

Die Architektur des Psychischen

Eine begrifflich-konzeptuelle Grundlegung der Psychologie und ihres Gegenstandes

Mike Lüdmann



20 Schriften aus der Fakultät Humanwissenschaften
der Otto-Friedrich-Universität Bamberg

Schriften aus der Fakultät Humanwissenschaften
der Otto-Friedrich-Universität Bamberg

Band 20



University
of Bamberg
Press

2015

Die Architektur des Psychischen

Eine begrifflich-konzeptuelle Grundlegung der Psychologie
und ihres Gegenstandes

von Mike Lüdman



Bibliographische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Informationen sind im Internet über <http://dnb.d-nb.de/> abrufbar.

Diese Arbeit hat der Fakultät Humanwissenschaften der Otto-Friedrich-Universität Bamberg als Dissertation vorgelegen.

1. Gutachter: Universitätsprofessor Dr. Dietrich Dörner

2. Gutachter: Universitätsprofessor Dr. Christian Illies

Tag der mündlichen Prüfung: 14. April 2015

Dieses Werk ist als freie Onlineversion über den Hochschulschriften-Server (OPUS; <http://www.opus-bayern.de/uni-bamberg/>) der Universitätsbibliothek Bamberg erreichbar. Kopien und Ausdrücke dürfen nur zum privaten und sonstigen eigenen Gebrauch angefertigt werden.

Herstellung und Druck: docupoint, Magdeburg

Umschlaggestaltung: University of Bamberg Press, Anna Hitthaler

Umschlagbild: Dirk Lilja

© University of Bamberg Press Bamberg 2015

<http://www.uni-bamberg.de/ubp/>

ISSN: 1866-8674

ISBN: 978-3-86309-304-4 (Druckausgabe)

eISBN: 978-3-86309-305-1 (Online-Ausgabe)

URN: urn:nbn:de:bvb:473-opus4-265358

Inhaltsverzeichnis

Zusammenfassung	15
1. Einleitung.....	17
1.1. Zum Aufbau der Arbeit	20
1.2. Eine grundlegende Begriffsklärung.....	35
2. Der dringende Bedarf einer begrifflich-konzeptuellen Grundlegung der Psychologie	45
2.1. Die aktuelle Debatte um die Biologisierung der Psychologie ...	46
2.1.1. Ebenenmodelle des Geistes: Begriffliche Vagheit und Potential	48
2.1.2. Gefährdung des Faches durch Neuroreduktionismus?	51
2.1.3. Wiederentdeckung der Bedeutung des Leib-Seele- Problems	54
2.1.4. Gehirn-Geist-Differenz als Garant psychologischer Eigenständigkeit?	56
2.2. Der Mythos der verspäteten Wissenschaft und seine Folgen ...	60
2.3. Die Notwendigkeit einer Theoretischen Psychologie	70

3. Der psych(olog)ische Ursprung des Leib-Seele-Problems.....	87
3.1. Das Leib-Seele-Problem und seine Fallgruben.....	90
3.1.1. Zur Intuitivität und Abwegigkeit des Interaktionismus.....	92
3.1.2. Epiphanomenalismus: Der Geist als bloße Begleiterscheinung	100
3.1.3. Der Eliminative Materialismus und die Nicht-Existenz des Geistes.....	103
3.2. Die Einheit und Differenz von Gehirn und Geist: Spinozas Dilemma	105
3.3. Die Gehirn-Geist-Differenz als sprachlich-deskriptives Phänomen.....	112
3.4. Die Gehirn-Geist-Differenz als Wahrnehmungsphänomen ..	120
3.4.1. Fechners Zwei-Perspektiven-Theorie	122
3.4.2. Perspektivität und Transparenz als psychische Strukturmerkmale	128
3.4.2.1. Perspektivität als divergierende Personenperspektiven	128
3.4.2.2. Perspektivität als divergierende Modi der Repräsentation.....	132
3.4.2.3. Phänomenale Transparenz und der naive Realismus.....	136
3.4.3. Eine wahrnehmungspsychologische Analyse von Multiperspektivität.....	143

3.4.4. Entwicklungspsychologische Einbettung bzw. Adaptation.....	154
3.5. Fazit: Ein Modell der geistigen Entzweigung von Gehirn und Geist.....	162
4. Psychologische Ebenenmodelle mentaler Prozesse	173
4.1. Der Begriff der Ebene	175
4.1.1. Begriffsgeschichtlicher Hintergrund	178
4.1.2. Wissenschaftstheoretische Begriffsbestimmung	181
4.2. Psychologische und kognitionswissenschaftliche Ebenenmodelle.....	186
4.2.1. Ebenenmodelle für kognitive Zustände und Prozesse	186
4.2.1.1. Informationsverarbeitende Modelle	187
4.2.1.2. Konnektionistische Modelle	194
4.2.2. Ebenenmodelle für phänomenale Zustände und Prozesse	201
4.2.3. Fazit.....	204
5. Die natürliche Organisation des Psychischen	207
5.1. Natürliche Organisationsebenen: Eine Systematisierung	209
5.1.1. Der quantitative Begriff natürlicher Organisationsebenen	210

5.1.2. Der kompositorische Begriff natürlicher Organisationsebenen	216
5.1.2.1. Mereologische Ebenen	217
5.1.2.2. Aggregative Ebenen.....	220
5.1.2.3. Mechanistische Ebenen	222
5.1.3. Perspektiven/analytische Ebenen versus natürliche Ebenen	229
5.2. Das Psychische als mereologisch-mechanistisch individuierte Organisationsebene.....	231
5.2.1. Die Gehirn-Geist-Differenz als divergierende mereologische Ebenen.....	233
5.2.2. Kategorienfehler und der „Mythos des Gegebenen“	240
5.2.3. Mereologisch-mechanistische versus aggregative Individuation	244
5.3. Spinozas fehlendes Puzzlestück: Gehirn und Geist als Trigger divergierender konzeptueller Formen	246
6. Licht und Schatten: Psychophysische Konzepte der gegenwärtigen Philosophie des Geistes.....	255
6.1. Die psychophysische Identitätstheorie (Type-Identity- Theory)	258
6.1.1. Ullin T. Place: Ist Bewusstsein ein Gehirnzustand?.....	261
6.1.2. Herbert Feigl: Zwischen neutralem und idealistischem Monismus	264

6.1.3. John J. C. Smart: Strikte psychophysische Identität.....	268
6.1.4. Stellenwert und Grenzen	272
6.1.4.1. Stärken und Vorzüge der Identitätstheorie	272
6.1.4.2. Multiple Realisierbarkeit	274
6.1.4.3. Starre Designatoren und der Anschein von Kontingenz	278
6.1.4.4. Die Erklärungslücke	286
6.1.4.5. Die problematische Interpretation der Identität	289
6.2. Die Identität von psychischen und physischen Ereignissen (Token-Identity-Theory)	297
6.2.1. Davidsons anomaler Monismus	297
6.2.2. Die problematische Identitätskonzeption der Token- Identitätstheorie	301
6.3. Supervenienz und der „Mythos des nicht-reduktiven Materialismus“	306
6.3.1. Psychische Supervenienz: Prinzip und Grenzen	307
6.3.2. Problematische Konsequenzen der Supervenienzkonzeption	314
6.4. Emergenz und das unerklärliche Auftauchen des Geistes	320
6.4.1. Konsistente und inkonsistente Emergenzkonzepte	322
6.4.1.1. Emergente Systemeigenschaften oder „moderate“ Emergenz.....	323
6.4.1.2. Radikale Emergenz	326

6.4.1.3. Starke (synchrone) Emergenz.....	328
6.4.2. Zu den Grenzen der Reichweite des Emergenzkonzeptes.....	331
6.5. Funktionalismus und physische Realisierung	336
6.5.1. Realisierung als Identität von Eigenschaftsinstantiierungen.....	340
6.5.2. Realisierung als kausale Teilmengenbeziehung.....	348
6.5.3. Eindimensionalität versus Multidimensionalität.....	350
6.5.4. Stellenwert und Grenzen.....	363
6.5.4.1. Problem der „Funktionalisierbarkeit“ phänomenaler Eigenschaften	363
6.5.4.2. Die problematische Konzeption der eindimensionalen Realisierung.....	369
6.5.4.3. Die problematische Konzeption der multidimensionalen Realisierung.....	374
6.5.4.4. Die Analyse von Eigenschaftsrelationen ist ontologisch verfehlt	379
6.6. Fazit.....	382
7. Die Architektur des Psychischen: Eine mereologisch-mechanistisch fundierte psychophysische Identitätskonzeption	389
7.1. Geist, Eigenschaft und Struktur	393
7.1.1. Der Strukturbegriff von Eigenschaften	393
7.1.2. Kims Ansatz mikrobasierter Eigenschaften.....	396

7.1.3. Mentale Eigenschaften als Struktureigenschaften	402
7.1.3.1. Kognitive Eigenschaften	403
7.1.3.2. Phänomenale Eigenschaften	409
7.2. Von psychophysischer Koextensivität zu lokaler Identität	413
7.2.1. Die Identifikation „mit“ statt der Identifikation „als“	417
7.2.2. Psychophysische Identität und systemische Emergenz	420
7.2.3. Die „Aufstapelungssillusion“ mereologischer Hierarchien	423
7.3. Die Einheit von Struktur und Form/Funktion	427
7.3.1. Funktionalismus versus Identitätstheorie	427
7.3.2. Aristoteles‘ Vermächtnis	431
7.3.3. Identität versus Realisierung	434
7.3.4. Identität versus Konstitution	437
7.4. Zur Realität psychischer bzw. psychophysischer Kausalität ...	443
7.4.1. Inter-Level-Kausalität als mechanistisch vermittelter Effekt	444
7.4.2. Mentale Verursachung als „ebenendurchdringende“ Kausalität	453
7.4.3. Ein Abriss zur Handlungs- und Entscheidungsfreiheit	467
7.5. Zur Möglichkeit typentheoretischer psychophysischer Identitäten	472
7.5.1. Heuristische Identitäten als Forschungsstrategie	477

7.5.2. Psychophysische Typen-Identitäten jenseits von Heuristiken?	482
7.5.2.1. Das exklusive Wesen des Mentalen – ein metaphysischer Abweg	483
7.5.2.2. Artspezifität und Granularitätsparität	489
7.5.2.3. Die scheinbare Kontingenz psychophysischer Identitätsaussagen	502
7.6. Was bleibt von der „Erklärungslücke“?	510
8. Zur Unabkömmlichkeit der Psychologie: Eine Entzauberung des (Neuro-)Reduktionismus	529
8.1. Entwirrung der begrifflichen Grundlagen	533
8.1.1. Reduktion und Reduktionismus	536
8.1.2. Formen der Reduktion und des Reduktionismus	538
8.1.2.1. Ontologische Reduktion und ontologischer Reduktionismus	539
8.1.2.2. Methodologische Reduktion und methodologischer Reduktionismus	541
8.1.2.3. Epistemische Reduktion und epistemischer Reduktionismus	543
8.2. Reduktion psychologischer auf (neuro-)biologische Theorien?	547
8.2.1. Die „klassische“ Theoriereduktion nach Ernest Nagel	547

8.2.2. Chancen und Grenzen der Nagel-Reduktion psychologischer Theorien	554
8.3. Ersetzung psychologischer durch (neuro-)biologische Theorien?	561
8.3.1. Die „neue Welle“ des Reduktionismus	561
8.3.2. Zur Persistenz der psychologischen Begriffs- und Theoriebildung	567
8.3.3. Die fundamentale Differenz zwischen Intra- und Inter-Level-Kontexten	573
8.4. Die reduktive Erklärung des Psychischen durch die Neurobiologie	579
8.5. Irrwege der „Autonomisierung“ der Psychologie	585
8.5.1. Der (metaphysische) Funktionalismus als Garant von „Autonomie“	585
8.5.2. Mikrodetermination statt psychophysischer Identität	591
8.6. Warum ein ontologisches Primat der Mikroebenen scheitert	599
8.6.1. Zur Vermengung von Physikalismus und Mikrophysikalismus	601
8.6.2. Zum Scheitern einer unidirektionalen Mikrodetermination	605
8.6.3. Zum Scheitern des Primats mikrophysikalischer Gesetze	609
8.6.3.1. Kein ontologischer Vorrang für Mikrogesetze	610

8.6.3.2. Sind Mikrogenetze in komplexen Systemen überhaupt instantiiert?	615
8.6.4. Zum Scheitern des Ansatzes einer Mikroverursachung ...	619
8.7. Warum ein explanatorisches Primat der Mikroebenen scheitert	624
8.7.1. Das explanatorische Defizit von Mikroerklärungen	625
8.7.2. Der Vorrang der Mikroebenen als kognitive Heuristik/Verzerrung.....	635
9. Leitlinien und Konsequenzen der begrifflich-konzeptuellen Grundlegung der Psychologie	643
9.1. Reduktivität als Eckpfeiler einer „autonom“-integren Psychologie	644
9.2. Der Gegenstand der Psychologie bzw. psychologischer Erklärungen	654
9.3. Die Realität und kausale Wirkmächtigkeit des Psychischen ..	665
9.4. Ontologische Einheit und explanatorischer Pluralismus	670
9.5. Das mechanistische Paradigma als Leitbild der Forschung und Theoriebildung.....	679
Literaturverzeichnis	687
Abbildungsverzeichnis	727

Zusammenfassung

Das zentrale Ziel dieser Arbeit ist die Bereitstellung eines begrifflich-konzeptuellen Fundaments der Psychologie in Form der Formulierung eines allgemeinen Modells einer Architektur des Psychischen. Es ist die Klärung einer längst überfälligen Frage, die Frage nach einer eindeutigen Bestimmung des *Gegenstandes der Psychologie* und seiner Einordnung in das (Gesamt-)System wissenschaftlicher Größen. Ihre Betrachtung ist nunmehr dringend notwendig geworden, da im Zuge der zunehmenden Erhebung biologischer Parameter zu psychologischen Konzepten die lange Zeit (vermeintlich) gewährte „Neutralität“ bei der Interpretation der Beziehung zwischen diesen Daten nicht mehr aufrecht erhalten werden kann. Auf der Basis der Teil-Ganzes-Beziehung in Form von Mechanismen unterschiedlicher Hierarchieebenen wird ein allgemeines Modell zur Einpassung von Daten unterschiedlichster Komplexitätsstufen etabliert. Dabei wird sich zeigen, dass das Psychische eine höherstufige Ebene in der natürlichen Organisation eines (menschlichen) Organismus bildet, zugleich aber dennoch mit seinen physischen Grundlagen eine voneinander unablösbare *Einheit* darstellt. Man könnte auch sagen, dass nach Maßgabe psychischer Eigenschaften aus dem diffusen Geflecht neuronaler Strukturen überhaupt erst funktionale Einheiten gebildet werden. Der Geist ist gewissermaßen das *strukturbildende* oder *formende* Prinzip auf allen Ebenen der natürlichen Organisation des Gehirns. Nicht nur trotz, sondern gerade *aufgrund* der Koextensivität bzw. Identität natürlicher Organisationsebenen handelt es sich beim Psychischen sowohl um ein *reales* als auch ein *kausal wirksames* Phänomen. Obgleich jede mentale Eigenschaft neben der psychischen Struktur, wie sie von der Psychologie beschrieben wird, eine Vielzahl (mikro-)struktureller Konfigurationen

aufweist (u.a. auf der neuronalen Ebene) und viele Aspekte psychischer Phänomene daher mit Rückgriff auf diese Mikroebenen von z.B. der Neurobiologie erklärt werden können, gibt es weder einen ontologischen noch einen explanatorischen Vorrang der Mikroebenen. Die Annahme eines solchen Primats ist lediglich das Produkt einer kognitiven Verzerrung, einer fehlgeleiteten Interpretation wissenschaftlicher Erklärungsmuster bzw. des Prozesses der wissenschaftlichen Erkenntnisgewinnung. Auch wenn sich das Psychische in das Gesamtsystem der Entitäten und Prozesse, wie sie von den Naturwissenschaften untersucht werden, einordnen lässt, und eine gewisse Reduktion psychischer Theorien und Konzepte auf solche der Biologie potentiell möglich sein mag, ist hierdurch keine Gefährdung des wissenschaftlichen Status‘ der Psychologie begründet. Ohne das begrifflich-konzeptuelle Gerüst der Psychologie bleibt auch die methodisch ausgefeiltste Neurowissenschaft der Zukunft bei der Erforschung mentaler Phänomene gänzlich blind.

1. Einleitung

Think of what consciousness feels like, what it feels like at this moment. Does it feel like billions of tiny atoms wiggling in place? (Carl Sagan, 1985, S. 255)



Abb. 1: Die blinden Männer und der Elefant.

Abbildung 1 illustriert ein altes asiatisches Gleichnis namens *Die blinden Männer und der Elefant* (wahrscheinlich aus dem Buddhismus oder Jainismus). Eine Gruppe blinder Männer untersucht einen Elefanten, um zu *begreifen* bzw. um sich einen *Begriff* zu machen, worum es sich bei diesem Tier handelt. Das Besondere bei diesem Gleichnis liegt nun darin, dass jeder der blinden Männer einen anderen Körperteil – aber jeder nur *einen* – des Elefanten ertastet. Bei einem Austausch über ihre gemachten Erfahrungen stellen sie fest, dass jeder der Weisen aufgrund seiner ganz individuellen Eindrücke zu einer eigenen, im Vergleich zu den anderen ganz unterschiedlichen Schlussfolgerung über das *Wesen* des ertasteten

Phänomens kommt. So heißt es beispielsweise in der buddhistischen Version, die im Udāna erzählt wird:

Da begab sich der König zu den Blinden und sprach zu ihnen: 'Ihr habt einen Elefanten erlebt, ihr Blinden?' - 'So ist es, Majestät. Wir haben einen Elefanten erlebt.' - 'Nun sagt mir, ihr Blinden: Was ist denn ein Elefant?' Da antworteten die Blindgeborenen, die den Kopf zu fassen bekommen hatten: 'Ein Elefant, Majestät, ist wie ein Kessel'; die das Ohr zu fassen bekommen hatten, antworteten: 'Ein Elefant, Majestät, ist wie ein Worfelkorb'; die einen Stoßzahn zu fassen bekommen hatten, antworteten: 'Ein Elefant, Majestät, ist wie der Stock eines Pfluges'; ein anderer, der den Rüssel erwischt hatte, antwortete: 'Ein Elefant, Majestät, ist wie ein Pflugbaum' [...] Und so prügeln sie aufeinander mit den Fäusten ein: 'So ist ein Elefant, nicht so! - Nein, so ist ein Elefant nicht; so ist er', und der König hatte seinen Spaß. [...] Weil sie nicht sehen, was Sinn und Unsinn ist, was Wahrheit und Unwahrheit ist, deshalb sind sie in Streit, Disput, Wortgefechte versunken und verletzen einander dauernd mit scharfen Worten: 'Das ist die Wahrheit, nicht das', - 'nein, das ist nicht die Wahrheit: So ist die Wahrheit!'" Aus diesem Anlaß tat der Erhabene [Buddha, M.L.] aus seiner Schau folgenden Anspruch:

Daran nun eben hängen sie,
die Pilger oder Geistlichen;
da disputieren, streiten sie,
als Menschen, die nur Teile seh'n. (zit. n. Schäfer, 2012)

Im Gleichnis steht die Blindheit (oder das im Dunkeln sein) dafür, dass man nicht in der Lage ist, etwas klar zu erkennen; der Elefant steht für eine Realität (oder eine Wahrheit). Die Geschichte soll aufzeigen, dass die Realität bzw. einer ihrer „Ausschnitte“ sehr unterschiedlich verstanden werden kann, je nachdem, welche Perspektive man einnimmt oder wählt. Eine vermeintlich absolute Wahrheit oder Ebene der Erkenntnis wird hierdurch ausgeschlossen. So wie es keine absolute Perspektive auf einen Gegenstand wie den Elefanten gibt, gibt es in der gesamten Erfahrungswirklichkeit immer nur unvollständige „Wahrheiten“. Wie Wimsatt (2007) es treffend ausdrückt:

Suppose that the five blind men of the legend, perceiving different aspects of the elephant, nonetheless recognize their common referent. "Good for them!" you might say. But not so fast: given the tremendous difficulties of reconciling their views of it, they nonetheless decide to treat their views as if they were of *different* objects! The net result is often not to talk about objects at all, but to emphasize predicates, or the systems of predicates grouped together as theories or models. (S. 180; Hervorh. im Orig.)

Im Rahmen dieser Arbeit soll aufgezeigt werden, dass die aktuelle Debatte um den Status des Psychischen und die Biologisierung der Psychologie im Grunde nichts anderes ist als das, was in diesem Gleichnis illustriert wird. Die Weisen wären in dem Fall die Psychologie, Biologie, Medizin und vielleicht die Philosophie, die sich um den primären Anspruch bei der Erforschung des Psychischen streiten. Allerdings gibt es einen Unterschied zum Gleichnis: Bei diesem sind es verschiedene Körperteile, die den verschiedenen Perspektiven auf den Elefanten zugrunde liegen. Im Falle der Wissenschaften, die um ihren vorrangigen Anspruch bei der Erforschung des Psychischen streiten, handelt es sich dagegen um *ein und denselben* Wirklichkeitsausschnitt, der allerdings auf unterschiedlichen repräsentationalen bzw. organisatorisch-strukturellen Niveaus *vergegenständlicht* wird.

1.1. Zum Aufbau der Arbeit

Das zentrale Ziel dieser Arbeit ist die Bereitstellung eines begrifflich-konzeptuellen Fundaments der Psychologie, wobei insbesondere die intertheoretische Einordnung psychischer Eigenschaften und Prozesse sowie psychologischer Konzepte und Theorien, die diese zum Gegenstand haben, vorgenommen werden muss. Es ist die Klärung einer längst überfälligen Frage, die Frage nach einer eindeutigen Bestimmung des *Gegenstandes der Psychologie* und seiner Einordnung in das (Gesamt-) System wissenschaftlicher Größen. Ihre Betrachtung ist nunmehr dringend notwendig geworden, da im Zuge der zunehmenden Erhebung biologischer Parameter zu psychologischen Konzepten die lange Zeit (vermeintlich) gewährte „Neutralität“ bei der Interpretation der Beziehung zwischen diesen Daten kaum mehr aufrecht gehalten werden kann. Das fehlende fachinterne Wissen um die (Grenzen der) Möglichkeit reduktiver oder gar reduktionistischer Bestrebungen erzeugt zudem bei vielen Psychologen den Eindruck einer – in ihrer „Erklärungshoheit“ psychischer Phänomene oder gar in ihrer bloßen Existenz – gefährdeten Disziplin, die sich gegenüber Vereinnahmungstendenzen seitens der Biologie erwehren muss.

Um direkt Missverständnissen vorzubeugen: Was diese Arbeit *nicht* leistet, ist eine Klärung der Frage, wie genau psychische Phänomene aus ihren physischen Basisstrukturen hervorgehen. Es wird kein Modell entwickelt, das die *Entstehung* oder das *Vorhandensein* des Psychischen selbst erklärt. Dies ist eine Aufgabe, die – wenn dies überhaupt in letzter Konsequenz möglich ist – primär durch die *empirische* Forschung in der Psychologie sowie den Kognitions- und Neurowissenschaften bewerkstelligt werden muss. Das, was die vorliegende Arbeit leisten soll, setzt

auf einer fundamentaleren Ebene an. Es ist die Klärung der Frage, was genau das Psychische überhaupt *ist*, in welcher Beziehung es zu seinen physischen Grundlagen, insbesondere zu den Eigenschaften und Prozessen des Gehirns, steht. Es muss sowohl ein inhaltlich kohärentes als auch empirisch plausibles begrifflich-konzeptuelles Instrumentarium zur Bestimmung der Geist-Gehirn-Relation entwickelt werden, auf dessen Basis es ermöglicht wird, psychologische Theorien auf eine solche Weise zu formulieren, dass diese bzw. die von ihnen verwendeten Konzepte eine adäquate ontologische Verankerung (resp. intertheoretische Einpassung) erfahren. Es werden alle erforderlichen Prinzipien für ein allgemeines Modell einer *Architektur des Geistes* bereitgestellt, das dann durchaus den *Rahmen* für die Klärung der oben genannten Frage nach der Entstehung von Geist oder Bewusstsein liefert. Zugleich wird aufgezeigt, inwiefern eine Reduktion des Geistes auf seine physischen Grundlagen, die Reduktion psychologischer Theorien auf solche der Neurobiologie ermöglicht wird und mit welchen Folgen dies für die Psychologie verbunden wäre. Hier nun eine Übersicht über die Struktur der Arbeit bzw. die Inhalte der einzelnen Kapitel: Nach der allgemeinen Einleitung und grundlegenden Begriffsklärung in Kapitel 1 wird in Kapitel 2 der dringende Bedarf einer begrifflich-konzeptuellen Grundlegung der Psychologie verdeutlicht. Hierzu wird zunächst die aktuelle Debatte um die Biologisierung der Psychologie primär anhand einer Ausgabe der *Psychologischen Rundschau* aus dem Jahr 2010 vorgestellt, wobei es vier Eckpfeiler sind, die diese Debatte kennzeichnen und hier näher beleuchtet werden (der verstärkte Gebrauch des Begriffes der *Ebene* zur Kennzeichnung der Gehirn-Geist-Relation, die Wiederentdeckung der Bedeutung des Leib-Seele-Problems, Befürchtungen hinsichtlich neuroreduktionistischer Bestrebungen sowie die Tendenz, der „reduktionistischen Bedrohung“ mit einem Verweis auf die fundamentale *phänomenale Differenz* zwischen psychischen und neu-

ronalen Prozessen zu begegnen). Es wird des Weiteren verdeutlicht, dass theoretische Arbeit in der Psychologie – vor allem in den letzten beiden Dekaden – zugunsten einer relativ unkoordinierten Akkumulation empirischer Detailbefunde stark vernachlässigt wurde. Während es zu vielen Neuerungen auf dem Gebiet der *Forschungsmethoden* gekommen ist, ist der *inhaltliche* Theoriefortschritt fast gänzlich zum Erliegen gekommen und die Schaffung eines kohärenten begrifflich-konzeptuellen Gerüsts der Psychologie in noch größere Ferne gerückt. Im Hintergrund dieser Entwicklung stehen zum einen wissenschaftssoziologische Fehlsteuerungsprozesse und zum anderen die gescheiterte „Vergangenheitsbewältigung“ in Bezug auf die vergleichsweise späte Loslösung der Psychologie von der Philosophie. Es wird für die Etablierung einer *Theoretischen Psychologie* argumentiert, der sowohl – ganz ähnlich wie der Theoretischen Physik – die Vereinheitlichung und Vereinfachung des theoretischen Fundaments der Psychologie als auch die Klärung erkenntnistheoretisch und ontologisch relevanter Fragen obliegt.

