materialis, formalis und efficiens zusammenfallen und somit hier die extrinsischen Eigenschaften z. B. neuronaler Netzwerke anzusiedeln sind. Zum anderen wurde bereits darauf hingewiesen, dass Neuronen mit rein intrinsischen Eigenschaften überhaupt nicht sensibel wären für ein Kausalgefüge mit äußeren Stimuli; deshalb müssen ihnen auch extrinsische Eigenschaften zufallen (vgl. Kap. 6.4.2). Eine Binnendifferenzierung kann auch unter dem Modell der finalen Kausalität durchgehalten werden, weil in Abstufung zur jeweiligen neuronalen Einheit mehr oder weniger intrinsische bzw. extrinsische Eigenschaften zugeschrieben werden können.

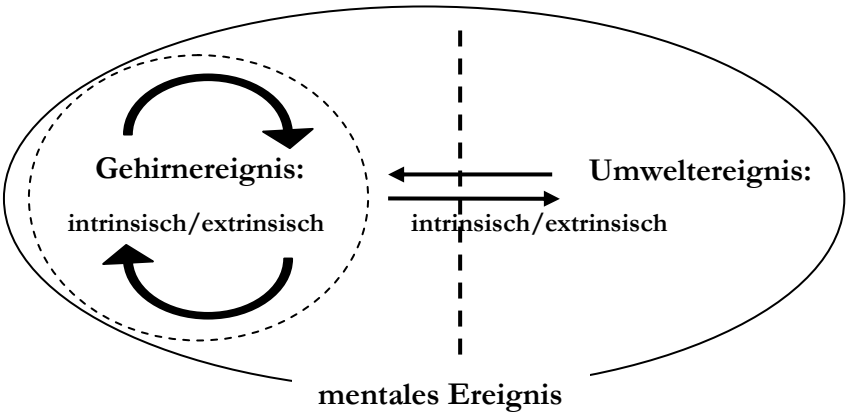
Schließlich kann über das Modell der finalen Kausalität ausgesagt werden, dass es die Ansprüche, die an intrinsische wie extrinsische Eigenschaften herangetragen werden, miteinander verbindet. Auf der Ebene der mentalen Ereignisse stellt sich dabei letztlich nicht mehr die Frage, ob sie intrinsische oder extrinsische Eigenschaften haben – hier ist ein „sowohl als auch“ möglich. Mentale Ereignisse bzw. Gedanken sind extrinsisch, da sie sich innerhalb einer gemeinsamen Umwelt auf andere (mentale) Ereignisse beziehen. Anders formuliert, mentale Ereignisse sind relational zu möglichen Umweltereignissen sowie zu weiteren mentalen Ereignissen. Der Gedanke „um

zwölf Uhr ist Mittagspause“ bezieht sich einmal auf ein konkretes normatives Ereignis innerhalb einer Unternehmenswelt und ist dabei in Relation zu anderen Regelungen zu setzen. Außerdem kann es in Relation zu weiteren Gedanken mit ähnlichem Inhalt, aber anderer Zeit o. ä. stehen. Mentale Ereignisse sind dagegen insofern intrinsisch, als sie sich auf ihr Bezugsobjekt, ein Umweltereignis, beziehen. Dabei umfasst das mentale Ereignis per definitionem die Schnittstelle zwischen Gehirneigenschaft und Umwelteigenschaft (vgl. Kap. 7.2.1). Außerdem sind sie intrinsisch, weil sie auch durch die intrinsischen Eigenschaften der Gehirneigenschaften respektive der physischen und mentalen Eigenschaften sowie der Umwelteigenschaften konstituiert werden.

Damit muss erklärt werden, inwiefern diese Eigenschaften auf der Detailebene intrinsisch oder extrinsisch sind. Grundsätzlich kann angenommen werden, dass sowohl die Umwelteigenschaften als auch die Gehirneigenschaften über eine intrinsische wie extrinsische Komponente verfügen. So kann für die Umwelteigenschaften bzw. -ereignisse jenseits der Perspektive der mentalen Ereignisse festgestellt werden, dass sie als Teil einer physischen Welt über üblicherweise als intrinsisch eingestufte Eigenschaften wie Masse, Größe, Form oder Ladung verfügen. Dieser Bereich der Welt ist es schließlich, in dem die Beobachtung kausaler Vorgänge unter das Modell der causa efficiens, der Bewegungskausalität, gefasst wird. Andererseits lassen sich die Ereignisse in diesem Bereich auch in Relation zueinander setzen: Beispielsweise gehört der Ball einem kleinen Jungen und liegt weniger als zwei Meter von einem Baum entfernt (vgl. Kap. 4.4). Ähnliches gilt auch für die Gehirneigenschaften. Einerseits verfügen sie über typische intrinsische Eigenschaften: das Gehirn hat ein bestimmtes Gewicht, eine bestimmte Form, etc. und gleiches trifft auch auf neuronale Netze oder einzelne Neuronen zu. Andererseits bezieht sich die Aktivität des Gehirns auf extramentale Ereignisse, wie hinsichtlich der Frage der Binnendifferenzierung angesprochen wurde (vgl. Kap. 6.4.3).[199] Das Gehirn wurde in der Diskussion deshalb bereits als essenziell reizverarbeitendes Organ gekennzeichnet – seine wesentliche Funktion ist extrinsisch (vgl. Kap. 6.4.2). Dieses Kennzeichen kann wiederum auf die eingangs angeführte Interpretation des subjektiven Erlebens als intrinsisch bezogen werden, da das Bezugsobjekt des Gehirns außerhalb seiner selbst liegt. Das Gehirn hat somit zwei Objektbereiche: sich selbst und die Umwelt. Insofern kann die Pointe aus dem Vergleich zwischen dem Kriterium der Relation und den empirischen Paradigmen auch unter dem Modell der finalen Kausalität gestützt werden, wonach sich intrinsische und extrinsische Eigenschaften wechselseitig bedingen. Denn auf der anderen Seite unterhält sich das Gehirn mit sich selbst in den zahlreichen unbewussten oder vorbewussten Prozessen, die möglicherweise rein intrinsisch sind. Einige Wahrnehmungen, Empfindungen

[199] An diesem Punkt kann nochmals darauf hingewiesen werden, dass auch sämtliche Körpereigenschaften „extramental“ sind. Unter dem Modell der finalen Kausalität wären sie auf der Ebene der mentalen Ereignisse einbezogen, auf der Detailebene wären sie Teil der differenzierten Betrachtung. Hier wäre genauer zu klären, inwiefern sie losgelöst vom Gehirn den Umwelteigenschaften oder doch zu den Gehirneigenschaften selbst zuzuordnen sind. Gegebenenfalls kommt ihnen auf einer rein physischen Ebene eine ähnliche Funktion als Schnittstelle zu, wie den mentalen Eigenschaften als Mittler zwischen Umwelt und Gehirn.

werden allein im Gehirn konstruiert – allerdings haben auch sie einen zeitlich zurückreichenden Einfluss. Also sind einige intrinsische Eigenschaften des Gehirns das Ergebnis der extrinsischen Eigenschaften. Aufgrund der mentalen Eigenschaften wird die neuronale Aktivität in ihrer physischen Eigenschaft geformt und erhält somit ihre spezifische Funktion. Da mentale und physische Eigenschaften der neuronalen Aktivität sich wechselseitig bedingen, ist letztlich auch einsichtig, warum es mitunter schwierig ist, auch hinsichtlich der intrinsischen bzw. extrinsischen Eigenschaften beide Bereiche eindeutig voneinander zu unterscheiden. So bleibt es problematisch, eindeutig festzulegen, welchem Bereich die emotionalen Bilder als extramentale Stimuli zuzuordnen sind. Einerseits handelt es sich dabei als Lichtwellen um rein physische Stimuli, andererseits jedoch, aus der Perspektive des subjektiven Erlebens, um psychologische oder normative Inhalte, die dem mentalen Bereich zuzuordnen sind (vgl. Kap. 6.1.1; Kap. 6.1.4). Dieser Punkt wurde im Rahmen der Nomologozität sowie der Multirealisierung bereits ausführlich dargestellt (vgl. auch Kap. 7.2.5).




Grafik 26: Das Gehirn als reizverarbeitendes Organ hat intrinsische wie extrinsische Eigenschaften. Diese korrespondieren mit gegebenen Eigenschaften von Umweltereignissen. Dabei hat das Gehirn zwei Objektbereiche: sich selbst und die Umwelt. Das mentale Ereignis beschreibt die Schnittstelle zwischen Gehirnereignis und Umweltereignis.

Die kausale Frage lässt sich nun insofern beantworten, als aufgrund der bisherigen Darstellung eine strenge Grenzziehung zwischen intrinsischen und extrinsischen Eigenschaften weder möglich noch notwendig ist. Dies wurde bereits in dem Umstand deutlich, dass maximale die minimalen Eigenschaften umfassen und deshalb „grenzsensitiv" sind (vgl. Kap. 4.4.2). Kausale Kraft ist auch nicht zwingend mit intrinsischen Eigenschaften verbunden, obwohl dies vielfach so diskutiert wird (vgl. Kap. 4.4.2; Kap. 6.4.4). Auf der Detailebene sind beide Eigenschaften sowohl Gehirneigenschaften als auch Umwelteigenschaften zuzuordnen. Hier findet die lokale Definition Anwendung, dass, sobald ein mentales Ereignis über eine intrinsisch-extrinsische Eigenschaft verfügt, es auch die intrinsische Eigenschaft P und die extrinsische Eigenschaft M umfasst (vgl. Kap. 4.4.2). Ferner wäre es im üblichen Gebrauch wenig sinnvoll, einem der beiden Berei-

che kausale Wirksamkeit abzusprechen, sowohl intern bzw. innerhalb des eigenen Grenzbereichs, als auch darüber hinaus.

Interaktion zwischen Gehirn und Umwelt scheint nicht nur möglich, sondern wesentlich zu sein. Diese Interaktion scheint mentale wie physische Eigenschaften als extrinsische Eigenschaften zu ermöglichen und somit die spezifische Funktion unterschiedlicher neuronaler Netzwerke zu begründen. Damit unterscheidet sich das Modell von einem funktionalistischen Ansatz, da nicht allein mentale Eigenschaften die kausale Rolle für ein höherstufiges übergeordnetes System festlegen. Dabei muss unter dem Modell der finalen Kausalität die Relation von den Umwelteigenschaften her gedacht werden. Diese sind als intrinsische Eigenschaften definierter Bestandteil der kausalen Struktur in der Welt. Zugleich sind sie als extrinsische Eigenschaften in Beziehung zu anderen Eigenschaften bzw. Ereignissen zu setzen (vgl. Kap. 7.2.1). Als Ausgangspunkt der finalen Kausalität bestimmen sie wesentlich den fortlaufenden Prozess; deshalb können in der Folge auch die Gehirneigenschaften als extrinsisch angesehen werden, da sie zwangsläufig in Relation zu anderen Eigenschaften stehen. Sie können auch als intrinsisch angesehen werden, da sie ebenfalls Bestandteil der kausalen Struktur der Welt werden. Entgegen der üblichen Dichotomie kann das Gehirn auch unter dem Modell der finalen Kausalität als intrinsisch-relational betrachtet werden.

### 7.2.6 Finale Kausalität und Intentionalität

Wie bereits dargestellt, ist Intentionalität eng mit dem Phänomen Inhalt verbunden. Während Inhalt sich hauptsächlich auf das mentale Erleben und die damit verbundene neuronale Aktivität eines Subjekts beziehen kann – als inneres Gehirnereignis – , wird mit Intentionalität die Spannung zwischen dem Subjekt bzw. dem Gehirn zu äußeren Ereignissen fokussiert. Mentale Zustände als intentionale Zustände repräsentieren konkrete oder sogar fiktionale Zustände in der Welt, sie sind auf ein intentionales Objekt gerichtet (vgl. Kap. 4.5).

Der Vergleich mit den neurowissenschaftlichen Paradigmen hat diese Sichtweise generell bestätigt. Sowohl auf der subjektiven als auch der neuronalen Ebene ist die Beziehung zu äußeren Objekten, wie z. B. emotionalen Bildern, als intentionale Beziehung beschreibbar. Auch die von Searle beschriebene Struktur intentionaler Zustände, insbesondere die Ausrichtung Welt-auf-Geist bzw. Geist-auf-Welt, konnte hierbei sinnvoll angewandt werden. Intentionale Verursachung als bidirektionales Modell mit kausaler Selbstbezüglichkeit wurde dabei am Phänomen der emotionalen Aufmerksamkeit diskutiert. Der Vergleich hat aber auch Schwierigkeiten von intentionalen Modellen verdeutlicht. So kann beispielsweise die geforderte Strenge einer hundertprozentigen Wahrscheinlichkeit von Dretskes Indikatorrelation nicht erfüllt werden. Ebenso ist die

Verlagerung der Poduzenten-Konsumenten-Rolle nach Millikans teleosemantischem Ansatz oder auch die evolutionär-intentionale Verknüpfung problematisch (vgl. Kap. 4.5.1; Kap. 6.5.1). Die Kernfrage lautet weiterhin, wie Intentionalität, insbesondere im Bezug auf den normativen Bereich, neuronal sein bzw. naturalisiert werden kann. In diesem Zusammenhang wurde die Disposition als Schlüsselbegriff für neuronale Intentionalität bestimmt.

Aufgrund der Definition der Gerichtetheit, der Bezugnahme zu äußeren Objekten ist Intentionalität ein Phänomen, das sich gut mit dem Modell der finalen Kausalität vereinbaren lässt. Wie bereits gezeigt wurde, wird auch bei der finalen Kausalität eine Relation zwischen Gehirnereignissen und Umweltereignissen beschrieben, die durchaus mit der intentionalen Spannung zwischen Geist und Objekt kompatibel ist. Beispielsweise werden emotionale Bilder als positiv oder negativ erlebt, indem sich die neuronale Aktivität bzw. der intentionale Bewusstseinszustand auf ein extramentales positives oder negatives Bild hin ausrichtet. Der intentionale Bewusstseinszustand wird final von dem äußeren Ereignis verursacht und prozessiert dieses in der Folge. Finale Kausalität favorisiert zunächst die Ausrichtung Geist-auf-Welt, da hier Wahrnehmungen und Erinnerungen die äußeren Ereignisse zwar repräsentieren, die Kausalität jedoch von der Welt auf den Geist verläuft.

Wie verhält es sich nun mit der Ausrichtung Welt-auf-Geist und finaler Kausalität? Hier verläuft die Kausalität vom Geist auf die Welt, betroffen sind z. B. volitive Zustände wie Wünsche oder Absichten (vgl. Kap. 4.5.2). Unter dem Modell der finalen Kausalität müsste allerdings auch hier die Kausalität final von den extramentalen Ereignissen her gedacht werden, da sich die Wünsche, so die Argumentation, eben auf diese konkreten oder durchaus auch fiktionalen Ereignisse bezögen. Wünsche oder Absichten werden zwar nicht hinreichend, aber zumindest notwendig durch zeitlich zurückliegende erlernte Ereignisse bestimmt. In der Bewegung verläuft die Kausalität als causa efficiens wiederum vom Gehirn her und prozessiert den Wunsch oder die Absicht. Insofern würde die finale Kausalität im Rahmen der Intentionalität zunächst nicht weiter zwischen Geist-auf-Welt- bzw. Welt-auf-Geist-Ausrichtung differenzieren. Damit ist jedoch nicht gesagt, dass nicht zwischen aktiven und passiven Gehirnereignissen bzw. Absichten oder Wahrnehmungen unterschieden werden kann. Zum einen werden hierbei vermutlich einfach andere Hirnareale aktiviert, wie auch verschiedene Ereignisse im Gehirn durch verschiedene Domänen kognitiv, motorisch oder emotional prozessiert werden (vgl. Kap. 6.5.2; Kap. 6.5.3). Zum anderen ist in der Beschreibung schlüssig davon zu sprechen, dass sich bei z. B. Absichten die Welt nach dem Geist richtet und bei Wahrnehmungen der Geist nach der Welt. Das Phänomen der Intentionalität ist dagegen durch die wechselseitige Bewegung von causa finalis und causa efficiens bestimmt, wobei die Zielursache die geistige Aktivität wesentlich beeinflusst. Insofern wird Searles Forderung nach einem bidirektionalen Kausalitätsmodell konsequent fortgeführt. Aller-

dings sind es im Modell der finalen Kausalität nicht Ursache und Wirkung, sondern zwei Ereignisbereiche – Gehirnereignis und Umweltereignis –, die sich wechselseitig ergänzen und somit das mentale Ereignis selbst verursachen.

Auch das Phänomen der emotionalen Aufmerksamkeit mit der Ausrichtung Geist-auf-Welt kann somit als finaler kausaler Prozess interpretiert werden. Das Pfeil-Signal bildet den finalen Stimulus als Umweltereignis, der die neuronale Aktivität hin zu dem passenden Gehirnereignis auslöst. Das neuronale Netzwerk ist als antizipatorisch gekennzeichnet, es nimmt das nachfolgende Gehirnereignis, auf das es sich zubewegt, final vorweg. Auf der Ebene der mentalen Ereignisse drückt sich das finale Moment dadurch aus, dass die subjektiv erlebte gesteigerte Aufmerksamkeit in dem anschließenden mentalen Ereignis des emotionalen Bildsehens als gesteigertes emotionales Bildsehen aufgeht. Die Erfüllungsbedingung für ein mentales Ereignis ist unter der finalen Kausalität somit nicht nur hinsichtlich der äußeren, extramentalen Ereignisse, sondern auch hinsichtlich der nachfolgenden mentalen Ereignisse zu betrachten (vgl. Kap. 4.5.2).

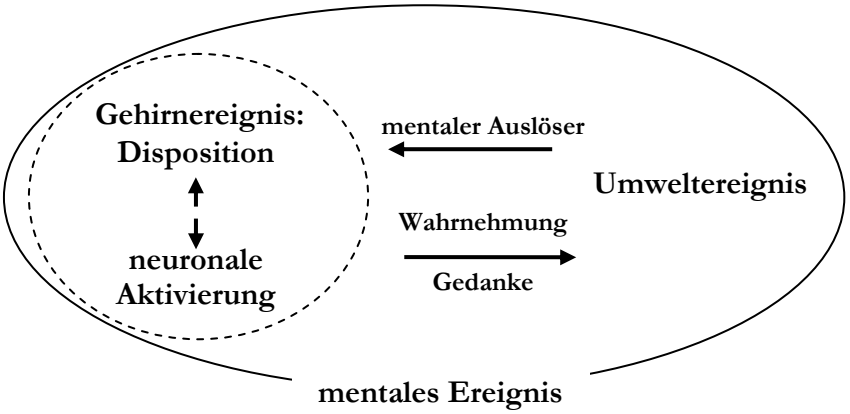
Problematisch ist dabei die Frage, wie die Normativität intentionaler Zustände auf die neuronale Ebene übertragen werden kann. Auf der neuronalen Ebene wurden im Vergleich mit den neurowissenschaftlichen Paradigmen zwar Korrelationen beobachtet und unterschiedliche Aktivität bei positiven oder negativen Bildern. So gibt es, vorsichtig interpretiert, neuronale Entsprechungen und zumindest getrennte Repräsentation von emotionaler und kognitiver Intentionalität (vgl. Kap. 6.5.2). Diese Prozesse müssen nicht zwingend als kausal interpretiert werden – es handelt sich bei der Prozessierung um eine zeitliche Abfolge. Dabei ging es zunächst um die Verbindung von normativen Standards wie Rationalität, Widerspruchsfreiheit oder Erfüllungsbedingungen einerseits und reine Kausalität andererseits. Searle hat versucht, diese Ansprüche unter dem Begriff der „intrinsischen Intentionalität" zusammenzuführen (vgl. Kap. 3.5.2).

Im Rahmen der finalen Kausalität muss für eine Erklärung der Fokus erneut auf die Ebene der mentalen Ereignisse gelegt werden. So wurde bereits im vorhergehenden Kapitel darauf hingewiesen, dass es eine wesentliche Eigenschaft des Gehirns ist, Ereignisse zu prozessieren. Diese Ereignisse sind zunächst vor-epistemisch, weder mental, noch physisch und damit im Grunde weder normativ noch kausal. Es sind zunächst Ereignisse für sich bzw. für das Gehirn. Während für Searle als Ideal eine Theorie des intentionalen Gehalts zweierlei Kausalbeziehungen – zwischen Menschen untereinander und zwischen Menschen und Sachverhalten – umfasst (vgl. ebd.), wird dies bei der finalen Kausalität unter dem Ereignisbegriff subsumiert. Hierbei würde eher zwischen Kausalbeziehungen innerhalb des Gehirns und Kausalbeziehungen des Gehirns mit extramentalen Ereignissen unterschieden – Menschen würden ebenfalls als Sachverhalte bzw. Umweltereignisse betrachtet werden.