In Kapitel 3 wird die Frage nach dem *Ursprung des Leib-Seele-Problems* aufgeworfen. Zunächst wird herausgestellt, warum die alltagstheoretisch plausibel erscheinenden substanzdualistischen Positionen (sowie ihr extremer Konterpart: der Eliminative Materialismus) bei genauerer Betrachtung unhaltbar sind. Wenn sie nicht zu halten sind, woher rührt dann das Leib-Seele-Problem und die große intuitive Anziehungskraft dualistischer Körper-Geist-Vorstellungen? Dies zu zeigen, wird die zentrale Aufgabe dieses Kapitels sein. Als Ausgangspunkt oder auch Rahmen der Besprechung werden hierbei bestimmte Feststellungen des Philosophen Spinoza herangezogen, die – obgleich sie vor über 350 Jahren getätigt wurden – eine ungeahnte Aktualität haben. So rang er mit der Frage, wie es einerseits sein kann, dass zwischen Gehirn und Geist aufgrund des Scheiterns des Dualismus keine „substantielle“ Differenz

bestehen kann, sie aber andererseits so unterschiedlich in Erscheinung treten. Seine Ausführungen stehen dabei in einem äußerst aufschlussreichen Zwiespalt zwischen einer „subjektivistischen“ (beobachterabhängigen) und „objektivistischen“ (beobachterunabhängigen) Auflösung dieses Problems (was ich als *Spinozas Dilemma* bezeichne). So scheint die Diskrepanz zwischen Gehirn und Geist etwas mit unserer (begrifflichen oder sprachlichen) Weltauffassung oder Wahrnehmung zu tun zu haben, gleichzeitig dürfte sie aber doch mehr sein als nur ein bloßes Artefakt unserer Subjektivität bzw. unserer arteigenen Vergegenständlichung einer an sich undifferenzierten Welt. In diesem Kapitel wird zunächst die These weiterverfolgt, dass sich die phänomenale Differenz zwischen Gehirn und Geist auf die unterschiedliche Art und Weise unseres Zugriffs auf ein und denselben Wirklichkeitsausschnitt zurückgeht und das Leib-Seele-Problem demnach einen psychischen oder psychologisch erklärbaren Ursprung hat. Dabei wird zunächst erörtert, inwiefern es sich hierbei um eine *sprachlich-deskriptive* Differenz im Sinne unterschiedlicher Sprach- oder Beschreibungssysteme handeln kann, deren Inkompatibilität uns zu der Annahme führt, dass auch ihre Bezugsgegenstände verschieden sein müssen. Auch wenn sich dies anhand einiger Beispiele (z.B. aus der Entwicklungspsychologie) untermauern lässt, spricht viel dafür, dass dies nicht die eigentliche Ursache des Leib-Seele-Problems ausmacht, sondern nur ein Ausdruck einer tiefer liegenden psychischen Dynamik ist. Um diese freizulegen, ist ein Ansatz von zentraler Bedeutung, den Gustav Theodor Fechner (in Anlehnung an Spinoza) maßgeblich propagiert hat. Demnach geht die Verschiedenheit von Gehirn und Geist auf unterschiedliche *Wahrnehmungsperspektiven* (Eigen- vs. Beobachterperspektive) sowie den Fakt zurück, dass die neuronalen Grundlagen psychischer Prozesse für gewöhnlich erlebnismäßig nicht einsichtig sind (Metzinger, 2006, spricht hier von *phänomenaler Transparenz*).

Beide Aspekte oder Strukturmerkmale unserer psychischen Verfasstheit, Perspektivität und phänomenale Transparenz, werden deshalb in begrifflicher, wahrnehmungs- und entwicklungspsychologischer Hinsicht genauer untersucht, erweitert und präzisiert, sodass am Ende des Kapitels eine begrifflich exakte und empirisch begründete Ausformulierung der These erfolgen kann, dass es sich beim Leib-Seele-Problem um ein Problem handelt, das durch unsere Wahrnehmung und Interpretation der Welt überhaupt erst in ebendiese getragen wird. Allerdings stellt sich die Frage – und dies ist es, womit auch Spinoza gerungen hat –, ob es sich beim Leib-Seele-Problem tatsächlich um ein Phänomen handelt, das *ausschließlich* auf eine psychische Dynamik zurückgeht oder ob die wahrgenommene Differenz nicht zugleich eine irgendwie geartete objektive (beobachterunabhängige) Entsprechung (nicht-substantieller Art) hat. Die Klärung dieser Frage wird der Gegenstand der zwei folgenden Kapitel sein.

In Kapitel 4 wird herausgestellt, dass nicht nur hierbei, sondern überhaupt für die in dieser Arbeit verfolgte Absicht der Etablierung eines allgemeinen Modells einer *Architektur des Geistes*, der Begriff der *Ebene* von ausschlaggebender Bedeutung ist. Nach einer Besprechung des Ebenenbegriffs wird aufgezeigt, dass die in der Psychologie und kognitiven Neurowissenschaft verwendeten Ebenenmodelle für kognitive und phänomenale Zustände und Prozesse als eine wissenschaftliche oder systematische Ausprägungsform des zuvor entwickelten „Perspektivitätsansatzes“ verstanden werden können. Es gibt demnach verschiedene (funktional konzipierte) *Abstraktionsebenen*, die das, was in unseren Gehirnen geschieht, abbilden können. Ein und dasselbe System lässt sich auf verschiedenen Ebenen mit einem unterschiedlichen Granularitätsniveau analysieren. In diesem Kontext wird der menschliche Geist typischerweise als ein informationsverarbeitendes System wie ein Computer

aufgefasst. Auch die menschliche Informationsverarbeitung lässt sich demnach wie ein Computer auf verschiedenen Ebenen analysieren. Sie kann als ein funktionales Geschehen verstanden werden, bei dem – mehr oder weniger bewusst – bestimmte Berechnungsschritte zur Lösung von Problemen getätigt werden. Sie kann aber auch als ein neuronales Geschehen betrachtet werden, denn auch hier bedarf jeder abstrakte Informationsverarbeitungsprozess offenbar einer bestimmten physischen Grundlage. Die psychologischen Ansätze funktionaler und konnektionistischer Modelle weisen hierbei für gewöhnlich eine klare Hierarchisierung der postulierten Ebenen auf, während sich die „Ebenen“ bei den ableitbaren Modellen für phänomenale Prozesse wie Emotionen in keine klare Hierarchiefolge einpassen lassen. Insgesamt bleiben die in der kognitionswissenschaftlichen Forschungsliteratur konzipierten Architekturen allerdings noch zu unbestimmt, wenn es um die Klärung des (ontologischen) Status des Psychischen selbst bzw. die Frage nach einer „objektivistischen“ Auflösbarkeit von Spinozas Dilemma geht.

Die Beantwortung der bisherigen Leitfrage nach dem *Ursprung des Leib-Seele-Problems* wird in Kapitel 5 zum Abschluss gebracht. Um zu klären, ob auch das Psychische etwas konstituiert, was man als natürliche Organisationsebene bezeichnen kann, muss zunächst im Allgemeinen geklärt werden, unter welchen Bedingungen die Vielfalt der Erscheinungsweise eines Gegenstandes den Schluss berechtigt, dass es sich um eine strukturelle Kennzeichnung des betrachteten Objektes selbst (und nicht nur um ein „Artefakt“ unseres Zugriffs auf ebendieses handelt). Woran liegt es also, dass wohl kaum ein Zweifel daran angemeldet werden kann, dass es sich bei den Ebenen von Atomen, Molekülen, Zellen, Zellverbänden etc. um Aspekte handelt, die nicht nur auf unsere Auffassung der Welt zurückgehen, sondern um eine natürliche Ordnung oder Organisation der Welt bzw. des betrachteten Gegenstandes selbst? Es wird gezeigt, dass

es primär die *Teil-Ganzes-Beziehung* zwischen den verschiedenen strukturellen Ebenen eines Objektes oder Organismus ist, die dies verständlich machen kann. Nach einer Besprechung der hiermit zusammenhängenden *mereologischen* und *mechanistischen* Individuationskriterien natürlicher Organisationsebenen wird nachgewiesen, dass sich auch das Psychische unter Heranziehung dieser Individuationskriterien als natürliche Organisationsebene verstehen lässt und somit die Möglichkeit einer „objektivistischen“ Auflösung von Spinozas Dilemma gegeben ist. Abschließend wird sich darüber hinaus zeigen, dass beide Stränge, die „subjektivistische“ und „objektivistische“ Auflösung des Dilemmas, miteinander versöhnt werden können und so nicht nur ein sehr kohärenter Erklärungsansatz zum Ursprung des Leib-Seele-Problems, sondern auch eine tragfähige Basis für die – im weiteren Verlauf der Arbeit verfolgte – Formulierung einer Theorie zur begrifflich-konzeptuellen Grundlegung der Psychologie gegeben ist.

Bevor die eigentliche Ausformulierung eines allgemeinen Modells einer Architektur des Geistes bzw. die Bestimmung des Gegenstands der Psychologie erfolgen kann, muss in Kapitel 6 eine intensivere Beschäftigung mit den Gegenwartskonzepten der psychophysischen Relation, wie sie in der Philosophie des Geistes diskutiert werden, erfolgen. Dies ist aus zwei Gründen erforderlich. Erstens weil – wenn in den vorgehenden Kapiteln von einer Gehirn-Geist-Einheit oder einer psychophysischen Identität die Rede ist – dies nur in Abgrenzung zu (substanz-)dualistischen Theorien des Geistes geschehen ist, ohne dass eine solche Identitätsannahme selbst weitergehend analysiert und auf möglicherweise vorhandene Schwierigkeiten hin untersucht worden ist. Dabei wird sich zeigen, dass die in der Philosophie des Geistes diskutierte *Identitätstheorie* mit einer höchstproblematischen Interpretationsweise des Identitätskonzeptes verbunden ist, die zu einer Gleichsetzung der Position mit einem

recht kruden *reduktiven Physikalismus* oder *Materialismus* geführt hat. Das Konzept einer psychophysischen Identität muss daher gewissermaßen aus seinen „begriffsgeschichtlichen Fesseln“ befreit werden, um im nächsten Kapitel eine gewisse Rehabilitation (auf der Basis einer mereologisch begründeten Koextensivität) erfahren zu können. Zugleich wird bei dieser Besprechung der Blick für Problemstellungen freigelegt, die auch für die hier verteidigte Konzeption einer psychophysischen Identität von großer Bedeutung sind. Vor allem handelt es sich hierbei um das so genannte Argument der *multiplen Realisierbarkeit*, die so genannte *Erklärungslücke* oder auch das Argument der *starrten Designatoren* von Saul Kripke. Zweitens muss der Frage nachgegangen werden, ob es denn überhaupt nötig ist, eine ontologisch so „starke“ These aus der Nicht-Dualität von Gehirn und Geist abzuleiten, wie sie eine Identitätsannahme darstellt. Ist eine psychophysische Identität nicht der partout zu vermeidende Extremfall von „ontologischer Nähe“, der die Möglichkeit zu einer Reduktion der Psychologie (bzw. ihrer Entitäten) auf die Biologie (bzw. ihre Entitäten) eröffnen würde? Es muss daher geklärt werden, ob nicht doch bereits auf der Basis von – gegenüber einer Identitätsannahme – ontologisch „schwächeren“ Konzepten eine gehaltvolle psychophysische Theorie begründet werden kann. Neben den Konzepten der *Emergenz* und *Supervenienz* wird hierbei vor allem die derzeit in der philosophischen Debatte dominierende *Realisierungstheorie* genauer untersucht. Grundsätzlich ist der Einbezug dieser Konzepte der psychophysischen Relation – vor allem im Sinne der Etablierung einer größeren begrifflichen Klarheit – durchaus gewinnbringend. Dennoch wird sich zeigen, dass sie keine eigenständige Theorie des Geistes begründen können, die als Alternative zu einer psychophysischen Identitätstheorie infrage kommt. Verweigern sie sich prinzipiell, mit Identitätsannahmen zu operieren, weisen sie innere Widersprüchlichkeiten oder zumindest doch

eine wenig hilfreiche Beliebigkeit in der Charakterisierung der Gehirn-Geist-Beziehung auf. Lassen sie hingegen Identitätsannahmen zu, verschwinden diese Schwierigkeiten zwar, sie können dann jedoch nicht mehr als echte Alternative – in der hier verfolgten Hinsicht – zu einer identitätsbasierten Gehirn-Geist-Theorie angesehen werden.

Kapitel 7 stellt das Herzstück der vorliegenden Arbeit dar. In ihm werden die Befunde der vorangegangenen Kapitel in eine begrifflich kohärente und empirisch plausible Theorie, ein allgemeines Modell einer Architektur des Geistes integriert. Es wird hierbei untersucht, ob es unter der Berücksichtigung kompositorischer Verhältnisse möglich ist, eine Theorie psychophysischer Identität zu postulieren, die zwar annimmt, dass zwischen mentalen Prozessen oder Eigenschaften und ihren physischen Grundlagen ein gewisses Identitätsverhältnis besteht, ohne jedoch ein krudes *Nichts-anderes-als-Verhältnis* unterstellen zu müssen, wie es die „klassische“ Identitätstheorie oder auch die Realisierungstheorie (in ihrer begrifflich konsistentesten Ausformulierung) suggerieren. Auch wenn in Kapitel 5 bereits aufgezeigt wurde, dass es grundsätzlich *möglich* ist, psychische Prozesse als etwas zu betrachten, das in eine natürliche Strukturgliederung des menschlichen Organismus eingepasst ist und hier eine eigene höherstufige natürliche Organisationsebene konstituiert, wurde noch nicht geklärt, *wie genau* man sich dies vorzustellen hat. Inwiefern ist dies also nicht nur *möglich*, sondern auch in begrifflich kohärenter und empirisch plausibler Weise nachzuweisen, dass eine solche „Einpassung“ auf der Basis der getätigten Feststellungen *tatsächlich gegeben* ist. Es wird nachgewiesen, dass mentale Eigenschaften als *mikrobasierte Struktureigenschaften* des menschlichen Organismus anzusehen sind. Neben ihrer psychischen (kognitiven und/oder phänomenalen) Struktur ist eine mentale Eigenschaft durch eine (potentiell unendliche) Anzahl von koextensiven Mikrokonfigurationen gekennzeichnet. Aufgrund der ungenügen-

den ontologischen Eindeutigkeit einer Koextensivitätsbeziehung, die eine weitergehende Bestimmung notwendig macht, wird argumentiert, dass man diese Koextensivität als eine bestimmte Form psychophysischer *Identität* betrachten sollte. Es wird zugleich verdeutlicht, warum diese – mereologisch-mechanistisch basierte – Art einer Identitätsrelation nicht mit den negativen Konsequenzen verbunden ist, die für die „klassische“ Identitätstheorie kennzeichnend sind. Es handelt sich hierbei um keine *Identifizierung* des Psychischen in seiner neuronalen Seinsweise, vielmehr stellt die neuronale Ebene nur eine unter vielen *Mikrostrukturen* einer mentalen Eigenschaft dar. Auch wenn mentale Eigenschaften mikrobasierte Eigenschaften sind, bleibt ihre psychische Struktur für sie maßgeblich. Diese *strukturelle* Bestimmung einer mentalen Eigenschaft führt keineswegs zu einer Diskrepanz zu funktionalistischen Theorien des Geistes, wie sie für die Kognitionswissenschaft charakteristisch sind. Wie bereits Aristoteles wusste, stellen Form (Funktion) und Materie (Struktur) eine unzertrennliche Einheit dar, was nun mittels des hier verfolgten mereologisch-mechanistischen Ansatzes belegt werden kann. Obgleich aufgezeigt wird, dass es genau genommen keine Top-down-Kausalität (aber genauso wenig Bottom-up-Kausalität) in kompositorisch verfassten Systemen gibt, sondern etwas, das man als „mechanistisch vermittelte Effekte“ bezeichnen kann, wird sich zeigen, dass das Psychische dennoch ein *kausal wirksames* Phänomen ist. Auch wenn dies nicht zu einem zentralen Themenschwerpunkt erhoben wird, soll außerdem in diesem Kontext – in Form eines kurzen Abrisses – verdeutlicht werden, warum die hier entwickelte Identitätstheorie eine *Handlungs- und Entscheidungsfreiheit* des Menschen keineswegs untergräbt, sondern untermauert. Wenn aus der psychophysischen Koextensivität eine Identitätsannahme abgeleitet wird, muss sich diese, egal, ob sie sich nun in vielerlei Hinsicht von der „klassischen“ Identitätstheorie absetzen mag,

mit den zentralen Einwänden gegen die Identitätstheorie, den Argumenten der multiplen Realisierbarkeit, der starren Designatoren und auch der explanatorischen Lücke, auseinandersetzen oder zumindest aufzeigen, warum das hier vorgestellte Identitätskonzept nicht von diesen Einwänden betroffen ist. Auch dies leistet die Darstellung in Kapitel 7.

In Kapitel 8 wird die Debatte um die Biologisierung der Psychologie und den Stellenwert neuroreduktionischer Tendenzen aufgegriffen und aufgeklärt. Es müssen hierfür zunächst die begrifflichen Grundlagen freigelegt werden, die eine Reduktion psychologischer Theorien oder auch des Psychischen selbst ermöglichen würden. Von zentraler Bedeutung ist an dieser Stelle die Differenzierung zwischen eher harmlosen *reduktiven* und den eigentlich subversiven *reduktionistischen* Ansätzen. Daraufhin werden die in der Wissenschaftstheorie bestehenden reduktiven und reduktionistischen Modelle einer kritischen Prüfung unterzogen. Nach dem klassischen Modell der Theoriereduktion nach Ernest Nagel ist die Frage der Reduzierbarkeit im Wesentlichen eine Frage der logischen *Ableitbarkeit* der Zieltheorie von der Basistheorie sowie der *Verknüpfbarkeit* der beiden Theorien bzw. ihrer Begrifflichkeiten. Obgleich die grundsätzliche Erfüllbarkeit dieser Bedingungen im Falle psychologischer versus neurobiologischer Theorien infrage gestellt werden kann, da in vielen Fällen von wissenschaftlichen Konzepten keine gute intertheoretische Passung vorliegt, wird herausgestellt, dass selbst im Zuge einer erfolgreichen Reduktion nach dem Schema Nagels keine (abschließende) *explanatorische* Reduktion der Psychologie auf die Neurobiologie herbeigeführt würde. Eine Theoriereduktion dieser Qualität hätte also keineswegs die befürchteten negativen Auswirkungen auf die Psychologie und ihre Bezugsgegenstände, sondern würde ganz im Gegenteil eher zu einer gewissen „Aufwertung“ ebendieser führen, da sich die Psychologie und ihre Entitäten, Eigenschaften und Prozesse so in das wissenschaftliche

Gesamtsystem einpassen lassen und ihren – für viele Wissenschaftler immer noch gegebenen – „mysteriösen“ Status verlieren. Ganz ähnlich ist der bereits im Zusammenhang der Realisierungstheorie besprochene Ansatz *reduktiver Erklärungen* zu bewerten. Hochproblematisch sind hingegen die reduktionistischen Ansätze, die aus der schlechten intertheoretischen Passung zwischen Psychologie und Neurobiologie schlussfolgern, dass neurobiologische Theorien psychologische Theorien zu *ersetzen* haben. Sie machen den eigentlichen Kern der Neuroreduktionismus-Debatte aus. Es kann allerdings systematisch nachgewiesen werden, dass eine solche Ersetzung psychologischer Theorien aus *prinzipiellen* Gründen scheitert. Auch die vollkommenste Neurowissenschaft der Zukunft wäre bei der Erforschung des Gehirns ohne die Führung durch das begrifflich-konzeptuelle System der Psychologie blind. Erst durch die Zuordnung von psychologisch zu beschreibenden Funktionen zu bestimmten Hirnbereichen bzw. deren Verhalten simulierenden neuronalen Netzwerkmodellen wird den indifferenten, zusammenhangslosen und ziellos verlaufenden neuronalen Geschehnissen überhaupt erst eine Bedeutung zuteil. Vor dem Hintergrund der gewonnenen Erkenntnisse soll dann aufgezeigt werden, in welche Sackgasse die Reduktionismusdebatte verläuft, wenn man – wie leider sowohl in der Philosophie als auch Psychologie sehr verbreitet – in einer psychophysischen Identitätsannahme den Kern der Gefahr für die Eigenständigkeit der Psychologie sieht. Ebendiese Überzeugung hat zu zwei problematischen Ansätzen der Verteidigung der „Autonomie“ der Psychologie geführt, die hier näher beleuchtet werden sollen: Einerseits ist es die Annahme, dass sich die Psychologie dem Funktionalismus zu verschreiben hat, weil (nur) dieser vermeintlich die – einer Identitätsannahme entgegenstehende – multiple Realisierbarkeit mentaler Eigenschaften gewährleistet. Andererseits ist es die Überzeugung, dass anstatt einer Identität eine scheinbar harmlose

Mikro-Determinierung des Psychischen angenommen werden könne, die dafür sorgen würde, dass das Psychische nicht aus dem Kreis wissenschaftlich respektabler Phänomene herausfällt, zugleich aber auch nicht – wie bei einer Identitätsannahme – auf seine physischen Grundlagen reduziert werden kann. Ironischerweise folgt aber nun gerade aus diesen Prämissen, dass die Ebene neurobiologischer Prozesse gegenüber der Ebene psychischer Prozesse ein *ontologisches* Primat und zugleich die Neurobiologie gegenüber der Psychologie ein *explanatorisches* Primat hätte. Schließlich wird in diesem Kapitel systematisch nachgewiesen, dass ein solches Primat der Mikroebenen oder Mikroerklärungen unter den Voraussetzungen der hier entwickelten mereologisch-mechanistischen Architektur des Geistes scheitert und dies nicht trotz, sondern *aufgrund* der dabei getätigten psychophysischen Identitätsannahme. Dass es dennoch den Eindruck eines Primats der Mikroebenen bzw. von Mikroerklärungen geben mag, muss auf kognitive Tendenzen bzw. Verzerrungen zurückgeführt werden, die sich aus bestimmten (unreflektierten) inkonsistenten Hintergrundannahmen oder einer fehlgeleiteten Interpretation wissenschaftlicher Erklärungsmuster bzw. des Prozesses der wissenschaftlichen Erkenntnisgewinnung ergeben.

Abschließend werden in Kapitel 9 sowohl die zentralen Leitlinien als auch die Konsequenzen der erfolgten begrifflich-konzeptuellen Grundlegung der Psychologie zusammengetragen und eine präzise Bestimmung des Gegenstandes der Psychologie vorgenommen. Es wird dafür argumentiert, dass der – die Gegenwartspsychologie immer noch stark dominierende – funktionalistische Ansatz in das umfassendere mechanistische Paradigma eingebettet werden sollte, das der Vielfalt an Forschungszugängen bzw. der sich in ihnen aufzeigenden Mikrostrukturen der Architektur des Psychischen gerecht wird.

Es muss angemerkt werden, dass eine solche Arbeit wie diese, die um begrifflich-konzeptuelle Klarheit und vor allem um einen *theoretischen* Erkenntnisfortschritt bemüht ist, vor der Schwierigkeit steht, dass über die hier getätigten Annahmen i.d.R. keine über die Empirie zu gewinnende Entscheidung zu erzielen ist. So ist z.B. kein Experiment entsinnbar, das als hinreichende Grundlage einer Entscheidung zwischen einer dualistisch und monistisch ausgerichteten Theorie des Geistes fungieren könnte (vgl. Smart, 1959). Ebenso kann nicht über eine *empirische* Entscheidung geklärt werden, welcher Begriff einer natürlichen Organisationsebene die Beziehung zwischen den Entitäten verschiedener Granularitätsstufen eines Objektes adäquat beschreibt. Dennoch ist die Entscheidung zugunsten eines bestimmten Begriffes und schließlich auch die Etablierung einer begrifflich-konzeptuellen Basistheorie der Psychologie keineswegs eine Angelegenheit willkürlicher philosophischer Spekulationen. Sie muss allerdings auf eine Weise gewonnen werden, wie sie in der Psychologie eher in Vergessenheit geraten ist: So muss zunächst jeweils der Spielraum an möglichen begrifflichen Lösungen abgesteckt werden und die einzelnen Konzepte oder Konzeptionen dann einer eingehenden Analyse entzogen werden. Dabei müssen Konsequenzen aus den jeweiligen Ansätzen oder Konzepten abgeleitet werden, die sich sehr stark darin unterscheiden, inwieweit sie mit wissenschaftlichen Grundannahmen (wie z.B. dem Energieerhaltungssatz oder der Evolutionstheorie) oder auch mit ganz konkreten empirischen Befunden bzw. so genannten *Constraints* (zumeist aus der Empirie abgeleitete, den möglichen begrifflichen Spielraum absteckende Vorannahmen) kompatibel sind.

Zugleich unterscheidet sich die Aussagekraft oder der *explanatorische Gehalt* bestimmter Begriffe im Hinblick auf in einem bestimmten Kontext bestehende Kriterien, sodass auch hierüber eine gut begründete Ent-

scheidung getroffen werden kann. Wenn z.B. dafür argumentiert wird, dass es der mereologische (bzw. mechanistische) Begriff natürlicher Organisationsebenen ist, der in dieser Arbeit gegenüber einem rein quantitativen Ebenenbegriff weiterverfolgt wirkt, liegt das daran, dass nur der erstere Begriff die hier als Kriterium gesetzte Erklärungsfunktion erfüllt, einer (beobachterunabhängigen) Vielheit der Erscheinungsweise eines Objektes trotz der Akzeptanz seiner ontologischen Einheit einen Sinn zu verleihen (denn ein komplexes Objekt ist über die Teil-Ganzes-Beziehung auf natürliche Weise strukturiert, sodass es einerseits eine Vielzahl an Strukturebenen aufweist, die jedoch eine voneinander untrennbare ontologische Einheit darstellen). Der quantitative Begriff ist damit nicht als selbstwidersprüchlich oder inkohärent herausgestellt worden, sondern einfach an dieser Stelle zu dem gesuchten Zweck nicht zu gebrauchen.

Aber auch im umgekehrten argumentativen Fall muss diese Vorgehensweise herangezogen werden: Denn auch im Hinblick auf die Ausräumung kruder neuroreduktionistischer Bestrebungen kann nicht einfach ein Konzept herausgegriffen und exemplarisch bzw. als Stellvertreter für alle übrigen möglichen Varianten widerlegt werden, sondern es muss zunächst auch hier der Spielraum an (in diesem Kontext relevanten) Bedeutungen abgeklärt werden, um dann sukzessive jede dieser Varianten näher zu beleuchten und ggf. auszuräumen, indem entweder nachgewiesen wird, dass sie bereits auf begrifflicher Ebene Widersprüche aufweisen oder aber mit empirischen Constraints in Konflikt stehen, die sie – auch im Vergleich zu anderen Interpretationsvarianten – unplausibel oder gar gänzlich unannehmbar erscheinen lassen.

1.2. Eine grundlegende Begriffsklärung

Im Folgenden wird auf eine Reihe basaler Konzepte zurückgegriffen. Obwohl der Gebrauch dieser Begriffe allgemein hin als Selbstverständlichkeit betrachtet werden mag, soll dennoch eine zumindest grobe Inhaltsannäherung bzw. ein Hinweis auf den Gebrauch, die Differenzierung oder die synonyme Verwendung bestimmter Begrifflichkeiten gegeben werden.

Zunächst lässt sich eine Quadriga gebräuchlicher ontologischer Basiskategorien herausstellen: Zustände – Prozesse – Ereignisse – Eigenschaften. *Ereignisse* und *Zustände* beziehen sich auf die in einem bestimmten Zeitabschnitt vorliegende Verfasstheit eines Systems, wobei Ereignisse¹ von geringerer (zeitlicher) Ausdehnung sind als Zustände. Ein Mensch kann sich z.B. in einem bewussten *Zustand* befinden und beim Anblick von Sonnenblumen wird sich eine Gelb-Empfindung *ereignen*. Es kann im Gehirn zu zeitlich eng umgrenzten physiologischen *Ereignissen* kommen (z.B. Feuern von Neuronen im Okzipitallappen) oder es treten zeitlich (und evtl. räumlich) weiter gefasste *Zustände* auf (z.B. δ -Wellen-Gehirnstromaktivität, Aktivitätsmuster eines ganzen Neuronenensembles etc.). Ereignisse können ebenfalls als Instantiierungen von Eigenschaften angesehen werden, das heißt, dass ein System zu einem bestimmten Zeitpunkt über eine ganz spezifische Eigenschaftsausprägung verfügt.

¹ Ereignisse lassen sich gemäß der Bestimmung von Kim (1993) als Instantiierung von Eigenschaften auffassen, die eine räumlich und zeitlich begrenzte Ausdehnung haben. Diese Bestimmung hat sich größtenteils gegenüber der von Davidson (1970/2007) durchgesetzt, der unter Ereignissen nichtwiederholbare, zeitlich bestimmte Individuen verstand und sie – anstatt den heute üblichen Eigenschaften – zum Gegenstand von Identitätsbehauptungen gemacht hat.

Prozesse sind Abläufe in einem bestimmten Zeitabschnitt, also Veränderungen von Zuständen dieses Systems. Für die bessere Lesbarkeit des Textes werde ich aber alternativ von Zuständen, Ereignissen oder Prozessen sprechen, ohne dass die Argumentation hiermit eine inhaltliche Verschiebung erfahren soll. Wenn es an bestimmten Stellen wichtig sein sollte, dass z.B. gerade die Prozesshaftigkeit eines Ereignisses relevant ist, werde ich hierauf gesondert hinweisen.

Besondere Beschaffenheiten oder Charakteristika eines Systems werden als *Eigenschaften* dieses Systems bezeichnet. Es ist z.B. die phänomenale Qualität der Gelb-Empfindung beim Anblick einer Sonnenblume, aber auch die Entladungsrate von bestimmten Neuronenpopulationen, Neurotransmitter-konzentrationen etc. Eigenschaften werden in der Philosophie wie Typen von Mengen an Einzeldingen behandelt. Z.B. umfasst die Eigenschaft „das Haben einer Rot-Empfindung“ in der hier zugrunde liegenden Logik und Sprechweise alle möglichen Fälle von Instantiierungen von Rot-Empfindungen. Ein Typ oder eine Eigenschaft beschreibt die allgemeinen Merkmale von Objekten; eine Instantiierung (einer Eigenschaft) bezieht sich dagegen auf einen konkreten Einzelfall, in dem ein System eine bestimmte Eigenschaft aufweist.

Da sich der Eigenschaftsbegriff sowohl auf die Verfasstheit von Prozessen als auch Zuständen anwenden lässt (vgl. z.B. Putnam, 1967), hat er sich – wahrscheinlich nicht zuletzt aus Praktikabilitätsgründen – in der gegenwärtigen Besprechung des Leib-Seele-Problems in der Philosophie des Geistes durchgesetzt. In diesem Sinne entspricht der Vorgang eines mentalen Abwägens der Eigenschaft, sich in einem Abwägungsprozess zu befinden, die Aktivität von Neuronen im Okzipitallappen dem Haben eines bestimmten neuronalen Aktivitätsmusters im Okzipitallappen usw. Eigenschaften können demnach als eine Art Gattungsbegriff betrachtet werden, der alle Aussagen über Prozesse und Zustände mit umfasst.

kann. Allerdings soll aufgezeigt werden, dass diese Herangehensweise bzw. die in der Philosophie übliche Besprechung des Leib-Seele-Problems anhand von Eigenschaftsrelationen nicht unproblematisch ist (vgl. Kap. 6.5.4.4).

Des Weiteren ist es in der Philosophie des Geistes gebräuchlich, *mentale* Aspekte (Eigenschaften, Zustände...) und *physikalische* Aspekte (Eigenschaften, Zustände...) gegenüberzustellen. Im Angelsächsischen gibt es nicht die im Deutschen gehaltvolle Unterscheidung zwischen *physisch* (für „körperlich“) und *physikalisch* (für „die Physik bzw. ihre Eigenschaften und Entitäten betreffend“). Beide Begriffe werden hier mit *physical* bezeichnet. Aufgrund der Dominanz der Diskussion in der englischsprachigen Literatur hat sich daher auch in der deutschsprachigen Literatur der Begriff *physikalisch* eingebürgert, obgleich an vielen Stellen eigentlich *physisch* der korrektere Begriff wäre.

Des Weiteren ist der Gebrauch von *physikalisch* in der Philosophie des Geistes auch nicht auf die von der Physik postulierten Entitäten, Eigenschaften etc. beschränkt, sondern hiermit ist der gesamte Bereich des *naturwissenschaftlichen* Zugriffs auf die Welt gemeint. Insofern wären neuronale oder biochemische Eigenschaften eine besondere Gruppe physikalischer Eigenschaften. Weil der Begriff des Physikalismus bzw. die hiermit verbundene Frage, ob mentale Eigenschaften *physikalische* Eigenschaften sind, nicht unproblematisch ist, wird dieser Begriff zu Gunsten von „physisch“ nach Möglichkeit vermieden werden. Jedoch wird der Rekurs auf bestimmte philosophische Arbeiten der Gegenwart – aufgrund der hier sehr starken Verbreitung – nicht gänzlich ohne ihn auskommen können. Wenn hierbei gelegentlich von physikalischen Eigenschaften (im weiteren Sinne) die Rede ist, dann sind hiermit sowohl Eigenschaften gemeint, die sich in physikalischen (im engeren Sinne), chemischen oder biologischen Vergegenständlichungen der Welt finden.

Im Hinblick auf die Bezeichnung menschlichen Erlebens lassen sich wiederum Begrifflichkeiten herausstellen, die eine zweite Quadriga formieren: psychisch – seelisch – mental – bewusst. Obgleich es zwischen den Begriffen *psychisch* und *mental* für die Diskussion bestimmter philosophischer Fragen wichtige Bedeutungsunterschiede geben mag, brauchen diese in der vorliegenden Arbeit keine Berücksichtigung finden, da nicht einsichtig ist, dass eine weitergehende Differenzierung für den *hier* verfolgten Zweck dienlich wäre. Da der Ausdruck *seelisch* im heutigen Sprachgebrauch bestimmte religiöse Konnotationen hat („unsterbliche Seele“ etc.), obgleich diese dem ursprünglichen Sinngehalt nicht gerecht werden mögen und der Nutzen der „Rehabilitierung“ dieses Begriffes aktuell debattiert wird (vgl. Mack, 2012), wird hier auf seine Verwendung weitestgehend verzichtet werden.

Im Hinblick auf *bewusste* Zustände muss hingegen angemerkt werden, dass diese über eine phänomenale Präsenz oder einen Erlebnischarakter² verfügen, was nicht zwangsläufig über alle mentalen Zustände gesagt werden kann. Insofern wird von *bewussten* Zuständen auch nur dann gesprochen werden, wenn dies tatsächlich in dieser Form intendiert ist. Eine weitere wichtige Unterscheidung ist die zwischen kognitiven und phänomenalen Prozessen (Eigenschaften...). Da insbesondere bei der

² Christian Wolff (1720/1740) hat auf der Grundlage des lateinischen Ausdrucks *conscientia* das substantivierte *Bewusst seyn* in den deutschen Sprachraum eingeführt und zwar in der ursprünglichen Bedeutung, dass man sich selbst zum Gegenstand seiner Überlegungen machen kann. Man muss jedoch konstatieren, dass eine Bedeutungsverschiebung dahingehend stattgefunden hat, Bewusstsein „lediglich“ als das *Gegebensein von Innerlichkeit* anzusehen. *Bewusst zu sein* in diesem Sinne setzt also nicht zwangsläufig voraus, dass es sich um eine *reflektierbare* Form von Innerlichkeit handelt (insofern kann auch von Kleinkindern und vielen Tierarten durchaus angenommen werden, dass sie über Bewusstsein verfügen). Da es sich bei der Differenzierung zwischen Bewusstsein und *Selbstbewusstsein* also um eine wichtige begriffliche Unterscheidung handelt, die bei einer Gleichsetzung der Bedeutungen der beiden Begriffe verloren ging, wird ihr auch hier Rechnung getragen.

Diskussion des Leib-Seele-Problems relevante Unterschiede zwischen diesen beiden Gruppen von psychischen Prozessen bestehen, wird diese in der Philosophie des Geistes gebräuchliche begriffliche Differenzierung übernommen.

Für *kognitive* Zustände ist kennzeichnend, dass sie einen repräsentationalen Charakter haben, d.h., dass sie bestimmte Bedeutungseinheiten oder Informationen in unserem „mentalen Raum“ darstellen. Sie sind auf kognitive Leistungen bezogen wie z.B. die Fähigkeit zur Aufmerksamkeitssteuerung, Gesichtererkennung, Gedächtnisprozesse, Sprachkompetenzen, Wissensrepräsentationen und dergleichen. Zentrales Kennzeichen ist dabei die Intentionalität kognitiver Zustände. Kognitive Zustände sind immer auf etwas gerichtet, sie beziehen sich auf etwas, sie haben ein intentionales Objekt bzw. eine Repräsentation der Welt. Paradigmatische intentionale Zustände sind Wünsche, Überzeugungen oder Absichten.

Phänomenale Zustände sind dadurch gekennzeichnet, dass sie einen bestimmten *subjektiven* Erlebnischarakter (qualitativer Gehalt) haben. Es fühlt sich für uns auf eine ganz charakteristische Art und Weise an, in solchen Zuständen zu sein. Wie bereits erwähnt, hat dies vor allem Nagel (1974/1997) auf den Punkt gebracht, indem er aufgezeigt hat, dass phänomenale Zustände genau diejenigen Zustände sind, für die sich die Frage sinnvoll stellen lässt: Wie ist es, sich in solchen Zuständen zu befinden? Zu ihnen gehören alle Bereiche der sinnlichen Wahrnehmung oder Exterozeption (visuell, auditiv, gustatorisch, olfaktorisch, taktil), der Somatosensorik bzw. Interozeption (propriozeptive Wahrnehmungen, also der Wahrnehmung von Körperlage und -bewegung im Raum, sowie Viszerozeption, der Wahrnehmung von Organtätigkeiten, Schmerz, Hunger, Durst, Hitze- und Kälteempfindungen) und gefühlshafter Erlebnisse (Emotionen, Stimmungen). In der Philosophie des Geistes spricht

man im Hinblick auf phänomenale Zustände auch von Quale (Singular) bzw. Qualia (Plural).

Die Unterscheidung zwischen kognitiven und phänomenalen Zuständen erscheint intuitiv verständlich und auch nützlich zu sein, auch wenn sie innerhalb der Psychologie in dieser Klarheit kaum vorzufinden ist. Natürlich ist die Bezeichnung „kognitiver Zustand“ ein ureigener psychologischer Terminus, die Bezeichnung „phänomenaler Zustand“ dagegen ist ungebräuchlich. Hier werden die Bereiche Wahrnehmung (Extero- und Interozeption) und Emotionen in der Regel getrennt gesprochen, ohne eine innere Verbindung zwischen diesen Zuständen herzustellen. All diesen Zuständen ist aber gemeinsam, dass sie eine bestimmte Erlebnisqualität haben, was man von kognitiven Zuständen nicht *per se* sagen kann (man denke z.B. an das Wissen des Faktes, dass Paris die Hauptstadt von Frankreich ist oder bloße „Kantendetektoren“ in Modellen der visuellen Wahrnehmung etc.).

Problematisch ist aber, wenn aus dieser an sich sinnvollen Unterscheidung abgeleitet wird, dass sich kognitive Zustände – da als Gegenstück zu phänomenalen Zuständen konzipiert – *gerade dadurch* auszeichnen, dass sie keine phänomenalen Eigenschaften haben. Natürlich dürften kognitive Zustände zwar vor allem dadurch gekennzeichnet sein, dass sie relativ abstrakte Repräsentationen bestimmter Bedeutungen sind. Nichtsdestotrotz ist ein Teilbereich kognitiver Zustände, den man als *kognitives Bewusstsein* bezeichnen kann, durch eine phänomenale Präsenz gekennzeichnet (damit bildet das kognitive Bewusstsein genau genommen einen Spezialfall phänomenaler Zustände). So ist einem z.B. das Nachdenken über etwas auf eine bestimmte Art und Weise selbst gegeben. Wie Heckmann und Walter (2001/2006) sagen, hat die Überzeugung, dass Kriegsverbrecher hart bestraft werden sollten, durchaus eine qualitative (emotionale) Komponente. Ebenso können phänomenale Zu-

stände durchaus intentionale Eigenschaften aufweisen, wie z.B. eine Schmerzempfindung auf eine Verletzung meiner Bandscheibe gerichtet sein kann. Während sich also Farbempfindungen auf das von bestimmten Gegenständen in Abhängigkeit von ihren Oberflächeneigenschaften reflektierte Licht *beziehen* und auf eine bestimmte Art und Weise in unserem Bewusstsein darstellen, beziehen sich Zustände des kognitiven Bewusstseins auf mentale Gegebenheiten mit intentionalem³ Charakter, also auf Wissensrepräsentationen, Überzeugungen, Einstellungen etc., und können ebenfalls einen *qualitativen* Gehalt haben.

Ein Gegenstand der Diskussion in der Philosophie bleibt, ob man von allen mentalen Zuständen sagen kann, dass sie einen phänomenalen *und* intentionalen Gehalt haben „oder ob wir es nicht doch eher mit einem Spektrum zu tun haben, an dessen Enden jeweils reine Fälle von intentionalen bzw. phänomenalen Zuständen liegen“ (ebd., S. 17; Hervorh. im Orig.). Für die zweite Alternative dürfte allerdings mehr sprechen, denn wie bereits gesagt wurde, gibt es Aspekte kognitiver Zustände (Kantendetektion, grammatikalische Ordnungsstrukturen bei der Analyse von Sprache etc.), die offenbar keine direkten phänomenalen Eigenschaften aufweisen.

Mit *Physikalismus* werden in der zeitgenössischen Philosophie des Geistes für gewöhnlich alle Positionen bezeichnet, die man klassischerweise als „Materialismus“ bezeichnet hat. Wenn man davon ausgeht, dass es – aufgrund der berechtigten Kritik an substanzdualistischen Vorstellungen – letztendlich nur eine substantielle Wirklichkeit gibt, dann ist

³ Hiermit ist nicht „Absichtlichkeit“ oder dergleichen gemeint, sondern im Sinne Franz Brentanos (1874/1971) das grundsätzliche Charakteristikum kognitiven Bewusstseins, das sich dieses immer *auf etwas bezieht* bzw. das dieses immer *auf etwas gerichtet* ist. Dabei ist es egal, ob es sich hierbei um reale oder nur vorgestellte Objekte oder Eigenschaften handelt (weshalb auch der Begriff der *intentionalen Inexistenz* in diesem Kontext verwendet wird).

diese in den Augen der Materialisten etwas Materielles. „Der Materialismus geht davon aus, daß die Materie die einzige die Wirklichkeit konstituierende Substanz ist. Er bestreitet die Existenz einer speziellen geistigen Substanz oder unreduzierbar geistiger Eigenschaften“ (Birke, 2004). Allerdings ist den Materialisten im Zuge der neueren naturwissenschaftlichen Entdeckungen ihr eigentlich als selbstevident eingeschätzter Materiebegriff zunehmend abhandengekommen, da sich die lebensweltlich gegebenen Eigenschaften des Gewichts, der Festigkeit bzw. Undurchdringlichkeit von Gegenständen auflöst, wenn man sie auf mikrophysikalischer Ebene untersucht. Die hier angesiedelten „Materie“-Teilchen weisen praktisch keine Masse auf, es herrscht eine Permeabilität vor usw., sodass die Eigenschaften, die den Begriff von Materie für gewöhnlich konstituiert haben, stark infrage gestellt worden sind. Aufgrund dieser Entwicklung haben die Vertreter des Materialismus die Kopplung an den Materiebegriff größtenteils fallen gelassen und bezeichnen ihre Position stattdessen als „Physikalismus“. Maßgeblich ist für diese Position, dass es die von der Physik postulierten Entitäten und Eigenschaften sind, die die Wirklichkeit konstituieren (welche dies genau sind, ist an den Forschungsstand der physikalischen Forschung gekoppelt und somit durchaus wandelbar), während alle anderen Entitäten und Eigenschaften in irgendeiner Form von denjenigen der Physik abgeleitet sind, durch diese determiniert werden oder identisch mit ihnen sind. Als eine starke Form des Physikalismus (teilweise gibt es auch einen synonymen Gebrauch) werden daher die verschiedenen Varianten einer Identitätstheorie (vgl. Kap. 6.1) aufgefasst, die eine psychoneuronale Identität und Reduzierbarkeit annehmen⁴.

⁴ Ob eine Identität allerdings zwangsläufig mit einer psychophysischen Reduktion einhergehen muss bzw. was eine solche Reduktion genau bedeutet (und was nicht), wird an

Eine letzte Anmerkung betrifft den nun schon häufig verwendeten Begriff *Leib-Seele-Problem*. Genau genommen stecken in ihm zumindest zwei potentielle Missverständnisse. Denn wie der Begriff des Seelischen weist auch der Begriff des Leibes bestimmte religiöse (christliche) Konnotationen auf, die nicht Gegenstand der vorliegenden Arbeit sind. Der Begriff *Körper-Geist-Problem* (wie im Englischen: *mind-body problem*) ist allerdings auch ungünstig, da hiermit genau genommen die Beziehung des gesamten Körpers zu mentalen Zuständen thematisiert würde, obgleich dies der empirischen Forschungslage nicht gerecht wird, die primär nur das Nervensystem mit „Geistigkeit“ in Verbindung bringt. Exakter wären eigentlich die Bezeichnungen *Gehirn-Geist-Problem* oder auch *Gehirn-Bewusstsein-Problem*, da diese sprachlich aber recht ungelenkt sind und sich (wohl nicht zuletzt deshalb) der traditionelle Begriff des „Leib-Seele-Problems“ in der Literatur weiterhin gehalten hat, werde auch ich ihn primär verwenden (wenn gelegentlich eine der anderen Bezeichnung verwendet wird, geschieht dies in synonyme Verwendung).

späterer Stelle diskutiert werden (vgl. Kap. 8 und 9).

2. Der dringende Bedarf einer begrifflich-konzeptuellen Grundlegung der Psychologie

Im Folgenden soll der dringende Bedarf einer begrifflich-konzeptuellen Grundlegung der Psychologie aufgezeigt werden. Als Ausgangspunkt wird hierbei die aktuelle Debatte um die Biologisierung der Psychologie aufgegriffen und zunächst anhand bestimmter Charakteristika dieser Debatte verdeutlicht, dass es profunde Fragen begrifflich-konzeptueller Natur sind, die die Psychologie im Zuge der zunehmenden Vernetzung der wissenschaftlichen Disziplinen und ihrer Forschungsmethoden beantworten muss. Es ist primär die Frage nach dem eigentlichen *Gegenstand der Psychologie*, die einer präzisen Klärung bedarf, um viele der weitergehenden Herausforderungen (oder gar Bedrohungen), denen sich die Psychologie ausgesetzt sieht (wie z.B. der Gefahr eines Neuroreduktionismus) begegnen zu können. Wenn – wie sich herausstellen wird – die profunde Auseinandersetzung mit dem Leib-Seele-Problem der entscheidende Schlüssel zur Klärung dieser Aspekte bzw. dieser Debatte ist, stellt sich die Frage, warum die Beschäftigung mit grundsätzlichen theoretischen bzw. begrifflich-konzeptuellen Angelegenheiten in der Psychologie zugunsten einer relativ unkoordinierten empirischen Detailforschung seit einigen Jahrzehnten stark vernachlässigt wurde. Es wird sich zeigen, dass eine solche Abwehr gegenüber theoretischer Arbeit neben bestimmten allgemeinen wissenschaftssoziologischen Fehlsteuerungsprozessen vor allem aus einer Art gescheiterten „Vergangenheitsbewältigung“ entspringt. Da Psychologen ihre Disziplin für gewöhnlich als *verspätete Wissenschaft* wahrnehmen, deren „Rückständigkeit“ gegenüber den klassischen Naturwissenschaften der späten Loslösung von der Philosophie geschuldet ist (was ich als *Mythos der verspäteten Wissenschaft*

bezeichne), scheint in der starken Abneigung gegenüber theoretischen Ansätzen, die nicht ins strenge Schema des Methodenkanons passen und nicht unmittelbar mit einer empirischen Erhebung verbunden sind, immer noch die Angst zu liegen, dass man einen Rückfall in die Umklammerung durch die Philosophie erleiden könnte.

Abschließend wird für die Etablierung einer *Theoretischen Psychologie* – analog zur Theoretischen Physik – argumentiert und anhand einer Reihe von Beispielen verdeutlicht, warum konzeptuelle Arbeit, die theoretische Fundierung oder auch die Schaffung eines möglichst einheitlichen theoretischen Rahmens für die Psychologie von substantieller Bedeutung sind und dies gerade im Hinblick auf eine adäquate *empirische* Forschung.

2.1. Die aktuelle Debatte um die Biologisierung der Psychologie

Die akademische Psychologie befindet sich in einer Phase des Umbruchs und der Unsicherheit. Die Ursache liegt in der in den 1990er Jahren einsetzenden Etablierung der so genannten kognitiven Neurowissenschaft, die aus einem Verbund aus Psychologie, Medizin, (Neuro-)Biologie, Informatik und der Philosophie (des Geistes) besteht. Einerseits erlebt man seitdem eine ungeheure Dynamik und Aufbruchstimmung, andererseits werden durch die Veränderungen auch eine Menge Ängste geschürt; und dies allem Anschein nach viel stärker in der Psychologie als in allen anderen in diesem Wissenschaftsverbund beteiligten Disziplinen. Die Ursache hierfür liegt in erster Linie in einem wahrgenommenen Kompetenz- bzw. Zuständigkeitsverlust der Psychologie. Vor allem die Biologie scheint in viele Bereiche vorzudrängen, die vormals genuin psychologisches Terrain waren. So werden viele Lehrstühle, die sich mit

der menschlichen Kognition beschäftigen, unter Etiketten wie „kognitive Neurowissenschaften“ oder „Verhaltensneurobiologie“ in die Biologie bzw. Medizin verlagert (Birbaumer, 2003). Gleichzeitig wird beklagt, dass es immer mehr Lehrstühle in der (biologischen) Psychologie gibt, die mit Nicht-Psychologen aus vor allem der Biologie besetzt werden (Jäncke & Petermann, 2010). Die Allgemeine Psychologie (und damit auch die kognitive Psychologie) ist innerhalb des Faches stark mit der biologischen Psychologie verschmolzen, sodass viele kognitionspsychologische Theorien und Modelle zunehmend durch neurowissenschaftliche Daten erweitert, wenn nicht (zumindest vermeintlich) sogar ersetzt werden. Aber auch in allen anderen Teildisziplinen der Psychologie lassen sich starke Einflüsse aus den Neurowissenschaften ausfindig machen. So gibt es auch in der Entwicklungspsychologie eine Dynamik in Richtung einer stärkeren Fokussierung auf eine Entwicklungsneuropsychologie (vgl. z.B. Kaufmann, 2007), die Etablierung einer *Psychoneuroimmunologie* sowie *Psychoneuroendokrinologie* (vgl. z.B. Schedlowski & Tewes, 1996; Ehlert & von Känel, 2011) und selbst in der pädagogischen Psychologie finden sich zunehmend Ansätze, die eine Anbindung an die neurobiologische Grundlagenforschung suchen (z.B. die Textsammlung von Herrmann, 2009: *Neuro-Didaktik*).

Die *Psychologische Rundschau* als Organ der *Deutschen Gesellschaft für Psychologie* (DGPs) hat im Jahr 2010 ein Themenheft hierzu unter dem Titel „Wie viel Biologie braucht die Psychologie?“ veröffentlicht, was die aktuelle Brisanz dieses Themas in der Psychologie unterstreicht. Die Autoren sind unterschiedlich stark über die oben skizzierten Zusammenhänge besorgt. Mancher nimmt bereits Auflösungserscheinungen des ganzen Faches wahr. Mancher sieht den aktuellen Entwicklungen eher gelassen entgegen und setzt auf eine Entlarvung der Schwemme an neuronal begründeten Pseudo-Erklärungen psychologischer Forschungs-

gegenstände (profunde Wissenschaftlichkeit ausstrahlende fMRI-Bilder als Garnitur zu fragwürdigen Forschungsergebnissen; Jäncke & Petermann, 2010; Strack, 2010) und einer damit einhergehenden Rückbesinnung auf die „klassischen“ psychologischen Forschungsmethoden und -modelle. Die Texte sind vor allem durch vier aufschlussreiche Charakteristika gekennzeichnet, die im Folgenden dargestellt und bewertet werden.

2.1.1. Ebenenmodelle des Geistes: Begriffliche Vagheit und Potential

Es zeichnet sich eine Tendenz ab, die sich in wissenschaftstheoretischen Ansätzen zur Psychologie schon lange vorfindet (vgl. Breuer, 1977), nun aber seit dem Aufkommen der kognitiven Neurowissenschaften eine virulente Verbreitung gefunden hat: Es ist die Anwendung des Begriffes der *Ebene* (engl.: *level*), die immer dann zum Zuge kommt, wenn eine Kennzeichnung des Verhältnisses der Gegenstände oder Theorien verschiedener (sowohl intra- als auch interdisziplinärer) Forschungsansätze vorgenommen werden soll. Auch wenn häufig der Eindruck erweckt wird, dass hiermit schon viel über das Verhältnis von Gehirn und Geist gesagt wäre, bleibt der bloße, inhaltlich nicht näher bestimmte, Rekurs auf den vieldeutigen Begriff der Ebene sehr vage.

Auch in der *Psychologischen Rundschau* wird begrifflich *unisono* eine psychologische Ebene von einer biologischen Ebene unterschieden und z.B. gefragt, ob und inwieweit man „mit den neuen biologischen Methoden psychologische Fragestellungen auf einer anderen Ebene bearbeiten [kann]“ (Jäncke & Petermann, 2010, S. 176). Die Biologie und die Psychologie beschäftigen sich also offenbar mit ein und denselben Bezugsgegenständen, untersuchen diese jedoch auf unterschiedlichen *Analyseebenen* bzw. entwerfen „ebenspezifische“ Theorien und Modelle. Durchleuch-

tet man die psychologische Fachliteratur, ist vor allem von Forschungs-, Untersuchungs- oder Betrachtungsebenen die Rede (vgl. z.B. Marr, 1982; Smolensky, 1988; Schandry, 2006; Goldstein, 2008; Schumacher, 2009). Hiermit scheinen zunächst so etwas wie unterschiedliche (wissenschaftliche oder konzeptuelle) *Perspektiven* auf ein und denselben Forschungsgegenstand gemeint zu sein. An anderen Stellen ist aber auch von (natürlichen) *Organisationsebenen* die Rede. So besitzen wir nach Mausfeld (2010) noch keine „theoretische Konzeption darüber, auf welcher physikalischen Ebene der Organisation des Gehirns die für eine Erklärung mentaler Phänomene relevanten Prinzipien verankert sind“ (S. 181). Während also mit den Analyseebenen die fachspezifischen modellhaften Vorstellungen und Theorien zu einem Gegenstandsbereich verbunden sind, scheint der Begriff der Organisationsebene zu implizieren, dass wir es hier mit einer echten (beobachterunabhängigen) strukturellen Beschaffenheit (eines Gegenstandes) zu tun haben. Die Wirklichkeit selbst (und nicht nur ihre Repräsentation in wissenschaftlichen Modellen) wird als geschichtet aufgefasst.

Eine aufschlussreiche Darstellung findet sich im international renommierten Lehrbuch der *Wahrnehmungspsychologie* von Goldstein (2008): „Die Idee hinter dem Konzept der Untersuchungsebenen ist, dass wir Prozesse auf unterschiedlichen Maßstäben betrachten können“ (S. 9), was er daraufhin am Beispiel der Beobachtung eines bestimmten Verkehrsgeschehens und hiermit verknüpften Erkenntnisinteresses beschreibt:

Wenn wir [...] die Funktionsweise eines individuellen Autos verstehen wollen, müssen wir das Bild bis auf die Ebene einzelner Straßen vergrößern, um ein Auto genauer zu betrachten. Vielleicht unternehmen wir damit sogar eine Probefahrt, um Leistungsmaße wie Beschleunigung, Bremsvermögen und Handling zu erheben. Durch einen Blick unter die Motorhaube können wir das Bild dann sogar noch

weiter vergrößern, um einen Einblick in die interne Funktionsweise des Automotors zu gewinnen. Entscheidend ist, dass die Untersuchung des Autos auf all diesen Ebenen uns ein viel tieferes Verständnis von Autos ermöglicht, als wir es aufgrund der Untersuchung nur einer einzigen Ebene erreichen könnten. Dieselbe Überlegung ist auch auf unsere Untersuchung der Wahrnehmung anwendbar. In diesem Buch werden wir ein tiefes Verständnis der Wahrnehmung dadurch gewinnen, dass wir sie auf zwei verschiedenen Untersuchungsebenen betrachten. Wir werden behandeln, wie die Wahrnehmung einer Person mit der Stimulation in der Umwelt zusammenhängt. Dies erreichen wir mittels der psychophysischen Untersuchungsebene, die auf den *Stimulus-Wahrnehmung-Zusammenhang* ausgerichtet ist [...]. Wir werden ebenfalls behandeln, wie die Wahrnehmung einer Person mit physiologischen Prozessen zusammenhängt, die innerhalb der Person ablaufen. Dies wird als physiologische Untersuchungsebene bezeichnet; auf dieser Ebene stehen die *Stimulus-Physiologie-Beziehung* [...] und die *Physiologie-Wahrnehmung-Beziehung* [...] im Mittelpunkt. (ebd., S. 9f; Hervorh. im Orig.)