Auf der Detailebene sind es die Gehirnereignisse, die auf Umweltereignisse gerichtet sind – intentional und damit eben final verursacht. Allein im physischen Kontext ließe sich anführen, dass ein passendes Neuronenensemble für einen extramentalen Stimulus, z. B. einen Lichtblitz, erst ausgebildet werden konnte, nachdem das Gehirn über die entsprechenden physischen Kanäle mit der notwendigen Information „Lichtblitz" versorgt wurde. Dieses „nachdem" ist zeitlich zu verstehen, aber auch kausal, da das Umweltereignis Lichtblitz ursächlich für den folgenden (neuro-) biologischen Prozess ist. Diese erste Prägung ist entscheidend für die folgenden intentionalen Ereignisse, die sich tatsächlich auf etwas beziehen können, ob es nun situativ da ist oder auch nur gedacht.

Die erste Prägung selbst ist dabei vielleicht nur eingeschränkt intentional, da das Gehirn sich nur intentional auf etwas beziehen kann, was es bereits kennt. Andererseits handelt es sich hierbei um ein intentionales Schlüsselereignis; das Gehirnereignis bezieht sich auf ein zunächst unbekanntes Umweltereignis, das zugleich zum bekannten Ereignis, dem prägenden mentalen Ereignis, wird. Im mentalen bzw. normativen Kontext verläuft dieser Prozess vergleichbar: Den intentionalen Ereignissen geht ein intentionales Schlüsselereignis voran. Beispielsweise ist ein Autounfall ein solches Schlüsselereignis, das in der ersten Wahrnehmung den Prototypen, die Prägung, für weitere Autounfälle oder Unfälle an sich bzw. die entsprechenden intentionalen Ereignisse darstellt. Normative Ereignisse sind als intentionale Ereignisse komplexer, weil hier verschiedene physische Eigenschaften zusammen mit den zusätzlichen abstrakten bzw. normativen Eigenschaften zu einem konkreten Ereignis zusammengeführt werden. Letztlich können die normativen Ereignisses nur aufgrund ihrer konkreten Realisierung, ihrer Exemplifikation, für das Gehirn bzw. dessen Träger greifbar und erkennbar werden. Dem ersten normativen Ereignis, dem Schlüsselereignis, folgen weitere konkrete normative Ereignisse, die vergleichbar sind (oder auch nicht), und somit der Schärfung des Normenverständnisses dienen.

Vor diesem Hintergrund ist auch das Phänomen der Disposition zu erklären, da in der Disposition für bestimmte Gedanken oder eine strukturelle Denkweise die mentalen und physischen Eigenschaften des Gehirnereignisses zusammenfallen. Die Disposition ist beispielhaft für Intentionalität und letztlich auch für finale Kausalität. Schließlich basiert das Modell der finalen Kausalität auch auf der Idee der Potenzialität. Intentionalität ist kein kausaler Vorgang per se, sondern die Gerichtetheit, d. h. Möglichkeit, aber zugleich Resultat des kausalen Prozesses. Dies wurde vom Prinzip her mit der Aktivierung neuronaler Kaskaden beschrieben (vgl. Kap. 6.5.3).




Grafik 27: Die Disposition als intentionaler Zustand ist die generelle Bereitschaft des Gehirns, auf ein Umweltereignis hin aktiv zu werden. Auf der Detailebene werden die erforderlichen neuronalen Populationen über einen mentalen Auslöser aktiviert und die Wahrnehmung oder ein Gedanke wird zielgerichtet auf das Umweltereignis prozessiert.

Im Fall einer Wahrnehmung oder auch einer Absicht wird eine dem normativen bzw. abstrakten Ereignis entsprechende neuronale Population aktiviert. Dabei ist aufgrund der Unbestimmtheit der einzelnen Neuronen und der sich daraus ergebenden Überlappung der neuronalen Populationen eine kombinatorische Vielfalt gewährleistet, die zudem die Flexibilität des Gehirns auf unterschiedliche oder sogar neue Ereignisse sichert. Man muss dabei zwischen potenzieller und aktueller Intentionalität unterscheiden, da die Gerichtetheit zumindest nicht immer im subjektiven Erleben bewusst ist. Umgekehrt ist es kaum vorstellbar und wohl auch neurophysiologisch nicht möglich, dass zu einem Zeitpunkt alle Dispositionen aktuell und bewusst würden – damit wäre eine konzentrierte Wahrnehmung außer Kraft gesetzt. Dagegen ist sehr wohl vorstellbar, dass einige begrenzte Ereignisse parallel aktuell und einige mehr potenziell intentional prozessiert werden. Dies ermöglicht eine wechselnde Fokussierung z. B. von dem Ereignis „Schreibtisch" auf das Ereignis „Kugelschreiber" und das Ereignis „Laptop". Insofern sich das Gehirn in einem absoluten Ruhezustand befände, könnte dies als größtmögliche Potenzialität oder absolute Disposition bestimmt werden. Dieser Zustand ist für jedes Gehirn individuell festzulegen, da es bereits geprägt ist oder neue Ereignisse prozessieren muss.

Für die Detailebene der mentalen und physischen Eigenschaften lässt sich nun sagen, dass die neuronalen Populationen auf den mentalen bzw. normativen Auslöser z. B. eines Umweltereignisses warten, um vom potenziellen Zustand der Disposition in den aktuellen intentionalen Zustand zu wechseln. Dabei scheint der extramentale Stimulus, das äußere Ereignis, entweder ein konkretes Neuronenensemble oder eine neue Kombination vorhandener Dispositionen auszulösen. Sobald das angestrebte Niveau des mentalen Ereignisses in der neuronalen Intensität, Breite und dem entsprechenden Abstraktionsgrad erreicht ist, wird in der Gegenbewegung durch diese Aktivierung eine entsprechende Wahrnehmung oder eine Absicht prozessiert – ein Gedanke gedacht (vgl. Kap. 6.5.3).

Insofern repräsentieren die aktuellen Gehirnereignisse extramentale Sachverhalte, die so oder ähnlich möglich sind. Dabei beinhaltet die (normative) Intentionalität aufgrund der Dispositionen auch immer die Möglichkeit einer Änderung und ist somit variabel. Deshalb darf Kausalität nicht von Normativität getrennt werden, weil Intentionalität in dem Bezug zum extramentalen Ereignis prinzipiell physische und mentale Eigenschaften vereinigt.[200] Dies gilt sowohl für die Detailebene als auch für die Ebene der mentalen Ereignisse. Einerseits sind normative Phänomene auch immer an konkrete physische Dinge bzw. Situationen gebunden. Andererseits werden deshalb im Gehirn auch die physischen Phänomene immer mit normativen Eigenschaften verbunden. Sie werden im Sinne des Hylëmorphismus als normativ-physische Ereignisse erlebt werden (vgl. Kap. 6.5.2). Das intentionale Objekt selbst ist dabei letztlich unscharf.

### 7.2.7 Finale Kausalität und Proportionalität

Das Modell der finalen Kausalität soll schließlich mit Bezug auf das Kriterium der Adäquatheit (vgl. Kap. 4.6) diskutiert werden. Die Herausforderung liegt hierbei in der Frage, inwiefern für finale Kausalität noch eine Differenzierung zwischen verschiedenen Abstraktionsebenen möglich und sinnvoll ist. In dem vorhergehenden Kapitel wurde aufgezeigt, dass intentionale Kausalität unter finaler Kausalität in der Verbindung von mentalen und physischen Eigenschaften sowie letztlich den mentalen Ereignissen selbst in Relation zur Umwelt zu suchen ist. Das Prinzip der Proportionalität sichert dagegen die kausale Wirksamkeit von Glaubensannahmen etc. durch die Unterscheidung zwischen einer hohen Abstraktionsebene und der Detailebene. Hier stellt sich die Herausforderung, diesen scheinbaren Widerspruch aufzulösen.

In dem Vergleich mit den neurowissenschaftlichen Paradigmen wurde deutlich, dass das Prinzip der Proportionalität empirisch dargestellt werden kann. Zum einen verbindet die Neurowissenschaft allein methodisch zwei Erklärungsbereiche, z. B. in der fMRT-Messung und den Post-hoc-Tests. Zum anderen konnten diese unterschiedlichen Erklärungsbereiche auch für den Versuch dargestellt werden: Die adäquate Wirkung der Aufmerksamkeitsstimuli, die adäquate Wirkung der emotionalen Stimuli und die adäquate Wirkung der Beurteilung können jeweils für die neuronale oder die subjektive Ebene entsprechend als Relation mit hohem oder geringerem Abstraktionsgrad dargestellt werden (vgl. Kap. 6.6.1; Kap. 6.6.3).

Was ergibt sich hieraus für das Modell der finalen Kausalität? Proportional ist, was angemessen repräsentiert bzw. prozessiert wird. Auf der Ebene mentaler Ereignisse kann dabei im

[200] In einer extremen Form eines intentionalen Zustands würde sich die Unterscheidung zwischen dem Subjekt, das einen Gedanken hat und dem intentionalen Objekt, auf das sich der Gedanke bezieht, aufheben. Es gäbe dann nur noch das pure mentale Ereignis. Als Erkenntnisinstrument sichert uns das Kriterium der Intentionalität die Unterscheidung zwischen Subjekt und Objekt, zwischen Gehirn- und Umweltereignis.

Prinzip beinahe tautologisch formuliert werden, dass jedes mentale Ereignis Ursache und Wirkung eines anderen mentalen Ereignisses ist. Mentale Ereignisse müssen sich als Ursache und Wirkung zueinander adäquat verhalten, sonst könnten sie nicht in einer kausalen Relation zueinander stehen. Entscheidend ist dabei die horizontale Ebene zwischen mentalen Ereignissen, z. B. von dem Ereignis der emotionalen Aufmerksamkeit hin zu dem Ereignis des emotionalen Bildsehens (vgl. Kap. 6.6.2; Kap. 6.6.3). Adäquate Verursachung bzw. Instanziierung bedeutet in diesem Kontext ferner, dass Gehirneigenschaft und Umwelteigenschaft, welche die mentalen Ereignisse umfassen, zueinander adäquat sind. Zu einer weiteren Bestimmung des Begriffs müssen deshalb die Zusammenhänge auf der Detailebene untersucht werden.

Auf der Detailebene ergeben sich Unterschiede zur strikten Trennung der Abstraktionsgrade. Zwar können aus einer epistemischen Perspektive ebenfalls Unterschiede für abstrakte bzw. detaillierte Ursachen und Wirkungen festgemacht werden. So können während der Phase der emotionalen Aufmerksamkeit für die physischen Eigenschaften der Gehirnereignisse adäquate Umweltereignisse, z. B. die Pfeilsignale, als rein physische Stimuli gelten, sowie für die mentalen Eigenschaften der Gehirnereignisse adäquate Umweltereignisse, z. B. die subjektiv erlebte Information „ich bin emotional aufmerksam“ (vgl. Kap. 6.6.3). Allerdings wurde hierbei bereits in dem Vergleich zwischen dem Kriterium Adäquatheit und den neurowissenschaftlichen Paradigmen die Schwierigkeit deutlich, den extramentalen Stimulus adäquat für die physische bzw. subjektiv mentale Wirkung zu beschreiben. Schließlich kann das antizipatorische Netzwerk auf der neuronalen Ebene mit dem geringeren Abstraktionsgrad eigentlich nur dann aktiviert werden, wenn das Gehirn das Pfeil-Signal als Aufmerksamkeitsstimulus interpretiert. Damit muss das Gehirn das Pfeil-Signal aber bereits auf einer Ebene mit höherem Abstraktionsgrad, der subjektiv mentalen Ebene, interpretieren können. Eine rein physische Interpretation würde lediglich zur Prozessierung der Form und Farbsignale ohne die Signalinformation führen. Im vorliegenden Fall vereint das extramentale Umweltereignis also verschiedene Abstraktionsgrade auf sich, die ebenfalls Wirkung mit verschiedenen Abstraktionsgraden erzielen. Nach Schröders Prinzip der Proportionalität müsste es sich dabei sowohl epistemisch als auch ontologisch um unterschiedliche Wirkungen handeln.

Nach dem Modell der finalen Kausalität kann dagegen zwar eine epistemische Unterscheidung getroffen werden, die Wirkung selbst ist jedoch auf der ontologischen Ebene eine einzige. Da sich physische und mentale Eigenschaft wechselseitig bedingen und auf das Umweltereignis hin ausgerichtet sind, kann auf der ontologischen Ebene keine Unterscheidung des Abstraktionsgrades zwischen diesen Eigenschaften getroffen werden. Insofern müsste eine Ursache, die für eine physische Eigenschaft adäquat ist, auch immer für eine entsprechende mentale Eigenschaft adäquat sein, vice versa. In der Frage der Proportionalität von Ursache und Wirkung würde auch bei finaler Kausalität zwischen verschiedenen Abstraktionsgraden unterschieden werden – aller-

dings stets bezüglich der Gehirnereignisse. So können entsprechend der vorgenommenen Ausdifferenzierung des Abstraktionsgrades höherstufige und niedrigstufige mental subjektive Eigenschaften unterschieden werden (vgl. Kap. 6.6.2). Als mental höherstufige Eigenschaften können komplexe Umweltsignale mit einem abstrakten Informationsgehalt gelten, der über einen bloßen Wahrnehmungsreiz hinausgeht und entsprechend vielschichtig prozessiert wird. Als mental niedrigstufige Eigenschaften können einfache Umweltsignale mit einem konkreten Informationsgehalt gelten, der sich lediglich auf eine Wahrnehmung bezieht und entsprechend in basalen neuronalen Regionen prozessiert wird.

Die Ausdifferenzierung des Abstraktionsgrades führt uns unter dem Modell der mentalen Kausalität somit zurück zur Unterscheidung zwischen determinates und determinables. Damit soll nicht die Unterscheidung zwischen verschiedenen Abstraktionsgraden, wie Schröder sie trifft, auf der epistemischen Ebene verworfen werden. Allerdings kann diese Unterscheidung auf der ontologischen Ebene für mentale Kausalität nicht aufrechterhalten werden (vgl. Kap. 4.6.2; Kap. 6.6.4). Konsequenz ist allerdings auch kein Reduktionsimus bzw. eine Stärkung des Ausschlussarguments in dem Sinne, dass mentale Eigenschaften epiphänomenal sind, sich auf physische Basen beziehen und diese die eigentlich kausale Arbeit verrichten. Deshalb soll an dieser Stelle nochmals betont werden, dass sich im Rahmen der finalen Kausalität mentale und physische Eigenschaften der Gehirnereignisse zielgerichtet auf das extramentale Umweltereignis hin wechselseitig realisieren (vgl. Kap. 7.2.1).

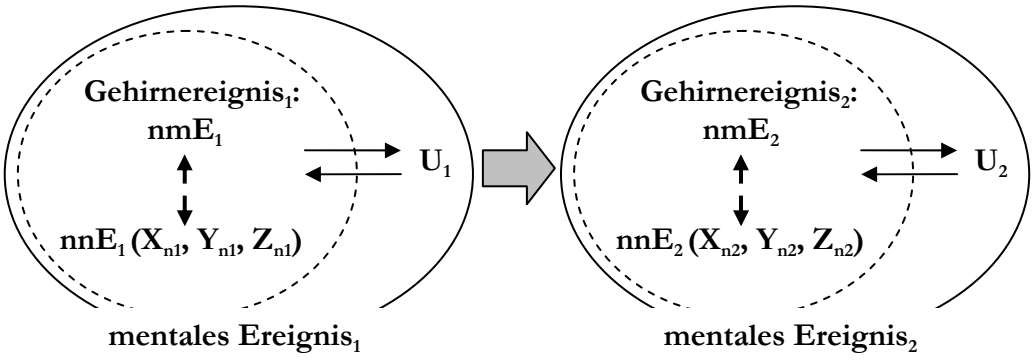
Die einfache Identifizierung der korrelierenden Eigenschaften ist dabei ausgeschlossen, da auf der epistemischen Ebene die beiden Eigenschaftsbereiche nicht deckungsgleich und auf der ontologischen Ebene wechselseitig multirealisierbar sind (vgl. ebd.). Wenn als extramentaler Stimulus ein Kinoplakat eine Person zum Kinobesuch bewegt, werden die niedrigstufigen mentalen Eigenschaften $nmE_1$ des Stimulus – dass ein Kinobesuch Freude bereitet o. ä. – in dem Gehirn adäquate niedrigstufige neuronale Eigenschaften $nnE_1$ wechselseitig hervorrufen bzw. realisieren. Die niedrigstufige neuronale Eigenschaft $nnE_1$ ist dabei Variable für unterschiedliche neuronale Instanziierungen $X_{n1}$, $Y_{n1}$ oder $Z_{n1}$ bei unterschiedlichen Gehirnen oder sogar innerhalb eines Gehirns (zu unterschiedlichen Konditionen). Die niedrigstufige mentale Eigenschaft $nmE_1$ „Kinobesuch amüsiert“ kann somit wechselseitig $X_{n1}$, $Y_{n1}$ oder $Z_{n1}$ realisieren. Entscheidend ist dabei, dass die niedrigstufige neuronale Eigenschaft nicht ohne die niedrigstufige mentale Eigenschaft aktiviert werden würde. Letztere Eigenschaft ist der „gatekeeper“ für die erstere. Die mentalen Eigenschaften sichern die Zielrepräsentation unter Berücksichtigung der Variabilität physischer Eigenschaften für mentale Repräsentationen (vgl. Kap. 4.6.2; Kap. 6.6.4). Fehlt das Wissen, was ein Kinobesuch bewirkt, kann keine entsprechende neuronale Aktivität entwickelt, sondern lediglich die Wahrnehmung prozessiert werden, wonach an einer bestimmten Stelle ein Plakat mit bunten Bildern steht. Erst die mentale Information kann die adäquate neuronale Akti-

vität und damit verbundene weitere Gehirnereignisse auslösen. Die mentale Eigenschaft kann aus diesem Grund auch als determinable der physischen Eigenschaft bestimmt werden, da die abstrakte Information die konkrete neuronale Aktivität umfasst. Umgekehrt sind die multirealisierenden physischen Eigenschaften determinates der mentalen Eigenschaft.

Die Schlüsselstellung der abstrakten Information der mentalen Eigenschaft kann noch einmal in einer zeitlichen kausalen Folge verdeutlicht werden. Denn nach dem Prinzip der Proportionalität verläuft mentale Kausalität horizontal, wobei eine kausale Relation zwischen zwei Korrelationen besteht (vgl. Kap. 6.6.3). Der niedrigstufigen mentalen Eigenschaft „Kino amüsiert“[201] $nmE_1$ folgt die niedrigstufige mentale Eigenschaft „jetzt ins Kino gehen“ $nmE_2$. Beide Eigenschaften realisieren wechselseitig zugehörige neuronale Eigenschaften $nnE_1$. und $nnE_2$. Unter finaler Kausalität kann nun das kontrafaktische Konditional formuliert werden, dass ohne $nmE_1$ weder $nmE_2$ noch $nnE_2$ erfolgen. Dies liegt im Grunde daran, dass unter finaler Kausalität ohne $nmE_1$ auch nicht $nnE_1$ möglich ist, da sich beide wechselseitig realisieren. In der Umkehrung kann auch das kontrafaktische Konditional formuliert werden, dass ohne $nnE_1$ weder $nmE_2$ noch $nnE_2$ ist. Als Gehirnereignisse $G_1\{nmE_1; nnE_1\}$und $G_2$ $\{nmE_1; nnE_2\}$sind beide final auf das Umweltereignis „Kinoplakat“ $U_1$ mit seinen physischen und mentalen Eigenschaften $\{pE_1; mE_1\}$ hin ausgerichtet.