Goldstein beschreibt Untersuchungsebenen hier zunächst als (wahrnehmungsbezogene) Perspektiven, die man einem Gegenstand gegenüber einnehmen kann, wobei sich unter der Einnahme ganz bestimmter Perspektiven – über die Zugriffsmöglichkeit auf strukturelle Organisationsprinzipien – die Funktionsweise dieser Gegenstände offenbaren würde, egal, ob es sich hierbei um ein technisches Objekt wie ein Auto oder das Phänomen der menschlichen Wahrnehmung selbst handelt. Einerseits scheinen Ebenen also ein eher subjektives oder beobachterabhängiges Phänomen zu sein, andererseits scheinen bestimmte Perspektiven die (beobachterunabhängigen) Organisationsprinzipien des betrachteten Gegenstands selbst zu enthüllen. Wie genau lässt sich dies aber nun verständlich machen? Was genau kann bzw. was sollte unter Ebenen (oder auch Perspektiven) verstanden werden? Wann genau ist es gerechtfertigt, diesen von Goldstein vollzogenen „Überstieg“ von einer bloßen (Wahrnehmungs-)Perspektive hin zu einer natürlichen Organisationsebene zu tätigen? Und – was für diese Arbeit eine ganz entscheidende Frage sein wird – lässt sich dieser Schritt auch im Hinblick auf das Psychische selbst

tätigen? Konstituiert es also selbst eine natürliche Organisationsebene des menschlichen Organismus?

2.1.2. Gefährdung des Faches durch Neuroreduktionismus?

Es ist eine schwerlich zu leugnende Tatsache, dass alle psychischen Prozesse eine physische Grundlage aufweisen. Daher liegt es in der Natur der Sache, dass wichtige Bestimmungsstücke des menschlichen Erlebens und Verhaltens nicht nur aus sich selbst heraus (d.h. über andere psychische Prozesse), sondern auch mit Rekurs auf ihre physischen (neuronalen) Grundlagen erklärt werden können. Dies hat bereits die vor allem von Gustav Theodor Fechner geprägte Psychophysik gewusst und zeigt sich nun verstärkt im Zuge der modernen Hirnforschung oder biologisch orientierten Psychologie (vgl. z.B. Prinz & Müsseler, 2008, S. 7ff). Insofern sagt auch Goldstein (2008), dass wir – wenn wir die Wahrnehmung wirklich verstehen wollen – beide der von ihm genannten Ebenen sowie deren genaue Beziehung untereinander untersuchen müssen:

Hierbei ist es wichtig, sich zu vergegenwärtigen, dass es trotz unserer Unterscheidung zwischen diesen beiden Ebenen ein großes Maß an Austausch oder, technisch formuliert, „Übersprechen“ zwischen ihnen gibt. Somit sind die Ebenen sehr eng miteinander verbunden. Genauso wie die mechanischen Eigenschaften eines Automotors die Leistung des Autos beeinflussen, so beeinflussen die physiologischen Eigenschaften des Wahrnehmungssystems einer Person deren Wahrnehmung. Eine der Konsequenzen der engen Beziehung zwischen der psychophysischen und der physiologischen Untersuchungsebene besteht darin, dass wir etwas über eine der beiden Ebenen lernen können, indem wir die andere untersuchen [...]. Betrachten Sie beispielsweise einen Patienten, der infolge einer Blockade einer der hirnzuführenden Arterien einen Schlaganfall erlitten hat. Wenn wir das Sehvermögen dieses Patienten testen, so stellen wir fest, dass er Schwierigkeiten hat, Objekte in einem Bereich zu seiner Linken zu sehen. Diese Messung der Beziehung zwischen Stimulus (einem Objekt zu seiner Linken) und Wahrnehmung (der Fähigkeit, das Objekt zu sehen) legt den Schluss nahe, dass in einem bestimmten Areal in der rechten Hirnhemisphäre des Patienten ein Problem besteht (da die rechte

Hemisphäre für das Sehen in der linken Hälfte des Gesichtsfeldes verantwortlich ist). Wir können diese Vermutung dann überprüfen, indem wir eine Gehirnuntersuchung mit einem bildgebenden Verfahren durchführen, durch das eine Darstellung der physiologischen Tätigkeit des Gehirns erzeugt wird. (S. 11)

Wenn aber nun psychische Phänomene – vermittelt über vor allem bildgebende Verfahren – zunehmend über den Rekurs auf die neuronale Untersuchungsebene erklärt werden können, stellt sich die Frage, was dies für die Psychologie bzw. den Stellenwert psychologischer Erklärungen bedeutet. Wie in der *Psychologischen Rundschau* herausgestellt wird, leiten viele Neurowissenschaftler und auch einige Philosophen bzw. Wissenschaftstheoretiker (z.B. Churchland, 1979, 1989; Bickle, 1998, 2003) hieraus ab, dass die Psychologie als Wissenschaft bald obsolet sein wird. So sei die (neuro-)biologische Forschungsebene die eigentlich wissenschaftlich respektable Erklärungsebene des Geistes, während die vermeintlich weniger exakten psychologischen Theorien nur den Charakter von temporären Hilfskonstruktionen hätten:

Warum rückt aber dann das Thema der Beziehung von Psychologie und Biologie gegenwärtig wieder in den Focus des Interesses? Der Grund liegt in dem in neurowissenschaftlich orientierten Bereichen geäußerten Anspruch, dass die ‚eigentliche‘ Erklärung psychologischer Phänomene auf einer ganz spezifischen Analyseebene der Biologie, nämlich der Ebene neuronaler Prozesse, zu suchen sei und dass psychologische Theorien bestenfalls vorübergehende Hilfskonstruktionen seien, bis man auf neuraler Ebene die ‚eigentliche‘ Erklärung für die betrachteten psychologischen Phänomene gefunden habe. (Mausfeld, 2010, S. 181)

Die im Hintergrund stehende Überzeugung ist hier offensichtlich, dass die Verfügbarkeit von Mikroerklärungen – wie sie die Neurobiologie in Bezug auf psychische Phänomene bereitstellt – eine grundsätzliche und einseitige Determination des Psychischen durch seine neurobiologischen Grundlagen widerspiegelt (vgl. Klee, 1984; Hüttemann, 2004). Aus der

bloßen Detektierbarkeit der physischen Grundlagen psychischer Prozesse im Gehirn wird abgeleitet, dass hier die *eigentlichen* Mechanismen installiert sind, die das Erleben und Verhalten bedingen. Weil wir also mehr und mehr wissen, wo unsere Emotionen, Entscheidungen oder moralischen Grundsätze im Gehirn verankert sind, muss es sich demnach bei all diesen Phänomenen um genuin (neuro-)biologische Größen handeln.

In diesem Sinne argumentiert z.B. der Neurowissenschaftler Wolf Singer (2004), wenn er sagt, dass wir erstens durch unsere *neuronalen* Verschaltungen vollkommen festgelegt (und daher „unfrei“) seien und dass zweitens das, was wir früher einmal als *psychische* Phänomene bezeichnet haben, in Wirklichkeit objektivierbare Verhaltensleistungen sind, die nun endlich im Zuge der neurowissenschaftlichen Neubegründung der Psychologie in kohärenter Weise durch naturwissenschaftliche Beschreibungssysteme erfasst werden könnten:

Die zunehmende Verfeinerung neurobiologischer Meßverfahren hat nunmehr die Möglichkeit eröffnet, auch die neuronalen Mechanismen zu analysieren, die höheren kognitiven Leistungen komplexer Gehirne zugrunde liegen. *Somit werden auch diese, oft als psychische bezeichneten Phänomene zu objektivierbaren Verhaltensleistungen, die aus der Dritten-Person-Perspektive untersucht und beschrieben werden können.* Zu diesen mit naturwissenschaftlichen Methoden untersuchbaren Leistungen zählen inzwischen auch solche, die uns bereits aus der Ersten-Person-Perspektive vertraut sind. Darunter fallen Wahrnehmen, Vorstellen, Erinnern und Vergessen, Bewerten, Planen und Entscheiden, und schließlich die Fähigkeit, Emotionen zu haben. Alle diese Verhaltensmanifestationen lassen sich operationalisieren, aus der Dritten-Person-Perspektive heraus objektivieren und im Sinne kausaler Verursachung auf neuronale Prozesse zurückführen. Somit erweisen sie sich als Phänomene, die in kohärenter Weise in naturwissenschaftlichen Beschreibungssystemen erfaßt werden können. (S. 35; Hervorh.: M.L.)

Muss sich die Psychologie im Angesicht solcher Aussagen tatsächlich vor einem *Neuroreduktionismus* fürchten? Um diese Frage zu beantworten,

muss zum einen der *ontologische* Status der Prozesse, Zustände und Eigenschaften geklärt werden, auf die sich die Psychologie bezieht. Was für einen Wirklichkeitsstatus haben die von der Psychologie postulierten Entitäten? Sind es in der Tat ausschließlich bloße „Konstrukte“, deren eigentliche Realität in ihren neuronalen Korrelaten zu suchen ist? Beziehen wir uns, wenn wir von Überzeugungen oder Emotionen sprechen, also *in Wirklichkeit* lediglich auf Vorgänge in unserem Gehirn? Zum anderen muss die *epistemische* Frage geklärt werden, ob es das hier behauptete Erklärungsprimat der Biologie gegenüber der Psychologie bzw. – allgemeiner formuliert – der „niedrigeren“ Forschungsebenen gegenüber den „höheren“ Forschungsebenen gibt, wie hier von Seiten der Neurowissenschaftler suggeriert wird. Ist die Furcht vor der Verdrängung psychologischer Erklärungen von Phänomenen also berechtigt, da diese zunehmend auf solche der Biologie *reduziert* oder gar durch ebendiese *ersetzt* werden können? Wird es zukünftig tatsächlich möglich sein, auf psychologische Beschreibungen und Erklärungen zu verzichten bzw. kann uns die Biologie dieselben Fragen zukünftig viel besser, exakter, d.h. mit einer höheren wissenschaftlichen Güte, beantworten? Ein solcher Vorrang der Biologie erscheint intuitiv abwegig und die Psychologie aus bestimmten Gründen unabkömmlich zu sein. Um dies allerdings nachzuweisen, bedarf es einer systematischen Untersuchung. Welche methodischen, empirischen oder gar prinzipiellen Gründe sind es, die solche neuroreduktionistischen Vereinnahmungstendenzen als Fehlannahmen und Trugschlüsse erweisen?

2.1.3. Wiederentdeckung der Bedeutung des Leib-Seele-Problems

Was damit offenkundig geworden ist, ist, dass sich die Psychologie der lange Zeit vernachlässigten Auseinandersetzung mit Frage nach der

Beziehung von Gehirn und Geist stellen muss. Tatsächlich wird auch in den Texten der *Psychologischen Rundschau* die Bedeutung des Leib-Seele-Problems für die Psychologie wiederentdeckt und seine Beantwortung sogar zu einer wesentlichen Aufgabe der „Psychologie der Zukunft“ erhoben. So kommt Jäncke (2010) zu dem Schluss, dass eine intensive Auseinandersetzung der Psychologie mit dem Gehirn keineswegs abwegig, sondern für die Psychologie förderlich sei und insbesondere einige grundlegende Fragen geklärt werden müssten:

Wie hängen psychische Leistungen mit bestimmten Hirnaktivitätsmustern zusammen? Kann man psychische Leistungen durch Veränderung der zugrunde liegenden neurophysiologischen Aktivierungsmuster verändern? [...] Anhand der Beispiele dürfte allerdings deutlich werden, dass diese Psychologie [der Zukunft] dann sehr nahe am Leib-Seele-Problem arbeitet. Demzufolge wäre letztlich die Beantwortung des Leib-Seele-Problems eine der wesentlichen Fragen, welche diese Psychologie zu lösen hätte. (S. 195)

So muss geklärt werden, ob es sich bei psychischen versus neuronalen Vorgängen um zwei distinkte Prozesse oder doch nur einen handelt, den wir aus irgendwelchen Gründen auf verschiedene Weise erfassen. Einen besonderen Stellenwert hat des Weiteren die Frage nach der Möglichkeit *mentaler Verursachung*. Wie kann es sein und wie ist es mit unserem naturwissenschaftlichen Weltbild vereinbar, dass Gedanken oder Gefühle einen Einfluss auf den physischen Seinsmodus haben (also z.B. die Verfolgung eines bestimmten Ziels dazu führt, dass ich meinen Körper bewege), obgleich jedes physische Ereignis bereits hinreichend durch andere physische Ereignisse determiniert ist?

2.1.4. Gehirn-Geist-Differenz als Garant psychologischer Eigenständigkeit?

Es lässt sich die Tendenz ausfindig machen, der (vermeintlichen) reduktionistischen Bedrohung dadurch zu begegnen, dass man eine solche allein schon aufgrund der offenkundigen qualitativen Verschiedenheit neuronaler und psychischer Gegebenheiten ausschließt. Pauen (1999) hat dies als das *Problem der phänomenalen Differenz* bezeichnet, das von allen nicht-dualistischen Leib-Seele-Theorien gelöst werden muss. Denn wenn man aus guten Gründen – wie wir sehen werden – annimmt, dass sich Psychologie und (Neuro-)Biologie letztendlich auf ein und denselben Forschungsgegenstand beziehen, bleibt ein zentraler Einwand gegen eine solche Vorstellung, dass unklar ist, warum zwei in unserer Wahrnehmung so verschiedene Dinge wie ein Vorgang in unserem Gehirn und ein bewusstes Erlebnis (zumindest in gewisser Hinsicht) ein und dasselbe sein können. So spricht in der *Psychologischen Rundschau* wiederum stellvertretend Mausfeld (2010) eine verbreitete Überzeugung aus: „Wie genau wir das Gehirn auch anschauen – sei es durch ein Mikroskop, durch moderne bildgebende Geräte oder zukünftig vielleicht mit noch genaueren Verfahren – wir finden stets nur physikalische Objekte der üblichen Art: Neuronen und Synapsen, Neurotransmitter, Ionen, Elektronen und Protonen“ (S. 181). Das Gehirn und seine strukturellen Einheiten weisen offenbar nichts auf, was in irgendeiner Hinsicht Eigenschaften hat, wie sie für unseren geistigen Raum charakteristisch sind (wie z.B. die spezifische Qualität von Gefühlen, Schmerzen etc.). In ähnlicher Form drückt es Lockwood (1993) aus: „For if the immediate objects of

introspective awareness just are states of, or events within, the brain, seen as they are themselves, why do they *appear to be* so radically different from anything that a knowledge of the physiology of the brain would lead one to expect?” (S. 273f; Hervorh. im Orig.).

Wenn hieraus – wie bei Mausfeld – abgeleitet wird, dass die biologische Forschung zwangsläufig von etwas anderem handeln muss als die psychologische Forschung, weil sich erstere offensichtlich mit ganz anderen strukturellen Einheiten oder gar einem völlig anderen Realitätsausschnitt befasst als letztere, ist hiermit ein Phänomen angesprochen, das bereits im so genannten *Mühlengleichnis* von Gottfried Wilhelm Leibniz (1720/2002) seinen Ausdruck gefunden hat und in dieser oder ähnlicher Form immer wieder im philosophischen und wissenschaftlichen Diskurs auftaucht:

Man muß im Übrigen eingestehen, daß die Perzeption und was davon abhängt, *durch mechanische Gründe*, d. h. durch Gestalten und Bewegungen, *unerklärbar* ist. Wollte man vorgeben, daß es eine Maschine gäbe, deren Struktur Denken, Empfindungen und Perzeptionen haben läßt, könnte man diese unter Bewahrung derselben Proportionen vergrößert begreifen, so daß man in sie wie in eine Mühle hineintreten könnte. Dies gesetzt, würde man beim Besuch im Innern nur einander stoßende Teile finden, niemals aber etwas, was eine Perzeption erklärt. (S. 117; Hervorh. im Orig.)

Nichts, was sich im Inneren der von Leibniz beschriebenen Mühle oder „Seelenmaschine“ vorfinden ließe, würde demnach darauf hindeuten, *dass* diese Maschine über Bewusstsein verfügt bzw. könnte erklären, *warum* diese Maschine über Bewusstsein verfügt. Während Peter Bieri (1994) das Argument in den Gegenwartskontext überträgt und einen „Rundgang durch das Gehirn“ beschreibt, der in gleicher Weise kein Bewusstsein auffinden lässt, war es vor allem Thomas Nagel, der eine

gewisse Erweiterung oder zumindest eine Anreicherung dieses Gedankens vorgenommen hat: Nach Ansicht Nagels würde eine naturwissenschaftliche (physikalistische) Theorie des Geistes, egal wie ausgeklügelt sie formuliert oder wie weit die technische und methodische Entwicklung weiterverlaufen mag, zwangsläufig immer einen ganz entscheidenden Aspekt auslassen müssen: das *subjektive* Erleben. So wirft Nagel (1974/1997) in seinem berühmt gewordenen und ebenso betitelten Aufsatz die Frage auf: *Wie ist es, eine Fledermaus zu sein?* Ein naturwissenschaftlicher Ansatz könne zwar darauf verweisen, dass die Tiere aufgrund ihres wissenschaftlich untersuchbaren Körperbaus offenbar über eine Art Radar- oder Echolot-Ortung verfügen, mit der sie sich orientieren⁵. Da wir Menschen aber nicht über ein solches Sinnessystem verfügen, sei uns prinzipiell verschlossen, jemals zu erfahren, wie es – aus der Perspektive der Fledermaus – ist, eine Fledermaus zu sein. Auch eine noch so genaue naturwissenschaftliche Untersuchung der physischen Struktur des Echolot-Systems würde an dieser Situation nichts ändern. Daraus schlussfolgert Nagel, dass alle naturwissenschaftlichen Theorien des (menschlichen) Geistes letztendlich unzureichend sein müssen, da sie dem subjektiven Charakter unseres Erlebens nicht Rechnung tragen könnten.

Während Dinge wie ein Blitz oder eine Wolke neben ihrer subjektiven Erscheinungsweise in unserem Bewusstsein auch eine objektive Natur haben (so besteht ein Blitz aus elektrischen Entladungen oder eine Wolke aus einer Masse winziger suspendierender Teilchen), die auch andere –

⁵ Sie registrieren dabei das von Objekten in ihrer Reichweite zurückgeworfene Echo ihrer rasch ausgestoßenen und einer starken Modulation unterworfenen Hochfrequenzschreie. Mittels der so gewonnenen Informationen sind Fledermäuse dazu befähigt, Unterscheidungen von Abstand, Größe, Gestalt, Bewegung und Struktur vorzunehmen, so wie wir dies über unseren Sehsinn bewerkstelligen (vgl. Nagel, 1974/1997).

potentiell existierende – intellektuelle Wesen untersuchen könnten, sei es dagegen im Falle unserer subjektiven Erlebnisse selbst...

...schwierig zu verstehen, was mit dem *objektiven* Charakter eines Erlebnisses gemeint sein könnte – unabhängig von der besonderen Perspektive, von der aus ein Subjekt sie erfährt. Was bliebe letzten Endes von der Weise übrig, wie es ist, eine Fledermaus zu sein, wenn man die Perspektive der Fledermaus entfernte? Wenn aber Erlebnisse nicht zusätzlich zu ihrem subjektiven Charakter eine objektive Natur haben, die von vielen verschiedenen Perspektiven aus erfährt werden kann, wie kann man dann annehmen, daß ein Marsmensch, der mein Gehirn untersuchte, physikalische Prozesse beobachten könnte, die meine mentalen Prozesse wären (so wie er physikalische Prozesse beobachten könnte, die Blitze wären), nur eben von einer anderen Perspektive aus? Wie schließlich, könnte ein menschlicher Physiologe sie von einer anderen Perspektive aus beobachten? (Nagel, 1974/1997, S. 267f; Hervorh. im Orig.).

Auch wenn der Weg des wissenschaftlichen Fortschritts zumeist darin bestehen mag, in einem immer größeren Maß von der Abhängigkeit unserer eigenen Perspektive auf ein Objekt zu abstrahieren, sei dies im Falle der Erfahrung selbst nicht anwendbar. Wenn zwei Dinge, die phänomenal so verschieden sind wie Gehirn und Geist, ist es nach Ansicht Nagels (1974/1997) unverständlich, wie sie miteinander identifiziert werden könnten. „Wir mögen dann nicht einmal eine vage Idee davon haben, wie oder in was die beiden referentiellen Pfade konvergieren könnten. Ein theoretischer Rahmen mag hinzugefügt werden müssen, damit wir dies verstehen können“ (S. 270).

Es wird ein zentrales Anliegen dieser Arbeit sein, einen solchen theoretischen Rahmen aufzuzeigen, der jedoch nicht die negativen Implikationen aufweist, die mit dem heraufbeschworenen Neuroreduktionismus assoziiert werden. Hiermit verbunden ist, wie sich klar gezeigt haben sollte, die zentrale Frage, die die Psychologie für sich beantworten muss: Was ist die „Psyche“, auf die wir uns als Psychologen mit großer Selbst-

verständlichkeit beziehen, überhaupt? Was ist also unsere mentale oder geistige Sphäre, deren Erforschung wir uns als Psychologen verschrieben haben? Kernstück der (Auflösung der) Biologisierungsdiskussion muss also die Frage nach dem „Gegenstand der Psychologie“ (Allolio-Näcke, 2011) sein. Dass eine Beschäftigung mit dieser doch an sich fundamentalen Frage für eine jede Wissenschaft, nämlich die, was ihr eigentlicher Forschungsgegenstand überhaupt ist, für lange Zeit in der Psychologie obsolet war, ist ein mehr als misslicher Umstand, der sich in der aktuellen Debatte schmerzlich bemerkbar macht. Im Zuge der nun notwendigen Bestimmung des Verhältnisses von Biologie und Psychologie kann ihr nicht mehr ausgewichen werden; im Gegenteil: sie ist – wie bereits gesagt – eine der zentralen Fragen einer „Psychologie der Zukunft“ (Jäncke, 2010).

2.2. Der Mythos der verspäteten Wissenschaft und seine Folgen

Böse Zungen behaupten, die Psychologie sei eine Wissenschaft, die Fragen beantwortet, die niemand gestellt habe, da entweder die Antworten sowieso längst bekannt sind oder aber die Fragen niemanden interessieren. Man hat nicht den Eindruck, daß diese „Analyse“ so völlig falsch ist. (Dörner, 1983, S. 13)

Obgleich diese „Analyse“ die Lage sicherlich etwas zuspitzt, liegt eine Gefahr unabweislich darin, dass die Psychologie zwar immer exakter und methodisch einflussreicher werden mag, dafür aber die konzeptuelle Weiterentwicklung, die begriffliche Konsistenz, die psychologisch-gesellschaftliche Bedeutung der untersuchten Fragestellungen und das Ideal der Bereitstellung eines integrativen theoretischen Gesamtrahmens, in den sich die vielzähligen empirischen Einzelbefunde einbetten lassen,

zunehmend aus dem Blick gerät (vgl. Slife & Williams, 1997). So ist geradezu eine Explosion der methodischen Vielfalt eingetreten, wenn man alleine an die vielen bildgebenden Verfahren denkt, die nun auch flächendeckend in der Psychologie eingesetzt werden, aber auch viele neue oder wiederentdeckte statistische Verfahren (Strukturgleichungsmodelle, Pfadanalysen usw.). Die Güte und der Ausdifferenzierungsgrad des bedeutungstragenden psychologischen Wissens blieben hiervon allerdings erstaunlich unberührt. Denn wie gleich eine ganze Reihe herausragender Psychologen der Gegenwart wie Klaus Fiedler, Ulman Lindenberger, Rainer Mausfeld oder auch Wolfgang Prinz in ihrer programmatischen Schrift zur *Psychologie im 21. Jahrhundert* festgestellt haben (vgl. Fiedler et al., 2008, S. 30), waren es die Methoden selbst, die die Weiterentwicklung von Psychologie und Hirnforschung in den letzten Jahrzehnten vorangetrieben haben, während Entwicklungssprünge auf dem Feld der Theoriebildung kaum zu verzeichnen waren.

Wenn die Methoden selbst aber so sehr in den Mittelpunkt rücken, besteht auf längere Sicht eine Gefahr für den eigentlichen⁶ Erkenntnisfortschritt in der Psychologie. Unterbleibt eine Weiterentwicklung der theoretischen bzw. begrifflich-konzeptuellen Arbeit in der Psychologie, können auch die ausgeklügeltsten neuen Methoden letztlich nur mit dem *bestehenden* konzeptuellen Gerüst der Psychologie „gefüttert“ werden. Wenn dann z.B. keine Passung eines Strukturgleichungsmodells zu erzielen ist, werden in erster Linie die *statistischen* Parameter verändert, anstatt dass die verwendeten psychologischen Konstrukte selbst einer *inhaltlichen* Prüfung oder Weiterentwicklung unterzogen werden. So

⁶ Denn am Anfang der Psychologie standen bestimmte *inhaltliche* (und nicht methodische) Fragestellungen, die ein besseres Verständnis des menschlichen Erlebens und Verhaltens zum Gegenstand hatten.

kann sich auf der Basis der bestehenden (wahrscheinlich häufig noch unzulänglichen oder inkonsistenten) Konzepte zwar womöglich ein Höchstmaß an mathematischer Präzision ergeben, bezüglich des inhaltlichen Erkenntnisstandes, der einer Ausdifferenzierung, Anpassung oder auch Vereinfachung des begrifflichen Netzwerkes zu einem psychologischen Forschungsgegenstand bedarf, tritt man so dagegen auf der Stelle. Diese Unabkömmlichkeit theoretischer Arbeit konstatieren nunmehr sogar biopsychologisch ausgerichtete Forscher wie Onur Güntürkün: „Die Neurowissenschaften bieten wertvolle neue Werkzeuge zur Erforschung des Geistes. [...] Dies enthebt uns nicht der Notwendigkeit, Theorien zu bilden, um die damit gewonnenen Daten zu verstehen“ (in Fiedler et al., 2008, S. 36).