Dabei ist nun das Problem zu berücksichtigen, dass sehr wohl unbewusste mentale Eigenschaften, gegebenenfalls sogar höherstufige, (bewussten) niedrigstufigen mentalen Eigenschaften vorangehen. Kann hier noch von adäquater Verursachung gesprochen werden, oder ist ein solcher Fall auszuschließen? Auf der physischen Ebene scheint uns dieser Gedanke zunächst keine großen Probleme zu bereiten, neuronale Aktivität in basalen Regionen kann durchaus zu neuronaler Aktivität in höherstufigen Regionen führen, und umgekehrt. Andererseits bedingen sich mentale und physische Eigenschaften wechselseitig, und neuronale Aktivität ist bereits Ausdruck dieser wechselseitigen Relation (vgl. Kap. 6.6.4). Also ist auch die Folge von niedrigstufigen mentalen Eigenschaften zu höherstufigen mentalen Eigenschaften nicht auszuschließen. Damit scheint nun aber die Adäquatheit ontologisch auf der horizontalen Ebene für mentale wie physische Eigenschaften aufgehoben zu sein, nachdem sie unter finaler Kausalität bereits ontologisch für die vertikale Ebene aufgehoben wurde.

[201] Problematisch ist, dass mentale Eigenschaften der Gehirnereignisse nicht formuliert werden können, da sie lediglich die „Form“ angeben. Der Satz „Kino amüsiert“ kommt prinzipiell dem mentalen Ereignis auf der umfassenden Ebene näher. Eine „Versprachlichung“ der mentalen Eigenschaften ist also grundsätzlich irreführend, erfolgt aber hier, um den Informationsgehalt der mentalen Eigenschaften zu veranschaulichen.




Grafik 28: Nach dem Prinzip der Proportionalität verläuft mentale Kausalität horizontal unter Berücksichtigung variabler Korrelation. Mit der niedrigstufigen mentalen Eigenschaft $nmE_1$ wird die niedrigstufige neuronale Eigenschaft $nnE_1$ wechselseitig realisiert. Letztere wiederum steht für verschiedene mögliche physische Eigenschaften $X_{n1}$, $Y_{n1}$ und $Z_{n1}$. Die mentale Eigenschaft sichert für das Gehirnereignis die Zielrichtung auf das Umweltereignis hin. Erst dadurch ist eine kontinuierliche Folge von dem mentalen Ereignis$_1$ zu dem mentalen Ereignis$_2$ möglich.

Tatsächlich muss man sich an dieser Stelle genau die Prozesse anschauen, um die es hier geht: nämlich um intramentale Prozesse. Diese verlaufen vielfach unbewusst, mitunter parallel zu bewussten mentalen Prozessen und werden aufgrund verschiedener Anlässe zu einem bewussten mentalen Prozess. Ein solcher ursächlicher Anlass kann eine innere Assoziationskette sein, bei der ein Gedanke verfolgt wird oder aber ein extramentaler Stimulus, ein Umweltereignis. Für den ersteren Fall kann auch auf extramentale Stimuli als ursächlichen Anlass verwiesen werden, die zu Beginn einer inneren Assoziationskette stehen können. Damit müsste der Zielpunkt einer solchen Gedankenfolge ebenfalls adäquat zu dem externen Umweltereignis sein. Aber auch innerhalb einer Gedankenfolge wird vermutlich eine adäquate Linie verfolgt, die allerdings zu Beginn und zum Ende hin einen Übergang zwischen niedrigstufigen und höherstufigen Abstraktionsebenen bzw. Prozessen umfasst – wobei jedoch parallel zu einem höherstufigen Prozess niedrigstufige Prozesse verlaufen, die gegebenenfalls in den höherstufigen Prozess münden. Im letzteren Fall kann auf die finale Kausalität verwiesen werden, wonach das Gehirnereignis adäquat das Umweltereignis prozessieren soll. In dieser Relation zwischen Gehirnereignis und Umweltereignis ist die eigentliche adäquate Verursachung zu finden. Erst durch die in der Intensität wie in der Breite hinreichende neuronale Aktivität auf das Umweltereignis hin wird das Gehirnereignis adäquate Wirkung, wobei die neuronale Aktivität selbst eben einer neuronalen Entwicklung unterliegt.

Damit wird aber auch deutlich, dass finale Kausalität einen Übergang von niedrigstufigen zu höherstufigen Prozessen bzw. von höherstufigen zu niedrigstufigen Prozessen ermöglichen muss – zumindest für intramentale Prozesse. Das Prinzip der Proportionalität kann dagegen nur Anwendung finden bei Prozessen, die direkt zwischen intramentalen und extramentalen Prozessen vermitteln. Andererseits gilt auch für intramentale Prozesse, dass kausale Wirkung adäquat sein

sollte, allerdings muss der Begriff der Adäquatheit für diese Prozesse neu definiert werden. Bei intramentaler Kausalität kann nicht davon ausgegangen werden, dass nur ein höherstufiger Prozess einen weiteren höherstufigen Prozess verursachen kann. Adäquate Verursachung muss hierbei in der Schwelle, im Übergang zwischen beiden Ebenen gesucht werden.

Die Frage, die sich dabei stellt, ist, wann neuronale Aktivität in basalen Hirnarealen zu neuronaler Aktivität in höherstufigen Hirnarealen führt. Ebenso kann auf der subjektiven mentalen Ebene gefragt werden, wann ein Übergang zwischen bloßem passivem Sehen von Gegenständen und gezielten Gedanken (dazu) stattfindet. Dieser Übergang wird aus der subjektiven Perspektive häufig mit dem Übergang zwischen unbewussten und bewussten Gedanken verbunden; diese Verbindung ist aber nicht notwendig. Es kann weiter zwischen bewusstem und unbewusstem Sehen von Gegenständen sowie bewusstem und unbewusstem Denken unterschieden werden, wobei hinter unbewusstem Denken eben nicht offen aktive Dispositionen stehen. Auf der neuronalen Ebene lässt sich darstellen, dass basale oder komplexe kognitive Funktionen nicht zwangsläufig (un)bewusst sind. Ein Beispiel ist hierbei der Unterschied zwischen dem ungeübten und dem geübten Autofahrer: während ersterer sich noch sehr stark bzw. bewusst auf sämtliche Signale und Ereignisse im Straßenverkehr konzentrieren muss, kann letzterer als Routinier entspannter fahren. Dabei macht dieses Beispiel auch deutlich, dass im Laufe der Wiederholung bzw. der Routine einer Tätigkeit diese anders abgespeichert wird, in tieferen kortikalen Regionen bzw. nicht mehr im Kurzzeit-, sondern im Langzeitgedächtnis. Auch für Dispositionen muss letztlich gelten, dass bei ihnen epistemisch zwischen mentalen und physischen Eigenschaften bzw. zwischen determinates und determinables unterschieden werden kann, während ontologisch diese Eigenschaften zusammenfallen.

Die besondere Herausforderung der Dispositionen, dass sie subjektiv wie neuronal nur erkennbar sind, wenn sie den rein dispositionellen Bereich verlassen, betrifft dabei auch die Frage der Adäquatheit. Der Herausforderung kann dabei unter Verweis auf die adäquate Aktivierung neuronaler Populationen bzw. Ensembles begegnet werden, beispielsweise bei der Wahrnehmung eines Gegenstandes (vgl. Kap. 6.6.4). Dabei hat es sich gezeigt, dass auf allen Verarbeitungsstufen die Neuronen auf mehr als nur ein Merkmal reagieren, weshalb verschiedene neuronale Populationen, die durch verschiedene Merkmale aktiviert werden, sich überlappen (vgl. ebd.; Singer 2002c, 149f.). Insofern kann man fragen, ob ein Neuron nicht auch verschiedene Dispositionen hat und über verschiedene physische wie mentale Eigenschaften verfügt. Natürlich manifestieren sich solche Einzeldispositionen nur in der Aktivierung eines gesamten Ensembles. Entscheidend für die finale Kausalität ist dabei, dass ein extramentaler Stimulus nicht direkt die einzelnen Neuronen aktiviert, sondern diese erst aufgrund einer niedrigstufigen oder höherstufigen Zuordnung aktiviert werden. Diese Zuordnung verbindet dabei ontologisch notwendig konkrete physische Eigenschaften mit abstrakten mentalen Eigenschaften. Sobald sie durch angemessene neuronale

Aktivität erreicht ist, kann von einer adäquaten finalen Kausalität gesprochen werden, unabhängig von dem Bewusstsein einer Person über diesen Zustand. Dabei scheint es auf einen äußeren Stimulus hin immer eine Vielzahl möglicher adäquater Gehirnereignisse zu geben, abhängig von der Situation, der Konstitution o. ä. der Person.

Die eingangs angesprochene Herausforderung, ob für finale Kausalität noch eine Differenzierung zwischen verschiedenen Abstraktionsebenen möglich und sinnvoll ist, kann abschließend geteilt beantwortet werden. Epistemisch können auch unter dem Modell der finalen Kausalität zwei verschiedene konkurrenzlose Erklärungsbereiche mit eigenständigem Informationsgehalt unterschieden werden. Sie können in ihrer kausalen Verknüpfung zwischen Gehirnereignis und Umweltereignis getrennt adäquat beurteilt werden. Dabei bleiben für die neurowissenschaftliche Beschreibung das Problem der Vagheit und das der Privatheit bestehen, während sich für die Psychologie das Problem der neuronalen Aktivität selbst stellt (vgl. Kap. 6.6.1).

Diese wechselseitigen epistemischen Limitationen verweisen letztlich bereits auf die wechselseitige ontologische Relation. Beide Aspekte liefern einen erkenntnistheoretischen Mehrwert, der mehr Wissen über das tatsächliche ontologische mentale Ereignis liefert. Dieses kann jedoch nicht weiter differenziert werden. Denn auch für adäquate finale Kausalität gilt, dass die Eigenschaften komplementär sind. Eine physische Eigenschaft mit einem hohen Detailgrad ist zudem stets als determinate einer mentalen Eigenschaft mit hohem Detailgrad anzusehen. Beide können aber nur bedingt adäquate Ursache bzw. Wirkung eines Gehirnereignisses mit hohem Abstraktionsgrad sein. Hierzu bedarf es eines weiteren, finalen Stimulus mit einem Eigenschaftsbereich von hohem Abstraktionsgrad, auf den sich die Aktivität hin bewegt. Mentale Eigenschaften führen dabei als „gatekeeper“ zu der Vorraussetzung, dass das nachfolgende Ereignis instanziiert werden kann.

## Kapitel 8

## Schlussbemerkung und Ausblick

Zielsetzung der vorliegenden Dissertation ist es, ein Modell der mentalen Kausalität unter neurophilosophischen Aspekten bzw. unter Anwendung einer neurophilosophischen Methodik zu entwickeln. Die spezifisch neurophilosophische Betrachtungsweise soll für die Diskussion über das traditionell philosophische Problem der mentalen Kausalität den Rahmen erweitern und neue, alternative Lösungswege aufzeigen. So kann durch die Verknüpfung von neurowissenschaftlichen Beobachtungssätzen und philosophischer Theorie einerseits die empirische Validität der philosophischen Aussagen und andererseits das theoretische Fundament der empirischen Aussagen geprüft werden. Entscheidender und für die Diskussion eigentlich fruchtbar sind aber neue Aspekte, Anknüpfungspunkte und Fragestellungen, die sich aus dem Vergleich zwischen Beobachtungssätzen und Theoriesätzen ergeben (vgl. Kap. 6). Deshalb sind der neurophilosophische Vergleich und die daraus abgeleiteten Ergebnisse wichtige Teilziele der Arbeit. Aufgrund dieser neuen Gesichtspunkte konnte schließlich als neurophilosophische Hypothese das Modell der finalen Kausalität für mentale Kausalität entwickelt werden, das sowohl theoretische als auch empirische Anforderungen berücksichtigt. Um dieses Ziel plausibel mit einem erkennbaren Bezug zur tatsächlichen zeitgenössischen Diskussion über mentale Kausalität zu erreichen, mussten innerhalb der vorliegenden Arbeit verschiedene Zwischenschritte durchgeführt werden:

- Zunächst erfolgte die Einordnung in den theoretischen Kontext anhand der Darstellung zentraler Positionen, die die Frage der mentalen Kausalität entscheidend beeinflusst und weiterentwickelt haben. Diese ermöglicht die Ableitung der relevanten Voraussetzungen wie auch die Bestimmung der spezifischen Kriterien für mentale Kausalität.

- Durch die anschließende Analyse der Voraussetzungen für mentale Kausalität konnten der „Spielraum“ bzw. die bestehenden Möglichkeiten der mentalen Kausalität eingegrenzt werden. Diese Verortung der mentalen Kausalität innerhalb der diskutierten Voraussetzungen verdeutlichte über eine negative Bestimmung, dass der Begriff der mentalen Kausalität und die damit verbundenen Anforderungen in einem nicht auflösbaren Widerspruch zu eben jenen Voraussetzungen steht, die eine im eigentlichen Sinne mentale Kausalität nicht oder allenfalls sehr eingeschränkt zulassen.

- Deshalb erfolgte in einem ersten Kernkapitel die Entwicklung eigenständiger Kriterien für mentale Kausalität. Damit wird der Begriff der mentalen Kausalität konkretisiert, um aufzuzeigen, welche Teilaspekte für dieses mentale Phänomen relevant sind, wenn es denn als solches ernsthaft diskutiert werden soll. Diese eigenständige Verortung der mentalen Kausalität wurde ebenfalls unter Berücksichtigung der zeitgenössischen Diskussion durchgeführt, um somit eine solide Begriffsbestimmung zu gewährleisten. Die Entwicklung der Kriterien für mentale Kausalität hatte darüber hinaus den methodischen Zweck, für den neurophilosophischen Vergleich geeignete empirische Explikationen der Theoriesätze zu bestimmen.

- Als Gegenstück zu den empirischen Explikationen wurden zwei neurowissenschaftliche Studien in ihrer Durchführung und ihren Ergebnissen vorgestellt. Sie ermöglichen eine Bezugnahme zu den Kriterien der mentalen Kausalität und ergänzen sich wechselseitig. Die ausgewählten Studien liefern insofern Anknüpfungspunkte für die Frage der mentalen Kausalität, als sie methodisch zwischen einer neuronalen und einer subjektiven Ebene differenzieren und innerhalb ihres Studiendesigns Wahrnehmungs- und Handlungsparameter be-obachten. Damit wurden auf der empirischen Ebene zentrale Anforderungen, die mit mentaler Kausalität verbunden sind, beschrieben.

- In einem weiteren Kernkapitel wurde der Vergleich zwischen neurowissenschaftlichen Beobachtungssätzen und philosophischen Theoriesätzen bzw. zwischen den theoretischen Implikationen der Empirie und den empirischen Explikationen der Theorie durchgeführt. Neben der Prüfung der empirischen Validität der theoretischen Kriterien sowie der theoretischen Aussagekraft der empirischen Beobachtungssätze erfolgte eine Modifizierung und Erweiterung des Begriffs der mentalen Kausalität. Aufgrund der Ergebnisse des Vergleichs konnten theoretische und empirische Anforderungen der mentalen Kausalität bestimmt werden.

- Auf dieser Basis konnte schließlich als neurophilosophische Hypothese ein eigenständiges Modell der finalen Kausalität für mentale Kausalität entwickelt werden. Es bezieht sich wesentlich auf die in den beiden Kernkapiteln herausgearbeiteten Kriterien der mentalen Kausalität und die Ergebnisse des neurophilosophischen Vergleichs. Deshalb wird das Modell mit Bezug auf die Kriterien und die Ergebnisse des Vergleichs geprüft und diskutiert. Die Entwicklung der neurophilosophischen Hypothese kann dabei auch als Reaktion auf die etablierten Voraussetzungen für mentale Kausalität verstanden werden. Sie erweitert be-

wusst – unter Einbeziehung der aristotelischen Ursachenlehre – den Rahmen entgegen der vermeintlichen Homogenität zwischen mentaler und physischer Kausalität hin zu einer spezifisch mentalen Kausalität. Dennoch, oder besser: Gerade dadurch kann das Modell der finalen Kausalität für mentale Kausalität theoretische und empirische Aspekte berücksichtigen.

In dem Vergleich zwischen den neurowissenschaftlichen Paradigmen und dem Kriterium der Regularität wurde deutlich, dass mentale Kausalität als „Diener zweier Herren" zwar nicht strikten Gesetzen, aber zumindest ceteris-paribus-Gesetzen unterliegen kann. Dabei muss betont werden, dass die neurowissenschaftliche Praxis implizit von psychophysischen Gesetzen ausgeht und diese sogar benötigt, um Aussagen über beobachtete Korrelationen zu machen. Es muss von einer Gesetzmäßigkeit der empirischen Beobachtungen ausgegangen werden; allerdings können für die biologischen Prozesse im Gehirn auf einer empirischen Ebene in der Praxis eigentlich nur ceteris-paribus-Gesetze ermittelt werden. Um hierbei überhaupt von einem mental-kausalen Geschehen sprechen zu können, ist es erforderlich, von der Detailebene des neurobiologischen Prozesses bzw. des subjektiven Erlebens auf die beide Ebenen umfassende Ereignisebene zu wechseln. Mentale Kausalität muss als Relation zwischen drei Komponenten, der neuronalen Aktivität, dem subjektiven Erleben und zusätzlich äußeren Stimuli bestimmt werden. Deshalb muss auch eine ceteris-paribus-Regularität der mentalen Kausalität unter Einbeziehung aller drei Komponenten bzw. Axiome erfolgen.

In dem Vergleich mit den neurowissenschaftlichen Paradigmen wurde das Kriterium der Multikausalität für die subjektive wie für die neuronale Ebene bestätigt. Das subjektive Erleben erweist sich als äußerst nuanciert; unterschiedliche Eindrücke können einen subjektiven Zustand verursachen. Hintergrund ist die Unschärfe mentaler Begriffe. Neuronal kann zwar eine funktionsspezifische Stabilität einzelner Regionen beobachtet werden. Allerdings gilt auch hierfür, dass unterschiedliche Stimuli dieselbe neuronale Aktivität verursachen können. Unter dem Argument der Multirealisierbarkeit wird zudem deutlich, dass auf einer Mikroebene Neuronen, die biologisch identisch sein sollten, dennoch an unterschiedlicher Aktivität in verschiedenen Netzwerken beteiligt sind. Auch für das Kriterium der Multikausalität für mentale Kausalität ist entscheidend, dass die Ereignisebene, einschließlich äußerer Ereignisse, insgesamt betrachtet werden muss. Eine Erklärung allein aufgrund der physischen Eigenschaften ist nicht möglich; dies lassen die individuellen Ereignisunterschiede nicht zu. Die Möglichkeit einer inversen Multikausalität, wonach eine Ursache unterschiedliche Wirkungen haben kann, unterstützt diese Position zusätzlich. Die Entwicklung mentaler wie physischer Eigenschaften kann deshalb als eine kausale Realisierung durch externale Eindrücke bzw. Stimuli verstanden werden.

Der Vergleich mit den neurowissenschaftlichen Paradigmen bestätigt die zweifache Anforderung an das Kriterium Inhalt: Es soll Extramentales repräsentiert werden, und die Repräsentationen selbst sollen wiederum kausal wirksam sein. Damit werden im Grunde zwei Kernanforderungen der mentalen Kausalität formuliert, die sich jedoch aufgrund der gegebenen Voraussetzungen zu widersprechen scheinen. Die Diskussion zwischen den empirischen Beobachtungssätzen und den philosophischen Theoriesätzen zeigt eine erweiterte Perspektive auf. Zunächst konnte gezeigt werden, dass die Language-of-Thought-Hypothese zumindest für einfache Begriffe plausibel ist. Allerdings sind die verschiedenen Strukturen im Gehirn kein Beleg für eine syntaktische Gehirnsprache Mentalese; dieses ist nicht beobachtbar. Ebensowenig konnte die Unterscheidung zwischen engem und weitem Inhalt empirische Bestätigung finden. Dagegen ist es sinnvoll, verschiedene Inhaltstypen zu bestimmen. Neben der Differenzierung zwischen konkreten und abstrakten Inhalten ist eine Differenzierung zwischen aktuellen, erlernten und assoziativen Inhalten möglich. Für sie alle gilt, dass eine mentale Kausalität nicht einseitig monokausal, sondern als wechselseitige Modulation internaler wie externaler Prozesse verläuft. Insofern ist Wahrnehmung als aktives Interpretieren kausal ebenso auf extramentale Prozesse hin ausgerichtet wie Gedanken bzw. Dispositionen.

Der Vergleich zwischen den neurowissenschaftlichen Paradigmen und dem Kriterium der extrinsischen Relation betonte im Ergebnis die wesentliche Funktion des Gehirns als reizverarbeitendes Organ. Das Gehirn wurde konsequent als intrinsisch-relational bestimmt. Diese Bestimmung ist darauf zurückzuführen, dass weder die subjektive noch die neuronale Ebene als ausschließlich extrinsisch bzw. intrinsisch gekennzeichnet werden kann. Die Gleichsetzung neuronaler Aktivität mit intrinsischen Eigenschaften ist nicht haltbar. Entscheidend ist die Bestimmung, dass neuronale Aktivität intrinsische und extrinsische Eigenschaften umfasst. Die intrinsischen Eigenschaften gelten dabei als Ergebnis der extrinsischen Eigenschaften; beide Eigenschaftsbereiche bedingen sich wechselseitig: intrinsische Eigenschaften ermöglichen mentale Kausalität, extrinsische Eigenschaften geben die Zielrichtung vor, wodurch der kausale neuronale Prozess erst ein mentaler wird. Innerhalb der neuronalen Aktivität ist zudem eine Binnendifferenzierung bzw. eine Abstufung erforderlich. Tatsächliche mentale Kausalität kann demnach lediglich auf einer höherstufigen neuronalen Ebene mit vernetzten, aktivierten neuronalen Regionen verortet werden.

Der Vergleich zwischen den neurowissenschaftlichen Paradigmen und dem Kriterium der Intentionalität hat zunächst Schwächen des Modells der Indikatorrelation und des konsumentenorientierten Modells bestätigt. Wie sich bereits im Vergleich mit dem Kriterium Regularität gezeigt hat, ist eine gesetzmäßige Strenge nicht erfüllbar. Ebenso ist es problematisch, komplexe Begriffe und mentale Ereignisse allein evolutionstheoretisch als erfolgreiche und deshalb vorhandene Begriffe zu erklären. Viele Begriffe und mentale Ereignisse entziehen sich einer direkten

evolutionstheoretischen Überprüfung. Dagegen hat sich die zentrale Rolle der Normen für das Kriterium der Intentionalität und damit für mentale Kausalität bestätigt. Es kann sogar die vorsichtige Interpretation gewagt werden, dass es für soziale Normen auch eine neurologische Entsprechung geben muss. Die Kausalität darf nicht von der Normativität getrennt werden, ansonsten geht das Phänomen der Intentionalität verloren. Das bedeutet auch für mentale Kausalität, dass sie Normen unterliegt. Dies kann durch das Modell der Geist-auf-Welt- bzw. Welt-auf-Geist-Verursachung beschrieben werden; neben neuronalen Dispositionen konnte die emotionale Aufmerksamkeit als extremes Beispiel identifiziert werden. Gegebenenfalls muss eine faktische wie begriffliche Differenzierung zwischen emotionaler, kognitiver und motorischer Intentionalität und den damit verbundenen neuronalen Prozessen erfolgen. Dabei sind die physischen subvenienten Eigenschaften abhängig von mentalen supervenienten Eigenschaften. Denn erst letztere schaffen eine Verbindung zwischen der normativen Umwelt und dem Gehirn.

Im Vergleich mit den neurowissenschaftlichen Paradigmen erwies sich das Prinzip der Proportionalität als empirisch valide. So können tatsächlich zwei unterschiedliche Erklärungsbereiche mit eigenständigen, spezifischen Aussagen festgemacht werden – dies zeigt auch die neurowissenschaftliche Praxis. Die Unterteilung zwischen Abstraktionsebene und Detailebene konnte noch weiter in verschiedene vertikale Erklärungsebenen und horizontale Erklärungstypen differenziert werden. Das Prinzip der Proportionalität hat nicht die einzelne Korrelation im Blick, sondern die kausale Relation zwischen zwei Korrelationen. Dabei sind beide Ebenen miteinander verwoben. Niedrigstufige Eigenschaften sind auch abhängig von höherstufigen Eigenschaften, da sie die Zielrepräsentation sicherstellen. Denn mentale Repräsentationen verfügen über Kriterien, physische Ereignisse in ihrer Variabilität zu beschreiben, einschließlich alternativer neuronaler Verknüpfungspfade. Mentale Eigenschaften sind wiederum Bestandteil von Generalisierungen, die nicht in physischen Begriffen beschrieben werden können. Für mentale Kausalität ist deshalb die Adäquatheit der Instanziierung entscheidend; auf der horizontalen Ebene muss das gesamte Ereignis betrachtet werden. Neuronale Populationen müssen sich proportional zu äußeren Ereignissen verhalten. Neuronale Aktivität ist die Basis für subjektives, bewusstes Erleben. Die Instanziierung subjektiver, bewusster Eigenschaften ist wiederum kausal notwendig für weitere mentale Ereignisse. Mentale und physische Eigenschaften lassen sich ebenso wie eine höherstufige Abstraktionsebene und eine niedrigstufige Detailebene empirisch und epistemisch unterscheiden. Aufgrund ihrer Verschränkung bilden sie dennoch eine gemeinsame ontologische Relation.

Tatsächlich führt uns der Vergleich zwischen den empirischen Explikationen und den theoretischen Implikationen in eine neue Richtung. Um mentale Kausalität zu bestimmen, ist eine absolute Trennung zwischen mentalen und physischen Eigenschaften nicht zielführend. Natürlich ist eine Abgrenzung der Begriffe in einem ersten Schritt notwendig und sinnvoll. Wie sonst kann

bestimmt werden, welche Bereiche für eine mentale Kausalität überhaupt eine Rolle spielen. Entsprechend erfolgt zunächst eine Bestimmung der physischen Eigenschaften, ihrer entscheidenden Funktion generell in kausalen Prozessen und auch speziell für mentale Kausalität. Entsprechend muss aber auch die Bestimmung mentaler Eigenschaften einschließlich ihrer Funktion für speziell mentale Kausalität berücksichtigt werden. Solange wir uns auf einer epistemischen Ebene bewegen, gibt es keinen zwingenden Grund, von vornherein einen der beiden Bereiche auszuschließen. Die Argumente gegen die Funktion mentaler Eigenschaften für mentale Kausalität ergeben sich aus den angegebenen Voraussetzungen. Diese selbst sind jedoch angreifbar und lassen Raum für gegenläufige Positionen. Außerdem konzentrieren sie sich in der Hauptsache auf ihren positiven Untersuchungsgegenstand – die physischen Eigenschaften in einer physischen Welt – unter bewusster Ausklammerung und negativer Abgrenzung der mentalen Eigenschaften und der mentalen Welt. Diese ist aber ebenso epistemisch erfahrbar, erlebbar und sogar messbar, wenn auch auf andere Weise. Das zeigen deutlich die Kriterien für mentale Kausalität auf.

Mentale Kausalität ist ein echtes Phänomen, zentral begegnet es uns im Zusammenhang mit mentalen Inhalten innerhalb unserer propositionalen Einstellungen. Mentale Kausalität ist fester Bestandteil der menschlichen Lebenswirklichkeit innerhalb seiner Gedankenwelt, aber auch innerhalb seiner Handlungsmöglichkeiten. Mentale Kausalität ist schließlich der Kernprozess hinsichtlich der Relation zwischen Gehirn, Person und Umwelt. Auch Selbstreflexion, die als eine typisch menschliche Eigenschaft angesehen wird, fußt letztlich auf dieser Relation. Ohne die Möglichkeit, den allein physischen Prozess zu überschreiten, wäre ein solcher Prozess wohl kaum möglich. Oder sind neue Gedanken nur ein kombinatorischer Zufall? Lediglich den physikalischen Maßstab an mentale Ereignisse anzulegen, kann nur zu einer Beschränkung auf die physikalische Position führen. Allein mit physischen Angaben ist keine umfassende Aussage zu mentalen Prozessen leistbar. Natürlich kann dargestellt werden, dass z. B. ein Proband, dessen Neuronen in einem Areal höhere Aktivität erreichen, eine bestimmte motorische oder kognitive Leistung erbringt. Dies ist jedoch nur ein Teilaspekt und nicht die ganze Geschichte über seinen mentalen Prozess. Aber ein System, welches wichtige Aspekte nicht berücksichtigt, ist selbst angreifbar. Deshalb können wir, was das Phänomen der mentalen Kausalität betrifft, nicht bei der Differenzierung stehen bleiben. Wir müssen in einem nächsten Schritt zu einer Verbindung mentaler und physischer Eigenschaften kommen und beide epistemischen Bereiche zusammenführen. Auch das hat der Vergleich zwischen den empirischen Explikationen und den theoretischen Implikationen bestätigt: Es gibt verschiedene Schnittpunkte, die es ermöglichen, mentale und physischen Eigenschaften zusammenzuführen. Es ist nicht sinnvoll, mit Davidson darauf zu beharren, dass das Mentale anomal sei und es strikte psychophysische Gesetze nicht geben kann. Das mag zwar der Fall sein, aber um das Phänomen der mentalen Kausalität zu betrachten, ist deshalb nicht die Konzentration auf diese Ungleichheit hilfreich, sondern vielmehr die Betrach-

tung der Schnittstelle. Denn jenseits strikter Gesetze ist eine Regelmäßigkeit unter ceteris-paribus-Gesetzen möglich. Entscheidend für eine Definition der mentalen Kausalität ist das Kriterium des Inhalts.

Wenn wir von mentaler Kausalität sprechen, geht es nicht darum, rein neuronale Abläufe und Funktionen zu beschreiben. Vielmehr soll die Gedankenwelt einer Person und ihr Handeln aus dieser heraus begründet werden. Damit beschreibt der Prozess der mentalen Kausalität einmal die Versorgung des Gehirns mit externen Inhalten. Zusätzlich muss mentale Kausalität die Entstehung individueller, persönlicher bzw. gehirnspezifischer Gedanken beschreiben. Damit muss die Dichotomie zwischen intrinsischen und extrinsischen Eigenschaften sowie internalem und externalem Inhalt überwunden werden. Aus der Diskussion innerhalb des neurophilosophischen Vergleichs folgt dem entgegen die Verbindung der scheinbar gegensätzlichen Begriffe. Mentale Kausalität selbst bezeichnet die Schnittstelle zwischen dem mentalen und physischen Bereich, zwischen Gehirn und Person und zwischen Faktischem und Normativem. Mentale Kausalität ist ein Prozess, der dies alles in sich zusammenführt.

Vor dem Hintergrund dieser Anforderungen ist es beinahe zwangsläufig, ein Modell der finalen Kausalität für mentale Kausalität zu entwickeln. Schließlich berücksichtigt finale Kausalität allein aufgrund der Verbindung verschiedener gegenläufiger kausaler Richtungen von vornherein unterschiedliche Aspekte. Finale mentale Kausalität resultiert aus den kritischen Überlegungen der Voraussetzungen, Kriterien und des Vergleichs der mentalen Kausalität. Insofern kann sie als Modifizierung des Begriffs und des Modells der mentalen Kausalität betrachtet werden. In diesem neuen Modell werden wichtige Prinzipien zusammengeführt. So wird auf der epistemischen Ebene der Blick auf verschiedene Detailebenen beibehalten. Es kann zwischen Relationen auf der Mikroebene mentaler wie physischer Eigenschaften und zwischen Relationen von Gehirn- und Umweltereignissen – dem eigentlichen mentalen Ereignis – auf der Makroebene differenziert werden. Für die Realisierungsrelation zwischen mentalen und physischen Eigenschaften auf der vertikalen Ebene kann eine wechselseitige Abhängigkeit beschrieben werden. Der eigentliche Prozess einer mentalen Kausalität wird im Kern nur auf der horizontalen Ebene festgemacht. Dabei fällt den mentalen Eigenschaften insofern eine bedeutsame Rolle zu, als sie innerhalb der finalen Kausalität die Verbindung zwischen Gehirn und Umwelt gewährleisten. Sie sind die „gatekeeper“ die das mentale Ereignis ermöglichen. Als Eigenschaften neuronaler Prozesse sind sie dabei physisch geerdet. Dass neuronale Populationen auf bestimmte Umweltstimuli sich in einer spezifischen Art und Weise anordnen und miteinander vernetzen, kann wiederum als deren spezifisch mentale Eigenschaft beschrieben werden. Neuronale Aktivität ist damit die Instanziierung mentaler und physischer Eigenschaften.

Mentale Kausalität ist insofern eine antizipierende Kausalität, wobei es wichtig ist, den kausalen Prozess als biologischen Prozess zu begreifen. Das Gehirn ist mit seinen Eigenschaften ein

biologisches Organ. Deshalb verfügt es über mental und physisch beschreibbare Eigenschaften. Diese erhalten in Bezug auf Umweltereignisse ihre eigentliche Funktion. Deshalb wird das Modell als ein finales kausales Modell beschrieben. Die wechselseitige Relation zwischen Gehirn und Umwelt kann als eine stetige Weiterentwicklung des mentalen Gesamtereignisses bezeichnet werden. Denn ein mentales Ereignis ist letztlich diese Relation. Insofern sich beide bedingen, ist auch hierfür ein wechselseitiger Fortschritt und eine Steigerung des Niveaus insgesamt möglich. Das bedeutet, dass Erkenntnisfortschritt und biologische Weiterentwicklung innerhalb des Gehirns auch Auswirkungen auf die Wahrnehmung der Umwelt, der Umwelt selbst für das Gehirn und die Person hat.

Für die ontologische Ebene ist die Etablierung des kausalen Prozesses auf der Ereignisebene entscheidend. Denn aus einer neutralen epistemischen Perspektive heraus kann weder der mentalen noch der physischen Beschreibung Vorrang gegeben werden. Das Modell der finalen mentalen Kausalität berücksichtigt das Prinzip des epistemischen Primats (vgl. Kap. 1.3). Über den Ereignisbegriff wird eine weitestgehende Neutralität gegenüber mentalen oder physischen Eigenschaften gewahrt. Andererseits ist mit dem Ereignis die Instanziierung der Eigenschaften in einem konkreten raum-zeitlichen Gefüge verbunden, wodurch eine konkrete ontologische Relation bezeichnet werden kann. Ontologische Existenzannahmen verweisen selbst auf epistemische Voraussetzungen, da nur die epistemischen Fähigkeiten einen tatsächlichen Zugang zu Realität und Existenz ermöglichen. Insofern hängen die ontologischen Annahmen von epistemischen Fähigkeiten ab. Deshalb ist ein epistemisches Primat notwendig, um die Verknüpfung zwischen diesen epistemischen Fähigkeiten und jenen ontologischen Aussagen zu untersuchen. Diese Untersuchung kann im Rahmen der finalen mentalen Kausalität geleistet werden: „'epistemic primacy' focuses on the epistemological conditions for the possibility of ontological assumptions" (Northoff 2004, 43).[202]

Anstelle festgeschriebener potenzieller ontologischer Annahmen im Rahmen des ontologischen Primats können verschiedene ontologische Annahmen mit verschiedenen epistemischen Fähigkeiten in Verbindung gebracht werden. Die Bestimmung ontologischer Annahmen ist insofern nicht mehr im Vorhinein festgelegt, sondern orientiert sich an den epistemischen Fähigkeiten – es sind viele ontologische Annahmen möglich. „Many different world versions are of

[202] Das epistemische Primat muss grundsätzlich als eine methodische Strategie aufgefasst werden und nicht als eine eigenständige epistemische Position. Es kann sich insofern auf naturalistische, wie nicht-naturalistische Positionen beziehen. Andererseits beinhaltet es einen methodischen Naturalismus, der verschiedene physische, biologische, phänomenlogische oder andere Beobachtungen umfasst. Methodologischer Naturalismus beinhaltet ferner die Inklusion empirischer Beobachtungen in epistemischen und ontologischen Untersuchungen (vgl. Koppelberg 2000). Die Untersuchung epistemischer Fähigkeiten benötigt deshalb auch Überlegungen über empirische Beobachtungen. Ohne solche empirischen Beobachtungen ist eine sinnvolle Verknüpfung zwischen epistemischen Fähigkeiten und ontologischen Annahmen über „Realität und Existenz" nicht möglich (vgl. Northoff 2004, 41f.).

independent interest and importance, without any requirement or presumption to a single base." (Goodman 1978, 4) Ein ontologischer Pluralismus impliziert somit weder Eliminierung noch Bevorzugung einer Ontologie, das Feld möglicher ontologischer Annahmen wird dagegen erweitert. Entsprechend ist auch ein epistemischer Pluralismus ohne epistemische Hierarchie möglich. Schließlich ist die menschliche Perspektive stets eine eingeschränkte Perspektive. Die Frage einer unabhängigen Existenz der Welt muss damit für den ontologischen Pluralismus jedoch offen bleiben. Stattdessen konzentriert er sich auf die Untersuchung der epistemischen Fähigkeiten, die notwendig zu dieser Frage führen. Durch diese Vorgehensweise ist auch der Fokus des Modells der finalen mentalen Kausalität bestimmt. Durch die Integration des mentalen wie des physischen Erklärungsbereiches kann auf der epistemischen Ebene ein Erkenntnisgewinn und ein erkenntnistheoretischer Mehrwert erreicht werden. Dadurch werden weiterführende Rückschlüsse auf die ontologische Ebene zumindest offen gehalten, die bei einer Priorisierung z. B. des physikalischen Erklärungsmusters verschlossen würden.

Weitere Forschungsmöglichkeiten ergeben sich aus der vorliegenden Arbeit aus der Diskussion der Voraussetzung, der Aufstellung von Kriterien für mentale Kausalität, aus dem Vergleich zwischen den neurowissenschaftlichen Paradigmen und den theoretischen Kriterien sowie aus dem Modell der finalen Kausalität selbst. So ist hinsichtlich der einzelnen Voraussetzungen deren Überprüfung sowie die der diskutierten neuen Teilergebnissen von Interesse. Beispielsweise kann die Abhängigkeit der Inhaltseigenschaften von neurophysiologischen Eigenschaften in Relation zu Umwelteigenschaften untersucht werden. Daneben ist von Bedeutung, inwiefern neuronale Prozesse als Ereignis oder als Eigenschaft beschrieben werden können. Eine weitere Frage betrifft das Verhältnis von Eigenschaftsinstanziierung und neuronaler Prozessierung. Inwiefern, müssen wir uns zudem fragen, ist die Unterscheidung zwischen physischen und mentalen Eigenschaften tatsächlich durchzuhalten und vor allem auf welche empirische Eigenschaft anzuwenden, wenn neuronale Aktivität beide Eigenschaftsbereiche umfassen soll? Ist entgegen eines homogenen Kausalmodells eine Neudefinition für eine biologische bzw. mentale Kausalität erforderlich? Damit muss auch geprüft werden, inwiefern mentale Ereignisse bzw. Eigenschaften tatsächlich aus dem physikalischen System herausfallen. Für die Physik mögen hier abstrakte Phänomene, wie z. B. der Informationsbegriff, eine Herausforderung darstellen. Ferner können Möglichkeiten neuronaler Mehrfachverursachung und ihre Implikationen für mentale wie physische Eigenschaften diskutiert werden. Schließlich wäre zu untersuchen, welche Rolle ceteris-paribus-Gesetze für neuronale Aktivität bzw. auf verschiedenen Ebenen neuronaler Aktivität (einzelnes Neuron, Neuronenensemble, neuronale Verknüpfung) oder für subjektive Wahrnehmung spielen.

Hinsichtlich der Kriterien für mentale Kausalität erscheint eine weiterführende und eingängige Untersuchung der einzelnen Kriterien unter empirischen Konditionen relevant. Dies wird auch in dem Vergleich zwischen empirischen Beobachtungssätzen und den theoretischen Kriterien deutlich, die daran anknüfpen. Beispielsweise ergibt sich die Frage, ob für einen biologischen Prozess ceteris-paribus-Gesetze zur Beschreibung der Regularität besser geeignet sind als so genannte strikte Gesetze. Zudem kann hierbei die Aufteilung in drei Gesetzesarten überprüft werden. Weitere Forschungsperspektiven ergeben sich aus der Unterscheidung zwischen verschiedenen Detailebenen für neuronale Aktivität, wie dem einzelnen Neuron, einer neuronalen Population oder der Vernetzung verschiedener Neuronenensembles. Inwiefern lässt sich die vorgenommene Binnendifferenzierung zusätzlich bestätigen? Hierfür ist unter anderem die Frage zu klären, ob es mögliche Grenzwerte gibt, ab denen von einem spezifisch neuronalen-mentalen Ereignis gesprochen werden kann. Ferner muss die Relation der drei Komponenten subjektives Erleben, neuronale Aktivität und Umwelt für mentale Kausalität eingehend untersucht werden. Ebenso ist eine weitergehende Untersuchung des Prinzips der Multirealisierbarkeit und Multikausalität für neuronale mentale Prozesse möglich. Inwiefern ist eine Abgrenzung individueller Gedanken gegenüber strukturell übergeordneten Merkmalen möglich und überhaupt notwendig?