Allerdings sind die begrifflich-konzeptuelle Arbeit in der Psychologie und dabei vor allem die Schaffung eines möglichst einheitlichen theoretischen Rahmens immer noch relativ defizitär. Die Folge ist, was sich vor allem im Zuge der dargestellten Biologisierungsdiskussion abzeichnet, dass die Psychologie im Gegensatz zu den meisten anderen wissenschaftlichen Disziplinen keine nötige (ontologische) Verankerung aufweist. Ihr fehlt ein klares begrifflich-konzeptuelles Fundament, vor allem solange sie keine Entscheidung über den Status des Psychischen und die Beziehung zu seinen physischen Grundlagen trifft. So stellt Rainer Mausfeld, der sonst eher als „Verteidiger“ der Psychologie gegenüber einer Vereinnahmung der Naturwissenschaften resp. Biologie auftritt, fest: „[...] leider verfügen wir bislang in der Psychologie nur in sehr wenigen Bereichen über etwas, das auch nur annähernd an das heranreicht, was wir in anderen Naturwissenschaften als *Theorie* bezeichnen“ (in Fiedler et al., 2008, S. 36; Hervorh.: M.L.). Vor diesem Hintergrund zeigt sich, dass es ein eklatanter Missstand ist, dass eine Theoretische Psychologie, die sich dieser und einiger weiterer notwendiger Aufgaben von grundsätzlicher

Bedeutung annehmen könnte, und gleichzeitig alles, was mit dem Etikett „Philosophie“ versehen werden könnte, zumeist beargwöhnt werden. Wenn nun die Frage im Raum steht, warum der Psychologie ein begrifflich-konzeptuelles Fundament im Sinne eines einheitlichen theoretischen Rahmens und einer klaren (ontologischen) Gegenstandsbestimmung des Faches weitestgehend fehlt, ist es aufschlussreich, wenn man sich die Umstände zum Zeitpunkt ihrer Entstehung als eigenständiger empirischer Wissenschaft vor Augen führt (vgl. Lüdmann, 2014). Wie letzten Endes (fast) alle Wissenschaften ist die Psychologie der Philosophie entsprungen. Bevor sie sich sukzessive ab dem ausgehenden 19. Jahrhundert zu einer selbstständigen akademischen Disziplin entwickelt hat, war die Psychologie – genau genommen seit der Antike – eine philosophische Disziplin, wie es heute immer noch die Ethik oder Logik ist. So sprach Wilhelm Wundt (1896/1911) in seiner bekanntermaßen wirkmächtigen Formulierung des *Grundrisses der Psychologie* einen Gedanken aus, der bis heute einen kollektiven Glaubenssatz der psychologischen Zunft darstellen dürfte (vgl. hierzu z.B. Lewin, 1931; Bischof, 1981, 2008). Demnach ist die Psychologie eine „verspätete“ Wissenschaft, die sich – wie die Physik – potentiell schon seit der Zeit Galileis aus dem philosophischen Kontext hätte lösen können. Weil sie aber seit der Renaissance „unter die Herrschaft der Metaphysik“ (Wundt, 1896/1911, S. 5) geraten sei bzw. sich der Umklammerung durch die Philosophie lange nicht erwehren konnte, hätte sie das ihr eigentlich inhärente Potential einer empirischen Wissenschaft nicht entfalten können. Mit ihrer Neuerfindung als empirische Disziplin ist sie nun – wie einstmals auch Physik, Chemie oder Biologie – auf dem Weg zu einer respektablen eigenständigen Wissenschaft. Vor allem sollte mit ihrer Emanzipation von der Philosophie ein fundamentaler Wandel ihrer fachlichen Methoden erfolgen. Statt einer Ableitung von Erkenntnissen aus spekulativen philoso-

phisch-metaphysischen Prämissen sollte sich die empirische Psychologie der Methoden der Naturwissenschaften, allen voran der Physik, bedienen:

Indem die Psychologie auf ihrem gegenwärtigen Standpunkt eine ‚Wissenschaft der unmittelbaren Erfahrung‘ für den gesamten Umfang der Erfahrung sein will, kann sie nun auch eine prinzipielle Verschiedenheit der psychologischen und der naturwissenschaftlichen Methoden nicht mehr anerkennen. Sie hat daher in erster Linie *experimentelle* Methoden auszubilden gesucht, die eine ähnliche, nur dem veränderten Standpunkt⁸ Rechnung tragende exakte Analyse der psychischen Vorgänge zustande bringen sollen, wie eine solche in bezug auf die Naturerscheinungen die erklärenden Naturwissenschaften unternehmen. (ebd., S. 9f; Hervorh. im Orig.)

Und so scheint es nahe zu liegen, dass Wundt mit der von ihm angestrebten Orientierung an den Methoden der Naturwissenschaften tatsächlich beabsichtigte, die Psychologie selbst zu einer solchen zu erheben⁹.

De facto standen sich im 19. Jahrhundert in der Psychologie (als philosophischer Disziplin) vor allem zwei „metaphysische“ Lager entgegen: Vertreter einer *spiritualistisch* ausgerichteten gegenüber einer *materialistisch* ausgerichteten „Ideologie“ einer Psychologie. Während die spiritualistisch ausgerichteten Denker (in der Folge vor allem Descartes‘) von einer

⁸ Hier wird von Wundt der von Gustav Theodor Fechner (1851/1922) übernommene Gedanke der *Perspektivität* angesprochen, wie er für eine Frage nach dem Ursprung des Leib-Seele-Problems von fundamentaler Bedeutung ist und auch in dieser Arbeit ausführlich besprochen wird (vgl. Kap. 3.4.1). Vereinfacht gesagt, greifen die Naturwissenschaften und die Psychologie auf ein und denselben Forschungsgegenstand aus verschiedenen Perspektiven zu. Fechners Position wird dabei von Wundt allerdings als ein bloßes *heuristisches* Prinzip verstanden.

⁹ Obgleich Wundt im *Grundriss der Psychologie* eine Orientierung an den Naturwissenschaften propagiert, ging es ihm hierbei wohl primär um eine Distanzierung gegenüber einer spekulativen Psychologie. Wie vor allem seine späteren Schriften zeigten, glaubte er aber nicht, dass die Psychologie eine Naturwissenschaft sei, obgleich sie zu den *empirischen* Wissenschaften zählen mag (vgl. Fahrenberg, 2011, S. 29).

eigenständigen nicht-materiellen Seelen-substanz ausgingen, behaupteten die Materialisten (unter ihnen auch viele Naturwissenschaftler und Mediziner), dass es sich im Falle von psychischen Vorgängen um eine bestimmte Art von physischen Prozessen handelt. Wie leider häufig schien sich auch hier die Diskussion in philosophischen Dickichten zu verfangen, ohne die Aussicht auf die Möglichkeit einer Klärung oder zumindest Versöhnung der konträren Standpunkte. Daher hat Wundt zu diesem Zeitpunkt aus guten Gründen dafür plädiert, dass sich die neu gegründete *empirische* Psychologie von solchen metaphysischen Spekulationen lösen sollte. Entgegen einer Ableitung von Erkenntnissen aus fragwürdigen metaphysischen Vorannahmen muss eine wissenschaftlich respektable empirische Psychologie im Sinne einer Naturwissenschaft handeln und die (systematische/methodengeleitete) *Erfahrung* zur Grundlage ihrer Überlegungen und Theorien machen. So sind sich „alle diese metaphysischen Richtungen darin einig, daß sie nicht die psychologische Erfahrung aus sich selbst zu interpretieren, sondern aus irgendwelchen Voraussetzungen über hypothetische Vorgänge eines metaphysischen Substrats abzuleiten suchen. Aus der Bekämpfung dieses letzteren Verfahrens ist die empirische Psychologie hervorgegangen“ (Wundt, 1896/1911, S. 7f).

Nach Ansicht Wundts sollte sich die Psychologie im Hinblick auf das Leib-Seele-Problem mit dem Postulat eines *psychophysischen Parallelismus*¹⁰ zufrieden geben, das lediglich als ein *heuristisches* Prinzip verstanden werden dürfe, mittels dessen man sich also in keiner Form ontolo-

¹⁰ „Den Satz, daß alle diejenigen Erfahrungsinhalte, die gleichzeitig der mittelbaren, naturwissenschaftlichen und der unmittelbaren, psychologischen Betrachtungsweise angehören, zueinander in Beziehung stehen, indem innerhalb eines jenes Gebiets jedem elementaren Vorgang auf psychischer Seite ein solcher auf physischer entspricht, bezeichnet man als das *Prinzip des psychophysischen Parallelismus*“ (Wundt, 1896/1911, S. 394; Hervorh. im Orig.).

gisch festzulegen hat (vgl. Lüdmann, 2012). Denn ansonsten müsste man sich zu Gunsten von Spiritualismus oder Materialismus entscheiden, was – vermeintlich – unausweichlich ein Abdriften in die spekulative Psychologie alter Tage mit sich brächte. Ob nun von Wundt intendiert oder nicht, war die Folge dessen, dass die Frage nach dem Verhältnis von Gehirn und Geist zu einem Phänomen gemacht wurde, das eine empirische, wissenschaftlich respektable Psychologie auszuklammern hat. Es ist eine Art (zumindest implizites) Dogma der Psychologie, dass eine Beschäftigung mit ontologischen oder erkenntnistheoretischen Fragen ein Rückfall in das Reich philosophischer Spekulationen wäre, sodass sie in der Psychologie im Laufe des letzten Jahrhunderts zunehmend als obsolet angesehen wurde. Das Leib-Seele-Problem wurde *de facto* den Philosophen überlassen.

Es ist allerdings aufschlussreich, dass sich Wundt selbst zunehmend von der strengen Orientierung an der Physik und sogar von der Überzeugung, dass es einer grundsätzlichen Trennung zwischen Psychologie von Philosophie bzw. ihrer Themenfelder bedarf, distanzierte (vgl. ebd., 1913, 1921). So sagt er hier, dass es ein grundlegendes Missverständnis wäre, davon auszugehen, dass es auf geistigem Gebiet „ein System von Gesetzen des Geistes geben [müsse], die von ähnlich exakter, mathematisch formulierbarer Beschaffenheit seien, wie die allgemeinsten Naturgesetze. [...] Diesem Mißverständnis begegnet nicht selten auch die experimentelle Psychologie“ (ebd., 1913, S. 89). Stattdessen spricht er sich in seiner – der experimentellen Psychologie gegenübergestellten – *Völkerpsychologie* (vgl. ebd., 1900-1920) für eine *kulturvergleichende Psychologie* aus, weil sich die komplexen Wechselbeziehungen zwischen den psychischen Elementarfunktionen gerade nicht in der experimentellen Individualpsychologie auf eine methodisch kontrollierte Art und Weise erfassen lassen würden. Komplexere Bewusstseinsinhalte und deren Wechselbeziehungen sind

„der Sache nach dem Experiment unzugänglich“ (Wundt, 1921, S. 537), da dieses zwangsläufig die Rahmenbedingungen, die für die Entstehung eines Phänomens konstitutiv sind, vernachlässigen muss (vgl. Wundt, 1907, S. 308; Laucken, 1998, S. 86). Stattdessen müsse man die Gesetzmäßigkeiten des geistesgeschichtlichen Nacheinanders sprachlicher, mythischer und sittlicher Kulturgebilde untersuchen und könnte – auch durch den Vergleich verschiedener Kulturen – zu einem adäquaten Verständnis des individuellen Geisteslebens kommen: „So werfen z.B. die Erscheinungen der Sprache, die an sich nur als eine Schöpfung des Gesamtgeistes [einer Kultur; M.L.] zu begreifen ist, doch zugleich ein helles Licht auf die psychologischen Gesetzmäßigkeiten des individuellen Denkens“ (Wundt, 1888, S. 21).

In der akademischen Psychologie hat man von dieser Kehrtwende oder zumindest diesen Klarstellungen Wundts offenbar kaum Notiz genommen. Denn wie bereits Koch (1951) und aktuell Dörner (2008) festgestellt haben, findet sich in weiten Teilen der Psychologie eine starke Abwehr gegenüber profunder theoretischer Arbeit wieder. Nach Slife und Williams (1997) herrscht dabei die Ansicht vor, dass die grundlegenden begrifflichen und theoretischen Arbeiten bereits durch Psychologen wie z.B. Piaget, Edelman, Neisser oder Sternberg erfolgt seien und nunmehr nur noch eine durch die Empirie zu leistende Verfeinerung oder Korrektur vorgenommen werden müsse (vgl. ebd., S. 118). Nicht selten werden Empirie und Theorie häufig sogar als „Gegenspieler“ aufgefasst. Diesem Denkmuster zufolge wirken Theorien – und insbesondere solche mit einem hohen Anteil an hypothetischen Konstrukten – „verdächtig“. Sie scheinen ein Behelfsmittel zu sein, das man in Kauf nehmen muss, solange man noch nicht die „eigentlichen Fakten“ kennt, die über die Empirie zu erheben sind (vgl. ebd.). Der Zusammenhang zwischen Empirie und Theorie bleibt in vielen wissenschaftlichen Arbeiten häufig

recht lose, sodass eine Weiterentwicklung oder Integration der bestehenden theoretischen Modelle (auf einem höheren Analyse-Niveau) oder der verwendeten Konzepte nur äußerst selten stattfindet. „The discipline has moved away from grand subsuming theories in the traditional sense and moved towards models, techniques, and microtheories in the more modern sense. [...] These models are rarely expanded to full-blown theories. And yet, [...] such models rest on a host of broader theoretical assumptions that are often never recognized and almost never examined” (ebd.).

Diese Schere zwischen Empirie und Theorie hat sich seit der Zulassung kumulativer Dissertationen noch deutlich verstärkt, denn durch die bei der Publikation von Zeitschriftenartikeln vorgeschriebene, zur Kürze verpflichtende Form besteht nur wenig Platz für theoretische Überlegungen und innovative Beiträge zum wissenschaftlichen Erkenntnisfortschritt (vgl. Witte & Brandt, 2010; Witte, 2013). Es entsteht so der Eindruck einer Disziplin, der die Überzeugung Eigen zu sein scheint, dass sich *allein* durch die *Anhäufung* von empirischen Daten mit einem zunehmenden Detailliertheitsgrad (und damit immer geringerer Anbindung an übergeordnete theoretische Rahmenkonzepte) sukzessive die gesuchten Erklärungsprinzipien zu psychologisch relevanten Fragestellungen etablieren lassen. Auch der von Mausfeld (2010) kritisch beleuchtete Neuroreduktionismus verdankt sich in dieser Hinsicht vermeintlich „einfacher“ und „objektiver“ Leistungsindikatoren wie sie seit den 1990er Jahren im ökonomisierten Wissenschaftsbetrieb installiert wurden:

Die Verbreitung neuroreduktionistischer Haltung wird in dem Maße begünstigt, in dem im Steuerungssystem der Forschungsorganisation und der individuellen Leistungsbewertung lokale Untersuchungen von Bedingungs- und Effektivvariablen im Rahmen etablierter inhaltlicher und experimenteller Paradigmen stärker belohnt werden als Beiträge zur Entwicklung eines Substrats kumulativer Theoriebildung. Genau dies ist aber eine Folge der in den vergangenen Jahrzehnten erfolgten zunehmend Ökonomisierung akademischer Karrieremechanismen, die sich

Hand in Hand mit den Bestrebungen entwickelt haben, auch die Institution der Universität betriebswirtschaftlichen Organisationsprinzipien und den Konzepten eines sogenannten *New Public Management* zu unterwerfen. [...] Die durch derartige Evaluationsparameter definierten Karrierefilter fördern massiv eine Fokussierung auf eng umgrenzte Studien in wohletablierten Paradigmen und eine Bearbeitung von Fragen, welche die Gewähr bieten, in einem möglichst kurzen Zeitrahmen zu einem Wertzuwachs auf relevanten Indikatoren [vor allem Publikations- und Zitationsindizes; M.L.] zu führen. [...] In dem Maße, wie Indikatoren, die Oberflächenvariablen eines komplexen Systems darstellen, zu *Zielvariablen* werden, die bei der Karriereplanung direkt optimiert werden können, wird tendenziell ein gleichsam egalitäres Element als Filtermechanismus etabliert, durch das ein jeder mit der nötigen Hartnäckigkeit in geradezu algorithmischer Weise gute Indikatorwerte erzielen kann. (S. 188f; Hervorh. im Orig.)

Hiermit sei nun der wissenschaftssoziologische Nährboden sowohl für die Verbreitung absurder neuro-reduktionistischer Annahmen gelegt als auch für die zu beobachtende Stagnation bei der Theoriebildung innerhalb der Psychologie. Die Auswirkungen betreffen also die gesamte Disziplin der Psychologie, denn...

...gegenwärtig wird der Theoriefortschritt der Psychologie durch einen zunehmenden Verlust derjenigen Leitideale bedroht, auf denen der Erfolg der neuzeitlichen Wissenschaft beruht. Durch die Art der internen akademischen Selektionsmechanismen des Faches ist die Bearbeitung schwieriger und ernsthafter Grundlagenprobleme, wie sie das Fach für eine kumulative Theorieentwicklung zu bewältigen hat, wegen des mit ihrer Bearbeitung verbundenen hohen Aufwandes für die akademischen Karrierechancen geradezu kontraproduktiv. Die derzeitige Tendenz zur Bearbeitung von rasch in Publikationen umsetzbaren Fragestellungen, die dann zwangsläufig trotz oftmals methodisch hoher Standards inhaltlich dürftig sind, ist eine natürliche Konsequenz der Ersetzung klassischer wissenschaftlicher Leitideale durch das der *'visibility'*. (ebd., S. 189; Hervorh. im Orig.)

Insgesamt lässt sich sagen, dass die Geringschätzung theoretischer Arbeit aus einer Mischung aus wissenschaftssoziologischen Fehlsteuerungsprozessen, aber auch – was hiervon keineswegs unabhängig ist – aus einer gescheiterten „Vergangenheitsbewältigung“ entspringt. Der

Mythos der verspäteten Wissenschaft nährt die Abwehr gegenüber profunder theoretischer Arbeit, die sich nicht ins strenge Schema des (statistischen) Methodenkanons einpassen lässt, weil hiermit vermeintlich ein Rückfall in die Umklammerung durch die Philosophie verbunden sein könnte sowie die Gefahr von den „anderen“ Naturwissenschaften nicht mehr bzw. noch weniger als ihresgleichen akzeptiert zu werden.

2.3. Die Notwendigkeit einer Theoretischen Psychologie

Im Angesicht der starken Orientierung an der Physik bei der Etablierung der Psychologie als eigenständige empirische Disziplin muss es verwundern, dass man eine wesentliche strukturelle Gliederung der Physik zu übersehen haben scheint. Die Physik unterteilt sich nämlich (neben der Aufspaltung in ihre inhaltlichen Teilbereiche) grundsätzlich in Experimentalphysik und Theoretische Physik, wobei die so genannte „Königsdisziplin“ der Physik interessanterweise nicht die Experimentalphysik, sondern gerade die *Theoretische* Physik ist. In ihr finden die entscheidenden Weiterentwicklungen des Faches statt, bei denen es sich in bedeutender Hinsicht um einen *begrifflich-konzeptuellen* Fortschritt handelt. Es ist nämlich keine Weiterentwicklung, die sich – wie von selbst – aus den empirischen Erhebungen ergibt, sondern eine, bei der aufgrund der Zusammenschau der empirischen Fakten (und der zwischen ihnen ggf. vorliegenden Diskrepanzen) neue integrative Theorien und Modelle konzipiert werden, die einen deutlichen Schritt über die bestehende Datenlage hinausgehen, um idealerweise eine neue Welle empirischer Forschung anzustoßen (vgl. Landau & Lifschitz, 1989). Es werden theoretische Rahmen- oder Leitkonzeptionen entwickelt, die für die Gesamtdisziplin der Physik oder wesentliche ihrer Teile wegweisend bzw. for-

schungsleitend sind. So hat Albert Einstein als Theoretischer Physiker seine allgemeine oder spezielle Relativitätstheorie entwickelt, die als Inbegriff einer Erkenntnis der modernen physikalischen Forschung verstanden wird. Ähnliches gilt für fast alle „großen“ wissenschaftlichen Theorien und Modelle der gegenwärtigen Physik (man denke an dunkle Materie oder auch das Elementarteilchenmodell der Mikrophysik). Zentral ist, dass bei diesen Modell- und Begriffsbildungen eine Vereinheitlichung und Kohärenz der bestehenden empirischen Fakten und der mit diesen zusammenhängenden theoretischen Konstrukte geschaffen werden soll. Die bestehende Informationsmenge soll hierbei auf eine möglichst geringe Menge an Grundannahmen oder hypothetischen Konstrukten reduziert werden (wobei man sich bei der Entscheidung zwischen verschiedenen mathematisch möglichen Modellen zumeist an den Prinzipien der Einfachheit und Symmetrie/Ästhetik orientiert; vgl. Bischof, 2008). Oder anders formuliert: Mittels bestimmter neuer Begrifflichkeiten oder Modellannahmen soll die Menge an bestehenden empirischen Daten und theoretischen Konzepten möglichst vereinheitlicht und vereinfacht werden. So wird versucht, bestehende Theorien in eine Beziehung zu setzen und sie möglichst auf einen „gemeinsamen Nenner“ zurückzuführen oder „höherstufige“ Theorien bzw. ihre Konstrukte in „niederstufigen“ Theorien bzw. ihren Konstrukte „aufzulösen“ (indem die Voraussagen der höherstufigen bzw. spezielleren Theorie durch die niederstufige bzw. allgemeinere Theorie erklärt werden können). Es ist also eine Mischung aus mathematischer und begrifflicher *Denkarbeit*, die von der Theoretischen Physik bewerkstelligt wird, um die Forschung voranzutreiben und einen einheitlichen wissenschaftlichen Gesamtrahmen zu bewahren.

Der Psychologie fehlt nun ein solches Analogon zur Theoretischen Physik aus den bereits herausgearbeiteten Gründen. Provokant formuliert

könnte man sagen, dass das Denken (als genuine Methode der Erkenntnisgewinnung) den Psychologen irgendwie fremd geworden ist, solange es nicht selbst Teil einer empirischen Erhebung ist bzw. zumindest im Zusammenhang mit statistischen Verfahren stattfindet. Im Hintergrund steht hierbei einerseits die irriige Annahme, dass „Denken“ bzw. die begrifflich-konzeptuelle Arbeit, die – im Zusammenhang mit der mathematischen Kohärenz – für eine Vereinheitlichung der bestehenden hypothetischen Konstrukte und Modellannahmen der Psychologie notwendig wäre, etwas ist, was irgendwie „un-empirisch“ oder gar philosophisch sein könnte. Andererseits wird alles, was auch nur ansatzweise mit dem Etikett „Reduktion“ versehen werden könnte, von „Spezialwissenschaften“ wie der Psychologie als etwas potentiell Gefährliches angesehen. Denn ließe man ein derartiges Unterfangen zu, könnte – vermeintlich – immer auch das Tor dafür geöffnet werden, bestimmte psychologische Theorien oder Konstrukte auf solche der (Neuro-)Biologie (oder gar der Chemie und Physik) zurückzuführen. Es ist in dieser Hinsicht also die von Mausfeld (2010) beschriebene Angst vor dem so genannten *Neuroreduktionismus*, die einer notwendigen begrifflich-konzeptuellen Vereinheitlichung oder theoretischen Integration entgegensteht.

Die Folge ist eine unüberschaubare Menge an Daten und Theorien zu nahezu jedem psychologischen Themenbereich. Nach Dörner (1983) ist es die unreflektierte Übernahme bestimmter, vermeintlich die Naturwissenschaften kennzeichnender Prinzipien in die Psychologie, die...

...zu jenem „weißen Rauschen“ der „Irgendwas“-Forschung geführt [hat], welches heute immer noch beobachtbar ist. Bestimmte, vielleicht notwendige Phasen der Wissenschaftsentwicklung hat die Psychologie glatt übersprungen. [...] Getreu dem galileischen Prinzip befaßte sich die akademische Psychologie mit Kleinstdetails, kam aber nie zu den Äquivalenzen der Newtonschen Gesetze und konnte dahin nicht kommen eben wegen des Systemcharakters psychischen Geschehens [...]. Er-

gebnis: eine unübersehbare Fülle unverbundener „theorunculae“ für dieses und jenes. (S. 24f)

So existieren bekanntlich über 100 verschiedene Theorien zum Thema der Entstehung von Gewalt/ Aggression und sicherlich gut 50 verschiedene Ansätze, die erklären wollen, warum ein Mensch die Motivation zu einer Handlung aufbringt oder eben nicht. Es mag zwar sein, dass die Psychologie nicht in dem gleichen Maß wie die moderne Physik Kohärenz und Einheitlichkeit erzwingen kann, denn es ist nicht möglich und – wie der „orthodoxe“ Behaviorismus gezeigt hat – vielleicht nicht einmal wünschenswert, flächendeckend strenge psychische Gesetze zu formulieren. Dennoch liegt zweifelsfrei ein großes, bislang kaum genutztes Potential vor, eine stärkere (paradigmenübergreifende) theoretische bzw. begrifflich-konzeptuelle Integration vorzunehmen (auch in Bezug auf die Verwendung dieser Konzepte in den verschiedenen psychologischen Teildisziplinen).

Bereits Johannes Lindworsky (1926) hat in seinem Werk *Theoretische Psychologie im Umriss* darauf hingewiesen, dass es überfällig ist, dass auch die Psychologie eine solche Teildisziplin benötigt, die sich als „Seitenstück zur theoretischen Physik“ (S. 1) versteht. Die Theoretische Psychologie soll daher wie die Theoretische Physik die zahllosen empirischen Einzelbefunde in ein überschaubares Gesamtsystem einordnen, diese Befunde auf eine relativ geringe Anzahl an Grundannahmen zurückführen und aus diesen theoretischen Grundannahmen wiederum bestimmte (noch nicht getätigte) Beobachtungen ableiten (vgl. Wolfradt, 2012; Brauns, 2012). Diese können dann durch Experimente bestätigt oder widerlegt werden, womit zugleich die Korrektheit der theoretischen Grundannahmen überprüft würde (vgl. Lindworsky, 1926, S. 2f). Das Ziel

müsse hierbei sein, „zu einem geschlossenen Bilde [...] aller seelischen Tatsachen und Gesetzmäßigkeiten [zu] gelangen“ (ebd., S. 1).

In ähnlicher Weise kritisiert Sigmund Koch (1951) die Theorieabneigung in der Psychologie („disaffection from theory“, S. 296), die zugunsten einer adäquaten Theoriebildung abgelegt werden müsse. Insbesondere müsse seiner Ansicht nach eine Analyse der methodologischen oder psychologischen Fundierungsprobleme sowie konfligierender theoretischer Formulierungen erfolgen, sodass neue (einheitlichere) theoretische Konzepte etabliert werden können (vgl. Brauns, 2012, S. 15f). Trotz eines also teilweise durchaus vorhandenen Problembewusstseins scheinen die Vorbehalte und Ängste gegenüber einer Verstärkung theoretischer Arbeit in der Psychologie, die – unberechtigterweise – eher mit einem Rückschritt als einem Fortschritt in der Entwicklung der Psychologie zu einer respektablen empirischen Wissenschaft assoziiert werden, zu überwiegen, denn Lindworskys oder Kochs Visionen blieben weitestgehend nur solche. Obgleich seit vielen Jahren – wie z.B. vom Präsidenten der Deutschen Gesellschaft für Psychologie Hans-Joachim Kornadt auf dem DGPs-Kongress im Jahr 1985 – betont wird, dass die Etablierung einer Theoretischen Psychologie notwendig ist und es auch in letzter Zeit an verschiedenen Stellen – mehr oder weniger euphorisch – heißt, dass wir nunmehr in das Zeitalter oder *Jahrhundert einer Theoretischen Psychologie* eingetreten wären (vgl. z.B. van Hezewijk, 2000), muss man doch konstatieren, dass es trotz einiger Bemühungen und Ausnahmen keine fest verankerte *Theoretische Psychologie* innerhalb der akademischen Psychologie gibt.

Es ist erstaunlich, wie prophetisch gerade Wundt (1913) – ganz im Sinne Goethes „die Geister, die ich rief, werd ich nun nicht los“ – in seinen späteren Werken die Folgen einer empirizistischen („theorie-feindlichen“) Ausrichtung der von ihm begründeten Wissenschaft bzw. einer

Loslösung von grundsätzlichen, Einzeldisziplinen übergreifenden theoretischen Fragestellungen vorausgesehen hat:

Jene allgemeineren und darum für die psychologische Bildung wichtigsten Fragen hängen aber so innig mit erkenntnistheoretischen und metaphysischen Standpunkten zusammen, daß gar nicht abzusehen ist, wie sie jemals aus der Psychologie verschwinden sollten. Eben das zeigt deutlich, daß die Psychologie zu den philosophischen Disziplinen gehört, und daß sie das auch nach ihrer Umwandlung in eine sogenannte selbstständige Wissenschaft bleiben würde, die am Ende doch nur von latenten und eventuell, wenn die aus der Philosophie ausgeschiedenen Vertreter der Psychologie der gründlichen philosophischen Bildung ermangeln, von unreifen metaphysischen Anschauungen getragen sein könnte. [...] Was heute, wie man wohl sagen darf, manche Philosophen irrtümlich gegen sie einwenden, sie sei mehr eine technische als eine rein wissenschaftliche Disziplin, das könnte in erschreckendem Maße zur Wirklichkeit werden. Sollte es dazu kommen, daß sich dann noch gar infolge der fortschreitenden Arbeitsteilung der eine der Psychologen etwa mit Gedächtnisversuchen [...] und der andere mit Reaktionsversuchen und ihren individuellen Verschiedenheiten, der dritte mit der Bestimmung von Unterschiedsschwellen, ein vierter mit Denkeperimenten ausschließlich beschäftigte, dann wäre die Zeit wirklich da, wo sich die Psychologen in Handwerker, aber nicht gerade in Handwerker der nützlichsten Gattung verwandelt hätten. (S. 24)

Auch wenn Wundt die Hoffnung hatte, dass diese Entwicklung der Psychologie erspart bliebe, dürfte seine Prognose durchaus eine recht valide Beschreibung der problematischen Fragmentierung bzw. Uneinheitlichkeit der Gegenwartspsychologie abgeben. Dies ist ein fundamentaler Unterschied zur Physik. Keineswegs kann davon die Rede sein, dass die Psychologie wie die Physik in einem auch nur annäherungsweise ähnlichen Verhältnis um eine Vereinheitlichung ihrer theoretischen Modelle oder begrifflich-konzeptuellen Grundlagen bemüht ist. Wie Staats (1983), Yanchar und Slife (1997) oder auch Slife und Williams (1997) beklagen, hat sich die ohnehin seit Anbeginn an vorhandene Fragmentierung des Faches in den letzten Jahrzehnten noch weiter ver-

schärft. Denn wer sollte auch eine weitergehende Zusammenschau, Zusammenführung und Vereinheitlichung, die Formulierung theoretischer Rahmenkonzepte etc. erbringen, wenn es hierfür in der Psychologie keine institutionalisierte Anlaufstelle wie eine Theoretische Psychologie gibt?