Insbesondere die zentrale Stellung des Kriteriums „Inhalt“ ist für weitere Forschung von Bedeutung: Die semantische Herausforderung für die syntaktische Struktur bleibt bestehen. Neuronale Aktivität muss letztlich beides umfassen. Aber auch die Ätiologie von Inhalt kann in diesem Zusammenhang weiter untersucht werden. Dabei hat sich gezeigt, dass die Unterscheidung zwischen intrinsischen und extrinsischen Eigenschaften sowie internalem und externalem Inhalt vom eigentlichen Phänomen weg führen. Dennoch kann deren Relevanz und natürlich die Position, wonach das Gehirn intrinsisch-relational ist, zusätzlich empirisch untersucht werden. Lassen sich die verschiedenen Inhaltstypen genauso wie die verschiedenen Arten von Intentionalität, die im Rahmen des Vergleiches entwickelt wurden, empirisch validieren? Schließlich ist auch eine nähere Untersuchung der wechselseitigen Abhängigkeit mentaler und physischer Eigenschaften als proportionale Eigenschaften auf einer hohen Abstraktions- bzw. niedrigen Detailebene notwendig. Grundsätzlich bietet diese Arbeit einen eher breiten Ausgangspunkt, aus dem, vor allem hinsichtlich der Kriterien, spezifische weitere tiefergehende empirische wie theoretische Forschung möglich ist.

Das Modell der finalen mentalen Kausalität ermöglicht nun einen größeren Beschreibungsrahmen als bisherige Ansätze zu Beschreibung mentaler Kausalität. Denn in dem Modell der finalen mentalen Kausalität werden nicht nur mentale und physische Eigenschaften, sondern darüber hinaus auch Gehirn, Person und Umwelt in dem kausalen Prozess integriert. Damit steht ein Modell zur Verfügung, das über einfache kausale Prozesse hinausgeht und eine Beschreibung

komplexer Prozesse ermöglicht. Mögliche Beispiele wären hierfür die klinischen Syndrome der Katatonie und der Schizophrenie.

So beschreibt die Katatonie als unspezifisches Syndrom sowohl neurologisch-organische als auch psychiatrisch-funktionelle Erkrankungen. Erstaunlicherweise liegen gerade für letztere Erkrankungen keine pathologischen Befunde vor (vgl. Northoff 1992, 23). Es ist vor allem der Zusammenhang zwischen motorischen, kognitiven und affektiven Komponenten, die das Syndrom der Katatonie kennzeichnen. Es besteht eine Verflechtung zwischen motorischen und psychischen Symptomen. Auch wenn hinsichtlich der Variante der motorischen Erkrankung biologische und hirnpathologische Erklärungen relevant werden, ist eine generelle Lokalisierung der Katatonie innerhalb einer bestimmten Hirnregion nicht möglich (vgl. Northoff 1997, 58). Die Erkrankung ist zudem situationsabhängig. Das Umfeld und äußere Faktoren haben entscheidenden Einfluss auf den Verlauf der Krankheit. Die Katatonie nimmt deshalb eine Zwischenstellung zwischen Neurologie und Psychiatrie ein (vgl. Kindt 1980, 60). Das Modell der finalen mentalen Kausalität bietet nun die Möglichkeit, solche Erkrankungen, wie das unspezifische Syndrom der Katatonie, zu beschreiben und dabei die Dichotomie zwischen psychischen und motorischen Störungen zu berücksichtigen. Zusätzlich können äußere Faktoren, wie Veränderungen in der Umwelt oder mit Stress verbundene Ereignisse, berücksichtigt werden. Motorische Funktionen werden nicht isoliert betrachtet, sondern im Zusammenhang eines psychologischen Kontexts.

Eine ähnliche Anwendung des Modells wäre auch mit Bezug auf die Erkrankung der Schizophrenie denkbar. Auch hierbei handelt es sich um eine Erkrankung mit einer mannigfaltigen Symptomatologie, für deren Auftreten sowohl genetisch-biologische als auch psychosoziale Ursachen im Wechselspiel Auslöser sein können (vgl. Bleuler 1978; Schneider 1992). Dabei stellt sich auch die Frage, ob biochemische Veränderungen Ursache oder Folge der Schizophrenie sind. Denkbar wäre ein Teufelskreis, wonach aufgrund einer genetischen Disposition durch ein akutes Umweltereignis die Krankheit ausgelöst wird und weitere biochemische Veränderungen verursacht (vgl. Arieti 2006, 102). Ähnlich wie bei der Katatonie scheinen extreme bzw. belastende Lebenssituationen zumindest Teilursache für eine akute schizophrene Phase zu sein. Hinsichtlich des Modells der finalen mentalen Kausalität ist dabei beispielsweise das Symptom der Projektion interessant. Der Patient projiziert auf seine (soziale) Umwelt ein System falscher Überzeugungen: Eigenschaften, Gefahren und Verschwörungen werden als real erlebt und sind für den Patienten quälende Fakten (vgl. ders., 55). Dabei entspringen diese Projektionen der Gedankenwelt des Patienten. Sie beinhalten insofern einen realen Kern, als sie für den Patienten eine Bedeutung haben (vgl. ders., 66). Ereignisse in der Umwelt werden deshalb in einem bestimmten, patientenspezifischen Sinn gedeutet – sie sind ein mentales Ereignis.

Auch hinsichtlich komplexer abstrakter Begriffe ermöglicht das Modell der finalen mentalen Kausalität einen erweiterten Zugang. Ein Begriff wie die Liebe, eine Ideologie oder eine

ethische Anweisung, wie z. B. Kants kategorischer Imperativ, muss nicht allein über evolutionstheoretische Aspekte begründet werden. Ein allein unidirektionales physisches Kausalitätsmodell kann diese Syndrome bzw. Krankheiten nicht angemessen beschreiben, da es sich hauptsächlich auf neuronale Beobachtungen und Messungen beziehen müsste. Durch das Modell der finalen mentalen Kausalität ist dagegen die Einbeziehung sowohl physischer als auch psychischer Gesichtspunkte möglich. Es ist schwer einsichtig, wie abstrakte Gedanken als mentale Ereignisse ausschließlich durch neuronale Verknüpfungen instanziiert werden können. Unter der Idee der finalen mentalen Kausalität kann dagegen beschrieben werden, wie neuronale Verknüpfungen im Hinblick auf soziale bzw. normative Ereignisse entwickelt werden und diese wechselseitig fortführen. Das Gehirn richtet sich nach dem normativen Umweltereignis hin aus, überschreitet dieses im Prozessverlauf möglicherweise und bildet schließlich seine eigene neuronal-mentale Wirklichkeit.

Auf einer empirischen Ebene ist ferner eine Verknüpfung des Modells der finalen mentalen Kausalität mit dynamischen Ansätzen innerhalb der Neurowissenschaft möglich. Solche Ansätze betonen die dynamische Komponente der Hirnaktivität und nehmen an, dass die entscheidenden Prozesse für Bewusstsein die Einheiten Gehirn-Körper-Umwelt umfassen und nicht hauptsächlich neuronalgebundene Gehirnereignisse sind. Bewusste mentale Prozesse werden dabei als emergente Prozesse kollektiver Variablen bzw. Parameter verstanden, die als Ursachen kontextsensitiv das gesamte Gehirn als System strukturieren. Damit haben sie als emergente Phänomene kausale Wirkung auf lokale neuronale Aktivität. Entscheidend für die Hirnaktivität ist dabei die Einbettung der neuronalen Prozesse in körperliche bzw. Umweltprozesse (vgl. Thompson/Varela 2001).

Schließlich kann das Modell der finalen mentalen Kausalität, wie oben dargestellt, weitere Anwendung auf empirischer wie theoretischer Ebene finden. Dabei sind auch weitere Anwendungsbereiche im soziologischen oder medienwissenschaftlichen Bereich jenseits der Neurophilosophie, der Neurowissenschaft oder der Philosophie denkbar. In diesen Forschungsbereichen könnte das Modell der finalen mentalen Kausalität in einem interdisziplinären Ansatz hinsichtlich der Auswirkung medialer Umwelteinflüsse oder sozialer Prägung auf Person und Gehirn Anwendung finden. Mögliche Fragestellungen könnten hierbei sein, inwiefern sich das Gehirn und damit die Persönlichkeit unter einem kontinuierlichen Einfluss spezifischer Medien verändert, einschließlich daraus resultierender weiterer Dispositionen. Da das Modell der finalen Kausalität zusätzlich berücksichtigt, dass sich das Gehirn gewissermaßen seine Umwelt sucht, wäre hierbei auch zu untersuchen, welche Konsequenzen sich aus einem spezifischen Medienkonsum hinsichtlich einer Spezialisierung und Individualisierung von Persönlichkeit, individuellem Nutzungsverhalten und Flexibilität gegenüber neuen Umweltsituationen ergeben.

Teilweise würden solche Forschungsansätze in den Bereich der Technikphilosophie fallen, soweit sie eine Veränderung einer Person und ihres Gehirns aufgrund der Anwendung einer Technik untersuchten. Verändert sich beispielsweise die Struktur des Gehirns durch die Benutzung eines Mobiltelefons, und verändert sich dadurch die Persönlichkeitsstruktur insgesamt? Jemand mit einem Mobiltelefon ist ein anderer Mensch als jemand ohne ein Mobiltelefon – in seiner neuronalen Struktur, in seiner Selbstwahrnehmung und schließlich für sein soziales Umfeld. In ähnlicher Weise könnte auch die Veränderung von Gehirn, Person und Umwelt durch Implantate untersucht werden. Daneben wäre auch eine Untersuchung der Auswirkung von besonderen Reiseerlebnissen bzw. dem konsequenten Wechsel des Umfeldes von Bedeutung. Weitere wichtige Fragen betreffen natürlich die Unterscheidung zwischen basalen und abstrakten Inhalten für die psychologische wie die neuronale Struktur einer Person.

In der zeitgenössischen analytischen Philosophie des Geistes konzentriert sich der vorherrschende Trend auf den physischen Bereich, das heißt auf die physische Beschreibung von Ereignissen und Eigenschaften mit der damit verbundenen ontologischen Priorisierung einer physischen Realität. Im Verlauf der vorliegenden Arbeit wurde aufgezeigt, dass diese Position hinsichtlich mentaler Ereignisse bzw. mentaler Kausalität nicht unstrittig ist. Zumindest auf einer epistemischen Ebene gibt es gute Gründe dafür, den mentalen Bereich zusätzlich zum physischen ernst zu nehmen und weiterhin zu untersuchen. Eine bloße Konzentration auf den physischen Bereich führt letztlich zu einem Verlust wichtiger Informationen über die menschliche Gedankenwelt. Ein System, welches diesen zentralen Bereich nicht einbezieht, kann nicht als vollständig und hinreichend erklärend betrachtet werden. Hinsichtlich der ontologischen Ebene bleibt diese Arbeit im Ergebnis zurückhaltend. Vor dem Hintergrund des epistemischen Primats sowie eines ontologischen Pluralismus sollte hinsichtlich mentaler Kausalität eine ontologische Neutralität gewahrt werden. Immerhin wird mit dem Modell der finalen mentalen Kausalität ein Modell entwickelt, dass ontologisch dieser Neutralität gerecht wird und epistemisch eine differenzierte Integration beider Bereiche verfolgt. Dabei wird berücksichtigt, dass die Entwicklung propositionaler Einstellungen, neuronaler Aktivität bzw. die Prozessierung von Gedanken ein biologischer Prozess ist.

Natürlich sind für diesen Prozess physische Eigenschaften notwendig und erforderlich. Dass wir Menschen in der tatsächlichen aktuellen Welt ohne unser Gehirn nicht denken könnten, kann nicht ernsthaft bestritten werden. Andererseits ist zweifelhaft, ob auf einer rein materiellen Ebene echtes menschliches Denken nachgebildet werden kann. Durch die immensen Fortschritte der Informationstechnologie sind solche Projekte inzwischen möglich. Beispielsweise wird an der Technischen Hochschule von Lausanne auf einem Rechner mit nahezu 10.000 Computerchips, die sich wie Nervenzellen verhalten, ein künstliches Gehirn nachgebaut. Ziel des so genannten

Projekts „Blue Brain“ ist die vollständige Simulation des menschlichen Gehirns und damit die Erforschung von Bewusstsein (vgl. http://bluebrain.epfl.ch/; letzter Abruf: 20.10.2008). Sicherlich wird man hieran die elektrophysischen Prozesse neuronaler Aktivität weiterführend untersuchen können. Dabei bleibt jedoch offen, ob es sich um tatsächliches Denken oder lediglich um die Simulation von Denken handelt. Genauso wenig wie wir sagen können, wie es ist, eine Fledermaus zu sein, können wir Auskunft darüber erhalten, wie es ist, ein Computergehirn zu sein. Der Kybernetiker Noel Sharkey geht ausdrücklich davon aus, dass zukünftig leistungsstarke, komplexe Roboter entwickelt werden können, die Menschlichkeit jedoch nur in der Anmutung simulieren können (vgl. Sharkey 2002). Aber auch für diese komplexen Computer- bzw. Robotergehirne bleibt die Frage der mentalen Kausalität virulent: Die Relation zwischen Siliziumchips, dem subjektivem Erleben und der Umwelterfahrung – die Entstehung und Prozessierung von Inhalt – wird nicht aufgehoben. Um mit Kant zu sprechen, ist ein organisiertes Wesen nicht bloß Maschine, da diese lediglich eine bewegende Kraft habe. Ein organisiertes Wesen besitzt zusätzlich in sich eine bildende Kraft, die es den Materien mitteilt (vgl. Kant 2006, 280f.). Die Untersuchung der Organisation als bildende Kraft bleibt für denkende Menschen, Roboter und andere Systeme zentral, um ihre mentale Kausalität zu bestimmen.

**Literaturverzeichnis**


Alexander, S. (1927): Space, Time, and Deity. 2 Bde., 2. Aufl., London.

Amedi, A./Jacobson, G./Hendler, T./Malach, R./Zohary, E. (2002): Convergence of visual and tactile shape processing in the human lateral occipital complex. Cereb. Cortex 12, 1202-1212.

Anders, S./Lotze, M./Erb, M./Grodd, W./Birbaumer, N. (2004): Brain activity underlying emotional valence and intensity: a response-related fMRI study. Hum. Brain Mapp. 23, 200-209.

Anderson, A. K./Christoff, K./Panitz, D./De Rosa, E./Gabrieli, J. D. (2003a): Neural correlates of the automatic processing of threat facial signals. J. Neurosci. 23, 5627-5633.

Anderson, A. K./Christoff, K./Stappen, I./Panitz, D./Ghahremani, D. G./Glover, G./Gabrieli, J. D./Sobel, N. (2003b): Dissociated neural representations of intensity and valence in human olfaction. Nat. Neurosci. 6, 196-202.

Anderson, A. K./Sobel, N. (2003): Dissociating intensity from valence as sensory inputs to emotion. Neuron 39, 581-583.

Antony, L. M. (1994): The inadequacy of Anomalous Monism as a Realist Theory of Mind. In: Preyer, G./F. Siebelt/A. Ulfig (Hrsg.): Language, Mind and Epistemology. On Donald Davidson's Philosophy, Kluwer 1994, 223-255.

Arieti, S. ((2006): Schizophrenie. Ursachen, Verlauf, Therapie, Hilfen für Betroffene, 9. Aufl., München.

Aristoteles (1978): Metaphysik. 1. Halbbd., Bücher I-VI, 1. Aufl., Hamburg.

Aristoteles (1980): Metaphysik. 2. Halbbd., Bücher VII-XIV, Hamburg.

Aristoteles (1987): Physik. 1. Halbbd., Bücher I-IV, Hamburg.

Aristoteles (1988): Physik. 2. Halbbd. Bücher V-VIII, Hamburg.

Armstrong, D. M. (1986): A Materialist Theory of Mind. New York.

Audi, R. (1999): Mental Causation: Sustaining and Dynamic. In: Heil, J./Mele, A. (Hrsg.): Mental causation. Oxford 1999, 53-74.

Baker, L. R. (1987): Saving the Belief: A Critique of Physicalism. Princeton.

Baker, L. R. (1993): Eliminativism and an argument from science. Mind and Language, 8, 180-88.

Baker, L. R. (1995): Propositional Attitudes. In: Guttenplan, S. (Hrsg.): A Companion to the Philosophy of Mind. Blackwell 1995, 488-493.

Baker, L. R. (1999): Metaphysics and Mental Causation. In: Heil, J./Mele, A. (Hrsg.): Mental causation. Oxford 1999, 75-95.

Barrett, L. F./Russell, J. A. (1999): The structure of current affect: controversies and emerging consensus. Curr. Dir. Psychol. Sci. 8, 10-14.

Beauchamp, T. (1991): Hume and the problem of causation. Oxford.

Becerra, L./Breiter, H. C./Wise, R./Gonzalez, R. G./Borsook, D. (2001): Reward circuitry activation by noxious thermal stimuli. Neuron 32, 927-946.

Beckermann, A. (1999): Analytische Einführung in die Philosophie des Geistes. Berlin.

Bennett, M. R./Hacker, P. M. S. (2003): Philosophical Foundations of Neuroscience. Blackwell Publishers.

Berlin, H. A./Rolls, E. T./Kischka, U. (2004): Impulsivity, time perception, emotion and reinforcement sensitivity in patients with orbitofrontal cortex lesions. Brain 127, 1108-1126.

Bermpohl, F./Pascual-Leone, A./Amedi, A./Merabet, L. B./Fregni, F./Gaab, N./Alsop, D./Schlaug, G./Northoff, G. (2006): Dissociable networks for the expectancy and perception of emotional stimuli in the human brain. NeuroImage 30, 588-600.

Bickle, J. (1998): Psychoneural Reduction: The New Wave. Cambridge.

Bickle, J. (2003): Philosophy and Neuroscience. A Ruthlessly Reductive Account, Dordrecht.

Bieri, P. (Hrsg.), (1981): Analytische Philosophie des Geistes. Königstein.

Birnbacher, D. (1990): Das ontologische Leib-Seele-Problem und seine epiphänomenalistische Lösung. In: Bühler, K-E. (Hrsg.): Aspekte des Leib-Seele-Problems. Philosophie, Medizin, Künstliche Intelligenz. Würzburg 1990, 59-79.

Birnbacher, D. (2006): Causal interpretations of correlations between neural and conscious events. In: Journal of consciousness studies, 13, 115-128.

Bleuler, E. (1978): Dementia praecox oder Gruppe der Schizophrenien. Un-veränd., autoris. Wiederabdruck. d. Ausg. Leipzig u. Wien 1911, München.

Block, N. (Hrsg.), (1980a): Readings in Philosophy of Psychology. Vol. 1, Cambridge.

Block, N. (1980b): Introduction: What is Functionalism? In: Block, N. (Hrsg.): Readings in Philosophy of Psychology. Vol. 1, Cambridge 1980, 171-184.

Block, N. (1980c): Troubles with Functionalism. In: Savage, C. W. (Hrsg.): Perception and Cognition. Issues in the Foundations of Psychology, Minnesota Stu-dies in the Philosophy of Science, vol. 9, Minneapolis 1978, 261-325. Wiederabdruck in: Block, N. (Hrsg.): Readings in Philosophy of Psychology. Vol. 1, Cambridge 1980, 268-307.

Block, N. (Hrsg.), (1981): Readings in Philosophy of Psychology. Vol. 2, Cambridge.

Block, N. (1986): Advertisement for a Semantics for Psychology. Midwest Studies in Philosophy, Vol. X. Minneapolis, 615-678.

Block, N. (1990): Can the mind change the world? In: Boolos, G. (Hrsg.): Meaning and method. Essays in Honor of Hilary Putnam, Cambridge 1990, 137-170.

Block, N. (1997): Anti-Reductionism Slaps Back. Philosophical Perspectives 11, 107-133.

Block, N. (2003): Do Causal Powers Drain Away? Philosophy and Phenomenological Research Vol. 67, 1, 133-150.

Block, N./Fodor, J. (1980): What Psychological States Are Not. Philosophical Review 81, 2 (1972), 159-181. Wiederabdruck in: Block, N. (Hrsg.): Readings in Philosophy of Psychology. Vol. 1, Cambridge 1980, 237-251.

Bontly, T. D. (2002): The Supervenience Argument Generalizes. Philosophical Studies 109 (2002), 75-96.

Braddon-Mitchell, D./Jackson, F. (1996): The Philosophy of Mind and Cognition. Oxford.

Brand, M. (Hrsg.), (1976): The Nature of Causation. Chicago.

Brandt, R./Kim, J. (1963): Wants as Explanations of Actions. Journal of Philosophy, Vol. 60, 425-435.

Brehm, J. W. (1966): A theory of psychological reactance. New York/London

Breiter, H. C./Aharon, I./Kahneman, D./Dale, A./Shizgal, P. (2001): Functional imaging of

neural responses to expectancy and experience of monetary gains and losses. Neuron 30, 619-639.

Breiter, H. C./Gollub, R. L./Weisskoff, R. M./Kennedy, D. N./Makris, N./Berke, J. D./Goodman, J. M./Kantor, H. L./Gastfriend, D. R./Riorden, J. P./Mathew, R. T./Rosen, B. R./Hyman, S. E. (1997): Acute effects of cocaine on human brain activity and emotion. Neuron 19, 591-611.

Brentano, F. (1874): Psychologie vom empirischen Standpunkt. Leipzig.

Broad, C. D. (1925): The Mind and Its Place In Nature. London.

Bröcker, W. (1935): Aristoteles. Frankfurt a. M.

Buchel, C./Dolan, R. J./Armony, J. L./Friston, K. J. (1999): Amygdala – hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. J. Neurosci. 19, 10869-10876.

Buchel, C./Morris, J./Dolan, R. J./Friston, K. J. (1998): Brain systems mediating aversive conditioning: an event-related fMRI study. Neuron 20, 947- 957.

Bunge, M. (1987): Kausalität, Geschichte und Probleme. Tübingen.

Burge, T. (1979): Individualism and the Mental. Midwest studies in Philosophy, Vol. 4, 73-121.

Burge, T. (1982): Other Bodies. In: Woodfield, A. (Hrsg.): Thought and Object. Oxford 1982.

Burge, T. (1986): Individualism and Psychology. Philosophical Review 45, 3-45.

Burge, T. (1988): Individualism and self-knowledge. Journal of philosophy 85, 11, 644-663.

Burge, T. (1989): Individualism and Causation in Psychology. Pacific Philosophical Quarterly 70, 303-322.

Burge, T. (1999): Mind-Body Causation and Explanatory Practice. In: Heil, A./J. Mele (Hrsg.): Mental causation. Oxford 1999, 97-120.

Burgess, A. (1975): A Clockwork Orange. London.

Bush, G./Luu, P./Posner, M. I. (2000): Cognitive and emotional influences in anterior cingulate cortex. Trends Cogn. Sci. 4, 215-222.

Bush, G./Vogt, B. A./Holmes, J./Dale, A. M./Greve, D./Jenike, M. A./Rosen, B. R. (2002): Dorsal anterior cingulate cortex: a role in reward-based decision making. Proc. Natl. Acad. Sci. U. S. A. 99, 523-528.

Canli, T./Desmond, J. E./Zhao, Z./Glover, G./Gabrieli, J. D. (1998): Hemispheric asymmetry for emotional stimuli detected with fMRI. NeuroReport 9, 3233-3239.

Carlsson, K./Petrovic, P./Skare, S./Petersson, K. M./Ingvar, M. (2000): Tickling expectations: neural processing in anticipation of a sensory stimulus. J. Cogn. Neurosci. 12, 691-703.

Carnap, R. (1966): Der logische Aufbau der Welt. 3. Aufl., Hamburg.

Cartwright, N. (1983): How the Laws of Physics Lie. Oxford.

Cartwright, N. (1999): The Dappled World. Cambridge.

Chalmers, D. (1994): On Implementing a Computation. Minds and Machines 4, 391-402.

Chalmers, D. (1997): The Conscious Mind. New York.

Chisholm, R. (1957): Perceiving. Ithaca, N. Y.

Churchland, P. M. (1981): Eliminative Materialism and propositional attitudes. Journal of Philosophy 78, 67-90.

Churchland, P. M. (1984): Matter and Consciousness. MIT Press.

Churchland, P. M. (1986): Some Reductive Strategies in Cognitive Neurobiology. Mind 95, 279-309.

Churchland, P. M. (1989): A neurocomputational perspective: The nature of mind and the structure of science. Cambridge.

Churchland, P. M. (1997): Die Seelenmaschine. Eine philosophische Reise ins Gehirn, Heidelberg.

Churchland, P. M./Churchland, P. S. (1981): Functionalism, qualia and intentionality. Philosophical Topics 12, 121-132.

Churchland, P. M./Churchland, P. S. (2001): Intertheoretic reduction: A neuroscientist's guide. In: Bechtel, W./Mandik, P./Mundale, J./Stufflebeam, R. S. (Hrsg.), (2001): Philosophy and the neurosciences. Oxford, 419-431.

Churchland, P. S./Sejnowski, T. J. (1992): The computational brain. Cambridge.

Colingwood, R. (1946): The Idea of History. Oxford.

Crane, T. (1991): Why indeed? Papineau on supervenience. Analysis 51, 32-37.

Crane, T. (1992): Mental causation and mental reality. Proceedings of the Aristotelian Society 66, 185-202.

Crane, T. (1995): The mental causation debate. Aristotelian Society Supplement 69, 211-236.

Crane T./Mellor, D. H. (1990): There is no question of physicalism. Mind 99, 185-206.

Crick, F. (1994): Was die Seele wirklich ist: die naturwissenschaftliche Erforschung des Bewusstseins. München.

Crick, F. (1995): The Astonishing Hypothesis. The Scientific Search for the Soul, London.

Critchley, H. D./Mathias, C. J./Dolan, R. J. (2001): Neural activity in the human brain relating to uncertainty and arousal during anticipation. Neuron 29, 537-545.

Cummins, R. (2000): Reply to Millikan. Philosophy and Phenomenological Research 60, 113-127.

Darwin, Ch. (1872): On the origin of species. London 1859. Neuauflage unter dem Titel: The Origin of Species. 6. Aufl., London.

David, M. (1997): Kim's Functionalism. Philosophical Perspectives 11, 133-149.

Davidson, D. (1963): Actions, Reasons and Causes. Journal of Philosophy 60. Deutsche Übersetzung: Handlungen, Gründe und Ursachen. In: Davidson, D. (1998): Handlung und Ereignis. 2. Aufl., Frankfurt a. M., 19-42.

Davidson, D. (1967): Causal Relations. Journal of Philosophy 64, 691-703.

Davidson, D. (1969): The Individuation of Events. In: Rescher, N. (Hrsg): Essays in Honor of Carl G. Hempel. Dordrecht 1969, 216-234. Deutsche Übersetzung: Zur Individuation von Ereignissen. In: Davidson, D. (1998): Handlung und Ereignis. 2. Aufl., Frankfurt a. M., 233-258.

Davidson, D. (1970): Mental Events. In: Foster, L./Swanson, J.: Experience and Theory. London 1970, 79-101. Deutsche Übersetzung: Geistige Ereignisse. In: Davidson, D.: Handlung und Ereignis. 2. Aufl., Frankfurt a. M. 1998, 291-317.

Davidson, D. (1974): Psychology as Philosophy. In: Brown, S. C. (Hrsg.): Philosophy of Psychology. London/New York. Deutsche Übersetzung: Psychologie als Philosophie. In: Davidson, D. (1998): Handlung und Ereignis. 2. Aufl., Frankfurt a. M., 321-342.

Davidson, D. (1982): Rational Animals. Dialectica 36, 317-328.

Davidson, D. (1986a): Wahrheit und Interpretation. Frankfurt a. M.

Davidson, D. (1986b): A Coherence Theory of Truth and Knowledge. In: LePore, E. (Hrsg.): Truth and Interpretation. Perspectives on the Philosophy of Donald Davidson, Oxford 1986, 307-319.

Davidson, D. (1987): Knowing One's Mind. Proceedings and Addresses of the American Philosophical Association 60, 441-458.

Davidson, D. (1989): The Myth of the Subjective. In: Krausz, M. (Hrsg.): Relativism: Interpretation and Confrontation. Notre Dame 1989, 159-172.

Davidson, D. (1990): Representation and Interpretation. In: Newton-Smith, W. H. (Hrsg.): Modelling the Mind. Oxford 1990, 13-26.

Davidson, D. (1993): Meaning, Truth, and Evidence. In: Der Mythos des Subjektiven. Stuttgart, 40-64.

Davidson, D. (1998): Handlung und Ereignis. 2. Aufl., Frankfurt a. M.

Davidson, D. (1999): Thinking Causes. In: Heil, J./Mele, A. (Hrsg.): Mental causation. Oxford 1999, 3-17.

Davidson, R. J./Irwin, W. (1999): The functional neuroanatomy of emotion and affective style. Trends Cogn. Sci. 3, 11-21.

Dawkins, R. (1987): Der blinde Uhrmacher. München.

Deecke, L. (1987): The natural explanation for the two components of the readiness potential. The Behavioral and Brain Sciences 10, 781f.

Dehmelt, H. (1989): Triton, Electron, Cosmon …: An Infinite Regression? Proceedings of the National Academy of Sciences 86, 8618-8619.

Dennett, D. (1978): Toward a cognitive theory of consciousness. Cambridge. In: Dennett, D.: Brainstorms: Philosophical Essays on Mind and Psychology. Cambridge 1978, 149-173.

Dennett, D. (1981): Three Kinds of Intentional Psychology. In: Healy, R. (Hrsg.): Reduction, Time, and Reality. Cambridge 1981, 37-61.

Dennett, D. (1987): Brainstorms: Philosophical Essays on Mind and Psychology, Cambridge.

Dennett, D. (1988): Quining qualia. In: Marcel, A./Bisiach, E. (Hrsg.): Consciousness in Contemporary Science. Oxford 1988, 42-77.

Dennett, D. (1991): Consciousness Explained. London.

Descartes, R. (1911): Über die Leidenschaften der Seele. Leipzig.

Descartes, R. (1977): Meditationen über die Grundlagen der Philosophie. 2. Aufl., Hamburg.

Devinsky, O./Morrell, M. J./Vogt, B. A. (1995): Contributions of anterior cingulate cortex to behaviour. Brain 118, Pt. 1, 279-306.

Di Salle, F./Formisano, E./Linden, D.E./Goebel, R./Bonavita, S./Pepino, A./Smaltino, F./Tedeschi, G. (1999): Exploring brain function with magnetic resonance imaging. Eur. J. Radiol. 30 84-94.

Dolcos, F./LaBar, K. S./Cabeza, R. (2004): Dissociable effects of intensity and valence on prefrontal activity indexing emotional evaluation and subsequent memory: an event-related fMRI study. NeuroImage 23, 64-74.

Dollard, J./Doob, L. W./Miller, N. E./Mowrer, O. H./Sears, R. R. (1939): Frustration and Agggression. New Haven 1939.

Dray, W. (1957): Laws and Explanation in History. Oxford.

Dretske, F. (1981): Knowledge and the Flow of Information. Cambridge.

Dretske, F. (1988): Explaining Behavior: Reasons in a World of Causes. Cambridge.

Dretske, F. (1994): Mind and Brain. In: Warner, R./Szubka, T. (Hrsg.): The mind-body problem: a guide to the current debate. Cambridge 1994, 131-136.

Dretske, F. (1995): Naturalizing the Mind. Cambridge.

Dretske, F. (1999): Mental Events as Structuring Causes of Behaviour. In: Heil, J./Mele, A. (Hrsg.): Mental causation. Oxford 1999, 121-136.

Drevets, W. C./Raichle, M. E. (1998): Reciprocal suppression of rCBF during emotion and affective style. Cogn. Emot. 12, 353-385.

Driver, J./Frith, C. (2000): Shifting baselines in attention research. Nat. Rev., Neurosci. 1, 147-148.

Duncan, J./Owen, A. M. (2000): Common regions of the human frontal lobe recruited by diverse cognitive demands. Trends Neurosci. 23, 475-483.

Dunn, J. M. (1990): Relevant Predication 2: Intrinsic Properties and Internal Relations. Philosophical Studies 60, 177-206.

Eccles, J. C. (1985): Mental summation: The timing of voluntary intentions by cortical activity. The Behavioral and Brain Sciences 8, 542-543.

Eccles, J. C. (1994): Wie das Selbst sein Gehirn steuert. München.

Edelmann, G. (1994): Bright Air, Brilliant Fire – On the Matter of Mind. Hamondsworth.

Ehring, D. (1997): Causation and Persistence. Oxford.

Eisler, R. (1904): Wörterbuch der philosophischen Begriffe. 2. Aufl., Berlin.

Ellis, B. (2001): Scientific Essentialism. Cambridge.

Emsley J./Mitchell, B./Kempermann G./Macklis J. (2005): Adult neurogenesis and repair of the adult CNS with neural progenitors, precursors and stem cells. Prog Neurobiol. 75, 321-341.

Evans, G. (1982): The Varieties of Reference. Oxford.

Feigl, H. (1958): The "Mental" and the "Physical". In: Feigl, H./Scriven, M./Maxwell, G. (Hrsg.): Concepts, Theories, and the Mind-Body Problem. Minnesota Studies in the Philosophy of Science, Vol. II, Minneapolis, 370-497.

Fink, M. (1996): Neuroleptic malignant syndrome and catatonia: one entity or two? Biol. Psychiatry 39, 1-4.

Fodor, J. (1968): Psychological Explanation. New York.

Fodor, J. (1974): Special Sciences (or: the disunity of science as a working hypothesis). Synthese 28, 77-115. Wiederabdruck in: Block, N. (Hrsg.): Readings in the Philosophy of Psychology. Vol. 1, Cambridge 1980, 120-133.

Fodor, J. (1975): The Language of Thought. New York.

Fodor, J. (1981): Representations: Philosophical Essays on the Foundations of Cognitive Science. Cambridge.

Fodor, J. (1982): Cognitive science and the twin-earth problem. Notre Dame Journal of Formal Logic 23, 98-119.

Fodor, J. (1987): Psychosemantics. Cambridge.

Fodor, J. (1989): Making Mind Matter More. Philosophical Topics 17, 59-80.

Fodor, J. (1990a): A Theory of Content and Other Essays. Cambridge.

Fodor, J. (1990b): Psychosemantics or: Where Do Truth Conditions Come From? In: Lycan, W. (Hrsg.): Mind and Cognition: A Reader. Oxford 1990, 312-337.

Fodor, J. (1995): Fodor, Jerry. In: Guttenplan, S. (Hrsg.): A Companion to the Philosophy of Mind. Blackwell 1995, 292-300.

Fodor, J. (1997): Special Sciences: Still Autonomous After All These Years. Philosophical Perspectives 11, 149-165.

Fodor, J. (1998): Concepts. Where Cognitive Science went wrong. Oxford.

Fodor, J./Pylyshyn, Z. (1988): Connectionism and cognitive architecture. Cognition Vol. 28, 3-71.

Fox, P. T./Raichle, M. E./Mintun, M. A./Dence, C. (1988): Nonoxidative glucose consumption during focal physiologic neural activity. Science 241, 462-464.

Franscecotti, R. (1999): How to define Intrinsic Properties. Noûs 33, 590-609.

Fried, I./Koch, Ch./Kraskov, A./Quiroga, R./Reddy, L. (2007): Local Field Potentials and Spikes in the Human Medial Temporal Lobe are Selective to Image Category Source. In: Journal of Cognitive Neuroscience, Vol. 19, 479-492.

Gadenne, V. (1997): Qualia ohne kausale Wirksamkeit. In: Logos, 23-39.

Gasking, D. (1976): Causation and Recipes. In: Brand, M. (Hrsg.): The Nature of Causation. Chicago, 213-223.

Geach, P. T. (1969): God and the Soul. London.

Giere, R. (1988): Explaining Science: A Cognitive Approach. Chicago.

Gigon, O. (Hrsg.), (1961): Aristoteles. Einführungsschriften. Zürich.

Gilboa, A. (2004): Autobiographical and episodic memory-one and the same? Evidence from prefrontal activation in neuroimaging studies. Neuropsychologia 42, 1336-1349.

Gillet, C. (2001): Does the Argument from Realization Generalize? Responses to Kim. Southern Journal of Philosophy 39, 79-98.

Ginet, C. (1990): On Action. Cambridge.

Glüer, K. (1993): Donald Davidson zur Einführung. Hamburg.

Goel, V./Dolan, R. J. (2003): Explaining modulation of reasoning by belief. Cognition 87, B11-B22.

Goldman, A. (1993): The Psychology of Folk Psychology. Behavioral and Brain Sciences 16, 15-28.

Goodale, M. A./Milner, A. D. (1992): Separate visual pathways for perception and action. Trends Neurosci. 15, 20-25.

Goodale, M. A./Westwood, D. A. (2004): An evolving view of duplex vision: separate but interacting cortical pathways for perception and action. Curr. Opin. Neurobiol. 14, 203-211.

Goodman, N. (1978): Ways of worldmaking. Hassocks.

Gorno-Tempini, M. L./Pradelli, S./Serafini, M./Pagnoni, G./Baraldi, P./Porro, C./Nicoletti, R./Umita, C./Nichelli, P. (2001): Explicit and incidental facial expression processing: an fMRI study. NeuroImage 14, 465-473.

Gottfried, J. A./O'Doherty, J./Dolan, R. J. (2002): Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. J. Neurosci. 22, 10829-10837.

Grimm, S./Schmidt, C. F./Bermpohl, F./Heinzel, A./Dahlem, Y./Wyss, M./Hell, D./Boesiger,

P./Boeker, H./Northoff, G. (2006): Segregated neural representation of distinct emotion dimensions in the prefrontal cortex – an fMRI study. NeuroImage 30, 325-340.

Gusnard, D. A./Akbudak, E./Shulman, G. L./Raichle, M. E. (2001): Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. Proc. Natl. Acad. Sci. U. S. A. 98, 4259-4264.

Guttenplan, S. (1995): Leibniz's Law. In: Guttenplan, S. (Hrsg.): A Companion to the Philosophy of Mind. Blackwell 1995, 431-432.

Haenschel, C./Uhlhaas, P. J./Singer, W. (2007): Synchronous oscillatory activity and working memory in schizophrenia. Pharmacopsychiatry, 40 Suppl. 1, 54-61.

Hardcastle, V. G. (1998): On the Matter of Minds and Mental Causation. Philosophy and Phenomenological Research 58, 1-25.

Hardcastle, V. G./Stewart, C. M. (2002): What do brain data really show? Philosophy of Science 69, 72-82.

Hardcastle, V. G./Stewart, C. M. (2003): The art of single cell recordings. Biology and Philosophy 18, 195-208.

Hardcastle, V. G./Stewart, C. M. (2005): Localization in the Brain and Other Illusions. In: Brook, A./Akins, K. (Hrsg.): Cognition and the brain: the philosophy and neuroscience movement. Cambridge 2005, 27-39.

Hare, R. (1952): The Language of Morals. Oxford.

Hariri, A. R./Bookheimer, S. Y./Mazziotta, J. C. (2000): Modulating emotional responses: effects of a neocortical network on the limbic system. NeuroReport 11, 43-48.

Hariri, A. R./Mattay, V. S./Tessitore, A./Fera, F./Weinberger, D. R. (2003): Neocortical modulation of the amygdala response to fearful stimuli. Biol. Psychiatry 53, 494-501.

Hart, W. D. (1988): The Engines of the Soul. Cambridge.

Haugeland, J. (1982): Weak Supervenience. American Philosophical Quaterly 19, 93-103.

Hawthorne, J. (2001): Intrinsic Properties and Natural Relations. Philosophy and Phenomenological Research 63, 399-403.

Haynes, J. (2008): Detecting deception from neuroimaging signals – a data-driven perspective. Trends Cogn Sci 12(4), 126-7.

Heil, J. (1992): The Nature of True Minds. Cambridge.

Heil, J. (1994): Minds and Bodies. In: Warner, R./Szubka, T. (Hrsg.): The Mind-Body Problem. A Guide to the Current Debate, Cambridge 1994, 156-164.

Heil, J. (1998): Philosophy of Mind: A Contemporary Introduction. London.

Heil, J. (1999): Multiple Realizability. American Philosophical Quarterly 36, 189-208.

Heil, J./Mele, A. (Hrsg.), (1999): Mental causation. Oxford.

Heil, J./Robb, D. (2003): Mental Properties. American Philosophical Quarterly 40, 175-96.

Heinzel, A. (2002): Gibt es eine Erklärungslücke beim emotionalen Erleben? Philosophische und empirische Argumente, Dissertation, Düsseldorf.