Aufgrund des fehlenden *erkenntnistheoretischen* Wissens findet in der Psychologie nicht selten eine Konfundierung begrifflich-apriorischer und empirischer Fragestellungen und Lösungswege statt. So zeigt Brandtstädter (1982, 1984) eine Vielzahl von Fällen auf, bei denen annäherungsweise begriffliche Wahrheiten empirisch untersucht wurden und „erstaunlicherweise“ zu überragenden empirischen Ergebnissen geführt haben. Apriorische bzw. begriffliche Wahrheiten sind solche, bei denen sich die Wahrheit einer Aussage bereits aus den verwendeten Begriffen selbst ergibt. Z.B. ist die Feststellung, dass alle Junggesellen unverheiratete junge Männer sind, notwendigerweise wahr, da das Unverheiratetsein eines jungen Mannes das zentrale Definitionsmerkmal eines Junggesellen ist. Eine empirische Untersuchung ist daher vollkommen überflüssig. Von solch einer Begriffswahrheit ist man nun aber nicht weit entfernt, wenn man z.B. im Rahmen der Attributionstheorie die Aussage empirisch untersucht: Je mehr Personen in einem bestimmten Setting (untereinander sichtbar) eine Aufgabe lösen, desto eher findet eine Attribution der Probanden auf die Leichtigkeit der Aufgabe statt (vgl. Weiner, 1976, S. 83). So mag eine solche Untersuchung zu erfreulich eindeutigen empirischen Ergebnissen führen, dennoch können Zusammenhänge dieser Art bereits aufgrund der „terminologischen Zusammenhänge bei sprachlich kompetenten Versuchsteilnehmern von vornherein erwartet werden. Insofern ist die Vorgabe, es handle sich hier um eine erfolgreiche Prüfung empirischer Hypothesen, irreführend“ (Brandtstädter, 1982, S. 269). Ein weiteres Beispiel wäre die im Rahmen entwicklungspsychologischer

Forschung getroffene Annahme, dass sozial-kognitive Kompetenzen eine notwendige Voraussetzung moralischer Urteilskompetenz sind (vgl. Selman, 1971; Kohlberg, 1976). Um diesen Zusammenhang nachzuweisen, sind eine Vielzahl von Untersuchungen durchgeführt worden, die die Hypothese bestätigen konnten (vgl. Selman, 1976). Obgleich es sich hierbei durchaus um eine gehaltvolle These handelt, war ihre *empirische* Überprüfung genau genommen überflüssig, da sich bereits aufgrund einer näheren Analyse der von den Forschern verwendeten Konzepte der sozial-kognitiven Kompetenzen und der moralischen Urteilskompetenz ergibt, dass die Annahme wahr sein *muss*. „Sozial-kognitive Kompetenzen, die mit der Erschließung der Orientierungsgrundlagen eigenen oder fremden Verhaltens zu tun haben, erweisen sich [...] in der Tat als notwendige, freilich *begriffs-notwendige* Voraussetzung kompetenten moralischen Urteils. [...] Empirische Untersuchungen zu diesem Zusammenhang ergeben dann allenfalls etwa als Methodenstudien zur Überprüfung der konzeptuellen Validität der eingesetzten Testinstrumente einen Sinn“ (Brandtstädter, 1982, S. 272; Hervorh.: M.L.). Man mag zwar darauf verweisen wollen, dass viele wichtige psychologische Erkenntnisse doch gerade eine Abweichung von der allgemeinen Erwartung kennzeichnet, doch kann hiervon bei den herangezogenen (annäherungsweise) Begriffswahrheiten wohl kaum die Rede sein. Erwartungswidrige Ergebnisse (bei annäherungsweise apriorischen Gewissheiten) könnten höchstens das Vorliegen von Begriffsverwirrungen auf Seiten der Probanden aufdecken (vgl. Brandtstädter, 1987). Jedenfalls scheint es diese Nähe zu reinen Begriffswahrheiten zu sein, die als entscheidende Grundlage für die von Dörner (1983) herausgestellte Selbstevidenz (oder Trivialität) vieler empirisch-psychologischer Erkenntnisse fungiert.

Ebenso gravierend hat sich der andere von Wundt vorausgesagte Umstand bewahrheitet, dass eine von der Philosophie abgekoppelte Psycholo-

gie „von unreifen metaphysischen Anschauungen“ oder – in heutiger Sprechweise – unreflektierten ontologischen Hintergrundannahmen getragen werden könnte. Wenn Wundt selbst in seinem *Grundriss der Psychologie* die Einnahme eines *Psychophysischen Parallelismus* als heuristisches Prinzip empfiehlt, so tat er dies auf der Grundlage einer profunden Auseinandersetzung mit dem Leib-Seele-Problem und der Frage der mentalen Verursachung bzw. psychischer Kausalität. Auch wenn dies selten expliziert wird, kann man aus der weitestgehenden Ausblendung des Leib-Seele-Problems als psychologisch relevantem Thema schließen, dass sich in der Psychologie vielmehr die Überzeugung herausgebildet hat, dass man sich gegenüber solchen philosophischen (ontologischen oder erkenntnistheoretischen) Fragen grundsätzlich *neutral* verhalten könne, worunter verstanden wird, dass man sich gar nicht erst mit dem Leib-Seele-Problem oder profunden erkenntnistheoretischen Belangen beschäftigen muss. Das Problem ist, dass man sich spätestens seit der verstärkten Öffnung der Psychologie für (neuro-)biologische Daten und Modelle im Hinblick auf die Körper-Geist-Problematik gar nicht mehr neutral verhalten *kann*. Denn wenn man – wie üblich geworden – zu jeder Untersuchung eines psychologischen Konstrukts neben den durch Experiment bzw. gegenwärtig vielmehr durch Fragebögen erhobenen Daten auch bildgebende Verfahren einsetzt, kann eine Interpretation der Daten grundsätzlich nicht gänzlich ohne eine (zumindest implizit vorhandene) Idee erfolgen, in welcher Beziehung die erhobenen psychologischen und neuronalen Parameter zueinander stehen. Aufgrund der fehlenden Ausbildung in begrifflich-konzeptuellen Belangen findet man nun die Tendenz vor, dass die impliziten *alltagstheoretischen* Vorstellungen bzw. Vorurteile unbemerkt Einzug in die Theoriebildung und Interpretation von Forschungsdaten halten (vgl. Slife & Williams, 1995).

So findet man in einer Vielzahl von psychologischen Studien, Modellen oder Erklärungsansätzen bestimmte psychophysische Interaktionsannahmen, die eine gewisse alltagstheoretische Plausibilität aufweisen mögen, sich bei genauerer Betrachtung aber als ein problematischer ontologischer Dualismus erweisen, obgleich dies von den Autoren gar nicht bemerkt zu werden scheint. Ein Beispiel wäre das *Transaktionale Stressmodell* von Lazarus und Folkman (1984), das eine *Beeinflussung* physiologischer Reaktionen durch Emotionen (bzw. Kognitionen) annimmt. In ähnlicher Weise postuliert das Entstehungsmodell der Posttraumatischen Belastungsstörung (PTBS) nach Rosenzweig, Breedlove und Leiman (2001), dass durch ein Trauma eine akute Aktivierung neurochemischer Reaktionen als auch eine Einwirkung auf die Amygdala stattfinden würde (vgl. Abb. 2).

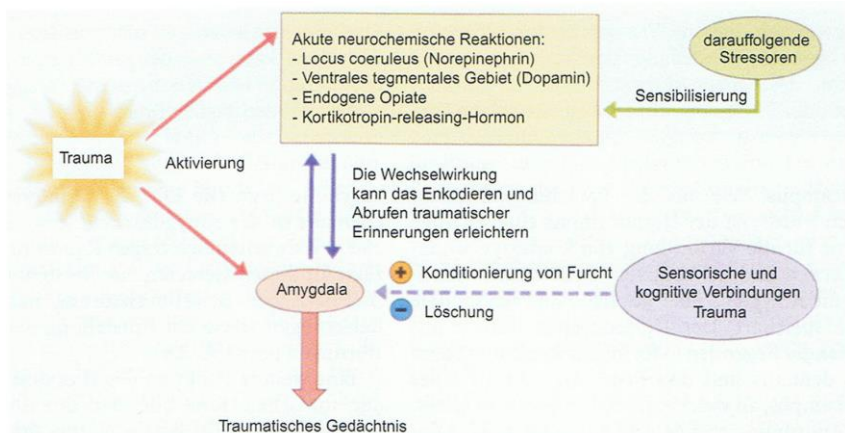


Abb. 2: Konditionierungsmodell PTBS (vgl. Heim & Meinschmidt, 2003, S. 38)

Selbst im ICD-10-F, dem weltweit gültigen Klassifikationssystem psychischer Störungen, finden sich (zumindest implizit) psychophysische Dua-

lismen. So ist bei der Diagnose „Psychologische Faktoren oder Verhaltensfaktoren bei anderenorts klassifizierten Krankheiten“ davon die Rede, dass es psychische Faktoren gäbe, die körperliche Störungen wie Asthma oder Dermatitis *bewirken* (vgl. Dilling, Mombour & Schmidt, 2013, F 54). Da gleichzeitig im ICD-10 streng zwischen psychischen Störungen und Störungen des Nervensystems unterschieden wird, bedeutet dies somit, dass eine Art körperloser Geist körperliche Veränderungen/Symptome auslöst, was ontologisch höchst problematisch ist.

Vieles scheint auf den ersten Blick harmlos zu sein. Problematisch werden solche (unreflektierten) Hintergrundannahmen oder – wie Wundt sagt – „unreifen metaphysischen Anschauungen“ aber vor allem dann, wenn die Ausprägung der hieraus getroffenen Ableitungen oder Voraussagen von eben diesen Annahmen abhängig sind, d.h., dass sich durch die Variation dieser Annahmen auch die Aussagekraft oder Voraussagen eines theoretischen Modells ändern. Dies könnte nicht nur ein Hemmnis für die Forschung darstellen, sondern auch zu gravierenden praxisbezogenen Fehleinschätzungen führen (vgl. Slife & Williams, 1995). In früheren Arbeiten habe ich z.B. nachgewiesen, dass das Bild, was man sich von der psychischen Störung Schizophrenie macht, stark in Abhängigkeit von der zugrunde gelegten psychophysischen Hintergrundtheorie variiert. So ergibt sich bei einer dualistischen „Auslegung“ der Beziehung zwischen der neuronal-biochemischen und phänomenologischen Störungsebene ein ganz anderes Verständnis des schizophrenen Syndroms, als wenn man ein identitätstheoretisches bzw. funktionalistisches Hintergrundmodell wählt (vgl. Lüdmann, 2009, 2009a, 2010). Nichtsdestotrotz sind sich die Forscher bzw. Theoretiker nur selten dieser unterschiedlichen Interpretierbarkeit ihrer Daten bewusst, was sich bereits dadurch zeigt, dass eine Besprechung der Leib-Seele-Problematik in diesem Kontext weitestgehend unterbleibt. Es konnte des Weiteren aufgezeigt wer-

den, dass es gerade die alltagstheoretisch plausiblen dualistischen Modelle sind, die nur eine äußerst geringe metatheoretische Konsistenz und Kompatibilität zu naturwissenschaftlichen Grundprinzipien aufweisen. Vielmehr scheinen bestimmte dualistische Varianten wie der Epiphänomenalismus sogar zu selbstwidersprüchlichen Ableitungen zu führen. Dennoch finden sich solche „naiv-dualistischen“ Überzeugungen nicht selten in der psychologischen Theoriebildung und damit auch in den neueren statistischen Verfahren wie z.B. Strukturgleichungsmodellen wieder. Denn wenn mit solchen Verfahren ein bestimmtes psychisches Geschehen (wie die Auslösung oder „pathologische Dynamik“ des schizophrenen Syndroms) modelliert wird, fließen die impliziten dualistischen Hintergrundannahmen wiederum unbemerkt in die postulierten Abhängigkeitsverhältnisse ein. Im Hinblick auf die Ätiologie bestimmter Symptome der Schizophrenie wird daher schnell geschlossen, dass z.B. die Volumenminderung im linkstemporalen Heschlschen Gyrus, im linken inferioren supramarginalen Gyrus und im mittleren inferioren rechten präfrontalen Gyrus (Gaser, Nenadic, Volz, Buchel & Sauer, 2004), die durch eine erhöhte Dopaminkonzentration verursachten Aktivierungsverstärkungen von Dopaminrezeptoren im Striatum der Basalganglien (Kandel, 2000) sowie die Aktivität in den primären auditorischen Arealen (Dierks et al., 1999) die *Ursachen* ebendieser schizophrenen Symptome sind. Aber was genau heißt das? Wäre es wirklich gerechtfertigt, ein Strukturgleichungsmodell mit einer Kausalbeziehung zwischen den genannten neuronal-biochemischen und mentalen Aspekten zu „füttern“, sodass z.B. die Aktivität im auditorischen Kortex (bei fehlender Reizgrundlage) die unmittelbare Ursache eines Stimmenhörens ist? Obgleich dies einleuchtend erscheinen mag, könnte es nicht auch sein, dass zwischen den beiden Störungsdimensionen ein Identitätsverhältnis besteht, sodass wir gewissermaßen *ein und dasselbe* Ereignis aus zweierlei

Perspektiven beschreiben? Wie sich im Rahmen dieser Arbeit zeigen wird, ist diese Annahme begrifflich-konzeptuell viel plausibler, metatheoretisch kohärenter als auch mit grundlegenden naturwissenschaftlichen Constraints konsistent. Damit wäre die zuvor getätigte Kausalitätsannahme aber falsch, denn sie würde nunmehr eine mysteriöse „Selbstverursachung“ postulieren. Das Perfide ist, dass sich der – in der Terminologie Wundts – statistisch-technische „Handwerker“ gerade in diesem begrifflich absurden Fall einer hohen statistischen Signifikanz seiner implizit dualistischen Modellannahmen erfreuen und diese als nunmehr *empirisch gesicherte* Fakten betrachten dürfte.

Um solche Missstände und Fehlschlüsse zu vermeiden oder zumindest zu reduzieren, besteht eine wichtige Aufgabe sowohl für die Psychologie als auch die Biologie darin, „zunächst die Grundlagen eines ernsthaften theoretischen Verstehens bereitzustellen, durch das wir dann [...] ein tieferes Verständnis dieser Beziehung [zwischen Gehirn und Geist; M.L.] erlangen können“ (Mausfeld, 2010, S. 182). Wir benötigen bei der psychologischen Forschungsarbeit ein „Theorieschema, ein konsistentes Bild von dem Hypothesengerüst, aus dem eine Theoretische Psychologie bestehen könnte“ (Dörner, 2008, S. 20). Es ist dieses Anliegen einer Theoretischen Psychologie, dem die vorliegende Arbeit gewidmet ist, obgleich es – wie aufgezeigt – noch eine ganze Reihe anderer Problemfelder gibt, denen sich Theoretische Psychologen zu stellen haben werden. Und zwar soll ein begrifflich-konzeptuelles Fundament geschaffen werden, das es ermöglicht, psychologische Konzepte auf eine Weise zu verstehen und zu bestimmen, die über die gängige Praxis einer reinen „Kreuzvalidierung“ psychologischer Konstrukte hinausgeht. Denn faktisch reicht es in der Psychologie derzeit zumeist aus, dass ein neu eingeführtes Konstrukt mit bereits bestehenden Konstrukten, die vorgeben, Ähnliches zu messen, korreliert, während es dies mit anderen bestehenden Konstruk-

ten, die vorgeben, anderes messen, nicht tut, damit man von einem *validen* psychologischen Konstrukt spricht (vgl. Cronbach & Meehl, 1955). Durch das Fehlen einer weitergehenden begrifflich-konzeptuellen Fundierung und einer Klärung der Frage nach der genauen Beziehung von psychologischen Größen zu ihren physischen bzw. neuronalen Grundlagen verbleiben sie im Allgemeinen im „luftleeren Raum“. Denn auch die bestehenden Konstrukte, die zur Rechtfertigung des neuen Konstrukts herangezogen werden, weisen keine weitergehende (ontologische) Verankerung als die genannte „Kreuzvalidierung“ (und vielleicht „praktische Bewährung“) auf¹³. Gleichmaßen soll auf der Basis dieser Grundlegung, dieses theoretischen Rahmens der Psychologie – wie bereits zuvor herausgestellt – eine grundsätzliche oder vielleicht sogar abschließende Klärung der Biologisierungs- bzw. Reduktionismusdebatte herbeigeführt werden.

Die Bedeutung zeigt sich auch z.B. im Rahmen der pädagogischen Psychologie, die sich fragt, inwieweit die Erkenntnisse der Hirnforschung für die Unterrichtsgestaltung und Wissensvermittlung von Belang sind. Hier führen die einen (z.B. Stern, Graber & Schumacher, 2005; Stern, 2005) ins Feld, dass Lernen ein aktiver individueller Konstruktionsprozess ist, in dem Wissensstrukturen verändert, erweitert, vernetzt, geordnet oder neu gebildet werden. Eine neurowissenschaftliche Forschung könne daher zur Klärung und Gestaltung (der Rahmenbedingungen) eines guten Unterrichts nichts beitragen. Die anderen halten dagegen und fragen, wer oder was denn bei diesem Konstruktionsprozess „aktiv“ ist: „das Gehirn, der lernende Schüler, beide zusammen, aber wie ge-

¹³ Verfolgt man die Reihe weiter, stößt man schnell auf bestimmte psychologische bzw. der Philosophie entlehnte Begriffe, deren Legitimation nicht mehr ist als ihre Selbstevidenz bzw. ihr alltagstheoretischer Begriffsgehalt.

nau?“ (Herrmann, 2009, S. 166). Wie hier werden häufig eine Person (bzw. ihre mentale Sphäre) und ihr Gehirn gegeneinander ausgespielt, obwohl offensichtlich ist, dass eine solche (explanatorische) Konkurrenzbeziehung unsinnig ist. Wenn sie aber offenkundig unsinnig ist, wie genau verhält es sich dann tatsächlich?

Bei dem Vorhaben einer begrifflich-konzeptuellen Fundierung der Psychologie lohnt es sich, auf den Erfahrungs- und Begriffsschatz der Philosophie und Wissenschaftstheorie zurückzugreifen. Da die Besprechung des Leib-Seele-Problems nun einmal primär hier lange Zeit verortet war, hat sich eine Vielzahl an ausdifferenzierten Begrifflichkeiten im Zusammenhang des Leib-Seele-Problems und eine präzise Analyse der Schwierigkeiten seiner Lösungsversuche etabliert, die als wichtige Leitlinien dienen können. Man denke z.B. an Emergenz, Supervenienz sowie (physische) Realisierung und ihre kritische Reflexion. Allerdings geht es hier nicht mehr darum, mittels solcher Begrifflichkeiten das (verborgene) „exklusive Wesen“ des Psychischen ergründen zu wollen (was sich immer noch viele Philosophen auf die Fahnen geschrieben haben), sondern nur gewisse Ansatzpunkte für eine praktikable, begrifflich kohärente und empirisch konsistente Ausformulierung der Gehirn-Geist-Beziehung zu haben. Auf der Basis eines tragfähigen begrifflich-konzeptuellen Fundaments der Psychologie muss eine Klärung der Beziehung zwischen der psychologischen und physischen (inkl. biologischen/neuronalen...) Ebene herbeigeführt bzw. überhaupt erst einmal bestimmt werden können, was genau diese Ebenen sind. Es muss das Problem der Pluralität der Entitäten und Eigenschaften der wissenschaftlichen Disziplinen vor dem Hintergrund der „substantiellen Einheit“ der Welt in einer Weise lösen, die sowohl natur- als auch geisteswissenschaftlich akzeptabel ist und darf dabei ihren Anspruch auf die Ergründung

oder Annäherung an eine subjektungebundene bzw. intersubjektiv zugängliche „Wahrheit“ nicht verlieren.

Eine anspruchsvolle, aber dennoch bzw. gerade deshalb aussagekräftige Zusammenstellung an eine kohärente psychologische Basistheorie¹⁸ formuliert Metzinger (1995/2005):

Um wirklich überzeugend zu sein, muß eine solche Theorie nicht nur begrifflich kohärent und empirisch plausibel sein: Wir müssen diese Theorie letztendlich auch als eine Theorie über unser *eigenes* inneres Erleben akzeptieren können. Sie muß der Subtilität und dem phänomenologischen Reichtum dieses Erlebens Rechnung tragen und die Innenperspektive des erlebenden Subjekts wirklich ernst nehmen. Vor allem muß sie uns erklären können, wie die Perspektive der ersten Person mit der Dritte-Person-Perspektive einer von außen operierenden Wissenschaft zusammenhängt. Sollte es zum Beispiel der Fall sein, daß sich die lebensweltlichen Intuitionen, die wir alle über unser eigenes Bewußtsein besitzen, und die Interpretation dieser Intuitionen durch unsere Alltagspsychologie als radikal falsch erweisen, dann müßte uns zumindest eine detaillierte Erklärung dafür geboten werden, *warum* wir alle uns so sehr über unser eigenes Bewußtsein täuschen. (S. 18; Hervorh. im Orig.)

¹⁸ Metzinger (1995/2005) hat hierbei allerdings mehr eine *Theorie des Bewusstseins* oder eine Wissenschaft bzw. einen autonomen Forschungsbereich vor Augen, die/der sich grundsätzlicher Fragen zum Thema Bewusstsein annimmt (vgl. S. 17).

3. Der psych(olog)ische Ursprung des Leib-Seele-Problems

At a certain stage, we will accept the existence of the entity or property as established – however corrigibly – and begin to use the differences observed through the diverse means of access to it as telling us still more about the object. [...] We will at the same time use these *differences* to tell about the means of access to the object (This *one* thing or property appears in these diverse ways *through these different means of access*). (William Wimsatt, 2007, S. 197; Hervorh. im Orig.)

Wie sich gezeigt hat, ist für das Vorhaben einer begrifflich-konzeptuellen Grundlegung der Psychologie eine eingehende Auseinandersetzung mit dem Leib-Seele-Problem unerlässlich. Metzinger (1995/ 2005) fordert zu Recht, dass hierbei dem Fakt unseres geistigen Erlebens Rechnung getragen und zugleich eine Erklärung dafür angeboten werden muss, wie es zu dem Auseindertreten unserer von der Psychologie erforschten Erlebnisperspektive und der „Dritte-Person-Perspektive“ der (Natur-)Wissenschaften kommt. Es ist die Frage nach dem *Ursprung des Leib-Seele-Problems*, die es zu beantworten gilt, um eine tragfähige Grundlage für die Etablierung eines allgemeinen Modells einer Architektur des Geistes, als einer praktikablen, begrifflich kohärenten und empirisch konsistenten Basistheorie der Psychologie, zu erhalten.

In diesem Kapitel soll daher der Frage nachgegangen werden, woher das Leib-Seele-Problem und die große intuitive Anziehungskraft dualistischer Körper-Geist-Vorstellungen rühren, wenn – wie eingangs aufgezeigt werden wird – ein ontologischer Gehirn-Geist-Dualismus im Sinne zweier eigenständiger Seinsbereiche oder grundlegender Substanzen vor dem Hintergrund unserer naturwissenschaftlichen Erkenntnisse und Grundsätze nahezu kategorisch ausgeschlossen werden kann. Wie Bieri (1981/1997) sagt, sind die Eindrücke, die zum intuitiven Dualismus

führen „nicht einfach aus der Luft gegriffen. Sie beruhen auf bestimmten unstrittigen Beobachtungen über mentale Phänomene, und was es zu zeigen gilt, ist, in welcher Weise der ontologische Dualismus diese Beobachtungen missversteht“ (S. 10). In ähnlicher Weise legt auch Lycan (1987/2007) dar, dass „eine adäquate Theorie des Geistes neben ihren anderen Aufgaben auch die Existenz des Leib-Seele-Problems selbst erklären [muss]. Man müsste in diesem Zusammenhang auch erklären, warum das Mentale sich so sehr vom Physischen zu unterscheiden *scheint*, dass naive Menschen zur Annahme des Cartesianismus neigen“ (S. 439; Hervorh. im Orig.). Dies zu zeigen, wird die zentrale Aufgabe dieses Kapitels sein. Mit dem Nachweis des psychischen bzw. psychologisch erklärbaren Ursprungs des Leib-Seele-Problems wird zugleich ein entscheidender Schritt in Richtung der (Auf-)Lösung dieses Problems gemacht. Denn wenn man weiß, wie ein Problem überhaupt erst in die Welt kommt, weiß man häufig auch, wie es aus der Welt zu schaffen ist oder warum es faktisch überhaupt nie in dieser war.

Als Ausgangspunkt oder auch Rahmen der Besprechung wird hierbei ein „alter“ Philosoph herangezogen, dessen Bestimmungen und Ringen um begriffliche Klarheit – sieht man von der dem damaligen Zeitgeist geschuldeten etwas sperrigen Terminologie ab – bei der Formulierung einer nicht-dualistischen Körper-Geist-Theorie ungeahnt höchste Aktualität haben: Baruch de Spinoza. Dabei steht Spinozas Lehre in einem äußerst aufschlussreichen Zwiespalt zwischen einer „subjektivistischen“ (beobachterabhängigen) und „objektivistischen“ (beobachterunabhängigen) Auflösung des Problems. Irgendwie schien ihm die Diskrepanz zwischen Gehirn und Geist etwas mit unserer (begrifflichen oder sprachlichen) Weltauffassung oder Wahrnehmung zu tun zu haben, gleichzeitig schien sie aber auch mehr zu sein als nur ein bloßes Artefakt unserer Subjektiv-

vität bzw. unserer arteigenen Vergegenständlichung einer „an sich selbst“ undifferenzierten Welt.

In diesem Kapitel wird – nach einer Skizzierung des Leib-Seele-Problems und seiner Fallgruben – der „subjektivistische“ Strang Spinozas, d.h. die These weiterverfolgt, dass es sich bei der Verschiedenheit von Gehirn und Geist um eine Differenz handelt, die ausschließlich auf unseren Zugriff auf die „Gehirn-Geist-Einheit“ zurückzuführen ist. Es wird zunächst erörtert, inwiefern es sich hierbei um eine *sprachlich-deskriptive* Differenz handeln kann, wie sie im Hinblick auf bestimmte Phänomene von der Entwicklungspsychologie beschrieben wird. Trotz einiger wichtiger Einsichten, die sich hierdurch gewinnen lassen, spricht viel dafür, dass die Tatsache, dass wir Gehirn und Geist unterschiedlich sprachlich beschreiben, nicht die eigentliche Ursache des Leib-Seele-Problems ausmacht, sondern dass dies nur ein Ausdruck einer tiefer liegenden psychischen Dynamik ist. Um diese freizulegen, ist ein Ansatz von zentraler Bedeutung, den Gustav Theodor Fechner (in Anlehnung an Spinoza) maßgeblich propagiert hat. Demnach geht die Verschiedenheit von Gehirn und Geist auf unterschiedliche Wahrnehmungsperspektiven (Eigen- vs. Beobachterperspektive) sowie den Fakt zurück, dass die neuronalen Grundlagen psychischer Prozesse für gewöhnlich erlebnismäßig nicht einsichtig sind (Metzinger, 2006, spricht hier von *phänomenaler Transparenz*). Beide Aspekte oder Strukturmerkmale unserer psychischen Verfasstheit, Perspektivität und phänomenale Transparenz, werden deshalb in begrifflicher, wahrnehmungs- und entwicklungspsychologischer Hinsicht genauer untersucht, erweitert und präzisiert, sodass am Ende des Kapitels eine begrifflich exakte und empirisch begründete Ausformulierung der These erfolgen kann, dass es sich beim Leib-Seele-Problem um ein Problem handelt, das durch unsere Wahrnehmung und Interpretation der Welt überhaupt erst in ebendiese getragen wird. Zugleich wird

allerdings problematisiert werden müssen, dass hiermit noch nicht aufgezeigt worden ist, dass es sich beim Leib-Seele-Problem um ein Phänomen handelt, das *ausschließlich* auf eine psychische Dynamik zurückgeht.

3.1. Das Leib-Seele-Problem und seine Fallgruben

Im Folgenden soll verdeutlicht werden, was genau das Leib-Seele-Problem überhaupt zu einem Problem macht, das Gelehrte seit den Ursprungstagen der Philosophie beschäftigt hat, ohne dass es möglich war, eine wirklich befriedigende Lösung zu formulieren. Dies zeigt sich vor allem bei der Betrachtung der verschiedenen „Fallgruben“, in die man bei der Suche nach einer Antwort auf die Frage nach der Beziehung zwischen Gehirn und Geist tappen kann. Es sind die Positionen des Interaktionismus, Epiphänomenalismus oder Eliminativen Materialismus, die jeweils eine bestimmte Facette des Problems in der Tat recht gut behandeln können, insgesamt aber als Lösungsansatz aufgrund fundamentaler Probleme scheitern und daher auch bei dem hier verfolgten Ziel der begrifflich-konzeptuellen Grundlegung der Psychologie ausscheiden.

Auch wenn bereits bei Platon (2004) dualistisches Gedankengut vorzufinden ist (z.B. im Dialog *Phaidon*) und die christlich geprägte abendländische Kultur für solche Vorstellungen sehr empfänglich war, so ist die primäre Grundlage dieser Problemstellung als eines Problems der philosophischen und wissenschaftlichen Reflexion primär die von Descartes (1641/1986) forcierte Weltsicht, die eine strikte Trennung zwischen der materiellen Welt der Dinge bzw. Körper und der immateriellen Welt des Geistes bzw. der Psyche vornimmt. Diese dualistische Überzeugung lieferte die philosophische Rahmenbedingung für die noch heute sehr verbreitete Differenzierung zwischen Geistes- und Naturwissenschaften

und steckt – obgleich die Trennlinien hier aufweichen – in vielen wissenschaftlichen Disziplinen die überhaupt infrage kommenden Arbeitshypothesen und somit auch die Grenzen ihrer Forschungsgebiete ab. Die Folge dieser Setzung ist aber auch, dass es sich bei Gehirn und Geist – vermeintlich – um zwei vollkommen in sich abgeschlossene Systeme handelt. Der materielle Kosmos der Naturwissenschaften ist seit der Entwicklung der modernen Physik im 17. Jahrhundert so konzipiert, dass nur Physisches auf Physisches einwirken kann, wie es das Energieerhaltungsgesetz impliziert. Dagegen wird von vielen Geisteswissenschaftlern reklamiert, dass sich ein geistiges (bzw. geistig verursachtes) Geschehen immer nur als Resultat wiederum *geistiger Gründe* adäquat verstehen lässt. Es scheint daher offenbar keinen Weg zu geben, der von dem einen „Weltentwurf“ in den anderen führt. Dies illustriert Metzinger (1995/2005) an folgendem Beispiel: Er fragt sich, wie das phänomenale Farberlebnis der „Bläue“ (zum Beispiel eines Buchumschlages) mit den physikalisch messbaren Daten auf seiner materiellen Seite zusammenhängt und ob es überhaupt „einen Berührungspunkt zwischen der Innenwelt des Bewusstseins und der Außenwelt der Physik“ (S. 29) geben kann. „In der physikalischen Außenwelt gibt es [...] nur elektromagnetische Schwingungen bestimmter Wellenlängen, aber nirgendwo die Eigenschaft ‚Blau‘“ (ebd.). Dies gelte ebenso für die Netzhaut, auf der nichts dem Erlebenseindruck Entsprechendes zu finden ist, sondern nur eine „Grauwertverteilung“ (ebd.). Des Weiteren ist „[d]as Feuern des optischen Nervs [...] genauso wenig blau, wie das Feuern der Neuronen in den für das Schmerzempfinden verantwortlichen Regionen unseres Gehirns schmerzhaft ist“ (ebd.). Es ist also theoretisch möglich, die Reaktion auf einen bestimmten Stimulus vom Innervieren des für ihn adäquaten Rezeptors bis zur Verarbeitung der mit ihm zusammenhängenden Information in einer bestimmten Hirnregion zu verfolgen, aber alles, was

sich dabei finden lässt, sind eben nur Verschaltungen bestimmter Zellen mit wiederum anderen Zellen, die mit wiederum anderen Zellen verschaltet sind usw.; es lässt sich aber an keinem Ende eines solchen Pfades plötzlich das phänomenale Erlebnis entdecken, welches der Betroffene just in diesem Moment hat. Andersherum betrachtet scheint es noch aussichtsloser, sich bei auch noch so sensibler Reflektion des eigenen Bewusstseinslebens einen Weg in das „Reich des Physischen“ bahnen zu wollen. Sicherlich gibt es Körperempfindungen, die das Erleben *beeinflussen*, aber es bleiben eben gerade nur Empfindungen in unserem Bewusstsein mit zwar womöglich lokalisierbaren physischen Infiltrationen; nie aber scheint es einen unmittelbaren, phänomalen Zugriff auf die gerade an einer unserer Synapsen zwischen zwei Neuronen ablaufenden Vorgänge geben zu können. Es ist diese fundamentale (phänomenale) Differenz zwischen mentalen und neuronalen Vorgängen bzw. die hiermit zusammenhängende grundverschiedene Konzeptualisierung von Gehirn und Geist, die die Gräben zwischen diesen (hiermit geschaffenen) „Seinsphären“ so unüberbrückbar erscheinen lassen.

3.1.1. Zur Intuitivität und Abwegigkeit des Interaktionismus

Intuitiv erscheint es also evident, dass Körper und Geist zu zwei unterschiedlichen, potentiell unabhängig voneinander existierenden Wirklichkeitsbereichen gehören. Schließlich *erleben* wir uns als geistige und damit zugleich in dieser Hinsicht auch als immaterielle Wesen. Unsere mentalen Zustände weisen keinerlei Eigenschaften auf, die wir für gewöhnlich Objekten der physikalischen Welt zuschreiben würden (Ort, Ausdehnung, Masse, Geschwindigkeit...). Gleichzeitig *haben* wir einen Körper, dem gerade diese Attribute zukommen, der aber für sich genommen keinerlei psychische Qualitäten (Empfindungen, Gedanken...) aufzuweisen

scheint, wie sie für unsere mentalen Zustände charakteristisch sind. Insofern scheint es plausibel, wie Descartes (1641/1986) davon auszugehen, dass unsere Welt aus zwei grundlegend verschiedenen *Substanzen* aufgebaut ist, aus einer geistigen oder psychischen Substanz und einer materiellen oder physischen Substanz. Man spricht daher auch von einem *Substanzdualismus*. In ihm wird der Intuition Ausdruck verliehen, dass es eine strikte ontologische Trennung zwischen einer Welt des Geistes, der denkenden Dinge (*res cogitans*), und einer Welt der Materie, der ausgedehnten Körper (*res extensa*), gibt. Da es für Descartes widerspruchsfrei *denkbar* ist, ein „reiner Geist“, also ein denkendes Etwas ohne Körper, zu sein, scheint ihm damit nicht nur eine metaphysische Möglichkeit, sondern sogar eine metaphysische *Notwendigkeit* einer körperlosen Seele als gegeben, die als etwas Nicht-Ausgedehntes dem Reich der ausgedehnten Körper gegenüber steht. Die Folge ist eine Art Realitätsverdopplung. Neben der geistigen Wirklichkeit, die uns durch unsere individuelle Erlebnisperspektive vertraut ist, gibt es immer auch eine von ihr zu trennende materielle Wirklichkeit. Wenn dem so ist, dann stellt sich natürlich unweigerlich die Frage, in was für einem Verhältnis diese beiden Wirklichkeitsbereiche zueinander stehen. Können sie sich gegenseitig beeinflussen? Ist es insbesondere möglich, dass psychische Aspekte, wie Gedanken oder Gefühle, auf physische Aspekte, wie Vorgänge in meinem Gehirn oder – weiter gefasst – auf Bewegungen meines Körpers – einwirken können? Wie ist also eine so genannte *mentale Verursachung* physischer Veränderungen möglich?

Der unseren Intuitionen Rechnung tragende und von Descartes selbst vertretene *interaktionistische Dualismus* nimmt an, dass sich Körper und Geist gegenseitig beeinflussen können. Mentale Phänomene wie Gedanken und Empfindungen können Nervenimpulse auslösen und weiterleiten, die schließlich in einer körperlichen Reaktion münden und umge-

kehrt ist es auch möglich, dass bestimmte neuronale Prozesse, zum Beispiel über die Wahrnehmung, mentale Vorgänge beeinflussen und auch bestimmen können. Diese dualistische Denkweise erscheint sehr vertraut, weil sie offensichtlich doch das widerspiegelt, was man im täglichen Leben erfährt. Nämlich, dass man mit seinem nichtstofflichen Willen eine Handlung initiieren kann, seinen stofflichen Körper also in Bewegung versetzt und dass auf der anderen Seite, wenn man sich mit einer scharfen Klinge in den Finger schneidet, ein Erleben von Schmerz einsetzt. Solche Wechselwirkungen lassen sich gut an einem Beispiel illustrieren (vgl. Abb. 3).

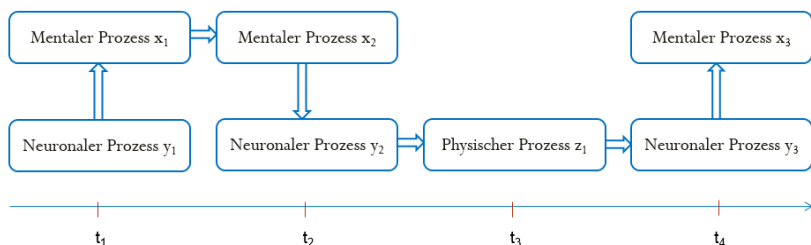


Abb. 3: Psychophysische Kausalität im Sinne des Interaktionismus

Nehmen wir an, dass ich zu einem bestimmten Zeitpunkt (t_1) Durst verspüre (mentaler Zustand x_1). Aus der Neurophysiologie ist bekannt, dass eine solche Durstempfindung durch osmosensible Neurone im vorderen Hypothalamus ausgelöst wird (neuronaler Prozess y_1), die hierbei auf eine Erhöhung der Salzkonzentration im Blut reagieren (vgl. Jänig, 2000, S. 200; Birbaumer & Schmidt, 2000, S. 456ff). Ich beschliesse daher, mir eine Cola aus dem Kühlschrank zu holen (mentaler Zustand x_2 zu t_2 : Bewegungsinitiation), wodurch ein neuronaler Gesamtzustand y_2 hervorgerufen wird, der u.a. die hierfür charakteristischen Aktivierungen im motorischen Kortex beinhalten wird. Zum Zeitpunkt t_3 bewege ich mich

deshalb zum Kühlschrank (physischer Prozess z_1), was von den hierfür relevanten neuronalen Zentren für Motorik, Gleichgewicht etc. kontrolliert wird²³. Dann trinke ich die Cola, was im weiteren Verlauf (t_3 zu t_4) eine Veränderung des neuronalen Gesamtzustandes (y_3) bewirkt. So kommt es durch die Flüssigkeitszufuhr zu einer Absenkung der Salzkonzentration im Blut, wodurch sich die Aktivität der osmosensiblen Zellen im Hypothalamus vermindert (evtl. kommt es zusätzlich zu einer Aktivierungsänderung/Dopaminausschüttung im limbischen System). Zusammengefasst bewirken diese neuronalen und physiologischen Aktivitäten nun wiederum den mentalen Zustand x_3 (Nachlassen des Durstes²⁴).

Was sich klar abgezeichnet haben sollte, ist, dass dieses Interpretationsmuster suggeriert, dass es sich in dem Beispiel um zwei distinkte Wirklichkeitsbereiche oder Substanzen handelt: einen mentalen bzw. psychischen Seinsmodus und einen physischen bzw. neuronalen Seinsmodus, zwischen denen kausale Einwirkungen in beiderlei Richtungen bestehen. Auch wenn dies unseren Intuitionen über die Beziehung zwischen Körper und Geist bzw. Gehirn und Bewusstsein sehr entgegen kommen dürfte und auch in vielen psychologischen Modellen mehr oder weniger explizit Verwendung findet, sind hiermit gravierende Probleme verbun-

²³ Ich verzichte hier und im Folgenden, wenn von Beispielen für neuronale Zustände die Rede ist, der Einfachheit halber auf die umständliche, aber empirisch natürlich angemessenere Formulierung, dass es nämlich ein neuronaler *Gesamtzustand* ist, der unter anderem die beschriebenen charakteristischen Merkmale beinhaltet. Dass zum Zeitpunkt t_3 kein mentaler Zustand eingezeichnet ist, soll nicht heißen, dass die Person gerade kein Bewusstsein aufweist. Natürlich findet auch hier eine Bewegungssteuerung usw. statt. Z_1 repräsentiert also einen Ausschnitt des neuronalen Gesamtzustands, der nicht an der Erzeugung von Bewusstsein beteiligt oder durch dieses beeinflusst ist.

²⁴ Es sei darauf hingewiesen, dass die empirische Korrektheit des hier verwendeten, sicherlich sehr vereinfachten Beispiels zu vernachlässigen ist. Es geht an dieser Stelle lediglich um die Illustration der prinzipiellen Möglichkeiten psychophysischer Kausalität.

den, die diesen Ansatz insgesamt diskreditieren. Das Hauptproblem liegt darin, dass die unterstellten wechselseitigen Verursachungsmöglichkeiten eine Verletzung physikalischer Grundprinzipien darstellen.

Wenn nämlich das Einwirken einer nicht-physikalischen Größe auf ein physikalisches Geschehen als prinzipiell möglich erachtet wird, liegt eine Verletzung der Geschlossenheit der physischen Welt bzw. des Energieerhaltungssatzes vor. Beide Prinzipien laufen auf die Annahme hinaus, dass wenn „ein physisches Ereignis zu einem Zeitpunkt t eine Ursache hat, dann hat es zu t eine *physikalische* Ursache“ (Metzinger, 2007, S. 14; Hervorh. im Orig.). Da dieses Grundpostulat dem gesamten naturwissenschaftlichen Gegenstandsentswurf zugrunde liegt, hätte eine geduldeten Verletzung schwerwiegende Konsequenzen. Denn wenn es für einen immateriellen Geist möglich ist, materielle Gegebenheiten (wie einen menschlichen Körper oder auch nur einzelne seiner Atome) in Bewegung zu setzen, dann ist fraglich, woher die Energie kommen soll, mit der dies vollbracht wird. Der immaterielle Geist ist per Definition des Substanzdualismus kein Teil der physischen Welt und kann daher auch keine physikalische Kraft besitzen (also eine solche, wie sie von der Physik oder den Naturwissenschaften insgesamt beschrieben wird).

Wenn man deshalb der geistigen Substanz eine besondere – vielleicht noch unentdeckte – nicht-physikalische Form von Energie zusprechen möchte, die eine solche Verursachung physischer Ereignisse vollbringen kann, dann taucht ein neues Problem auf: Da sich gemäß dem Prinzip der Geschlossenheit der physischen Welt neben der postulierten mentalen Verursachung immer auch ein ursächliches physisches Ereignis ausfindig machen lassen muss, würde sich eine zweifelhafte *Überdetermination* des neuronalen Ereignisses y_2 ergeben. Denn da für jedes physische Ereignis gilt, dass es bereits hinreichend durch vorhergehende physische Ereignisse kausal determiniert wird (kein physisches Ereignis ohne

physische Ursache), muss angenommen werden, dass y_2 durch den zu t_1 vorausgehenden physischen bzw. neuronalen Prozess (y_1) verursacht wurde. Es ist also die Reaktion der osmosensiblen Zellen selbst, die den Anstoß zu den nachfolgenden neuronalen und physischen Prozessen gibt und letztendlich zur Aufnahme von Flüssigkeit führt (vgl. Abb. 4).

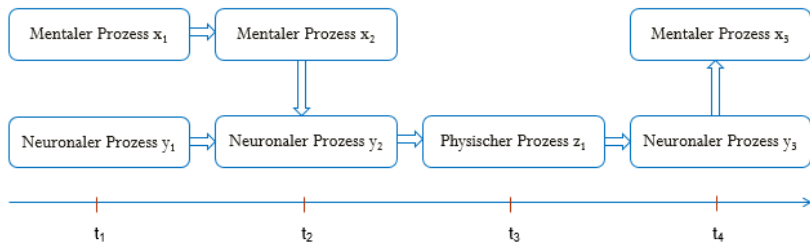


Abb. 4: Kausale Überdetermination im Interaktionismus

Welchen kausalen Beitrag kann dann aber der mentale Prozess x_2 geleistet haben? Muss man sich bezüglich der Verursachung von y_2 zwischen x_2 und y_1 entscheiden, spricht wohl kaum etwas dafür, eine ominöse mentale, nicht-physikalische Kraft hierbei zu bevorzugen, wenn gleichzeitig der Preis hierfür ist, die kausale Geschlossenheitsannahme der physischen Welt und den Energieerhaltungssatz aufgeben zu müssen.

Um dieses Dilemma zu beheben, gibt es natürlich die Möglichkeit, die Gültigkeit der physikalischen Grundprinzipien infrage zu stellen. So weist z.B. Popper (Popper & Eccles, 1982) darauf hin, dass in den Augen vieler prominenter Physiker der Energieerhaltungssatz und auch die kausale Geschlossenheitsannahme der physischen Welt (zumindest im Sinne eines durchgehenden Determinismus) im Angesicht *quantentheoretischer* Implikationen keine apriorische Notwendigkeit mehr besitzen. In diesem Sinne könnte unsere geistige Substanz auf „quantenphysika-

lische Art“ die Wahrscheinlichkeit des Auftretens physischer (neuronaler) Ereignisse beeinflussen. Der zentrale Gedanke scheint hierbei zu sein, dass es unter der Annahme einer indeterminierten oder zumindest nicht vollständig determinierten Welt keine strenge Abfolge kausaler Verkettungen mehr gibt, wie sie noch von der klassischen Physik postuliert wurde.

Für jedes Ereignis im Sinne der klassischen (Newtonschen) Physik gilt, dass es sich restlos aus seinen Vorbedingungen ergibt. Ein Systemzustand oder -prozess (z.B. ein Feuern von Neuronen im motorischen Kortex) ist eine Folge vorheriger physikalischer Ereignisse oder Systemzustände, auf die er sich restlos zurückführen lässt; er wird also durch diese Vorbedingungen vollständig festgelegt. Da dem Argument zufolge die klassische Physik keine Gültigkeit mehr besitzt, muss nun von diesem Systemzustand gesagt werden, dass er sich nicht mehr restlos aus den auf ihn wirkenden Vorbedingungen ergibt, sondern es könne nur noch gesagt werden, dass dieses Ereignis – je nach gegebener Vorbedingungen – mit einer bestimmten *Wahrscheinlichkeit* auftritt. Und diese unvollständige Determination eines Systemzustandes ist es, was Raum für das Einwirken eines immateriellen Geistes schaffen soll. Der Geist legt sich gewissermaßen „um“ oder „über“ das Gehirn und erhöht durch die in ihm ablaufenden Willensbildungen die Auftretenswahrscheinlichkeit eines bestimmten physikalischen Zustandes. In diesem Sinne werden von dualistisch ausgerichteten Psychologen und Neurowissenschaftlern wie Eccles (Popper & Eccles, 1982) und Libet (2007) Theorien vertreten, in denen geistige Prozesse als „bewusste mentale Felder“ aufgefasst werden. Ähnlich wie Magnetfelder würden sie so Einfluss auf physische Gegebenheiten nehmen können. Durch meinen Wunsch, mich in Bewegung zu setzen, wird so die Wahrscheinlichkeit erhöht, dass es zum Feuern der Neuronen im motorischen Kortex kommt und ich mich

tatsächlich aus dem Sessel erhebe. In diesem Sinne wird der Versuch unternommen, geistige Prozesse in unser physikalisches Weltbild zu integrieren. Sie wären eine neue, besondere Art von (physikalischen) Eigenschaften, um die unser jetziges naturwissenschaftliches Verständnis der Welt erweitert werden müsste.

Auch wenn dies in der Tat interessante Vorschläge sind, die zumindest eine prinzipielle Möglichkeit aufzeigen, wie ein interaktionistischer Dualismus mit unserem wissenschaftlichen Weltbild in Einklang gebracht werden könnte, muss doch konstatiert werden, dass hierfür in der Empirie kaum oder besser gesagt: *keinerlei* Evidenz besteht. Es ist nicht einmal aufgezeigt worden, dass es überhaupt möglich ist, von der evtl. probabilistisch verfassten (mikrophysikalischen) Quantenphysik auf Ereignisse in unserer makroskopischen Welt zu schließen²⁵. Dass unsere gesamte (makroskopische) Welt nach Gesetzmäßigkeiten funktioniert, wie sie für den Bereich der Quantenphysik Gültigkeit haben, ist also noch gar nicht ausgemacht. Ein Interaktionismus kann aufgrund dieser fundamentalen Schwierigkeiten nicht als die gesuchte psychologische bzw. psychophysische Basistheorie fungieren.

²⁵ Und selbst in der Physik scheint die Gruppe derjenigen, die die Kopenhagener Deutung der Quantenphysik und ihr Postulat einer *indeterministischen Welt* (zumindest im Bereich quantenphysikalischer Vorgänge) verfechten, kleiner zu werden. Vor allem der *Viele-Welten-Ansatz* (oder auch *Multiversumstheorie*) gewinnt als konkurrierende Interpretation der Quantenphänomene zunehmende Bedeutung. Dieser ermöglicht die Aufrechterhaltung einer deterministischen Weltsicht, allerdings auf der Basis der Annahme einer unendlichen Anzahl von verschiedenen „Welten“, in denen die unterschiedlichen Zustände eines Quantensystems nach einer Messung auftreten. Die vielen „Welten“ sind dabei allerdings keine räumlich getrennten Paralleluniversen, wie es häufig heißt, sondern lediglich getrennte Zustände im jeweiligen Zustandsraum (Everett, 1957; DeWitt, 1970; v. Weizsäcker, 1985).

3.1.2. Epiphänomenalismus: Der Geist als bloße Begleiterscheinung

Möchte man an seinen dualistischen Intuitionen und an den genannten naturwissenschaftlichen Grundannahmen festhalten, ohne in die Sackgasse einer kausalen Überdetermination zu geraten, führt der Weg nahezu zwangsläufig in die Position des so genannten *Epiphänomenalismus*, der die Möglichkeit eines mental verursachten Eingreifens in den physischen Seinsmodus einfach abstreitet. Dieser Auffassung nach begleiten die geistigen Vorgänge die physikalischen Aktivitäten lediglich, ohne sie beeinflussen zu können. Bei den resultierenden mentalen Epiphänomenen handelt es sich daher „um die absoluten Endpunkte kausaler Ketten“ (Kim, 1996/1998, S. 143).

Wenn wir beim herangezogenen Beispiel bleiben und in der Abbildung 5 diesmal beim Zeitpunkt t_2 ansetzen, finden wir hier weiterhin die Bewegungsinitiation (mentaler Zustand x_2) vor, mittels derer ich meinen Gang zum Kühlschrank zu starten glaubte.

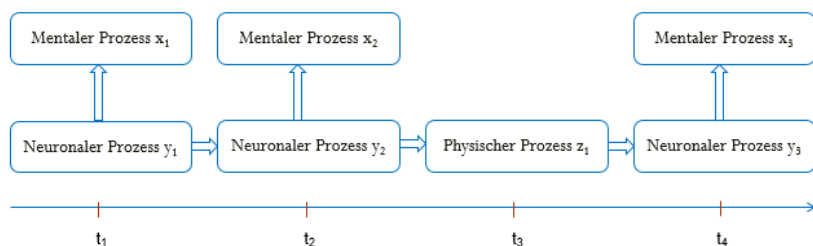


Abb. 5: Epiphänomenalistischer Dualismus

Im Gegensatz zum vorherigen Modell ist dieser mentale Zustand *de facto* aber nicht die Ursache der Aktivierung im motorischen Kortex und damit der körperlichen Bewegung, sondern ganz im Gegenteil: x_2 wird durch die neuronale Aktivierung y_2 hervorgebracht; das Erleben einer Bewegungsinitiation ist eine *Folge* und nicht die Ursache bestimmter Gehirn-

aktivierungen (im motorischen Kortex). Wenn man dies zugesteht, aber darauf insistiert, dass es doch mein Durst bzw. Wunsch war, mir eine Cola aus dem Kühlschrank zu holen (also der zeitlich vorgeordnete mentale Zustand x_1), der mich dazu veranlasst hat, den Beschluss zu fassen, mich aus dem Sessel zu erheben (x_2), dann ist dies gemäß den Prämissen des Epiphänomenalismus nur insofern richtig, als dass es der neuronale Zustand y_1 zum Zeitpunkt t_1 war, der den Durst und den Wunsch nach Flüssigkeitszufuhr bewirkt hat. Diese Aktivität y_1 ist es also, die sowohl die Aktivierung der motorischen Areale (y_2) hervorruft als auch – als reine Begleiterscheinung – meinen subjektiv erlebten Durst und Wunsch, eine Cola zu trinken (x_1). Mein Wunsch selbst ist dabei wirkungslos. Alles hätte sich so ohne ihn genauso zugetragen. Wichtig ist hier, dass das Mentale zwar weiterhin immaterieller Natur sein soll, dass es aber im Gegensatz zum Interaktionismus immer einer physischen Verursachung mentaler Zustände bedarf. Ohne physische Grundlage kann es kein Bewusstsein geben (im Interaktionismus dagegen schon).

Es ist also zwar so, dass ich den Eindruck haben mag, dass es mein Wunsch nach einer Cola war, der mich dazu veranlasst hatte, aufzustehen und zum Gefrierschrank zu gehen. In Wirklichkeit waren es aber physische resp. neuronale Aspekte, die die eigentlichen kausalen Mechanismen beinhalten. Bestimmte Bewusstseinsinhalte wie Gedanken und Empfindungen können nicht selbst handlungswirksam werden. Man könnte auch sagen, dass wir in der Abbildung 5 den zeitlichen Verlauf eines *neuronalen* Geschehens sehen, bei dem aus irgendwelchen Gründen zu den Zeitpunkten t_1 , t_2 und t_4 zugleich mentale Ereignisse als eine Art „Nebenprodukt“ dieses neuronalen Prozesses aufflackern.

Diese Sichtweise, die im 19. Jahrhundert mit dem Biologen Thomas Henry Huxley (1866) ihren prominentesten Vertreter fand und heute von Philosophen wie Frank Jackson (1982) vertreten wird, lässt sich nicht

selten – wenn auch zumeist nur in impliziter Form – in aktuellen Spekulationen einiger Neurowissenschaftler ausfindig machen. Zwar stellen sich die Forscher selbst oft als eher physikalistisch orientiert vor oder verweigern sich systematisch einer expliziten Stellungnahme zur psychophysischen Problematik; bei genauerer Analyse ihrer Erklärungsmuster und Beispiele sind trotzdem mitunter eindeutig dualistisch-epiphänomenalistische Elemente zu erkennen. So übt z.B. nach Roth (2002) „das bewusste Ich [...] keinen entscheidenden, sondern nur einen beratenden Einfluss auf diejenigen Handlungen [aus], die es als selbstveranlasst empfindet“ (S. 42).