Heisenberg, W. (1970): Die Entwicklung und Deutung der Quantentheorie. In: Krüger, L. (Hrsg.): Erkenntnisprobleme der Naturwissenschaften. Texte zur Einführung in die Philosophie der Wissenschaft, Berlin 1970, 412-428.

Hellman, G./Thompson, F. (1975): Physicalism: ontology, determination and reduction. Journal of Philosophy, 72, 551-64.

Hempel, C. G. (1942): The Function of General Laws in History. Journal of Philosophy 9, 35-48.

Hempel, C. G. (1963): Reasons and Covering Laws in Historical Explanation. In: Hook, S. (Hrsg.): Philosophy and History. New York 1963, 143-163.

Honderich, T. (1998): Consciousness as Existence. In: O'Hear, A. (Hrsg.): Current Issues in Philosophy of Mind. Cambridge 1998, 137-155.

Honderich, T. (1999): The Union Theory and Anti-Individualism. In: Heil, A. /Mele, J. (Hrsg.): Mental causation. Oxford 1999, 137-160.

Honderich, T. (2004): On Consciousness. Edinburgh.

Honderich, T. (2008): Consciousness as Existance. http://www.ucl.ac.uk/~uctytho/ted4.htm. Letzter Aufruf 21.09.2008.

Hooker, C. A. (1981): Towards a General Theory of Reduction. Dialogue 20, 38-60.

Horgan, T. (1997): Kim on Mental Causation and Causal Exclusion. Philosophical Perspectives 11, 165-184.

Horgan, T. (2001): Causal Compatibilism and the Exclusion Problem. Theoria 16, 95-116.

Hornak, J. et al. (2003): Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. Brain 126, 1691-1712.

Hornsby, J. (1986): Bodily movements, actions, and mental epistemology. Midwest Studies in Philosophy, 10, 275-286.

Hornsby, J. (1997): Simple mindedness: In defence of naive naturalism in the philosophy of mind. Cambridge.

http://bluebrain.epfl.ch/; letzter Abruf 20.10.2008.

Humberstone, I. L. (1996): Intrinsic/Extrinsic. Synthese 108 (1996), 205-67.

Hume, D. (1748): An inquiry concerning human understanding. With a supplement, an abstract of a treatise of human nature, Indianapolis.

Hume, D. (1984): Eine Untersuchung über den menschlichen Verstand. 11. Aufl., Hamburg.

Ingvar, D. H. (1985): "Memory of the future": an essay on the temporal organization of conscious awareness. Hum. Neurobiol. 4, 127-136.

Jackson, F. (1986): What Mary didn't know. Journal of Philosophy 83, 291-295.

Jackson, F. (1995): Essentialism, Mental Properties and Causation. Proceedings of the Aristotelian Society 95, 253-268.

Jackson, F. (1996): Mental causation: The state of the art. Mind 105, 377-413.

Jackson, F. (1998): Mental Causation I. In: Jackson, F.: Mind, Method, and Conditionals: Selected Essays. London/New York 1998, 80-112.

Jackson, F./Pettit, P. (1988): Functionalism and Broad Content, Mind 387, 381-400.

Jackson, F./Pettit, P. (1990): Causation in the Philosophy of Mind. Philosophy and Phenomenological Research 50, 195-214.

Jacob, P. (2005): Grasping and Perceiving Objects. In: Brook, A./Akins, K. (Hrsg.): Cognition and the brain: the philosophy and neuroscience movement. Cambridge 2005, 241-283.

James, W. (1892): Psychology: Briefer Course. New York.

Jensen, J./McIntosh, A. R./Crawley, A. P./Mikulis, D. J./Remington, G./Kapur, S. (2003): Direct activation of the ventral striatum in anticipation of aversive stimuli. Neuron 40, 1251-1257.

Johnson, M. (1993): Brain Development and Cognition: A Reader. Oxford.

Johnston, M. (1985): Why having a Mind Matters. In: LePore, E./McLaughlin, B. (Hrsg.): Actions and Events. Perspectives on the Philosophy of Donald Davidson, Oxford 1985, 408-427.

Joseph, G. (1980): The Many Sciences and the One World. Journal of Philosphy 77, 773-790.

Kant, I. (1998): Kritik der reinen Vernunft. Hamburg.

Kant, I. (2006): Kritik der Urteilskraft. Hamburg.

Kastner, S./Pinsk, M. A./De Weerd, P./Desimone, R./Ungerleider, L. G. (1999): Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron 22, 751-761.

Keenan, J. P./Wheeler, M. A./Gallup Jr., G. G./Pascual-Leone, A. (2000): Self-recognition and the right prefrontal cortex. Trends Cogn. Sci. 4, 338-344.

Keightley, M. L./Winocur, G./Graham, S. J./Mayberg, H. S./Hevenor, S. J./Grady, C. L. (2003): An fMRI study investigating cognitive modulation of brain regions associated with emotion processing of visual stimuli. Neuropsychologia 41, 585-596.

Kempermann, G. (2006): Adult Neurogenesis. Stem Cells and Neuronal Development in the Adult Brain. Oxford.

Kienzle, B. (1997): Von der Vorstellung der notwendigen Verknüpfung. In: Kulenkampff, J. (Hrsg.): David Hume, Eine Untersuchung über den menschlichen Verstand. Berlin 1997, 115-133.

Kim, J. (1976): Events as Property Exemplifications. In: Brand, M./Walton, D. (Hrsg.): Action Theory. Dordrecht 1976, 159-177.

Kim, J. (1980): Phenomenal Properties, Psychophysical Laws, and the Identity Theory. The Monist*,* Vol. 56 (1972), 177-92. Wiederabdruck als Auszug unter dem Titel "Physicalism and the Multiple Realizability of Mental States" in: Block, N. (Hrsg.): Readings in Philosophy of Psychology*,* Vol. 1, Cambridge 1980, 234-236.

Kim, J. (1982): Psychophysical Supervenience. Philosophical Studies, Vol. 41, 51-70.

Kim, J. (1984): Concepts of Supervenience. Philosophy and Phenomenological Research 45, 153-176. Wiederabdruck in: Kim, J. (Hrsg.): Supervenience and Mind. Selected Philosophical Essays, Cambridge 1993, 53-78.

Kim, J. (1985): Psychophysical Laws. In: LePore, E./McLaughlin, B. (Hrsg.): Actions and Events: Perspectives on the Philosophy of Donald Davidson. Oxford 1985, 369-386.

Kim, J. (1987): "Strong" and "Global" Supervenience Revisited. Philosophy and Phenomenological Research 48, 315-326. Wiederabdruck in: Kim, J. (Hrsg.): Supervenience and Mind. Selected Philosophical Essays, Cambridge 1993, 79-91.

Kim, J. (1988): Mind in a physical world. An Essay on the Mind-Body Problem and Mental Causation, Cambridge.

Kim, J. (1989): The Myth of Nonreductive Materialism. Proceedings and addresses of the American Philosophical Association 63, 31-47. Wiederabdruck in: Kim, J. (Hrsg.): Supervenience and Mind. Selected Philosophical Essays, Cambridge 1993, 265-284.

Kim, J. (1990): Supervenience as a Philosophical Concept. Metaphilosophy 21, 1-27. Wiederabdruck in: Kim, J. (Hrsg.): Supervenience and Mind. Selected Philosophical Essays, Cambridge 1993, 131-160.

Kim, J. (1992): Multiple Realization and the Metaphysics of Reduction. Philosophy and Phenomenological Research 52, 1-26. Wiederabdruck in: Kim, J. (Hrsg.): Supervenience and Mind. Selected Philosophical Essays, Cambridge 1993, 309-335.

Kim, J. (Hrsg.), (1993a): Supervenience and Mind: Selected Philosophical Essays. Cambridge.

Kim, J. (1993b): Epiphenomenal and Supervenient Causation. In: Kim, J. (Hrsg.): Supervenience and Mind. Selected Philosophical Essays, Cambridge 1993, 271-283.

Kim, J. (1995): Artikel „Supervenience". In: Guttenplan, S. (Hrsg.): A Companion to the Philosophy of Mind. Blackwell 1995, 575-583.

Kim, J. (1998): Philosophie des Geistes. Wien.

Kim, J. (1999a): Can Supervenience and 'Non-Strict Laws' Save Anomalous Monism? In: Heil, J./Mele, A. (Hrsg.): Mental causation. Oxford 1999, 19-26.

Kim, J. (1999b): The Non-Reductivist's Troubles with Mental Causation. In: Heil, J./Mele, A. (Hrsg.): Mental causation. Oxford 1999, 189-210.

Kim, J. (2001): Mental Causation and Consciousness: The Two Mind-Body Problems for the Physicalist. In: Gillett, C./Loewer, B.: Physicalism and Its Discontents. Cambridge 2001, 271-283.

Kim, J. (2003): Blocking Causal Drainage and other Chores with Mental Causation. Philosophy and Phenomenological Research Vol. 67, 1, 151-156.

King, A. (1999): Sensory Expereince and the Formation of a Computational Map of Auditory Space in the Brain. BioEssays 21, 900-911.

Kirsch, P./Schienle, A./Stark, R./Sammer, G./Blecker, C./Walter, B./ Ott, U./Burkart, J./Vaitl, D. (2003): Anticipation of reward in a nonaversive differential conditioning paradigm and the brain reward system: an event-related fMRI study. NeuroImage 20, 1086-1095.

Kistler, M. (1999): Causalité et lois de la nature. Paris.

Knutson, B./Fong, G. W./Adams, C. M./Varner, J. L./Hommer, D. (2001): Dissociation of reward anticipation and outcome with event-related fMRI. NeuroReport 12, 3683-3687.

Knutson, B./Fong, G. W./Bennett, S. M./Adams, C. M./Hommer, D. (2003): A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. NeuroImage 18, 263-272.

Kripke, S. (1971): Identity and necessity. In: Munitz, M. (Hrsg.): Identity and Individuation. New York 1971, 135-164.

Kripke, S. (1980): Naming and Necessity. Rev. Fassung, Cambridge MA. Deutsche Übersetzung: Name und Notwendigkeit. Frankfurt a. M. 1981.

Kuhn, T. (1996): Die Struktur wissenschaftlicher Revolutionen. 13. Aufl., Frankfurt a. M. 1996.

Kutschera, F. von (1982): Grundfragen der Erkenntnistheorie. Berlin.

Kutschera, F. von (2003): Jenseits des Materialismus. Paderborn.

LaBar, K. S./Gatenby, J. C./Gore, J. C./LeDoux, J. E./Phelps, E. A. (1998): Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. Neuron 20, 937-945.

Lang, P. J./Bradley, M. M./Cuthbert, B. N. (1999): International Affective Picture System (IAPS). Instruction Manual and Affective Ratings (Rep. No. A-4). The Center for Research in Psychophysiology, University of Florida, Gainsville.

Lang, P. J./Greenwald, M. K./Bradley, M. M./Hamm, A. O. (1993): Looking at pictures: affective, facial, visceral, and behavioral reactions. Psychophysiology 30, 261-273.

Lange, K./Williams, L. M./Young, A. W./Bullmore, E. T./Brammer, M. J./Williams, S. C./Gray, J. A./Phillips, M. L. (2003): Task instructions modulate neural responses to fearful facial expressions. Biol. Psychiatry 53, 226-232.

Langton, R./Lewis, D. (1998): Defining 'Intrinsic'. Philosophy and Phenomenological Research 58, 333-45.

Lee, T. M./Liu, H. L./Hoosain, R./Liao, W. T./Wu, C. T./Yuen, K. S./Chan, C. C./Fox, P. T./Gao, J. H. (2002): Gender differences in neural correlates of recognition of happy and sad faces in humans assessed by functional magnetic resonance imaging. Neurosci. Lett. 333, 13-16.

Leibniz, G. W. (1985): Metaphysische Abhandlung. 2. Aufl., Hamburg.

Leibniz, G. W. (2007): Der Briefwechsel mit Bartholomäus Des Bosses. Hamburg.

LePore, E./Loewer, B. (1987): Mind Matters. Journal of Philosophy 84, 630-642. Wiederabdruck in: Warner, R./Szubka, T. (Hrsg.): The Mind-Body Problem. A Guide to the Current Debate, Cambridge 1994, 261-273.

LePore, E./Loewer, B. (1989): More on Making Mind Matter. Philosophical Topics 17, 175-191.

Levine, J. (1983): Materialism and qualia: the explanatory gap. Pacific Philosophical Quarterly, 64, 354-361.

Levine, J. (1993): On leaving out what it's like. In: Davies, M./Humphreys, G. (Hrsg.): Consciousness: Psychological and Philosophical Essays. Oxford 1993, 121-136.

Levine, J. (2001): Purple Haze. The puzzle of consciousness. Oxford.

Lewis, D. (1972): Psychophysical and Theoretical Identifications. Australasian Journal of Philosophy 50 (1972), 249-258. Wiederabdruck in: Block, N. (Hrsg.): Readings in Philosophy of Psychology. Vol. 1, Cambridge 1980, 207-215.

Lewis, D. (1980): Mad Pain and Martian pain. In: N. Block (Hrsg.): Readings in the Philosophy of Psychology, Vol. 1. Cambridge 1980, 216-222.

Lewis, D. (1983a): An argument for the identity theory. In: Journal of Philosophy 63, 17-25. Wiederabdruck in: Lewis, D.: Philosophical Papers Vol. 1. Oxford 1983, 99-107.

Lewis, D. (1983b): Extrinsic Properties. Philosophical Studies 44, 197-200.

Lewis, D. (1986): Causation. In: Lewis, D. (Hrsg.): Philosophical Papers. Vol. 2, Oxford 1986, 159-172.

Libet, B. (1985): Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action. The Behavioral and Brain Sciences 8, 529-566.

Libet, B. (Hrsg.), (1993): Neurophysiology of Consciousness. Selected Papers and new Essays, Boston.

Libet, B./Wright, Jr. E. W. /Gleason, C. A. (1982): Readiness-Potentials Preceding Unrestricted "Spontaneous" vs. Pre-Planned Voluntary Acts. Electroencephalography and Clinical Neurophysiology 54, 322-335.

Lloyd, P. B. (2006): Mental Monsim Considered as a Solution to the Mind-Body problem. In: Batthyany, A. (Hrsg.): Mind and its place in the world. Non-reductionist approaches to the ontology of consciousness, Frankfurt a. M. 2006, 101-144.

Loar, B. (1987): Subjective Intentionality. Philosohical Topics 15, 1, 89-124.

Loar, B. (1991): Social Content and Psychological Content. In: Rosenthal, D. (Hrsg.): the Nature of Mind. New York/Oxford 1991, 568-575.

Locke, J. (2006): Versuch über den menschlichen Verstand 1. Buch 1 und 2, Hamburg.

Logothetis N. K./Pauls J./Augath M./Trinath T./Oeltermann A. (2001): Neurophysiological investigation of the basis of the fMRI signal. Nature 412(6843), 150-7.

Lombard, L. B. (1986): Events. A Metaphysical Study, London/Boston.

Lou, H. C./Luber, B./Crupain, M./Keenan, J. P./Nowak, M./Kjaer, T. W./Sackeim, H. A./Lisanby, S. H. (2004): Parietal cortex and representation of the mental self. Proc. Natl. Acad. Sci. U. S. A. 101, 6827-6832.

Lowe, E. J. (2000): Causal Closure Principles and Emergentism. Philosophy 75, 571-585.

Lowe, E. J. (2003): Physical Causal Closure and the Invisibility of Mental Causation. In: Walter, S./Heckmann, H.-D.: Physicalism and Mental Causation. Thorverton 2003, 137-154.

Lycan, W.G. (1996): Consciousness and Experience. Cambridge.

Macdonald, C./Macdonald, G. (1995): How to Be Psychologically Relevant. In: Ders. (Hrsg.): Philosophy of Psychology: Debates on Psychological Explanation, Vol. 1, Oxford 1995, 60-77.

Mackie, J. L. (1976): Causes and Conditions. In: Brand, M. (Hrsg.): The Nature of Causation. Chicago 1976, 307-344.

Maddock, R. J. (1999): The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. Trends Neurosci. 22, 310-316.

Maddock, R. J./Garrett, A. S./Buonocore, M. H. (2003): Posterior cingulated cortex activation by emotional words: fMRI evidence from a valence decision task. Hum. Brain Mapp. 18, 30-41.

Mandik, P. (2008): Fine-grainde supervenience, cognitive neuroscience, and the future of functionalism. http://www.petemandik.com/philosophy/papers/fgsff.html letzter Abruf 24.06.2008.

Manzotti, R.: A. (2006): Radical Externalist Approach to Consciousness: The Enlarged Mind. In: Batthyany, A. (Hrsg.): Mind and its place in the world. Non-reductionist approaches to the ontology of consciousness, Frankfurt a. M. 2006, 197-223.

Marr, D. (1982): Vision. New York.

Marras, A. (2003): Methodological and Ontological Aspects of the Mental Causation Problem. In: Walter, S./Heckmann, H.-D.: Physicalism and Mental Causation. Thorverton 2003, 243-264.

Mayberg, H. S. (2003): Modulating dysfunctional limbic-cortical circuits in depression: towards development of brain-based algorithms for diagnosis and optimised treatment. Br. Med. Bull. 65, 193-207.

Mayberg, H. S./Liotti, M./Brannan, S. K./McGinnis, S./Mahurin, R. K./Jerabek, P. A./Silva, J. A./Tekell, J. L./Martin, C. C./Lancaster, J. L./Fox, P. T. (1999): Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. Am. J. Psychiatry 156, 675-682.

McDowell, J. (1994): Mind and World. Cambridge.

McGinn, C. (1977): Charity, interpretation, and belief. In: Journal of Philosophy, 74, 521-535.

McGinn, C. (1989): Mental Content. Oxford.

McGinn, C. (1991): Can we solve the Mind-Body Problem? In: McGinn, C. (Hrsg.): The Problem of Consciousness. Cambridge 1991, 1-22.

McLaughlin, B. (1984): Event Supervenience and Supervenient Causation. Southern Journal of Philosophy 22, 71-92.

McLaughlin, B. (1985): Anomalous Monism and the Irreducibility of the Mental. In: LePore, E./McLaughlin, B.: Actions and Events. Perspectives on the Philosophy of Donald

Davidson, Oxford 1985, 331-369.

McLaughlin, B. (1989): Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical. Philosophical Perspectives 3, 109-135.

McLaughlin, B. (1997): Supervenience, Vagueness, and Determination. Philosophical Perspectives 11, 209-231.

McLaughlin, B. (1999): On Davidson's Response to the Charge of Epiphenomenalism. In: Heil, J./Mele, A. (Hrsg.): Mental causation. Oxford 1999, 27-40.

Meixner, U. (2004): The two sides of being. A reassessment of psycho-physical dualism. Paderborn.

Melden, A. I. (1961): Free Action. London.

Menzies, P. (2003): The Causal Efficacy of Mental States. In: Walter, S./Heckmann, H.-D.: Physicalism and Mental Causation. Thorverton 2003, 195-223.

Millikan, R. (1984): Language, Thought, and Other Biological Categories. Cambridge.

Millikan, R. (1989a): In Defense of Proper Functions. Philosophy of Science, 56, 288-302.

Millikan, R. (1989b): Biosemantics. In: Journal of Philosophy, 86, 281-97.

Millikan, R. (2000): On Clear and Confused Ideas: An Essay about Substance Concepts. Cambridge.

Montero, B. (2003): Varieties of Causal Closure. In: Walter, S./Heckmann, H.-D.: Physicalism and Mental Causation. Thorverton 2003, 173-187.

Moore, G. E. (1922): Philosophical Studies. London.

Morris, J. S./Friston, K. J./Buchel, C./Frith, C. D./Young, A. W./Calder, A. J./Dolan, R. J. (1998a): A neuromodulatory role for the human amygdala in processing emotional facial expressions. Brain 121, Pt. 1, 47-57.

Morris, J. S./Ohman, A./Dolan, R. J. (1998b): Conscious and unconscious emotional learning in the human amygdala. Nature 393, 467-470.

Murphy, F. C./Nimmo-Smith, I./Lawrence, A. D. (2003): Functional neuroanatomy of emotions: a meta-analysis. Cogn. Affect. Behav. Neurosci. 3, 207-233.

Nagel, E. (1961): The Structure of Science. New York.

Nagel, T. (1974): What Is It Like to Be a Bat? In: Philosophical Review 83, 435-450.

Nagel, T. (1981): Physikalismus. In: Bieri, P. (Hrsg.): Analytische Philosophie des Geistes. Königstein 1981, 56-72.

Nakamura, K./Kawashima, R./Ito, K./Sugiura, M./Kato, T./Nakamura, A./Hatano, K./Nagumo, S./Kubota, K./Fukuda, H./Kojima, S. (1999): Activation of the right inferior frontal cortex during assessment of facial emotion. J. Neurophysiol. 82, 1610-1614.

Neander, K. (1995): Malfunctioning and Misrepresenting. Philosophical Studies, 79, 109-141.

Noordhof, P. (1999): Micro-Bases Properties and the Supervenience Argument. Proceedings of the Aristotelian Society 99, 109-114.

Northoff, G. (2000): Das Gehirn. Eine neurophilosophische Bestandsaufnahme, Paderborn.

Northoff, G. (2004): Philosophy of the Brain. The Brain Problem, Amsterdam.

Northoff, G./Heinzel, A./Bermpohl, F./Niese, R./Pfennig, A./Pascual-Leone, A./Schlaug, G. (2004): Reciprocal modulation and attenuation in the prefrontal cortex: an fMRI study on emotional–cognitive interaction. Hum. Brain Mapp. 21, 202-212.

O'Doherty, J. P./Deichmann, R./Critchley, H. D./Dolan, R. J. (2002): Neural responses during anticipation of a primary taste reward. Neuron 33, 815-826.

Ochsner, K. N./Bunge, S. A./Gross, J. J./Gabrieli, J. D. (2002): Rethinking feelings: an fMRI study of the cognitive regulation of emotion. J. Cogn. Neurosci. 14, 1215-1229.

Papineau, D. (1993): Philosophical Naturalism. Oxford.

Papineau, D. (1998): Mind the Gap. In: Philosophical Perspectives, Vol. 12, 373-388.

Papineau, D. (2001): The Rise of Physicalism. In: Gillet, C./Loewer, B. (Hrsg.): Physicalism and its Discontents. Cambridge 2001, 3-36.

Papineau, D. (2004): Thinking about Consciousness. Oxford.

Parkinson, J. A./Robbins, T. W./Everitt, B. J. (2000): Dissociable roles of the central and basolateral amygdala in appetitive emotional learning. Eur. J. Neurosci. 12, 405-413.

Pauen, M. (2001): Grundprobleme der Philosophie des Geistes und die Neurowissenschaften. In: Pauen, M. (Hrsg.): Neurowissenschaft und Philosophie. München 2001, 83-122.

Paus, T. (2001): Primate anterior cingulate cortex: where motor control, drive and cognition interface. Nat. Rev., Neurosci. 2, 417-424.

Pavlov, I. P./Anrep, G. V. (1927): Conditioned Reflexes; An Investigation of the Physiological Activity of the Cerebral Cortex. London.

Peacocke, C. (1992): Scenarios, Concepts and Perception. In: Crane, T. (Hrsg.): The Contents of Experience. Cambridge 1992, 105-135.

Phan, K. L./Wager, T./Taylor, S. F./Liberzon, I. (2002): Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. NeuroImage 16, 331-348.

Piaget, J. (1927): La causalité physique chez l'enfant. Paris.

Piefke, M./Weiss, P. H./Zilles, K./Markowitsch, H. J./Fink, G. R. (2003): Differential remoteness and emotional tone modulate the neural correlates of autobiographical memory. Brain 126, 650-668.

Place, U. T. (1956): Is Consciousness a Brain Process? British Journal of Psychology 47, 44-50.

Platek, S. M./Keenan, J. P./Gallup Jr., G. G./Mohamed, F. B. (2004): Where am I? The neurological correlates of self and other. Brain Res. Cogn. Brain Res. 19, 114-122.

Ploghaus, A./Tracey, I./Gati, J. S./Clare, S./Menon, R. S./Matthews, P. M./Rawlins, J. N. (1999): Dissociating pain from its anticipation in the human brain. Science 284, 1979-1981.

Popper, K. (1959): The Logic of Scientific Discovery. London.

Popper, K./Eccles, J. (1977): Das Ich und sein Gehirn. München.

Porro, C. A./Baraldi, P./Pagnoni, G./Serafini, M./Facchin, P./Maieron, M./Nichelli, P. (2002): Does anticipation of pain affect cortical nociceptive systems? J. Neurosci. 22, 3206-3214.

Porro, C. A./Cettolo, V./Francescato, M. P./Baraldi, P. (2003): Functional activity mapping of the mesial hemispheric wall during anticipation of pain. NeuroImage 19, 1738-1747.

Price, C. J./Friston, K. J. (1997): Cognitive conjunction: a new approach to brain activation experiments. NeuroImage 5, 261-270.

Pruessmann, K. P./Weiger, M./Scheidegger, M. B./Boesiger, P. (1999): SENSE: sensitivity encoding for fast MRI. Magn. Reson. Med. 42, 952-962.

Putnam, H. (1967a): The Mental Life of Some Machines. In: Castaneda, H. N. (Hrsg.): Intentionality, Minds, and Perception. Detroit 1967, 177-200. Wie-derabdruck in Putnam, H.: Mind, Language, and Realitiy. Philosophical Papers, Vol. 2, Cambridge 1975, 408-428.

Putnam, H. (1967b): Psychological Predicates. In: Capitan, W. H./Merrill, D. D. (Hrsg.): Art, Mind, and Religion. Pittsburgh 1967. Wiederabdruck unter dem Titel "The Nature of Mental States" in: Putnam, H.: Mind, Language, and Realitiy. Philosophical Papers, Vol. 2, Cambridge 1975, 429-440 und Block, N. (Hrsg.): Readings in Philosophy of Psychology. Vol. 1, Cambridge 1980, 223-231. Deutsche Übersetzung: Die Natur mentaler Zustände. In: Bieri, P. (Hrsg.): Analytische Philosophie des Geistes. Königstein 1981, 121-135.

Putnam, H. (1975a): Mind, Language, and Realitiy. Philosophical Papers, Vol. 2, Cambridge.

Putnam, H. (1975b): Minds and Machines. In: Hook, S. (Hrsg.): Dimensions of Mind. New York 1960, 138-164. Wiederabdruck in: Putnam, H.: Mind, Language, and Realitiy. Philosophical Papers, Vol. 2, Cambridge 1975, 362-385.

Putnam, H. (1975c): The meaning of 'meaning'. In: Gunderson, K. (Hrsg.): Language, Mind, and Knowledge. Minneapolis 1975. Wiederabdruck in: Putnam, H.: Mind, Language, and Reality. Philosophical Papers, Vol. 2, Cambridge 1975, 215-271.

Putnam, H. (1981): Die Natur mentaler Zustände. In: Bieri, P. (Hrsg.): Analytische Philosophie des Geistes. Königstein 1981, 121-135.

Quine, W. V. (1951): Two dogmas of empiricism. Philosophical Review 60, 20-43.

Quine, W. V. (1960): Word and Object. Cambridge.

Quine, W. V. (1982): Theories and Things. Cambridge.

Quine, W. V. (1995): Naturalism or living within one's means. Dialectica 49, 251-261.

Quine, W. V. (1963): On what there is. In: Quine, W. V.: From a logical point of view. 2. überarbeitete Aufl. New York, 1-19.

Raichle, M. E./MacLeod, A. M./Snyder, A. Z./Powers,W. J./Gusnard, D. A./Shulman, G. L. (2001): A default mode of brain function. Proc. Natl. Acad. Sci. U. S. A. 98, 676-682.

Rainville, P./Duncan, G. H./Price, D. D./Carrier, B./Bushnell, M. C. (1997): Pain affect encoded in human anterior cingulate but not somatosensory cortex. Science 277, 968-971.

Rauschecker, J. (1995): Compensatory Plasticity and Sensory Substitution in the Cerebral Cortex. Trends in Neurosciences 18, 36-43.

Raymont, P. (2003): Kim on Closure, Exclusion and Nonreductive Physicalism. In: Walter, S./Heckmann, H.-D.: Physicalism and Mental Causation. Thorverton 2003, 225-242.

Reiman, E. M./Lane, R. D./Ahern, G. L./Schwartz, G. E./Davidson, R. J./Friston, K. J./Yun, L. S./Chen, K. (1997): Neuroanatomical correlates of externally and internally generated human emotion. Am. J. Psychiatry 154, 918– 925.

Rescher, N. (1985): Die Grenzen der Wissenschaft. Stuttgart.

Rorty, R. (1965): Mind-Body Identity, Privacy, and Categories. Review of Metaphysics 19 (1965), 24-54. Deutsche Übersetzung „Leib-Seele Identität, Privatheit und Kategorien. In: Bieri, P. (Hrsg.): Analytische Philosophie des Geistes. Königstein 1981, 93-120.

Rorty, R. (1979): Philosophy and the Mirror of Nature. Princeton.

Rosen, B. R./Aronen, H. J./Kwong, K. K./Belliveau, J. W./Hamber, L. M./Fordham, J. A. (1993): Advances in clinical neuroimaging: functional MR imaging techniques. Radiographics 13, 889-896.

Rosenberg, A. (2001): On Multiple Realization and the Special Sciences. Journal of Philosohy 98, 365-373.

Rosenthal, D. M. (Hrsg.), (1991): The Nature of Mind. New York/Oxford.

Ross, W. D. (1949): Aristotle. 5. Aufl., London.

Roth, G. (1996): Das Gehirn und seine Wirklichkeit. Kognitive Neurobiologie und ihre philosophischen Konsequenzen, 5. Aufl., Frankfurt a. M.

Roth, G. (2001): Die neurobiologischen Grundlagen von Geist und Bewusstsein. In: Pauen, M. (Hrsg.): Neurowissenschaft und Philosophie. München 2001, 155-209.

Roth, G. (2003): Fühlen, Denken, Handeln. Wie das Gehirn unser Verhalten steuert, Frankfurt a. M.

Runggaldier, E./Kanzian, C. (1998): Grundprobleme der Analytischen Ontologie. Paderborn, München, Wien, Zürich.

Russell, J. A. (1980): A circumplex model of affect. J. Pers. Soc., 1161-1178.

Ryle, G. (1949): The Concept of Mind. London. Deutsche Übersetzung: Der Begriff des Geistes. Stuttgart 1969.

Salmon, W. C. (1984): Scientific Explanation and the Causal Structure of the World. Princeton.

Schiffer, S. (1981): Indexicals and the Theory of Reference. Synthese 49, 43-100.

Schiffer, S. (1987): Remnants of Meaning. Cambridge.

Schneider, K. (1992): Klinische Psychopathologie, 14. Auflage, Stuttgart.

Schröder, J. (2002): The Supervenience Argument and the Generalization Problem. Erkenntnis 56, 319-328.

Schröder, J. (2003): Mental Causation: The Supervenience Argument and the Proportionality Constraint. In: Maasen, S./Prinz, W./Roth, G. (Hrsg.):Voluntary action. Brains, minds, and sociality, Oxford 2003, 172-187.

Schröder, J. (2004): Einführung in die Philosophie des Geistes. Frankfurt a. M.

Schröder, J. (2007): Mental Causation and the Supervenience Argument. Erkenntnis, Vol. 67, 2, 221-237.

Searle, J. (1980): Minds, Brains and Programs. The Behavioral and Brain Sciences 3, 417-424.

Searle, J. (1984): Minds, Brains and Science. London. Deutsche Übersetzung: Geist, Hirn und Wissenschaft. Frankfurt a. M. 1986.

Searle, J. (1993): Die Wiederentdeckung des Geistes. München. Englische Ausgabe: The Rediscovery of Mind. Cambridge 1992.

Searle, J. (2001): Geist, Sprache und Gesellschaft. Philosophie in der wirklichen Welt, Frankfurt a. M.

Segal, G. (2000): A Slim Book about Narrow Content. Cambridge.

Segal, G./Sober, E. (1991): The Causal Efficacy of Content. Philosophical Studies 61, 1-30.

Shaffer, J. (1963): Mental events and the brain. Journal of Philosophy 60, 160-66.

Sharkey, N.E. (2002): Biologically Inspired Robotics. In: Arbib, M. (Hrsg.): Handbook of Brain Theory and Neural Networks, Cambridge, 160-175.

Shinkareva, S. V./Mason, R. A./Malave, V. L./Wang, W./Mitchell, T. M. et al. (2008): Using fMRI brain activation to identify cognitive states associated with perception of tools and dwellings. PLoS ONE 3(1), e1394. doi10.1371/journal.pone.0001394.

Shoemaker; S. (1980a): Functionalism and Qualia. In: Block, N. (Hrsg.): Readings in Philosophy of Psychology. Vol. 1, Cambridge 1980, 251-268.

Shoemaker, S. (1980b): Causality and properties. In: van Inwagen, P. (Hrsg.): Time and Cause. Dodrecht 1980, 109-135.

Shoemaker, S. (1981): Some Varieties of Functionalism. Philosophical topics 12.1, 83-118.

Shoemaker, S. (2001): Realization and Mental Causation. In: Gillet, C./Loewer, B. (Hrsg.): Physicalism and its Discontents. Cambridge 2001, 74-98.

Sider, Th. (2001): Maximality and Intrinsic Properties. Philosophy and Phenomenological Research 63, 357-64.

Sider, Th. (2003): Maximality and Microphysical Supervenience. Philosophy and Phenomenological Research 66, 139-149.

Singer, T./Seymour, B./O'Doherty, J./Kaube, H./Dolan, R. J./Frith, C. D. (2004): Empathy for pain involves the affective but not sensory components of pain. Science 303, 1157-1162.

Singer, W. (1999): Neuronal synchrony: A versatile code for the definition of relations? Neuron 24, 49-65.

Singer, W. (Hrsg.), (2002a): Der Beobachter im Gehirn. Essays zur Hirnforschung, Frankfurt a. M.

Singer, W. (2002b): Hirnentwicklung oder die Suche nach Kohärenz. In: Krämer, S. (Hrsg.): Geist – Gehirn – Künstliche Intelligenz. Ringvorlesung an der Freien Universität Berlin, Berlin 1994, 165-188. Wiederabdruck in: Singer, W. (Hrsg.): Der Beobachter im Gehirn. Essays zur Hirnforschung, Frankfurt a. M. 2002, 120-143.

Singer, W. (2002c): Der Beobachter im Gehirn. In: Meier, H./Plog, D.: Der Mensch und sein Gehirn. München 1997, 35-65. Wiederabdruck in: Singer, W. (Hrsg.): Der Beobachter im Gehirn. Essays zur Hirnforschung, Frankfurt a. M. 2002, 144-170.

Singer, W. (2002d): Hirnentwicklung und Umwelt. In: Gehirn und Denken. Kosmos im Kopf, hrsg. vom Deutschen Hygienemuseum, Ostfildern-Ruit, 108-115. Wiederabdruck in: Singer, W. (Hrsg.): Der Beobachter im Gehirn. Essays zur Hirnforschung, Frankfurt a. M. 2002, 112-119.

Singer, W. (2002e): Im Grunde nichts Neues. Rechtshistorisches Journal 19, 41-51. Wiederabdruck in: Singer, W. (Hrsg.): Der Beobachter im Gehirn. Essays zur Hirnforschung, Frankfurt a. M. 2002, 171-180.

Singer, W. (2004): Synchrony, oscillations, and relational codes. In: Chalupa, L. M./Werner, J. S. (Hrsg.): The Visual Neurosciences. Cambridge 2004, 1665-1681.

Small, D. M./Gregory, M. D./Mak, Y. E./Gitelman, D./Mesulam, M. M./Parrish, T. (2003): Dissociation of neural representation of intensity and affective valuation in human gustation. Neuron 39, 701-711.

Smart, J. (1959): Sensations and Brain Processes. Philosophical Review 68, 141-156.

Smart, J. (1962): Sensations and brain processes. V. C. Chappell (Hrsg.): Philosophy of Mind. Englewood, 161-162.

Soon, C. S./Brass, M./Heinze, H. J./Haynes, J. D. (2008): Unconscious determinants of free decisions in the human brain. Nat Neurosci. 11 (5), 543-545.

Sosa, E. (1984): Mind-Body Interaction and Supervenient Causation. Midwest Studies in Philosophy, 9, 271-281.

Sosa, E. (1999): Davidson's Thinking Causes. In: Heil, J./Mele, A. (Hrsg.): Mental causation. Oxford 1999, 41-50.

Stalnaker, R. (1991): On what's in the head. In: Rosenthal, D. (Hrsg.): The Nature of Mind. New York/Oxford 1991, 576-589.

Stark, C. E./Squire, L. R. (2001): When zero is not zero: the problem of ambiguous baseline conditions in fMRI. Proc. Natl. Acad. Sci. U. S. A. 98, 12760-12766.

Steele, J. D./Lawrie, S. M. (2004): Segregation of cognitive and emotional function in the prefrontal cortex: a stereotactic meta-analysis. Neuro-Image 21, 868-875.

Stegmüller, W. (1983): Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie. Bd. 1, Erklärung – Begründung – Kausalität, 2. Aufl., Berlin.

Stich, S. (1983): From Folk Psychology to Cognitive Science: The Case Against Belief, Cambridge.

Stich, S. (1996): Deconstructing the Mind. New York.

Stoutland, F. (1970): The Logical Connection Argument. American Philosophical Quarterly 4, 117-121.

Suppe, F. (1989): The Semantic Conception of Theories and Scientific Realism. Chicago.

Swinburne, R. (1994): Body and Soul. In: Warner, R./Szubka, T. (Hrsg.): The Mind-Body Problem. A Guide to the Current Debate, Cambridge, 311-316.

Taylor, S. F./Phan, K. L./Decker, L. R./Liberzon, I. (2003): Subjective rating of emotionally salient stimuli modulates neural activity. NeuroImage 18, 650–659.

Teller, P. (1983): Quantum Physics, the Identity of Indiscernibles and Some Unanswered Questions. Philosophy of Science 50, 309-319.

Thomasson, A. (1998): A nonreductivist solution to mental causation. Philosophical Studies 89-90, 181-195.

Titze, H. (1992): Das philosophische Gesamtwerk. Bd. 2, Kausalität in Physik und Philosophie, Berlin.

Turing, A. (1936): On Computable Numbers with an Application to the Entscheidungsproblem. In: Proceedings of the London Mathematical Society 42, 230-265.

Tye, M. (1995): Ten Problems of Consciousness. A Representational Theory of the Phenomenal Mind, Cambridge.

Tye, M. (2000): The Imagery Debate. Cambridge.

Ueda, K./Okamoto, Y./Okada, G./Yamashita, H./Hori, T./Yamawaki, S. (2003): Brain activity during expectancy of emotional stimuli: an fMRI study. NeuroReport 14, 51-55.

Van Gulick, R. (1999): Who's in charge here? And who's doing all the work? In: Heil, A. /Mele, J. (Hrsg.): Mental causation. Oxford 1999, 233-256.

Velmans, M. (2000): Understanding Consciousness. London.

Vuilleumier, P./Armony, J. L./Driver, J./Dolan, R. J. (2001): Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. Neuron 30, 829-841.

Wagner, T. D./Jonides, J./Reading, S. (2004): Neuroimaging studies of shifting attention: a meta-analysis. NeuroImage 22, 1679-1693.

Walker, G. (1996): The Secret Heart of a Quark. New Scientist 17, 17.

Walter, H. (1999): Neurophilosophie der Willensfreiheit. Von literarischen Illusionen zum Konzept natürlicher Autonomie, 2. Aufl., Paderborn.

Warner, R./Szubka, T. (Hrsg.), (1994): The Mind-Body Problem. A Guide to the Current Debate, Cambridge.

Weiss, H. (1942): Kausalität und Zufall in der Philosophie des Aristoteles. Basel.

Wieland, W. (1962): Die aristotelische Physik. Untersuchung über die Grundlegung der Naturwissenschaft und die sprachlichen Bedeutungen der Prinzipienforschung bei Aristoteles. Göttingen.

Wilczek, F. (1996): A Crack in the Standard Model? Nature, 380, 19.

Wittgenstein, L. (1953): Philosophische Untersuchungen. Oxford.

Woldorff, M. G./Matzke, M./Zamarripa, F./Fox, P. T. (1999): Hemodynamic and electrophysiological study of the role of the anterior cingulate in target-related processing and selection for action. Hum. Brain Mapp. 8 (1999), 121-127.

Yablo, S. (1992): Mental causation. Philosophical Review 101, 245-280.

Yablo, S. (1997): Wide causation. Philosophical Perspectives 11, 251-281.

Yablo, S. (1998): Intrinsicness. Philosophical Topics 26, 479-504.

Yablo, S. (2003): Causal Relevance. Philosophical Issues, 13, 316-328.