Auch wenn wir also mit der Einnahme der epiphänomenalistischen Position das Problem der mentalen Verursachungsfähigkeit *aufgelöst* haben (da es diese *per definitionem* nicht gibt), tun sich nun Probleme ganz andere Art auf: Die Frage, die sich natürlich unweigerlich aufdrängt, ist, warum die Natur so ein komplexes und Energie verschlingendes System ins Leben gerufen hat, wenn es doch vollkommen nutzlos ist. Das Subjekt scheint nach dieser Denkweise zu einem bloßen Zuschauer zu verkommen. Alles würde in der physischen Welt, in der Welt der ausgedehnten Körper, in der die eigentliche Kausalität stattfindet, genauso ablaufen, wenn es die mentale Welt der denkenden Dinge nicht geben würde. Neben dem extremen Widerspruch zu unserem alltagstheoretischen Verständnis des Mentalen erscheint der Epiphänomenalismus vor allem im Lichte der Darwinschen Evolutionslehre mit ihren strengen Selektionsprinzipien wenig überzeugend. Aufgrund dieser Schwierigkeiten haben sich die meisten Philosophen des Geistes und Wissenschaftstheoretiker von substanzdualistischen Annahmen abgewandt, sodass solche Positionen in der (philosophischen) Fachwelt heutzutage keine große Rolle mehr spielen und daher auch hier nicht weiterverfolgt werden.

3.1.3. Der Eliminative Materialismus und die Nicht-Existenz des Geistes

Lässt sich daraus aber nun ableiten, dass mit dem Ablehnen eines eigenen „unabhängigen Wirklichkeitsbereiches“ des Mentalen, einer eigenen „psychischen Substanz“, diejenigen Recht bekommen, die dem Geistigen die Existenz absprechen? Teilweise wurde und wird diese Schlussfolgerung tatsächlich gezogen, insbesondere mit Verweis auf die Alltagstheorien, die wir über unseren Geist haben. So konstituieren gemäß den Vertretern des so genannten *eliminativen Materialismus* Annahmen, die wir mit dem „gesunden Menschenverstand“ über die Welt machen, eine Art primitiver Theorie (eine „Alltagspsychologie“, wie sie auch im Zusammenhang des *Theory-of-Mind-Ansatzes* postuliert wird). Doch wie bei jeder Theorie sei es so, dass die von ihr postulierten Entitäten nur gerechtfertigt sind, wenn die Theorie wahr ist, und nach Meinung der eliminativen Materialisten wird sich eben irgendwann herausstellen (und hierfür würden sich bereits jetzt die Hinweise mehren), dass die Alltagspsychologie falsch ist und somit auch unsere geistigen Vorgänge, auf die sie sich vermeintlich bezieht, in Wirklichkeit nicht existieren. Beispiele hierfür finden sich tatsächlich in der Wissenschaftsgeschichte (vgl. Kuhn, 1962/1976). In der Chemie hielt sich z.B. bis ins 18. Jahrhundert die Phlogiston-Theorie der Verbrennung. Phlogiston ist eine hypothetische Substanz, von der man vermutete, dass sie bei Erwärmung in alle brennbaren Körper eindringt sowie bei der Verbrennung aus diesen austritt. Heute weiß man, dass es Sauerstoff ist, den die Körper bei der Verbrennung binden. Statt des angenommenen Austritts von Phlogiston wird die Luft bzw. ihre Bestandteile absorbiert (Oxidationstheorie der Verbrennung). Mit der Widerlegung der Phlogiston-Theorie der Verbrennung ist nun aber auch die Substanz Phlogiston selbst von der Weltbühne verschwunden. Und genau solch eine Entwicklung prophezeien die Vertreter des

eliminativen Materialismus auch dem Psychischen, das in ihren Augen nur den Status eines antiquierten theoretischen Konstrukts hat.

Zum Beispiel bezweifelt Dennett (1991/1994) die Existenz von phänomenalen Eigenschaften, da wir seiner Meinung nach aus der Perspektive der ersten Person keine zuverlässigen Angaben über unsere phänomenalen Vorgänge machen könnten. Die vermeintlichen phänomenalen Eigenschaften seien in Wirklichkeit – ganz im Sinne des *logischen Behaviorismus* (vgl. Kim 1996/1998) bzw. semantischen Physikalismus (vgl. Beckermann, 1999) – lediglich komplexe Handlungsdispositionen, die aus der Perspektive der dritten Person genauer zu erfassen sind als aus der Perspektive des Erlebenden. Patricia Churchland (1986) argumentiert daher, dass alltagspsychologische Begriffe wie „Glauben“ oder „Empfindung“ inkohärent sind und durch eine neurowissenschaftliche Terminologie ersetzt werden sollten.

Der eliminative Materialismus mag theoretisch viele der hier angesprochenen Schwierigkeiten in Bezug auf das Verhältnis von physischen und mentalen Ereignissen lösen (denn durch die angenommene Nicht-Existenz Letzterer können diese auch keine Schwierigkeiten mehr anrichten). Doch zu welchem Preis? Mit der Leugnung der Existenz mentaler Prozesse widerspricht er in so eklatanter Weise der Realität des Bewusstseins und somit unserem Selbstverständnis als denkende und fühlende Subjekte, dass diese „Lösung“ des Leib-Seele-Problems wohl nur für die wenigsten Menschen akzeptabel sein dürfte. Aber auch unter Betrachtung logischer Gesichtspunkte lässt sich dieser Denkrichtung nicht viel abgewinnen. So weist die aufgestellte Behauptung, es gäbe keine mentalen Zustände, einen inhärenten Selbstwiderspruch auf, nimmt sie doch mit ihrem Geltungsanspruch selbst Bezug auf mentale Zustände wie „Bedeutung“ oder „Wahrheit“. Wenn es in der Welt aber keine bedeutungsvollen Zustände gibt, die wahr oder falsch sein können, sondern nur

„blindes“ materielles bzw. neuronales Geschehen, dann dürfte es auch diese Behauptung nicht geben. Mit ihrem Grundpostulat setzen eliminative Materialisten also etwas implizit voraus, was sie eigentlich bestreiten wollen: mentale Zustände (vgl. Beckermann, 1999).

Wenn man sich aus guten Gründen von der Annahme verabschiedet, dass Körper und Geist so etwas wie eigenständige Substanzen oder ontologisch distinkte Wirklichkeitsbereiche sind, wie kann man sich dann ihre Existenz und ihre Beziehung zueinander begreiflich machen? Wenn man einen Substanzdualismus (oder -pluralismus) ablehnt, kann es nur eine einzige „substantielle Wirklichkeit“ geben. Wie lässt sich dies aber mit der gleichzeitig offenkundigen Existenz von mentalen und physikalischen Gegebenheiten in Einklang bringen, ohne in einen eliminativen Materialismus zu geraten? Es scheint angenommen werden zu müssen, dass es sich bei Gehirn und Geist um eine gewisse (ontologische) *Einheit* handelt, die offenbar auf verschiedene Art und Weise in Erscheinung tritt bzw. sich für uns auf unterschiedliche Art und Weise darstellt. Aber was genau bedeutet das? Wird hier eine Differenz in die Welt getragen, die so in dieser selbst gar nicht vorhanden ist?

3.2. Die Einheit und Differenz von Gehirn und Geist: Spinozas Dilemma

Zur Vermeidung einer Kollision der verschiedenen Sorten von „Realitäten“, von denen Physik und Psychologie handeln, hat Spinoza bzw. Fechner die Lehre des psychophysischen Parallelismus erfunden, welche mich offen gestanden völlig befriedigt. (Albert Einstein, zit. nach Heidelberg, 2002, S. 53)

Auch wenn man die Phrase, dass von früheren Denkern und Forschern bereits das meiste der heutigen Diskussion und Problemlösungsvorschlä-

ge zu philosophischen Fragen wie dem Leib-Seele-Problem vorweg genommen wurde, überstrapazieren und gegenwärtige Weiterentwicklungen unterschätzen kann, lohnt es sich an dieser Stelle dennoch, in einem ersten Schritt einen „alten“ Philosophen als Ausgangspunkt heranzuziehen, der einen Ansatz entwickelt hat, der den Rahmen einer gehaltvollen und empirisch plausiblen Erklärung für den Ursprung des Leib-Seele-Problems trotz des faktischen Bestehens einer psychophysischen Einheit/Identität abgesteckt hat: Baruch de Spinoza²⁶. So formen seine Bestimmungen, aber auch gerade die „Grenzkpunkte“ seines Werkes, die Leitlinien für eine – auch vor dem Hintergrund der naturwissenschaftlichen Constraints der Gegenwart – gehaltvolle nicht-dualistische Theorie des Geistes bzw. eine Erklärung für den Ursprung des Leib-Seele-Problems. Spinoza sieht den Geist als ein natürliches (nicht von seinen physischen Grundlagen abtrennbares) und dennoch eigenständiges, wissenschaftlich respektables Phänomen an.

Mit Spinoza beginnt in der Philosophie der Neuzeit eine Entwicklungslinie, die man häufig – wie im obigen Zitat auch Einstein – etwas missverständlich als *Psychophysischen Parallelismus*²⁷ bezeichnet hat, die sich ge-

²⁶ Es wird an dieser Stelle ganz bewusst darauf verzichtet, die Identitätstheorie wie sie von Place (1956), Smart (1959) oder Feigl (1958/1967) vertreten wird, zum Ausgangspunkt der Besprechung der Möglichkeit einer Gehirn-Geist-Einheit oder eben -*Identität* zu machen (obgleich dies in der Philosophie des Geistes üblich sein mag). Der Grund hierfür ist, dass diese Identitätskonzeption eine sehr problematische Auslegung erfahren hat, bei der das Psychische *nichts anderes als* ein Vorgang in unserem Gehirn ist. Da dies für die Zielstellung dieses Kapitels zu vernachlässigen ist, wird der Begriff einer Einheit/Identität zunächst nur in etwas vager Form eingeführt, um dann im weiteren Verlauf der Arbeit (vor allem in Kapitel 6 und 7) klarer herausgearbeitet zu werden.

²⁷ Die Bezeichnung *Psychophysischer Parallelismus* geht wahrscheinlich auf Wilhelm Wundt oder Alexander Bain zurück (vgl. Heidelberger, 2002, S. 44). Sie ist etwas irreführend, da sie das Parallelgehen zweier Wirklichkeitsbereiche zu suggerieren scheint, wie es in der Leibnizschen *Theorie der prästabilierten Harmonie* der Fall ist (zumindest auch diese Theorie oft diese Bezeichnung beigelegt bekommt). Dabei ist der Psychophysische Parallelismus Spinozas,

zielt vom kartesischen Dualismus absetzt. Er hat einen metaphysischen *Monismus* vertreten, also angenommen, dass die Welt nur aus *einer* grundlegenden Substanz²⁸ besteht. Anschauliche Ideen (d.h. bewusste geistige Vorgänge/Inhalte) sind nach Spinoza immer unmittelbar auf den eigenen Körper bezogen. Die äußere Welt ist uns niemals direkt gegeben, sondern wir müssen sie immer über unsere Physis (vor allem über unsere Sinnessysteme und durch unser Nervensystem) erfassen. Daher ist das „Objekt der Idee, die den menschlichen Geist ausmacht, [...] der Körper“ (Spinoza, 1677/1977, S. 141). Modern interpretiert könnte man sagen, dass der Körper bzw. das Gehirn *Träger von Informationen* ist. Jeder Bewusstseinsinhalt hat zwangsläufig ein körperliches Gegenstück, die den Geist konstituierenden Ideen stehen immer in Verbindung bzw. sind identisch mit bestimmten körperlichen Zuständen. Wenn nun also die Frage gestellt würde, was der Mensch seinsmäßig ist, dann ließe sich mit Spinoza darauf antworten, dass er weder nur ein körperloser Geist oder geistloser Körper ist, vielmehr ist der Mensch eine „Einheit von Geist und Körper“ (ebd., S. 143), die sich aus der Zusammengehörigkeit in der einen und einzigen Substanz ergibt.

Die von Descartes (1641/1986) noch als zwei verschiedene Substanzen bezeichneten Größen *extensio* und *cogitatio* sind daher für Spinoza nicht von substantieller Beschaffenheit, sondern sie sind *Modi* (Zuständlichkeiten) der einzig existierenden Substanz²⁹, die sich uns in Form be-

Fechners oder auch Wundts keiner, der einen ontologischen, sondern – wenn überhaupt – nur einen *epistemischen* Dualismus annimmt.

²⁸ Die all-eine Substanz ist bei Spinoza wiederum mit Gott gleichzusetzen, weshalb man seine Position auch als *Pantheismus* bezeichnet. Gott ist die Gesamtheit alles Existierenden. Man könnte in diesem Sinne auch sagen, *das* größte „Wunder“ der Welt ist bereits deren bloße Existenz. Für den hiesigen Kontext ist diese Annahme allerdings nicht weiter relevant.

²⁹ Hiermit ist nicht gemeint, dass es nur einen bestimmten Grundstoff der Welt gibt wie die Annahme, dass z.B. alles aus Kohlenstoff bestehen würde. Vielmehr ist lediglich ge-

stimmter *Attribute* präsentiert. Obgleich wir verschiedene Dinge über das Wesen der Substanz aussagen können mögen, bezieht sich dies mit Notwendigkeit auf die *all-eine* Substanz, sodass „folglich die denkende Substanz und die ausgedehnte Substanz ein und dieselbe Substanz ist, die bald unter diesem, bald unter jenem Attribut aufgefaßt wird“ (ebd., 1677/1977, S. 123). Die von ihm getätigten ontologischen Unterscheidungen führen zu einem – entgegen Descartes – durchaus naturwissenschaftlich konsistenten Bild: Da Körper und Geist Modi ein und derselben Substanz sind, sind auch ein Eindruck in unserem Geist (z.B. die Wahrnehmung eines Geräusches) und der mit ihm assoziierte neurophysiologische Prozess *substantiell* letztendlich ein und dasselbe. Ihre Verschiedenheit rührt nur daher, dass sich diese Modi der Substanz unter verschiedenen Attributen manifestieren. Trotz dieser gewissen Identitätsannahme besteht Spinoza auf die Verschiedenheit der Attribute des Psychischen und Physischen, die nicht aufeinander reduziert werden können.

Was genau sind aber nun Attribute? Hier bleibt das Spinozistische System etwas vage und auch die Spinoza-Interpreten sind sich uneins. Per Definition bezieht es sich auf „dasjenige an der Substanz, was der Verstand als zu ihrem Wesen gehörig erkennt“ (ebd., S. 5). Insofern scheint gemeint zu sein, dass die Verschiedenheit von Körper und Geist auf die Art der Bezugnahme (z.B. über unser Wahrnehmungssystem) zur all-einen Substanz hervorgerufen wird und nicht durch deren Verschiedenheit „an sich selbst“. Es ist die These der *geistigen Entzweiung der Gehirn-Geist-Einheit*, die vor allem Schelling in seinem Frühwerk und auch Fechner aufgegriffen haben. In diese Richtung interpretiert ebenfalls

meint, dass es nicht zwei fundamentale Substanzen gibt, aus denen die Welt besteht, eine geistige und eine materielle Substanz, die irgendwie aufeinander bezogen werden müssen. Statt eines Dualismus wird hier also ein Monismus vertreten (was aber keineswegs heißen muss, dass es in Wirklichkeit nur ein materielles Sein gibt).

Rang (2000) Spinoza. Er beruft sich auf einen Brief, den Spinoza an Simon de Vries, einem Freund und Schüler, geschrieben hat, in dem er darauf hinweist, „daß man doch auch sonst ein und dieselbe Sache verschieden kennzeichnen könne, ohne daß die verschiedenen Kennzeichnungen eine Verschiedenheit in die gekennzeichnete Substanz hineinbrächten“ (S. 37).

Es gibt jedoch auch Passagen in seinem Werk, die etwas anderes nahe legen. Wenn er z.B. sagt: „Substanzen oder, was dasselbe ist [...] ihre Attribute“ (Spinoza, 1677/1977, S. 11), scheint er keine solche Kluft zwischen der all-einen Substanz und den Attributen zuzulassen. Im Gegenteil: die Attribute sollen sogar das Wesen der Substanz konstituieren. Auch die meisten Spinoza-Interpreten nehmen an, dass die essentielle Verschiedenheit von Geist und Körper (Gehirn), so wie sie sich uns präsentiert, darauf zurückzuführen ist, dass auch die sie hervorbringende Substanz über Attribute essentiell gegliedert ist (vgl. z.B. Wahle, 1889, S. 12f; Bartuschat, 1996, S. 69ff). Da der Mensch keinen Zugriff auf die Substanz selbst hat, erscheint sie ihm stets nur in Form von Attributen, die allerdings (jedes für sich) die Substanz durchaus korrekt widerspiegeln und keine bloßen Artefakte des menschlichen Wahrnehmungssystems bzw. unserer arteigenen Auffassungsgabe sind.

Wie es scheint, ist sich Spinoza selbst über den Status der Attribute nicht abschließend klar geworden. Hier hat sein ansonsten sehr ausgeklügeltes System gewissermaßen eine Achillesferse. Oder wie Rang (2000) *Spinozas Dilemma* – etwas zugespitzt – beschreibt:

Einerseits ist ein Attribut der Substanz als etwas, das ihr Wesen ausmacht, so nahe an die Substanz herangerückt, daß es mit ihr zu verschmelzen scheint, andererseits ist es von ihr so weit entfernt, daß Spinoza Grund hatte, das Verhältnis eines Attributs zur Substanz mit dem Verhältnis eines Namens zu dem von ihm Benannten zu vergleichen. Auf der einen Seite droht der Rückfall in die Zweisubstanzenlehre

Descartes', auf der anderen Seite die totale Bestimmungslosigkeit der Substanz an ihr selbst. (S. 39)

Spinoza hat die Attribute eingeführt, um der Tatsache Rechnung zu tragen, dass wir keinen unmittelbaren Zugriff auf die Substanz selbst haben. Insofern hat er zunächst *de facto* einen Schritt in Richtung „Attribute als Determinanten subjektspezifischer Repräsentationsmodi“ getan. Gleichzeitig schien es ihm aber – aus guten Gründen – abwegig zu sein, von einem lediglich anthropomorphen Phänomen zu sprechen. Denn die Art und Weise, in der sich uns die Substanz darstellt, muss doch auch etwas mit dieser selbst zu tun haben und nicht nur auf unsere – durch unsere eigene Beschaffenheit determinierte – Zugriffsweise zurückgehen. Wie aber kann man sich diese strukturelle Gliederung der Substanz selbst dann vorstellen? Diesen Widerspruch aufzulösen, gelang Spinoza nicht und er konnte – meines Wissens – auch im Zuge der Spinoza-Rezeption nicht aufgeklärt werden.

Mit seinen Bestimmungen liefert er dennoch den Rahmen für eine weitergehende Klärung der Frage nach dem Ursprung des Leib-Seele-Problems, denn offensichtlich gibt es zunächst zwei logische Möglichkeiten der Auflösung dieses Dilemmas: Entweder sind die Attribute ein Ausdruck unseres *subjektiven* (beobachterabhängigen) Zugriffs auf die Welt und die wahrgenommene Gehirn-Geist-Differenz bzw. der subjektive Dualismus entspringt unserer arteigenen Verfasstheit, wobei dann die psychischen Mechanismen freigelegt werden müssten, die dieser Entzweiung zugrunde liegen; oder aber die Attribute beziehen sich auf *objektive* (beobachterunabhängige) Tatsachen. Dann wäre zu klären, wie genau es eine solche Vielgestaltigkeit geben kann, wenn es sich hierbei doch um keine „substantielle“ Pluralität handeln kann. Während für den ersten Ansatz der Begriff der *Perspektive* von zentraler Bedeutung ist, ist es der

der *Ebene* für den zweiten Ansatz. An späterer Stelle wird sich zeigen, dass sich beide Ansätze – entgegen dem ersten Anschein – unter Bezugnahme auf den gegenwärtigen Stand der Wissenschaft bzw. philosophischen Begriffsbildung miteinander versöhnen lassen, was nicht nur eine Klärung der Frage nach dem Ursprung des Leib-Seele-Problems ermöglicht, sondern auch entschieden zur Etablierung eines allgemeinen Modells einer Architektur des Geistes beiträgt.

Die im Folgenden zunächst weiterverfolgte Idee besteht darin, dass es sich bei der Gehirn-Geist-Differenz um ein Resultat unterschiedlicher *Perspektiven* gegenüber ein und derselben Sache handelt, d.h., dass die Verschiedenheit der Erscheinungsweise von mentalen und physischen bzw. neuronalen Phänomenen darauf zurückgeht, dass wir auf *ein und denselben* Wirklichkeitsausschnitt auf unterschiedliche Art und Weise sprachlich-konzeptuell oder perzeptuell zugreifen. Das, was wir in unserer subjektiven Perspektive als mentales Erlebnis erfahren oder beschreiben, wird in der Perspektive der Hirnforschung als neuronaler Zustand bzw. neuronales Aktivitätsmuster erfasst. Das Leib-Seele-Problem (bzw. der intuitive Dualismus) wäre somit etwas, was auf unsere mentale Verfasstheit zurückgeht, die keine „absolute“ Perspektive auf einen Gegenstand zulässt, sondern immer nur eine, die *relativ* zu gegebenen perzeptuellen und/oder theoretischen Rahmenbedingungen ist. Das Leib-Seele-Problem hätte demnach eine psychische bzw. psychologisch erklärbare Ursache.

Die Idee, dass es sich bei Gehirn und Geist um divergierende *Erscheinungsformen* einer einzigen, diesen zugrunde liegenden Realität handelt, ist allerdings keineswegs neu. Es war insbesondere Friedrich Wilhelm Joseph Schelling (1775, 1802/2005), der für diese Variante der Auflösung des Spinozistischen Dilemmas argumentiert hat und mit seiner so genannten *Identitätsphilosophie* einen wichtigen Mittelsmann zwischen

Spinoza und dem – später besprochenen und für die Zielstellung dieses Kapitels zentralen – Ansatz Gustav Theodor Fechners darstellt³⁰. Weitere namhafte Vertreter dieser Ansicht sind Friedrich Albert Lange (1866/1908), Ernst Mach (1886/1903, 1905/1968), Moritz Schlick (1918/1979) oder Michael Pauen (1999, 2005, 2007) sowie teilweise auch Wilhelm Wundt (1894, 1911), Bertrand Russell (1921/2000), Carl Friedrich von Weizsäcker (1971/1989) oder auch Jochen Fahrenberg (2013).

3.3. Die Gehirn-Geist-Differenz als sprachlich-deskriptives Phänomen

Eine Möglichkeit, die sich bei Spinoza selbst in dem genannten Brief angedeutet hat, ist also, dass die Differenz zwischen Gehirn und Geist auf divergierende oder konfligierende *Beschreibungen* ein und desselben Systems zurückgehen. Es sind demnach verschiedene *konzeptuelle* oder *sprachliche* Zugangsweisen, die bestimmen, ob der herangezogene Gegenstand als Gehirn oder Geist erfasst wird.

Die hier angesprochene und auch in der Gegenwartsphilosophie verbreitete Überzeugung, dass die Wurzel philosophischer Probleme auf einer sprachlichen Ebene zu suchen sei und dass diese Probleme daher auch auf der sprachlichen Ebene gelöst oder aufgelöst werden müssen, geht vor allem auf die *analytische* Tradition der Philosophie (des Geistes) im 20. Jahrhundert zurück. Nach deren Grundüberzeugung handelt es sich auch beim Leib-Seele-Problem um ein sprachbezogenes Problem. So begann, wie Bieri (1981/1997) aufzeigt, vor allem mit Gilbert Ryles (1949)

³⁰ Behandelt wird Schelling in dieser Funktion in Lüdmann (unveröffentlichtes Manuskript).

The Concept of Mind eine neue Ära in der Philosophie des Geistes. Nach Ryle und seinen Anhängern sind unsere dualistischen Intuitionen „das Produkt einer unterlassenen oder falschen Analyse des mentalistischen Vokabulars“ (Bieri, 1981/1997, S. 12).

Kennzeichnend für die hieraus erwachsene *ordinary language philosophy* ist, dass statt der Ausformulierung und Verteidigung ontologischer Gehirn-Geist-Konzeptionen eine *Begriffsanalyse* betrieben werden sollte, wobei Begriffe als rein sprachliche Entitäten aufgefasst werden und die erkenntnistheoretische Prämisse akzeptiert wird, dass „wir keinen epistemischen Zugang zu den Phänomenen haben, der unabhängig von Sätzen ist“ (ebd., S. 13). Dies bedeutete dann aber auch, dass eine Lösung für Probleme wie das des Psychophysischen durch eine genauere Untersuchung erreicht werden kann, wie wir für gewöhnlich über Gehirn und Geist sprechen. Wenn die Entlarvung des ontologischen Dualismus als Abweg feststeht, muss die Aufgabe einer solchen Philosophie daher sein nachzuzeichnen, wie es historisch in der sprachlichen Begriffsbildung oder Definition von Begriffen dazu kommen konnte, dass sich Mentales und Physisches prinzipiell auszuschließen scheinen.

Auch wenn die „Philosophie der gewöhnlichen Sprache“ trotz einiger Erfolge schnell ins Wanken geraten ist, nicht zuletzt, da sich unter dem Deckmantel der Begriffsanalyse „Berge von Trivialitäten“ (ebd.) aufgetürmt haben und – infolge der Begriffsanalyse als Methode der Wahl – eine merkwürdige Abkapselung von der empirischen (psychologischen) Forschung stattgefunden hat, ist der Sprachduktus aus jenen Tagen in der Philosophie des Geistes immer noch allgegenwärtig (z.B. im Hinblick auf Formulierung wie: „Es lässt (nicht) sich sagen, dass...“; der Verweis auf so genannte *Kategorienfehler* etc.). Des Weiteren ist die in diesem Zusammenhang ableitbare These, dass das Leib-Seele-Problem bzw. die

Gehirn-Geist-Differenz auf unsere Sprachgewohnheiten zurückgeht, nicht uninteressant.

So ist es nicht von der Hand zu weisen, dass wir im Laufe unserer kulturellen und wissenschaftlichen Entwicklung völlig unterschiedliche sprachlich-konzeptuelle Systeme ausgebildet haben, um uns einerseits auf körperliche Prozesse, zu denen auch Vorgänge in unserem Nervensystem gehören, oder andererseits auf unser mentales Innenleben zu beziehen. Es sind die intuitiven Differenzen zwischen Gehirn und Geist, denen vor allem Descartes (1641/1986) Ausdruck verliehen hat, die sich in unserer Sprache niedergeschlagen haben und hierüber wiederum unser Denken, unsere Weltauffassung, beeinflussen bzw. bereits vorhandene Intuitionen verfestigen: während der Geist als *res cogitans* („denkendes Ding“) offenbar keine Ausdehnung besitzt und damit nicht im Raum lokalisierbar ist, kommen materiellen Gegebenheiten der *res extensa* wie dem Gehirn und seinen strukturellen Einheiten genau die gegenteiligen Attribute zu (ausgedehnt und räumlich lokalisierbar, nicht-denkend). Und so werden auch die Wissenschaften klassischerweise in Geistes- und Naturwissenschaften eingeteilt, die über unterschiedliche Gegenstandsbereiche und (damit) verschiedene Sprachsysteme zur Beschreibung und Erklärung ihrer Forschungsobjekte verfügen. Während Anziehung, Abstoßung, Ladung, Dichte, Diffusion oder Aggregatzustände klassische Fachtermini naturwissenschaftlicher (nicht-intentional konzipierter) Theorien sind, sind diese offensichtlich ungeeignet, um Gegenstände der Geisteswissenschaften zu beschreiben. Historiker (re-)konstruieren hingegen – sehr verallgemeinert – den Verlauf der Ereignisse im Hinblick auf Staaten, Strömungen oder auch der gesamten Menschheit vor dem Hintergrund bestimmter zumeist absichtsvoller Handlungen politischer, nationalstaatlicher oder religiöser Akteure, wobei das hier verwendete

sprachlich-konzeptuelle System ungeeignet zu sein scheint, um naturwissenschaftliche Fragestellungen zu erhellen.

Mit dieser klaren Grenzziehung war es allerdings spätestens vorbei, als geistige Eigenschaften und Vorgänge im Zuge der voranschreitenden technischen Entwicklung zunehmend auch Gegenstand der naturwissenschaftlichen Forschungsmethoden wurden. Vor allem mit dem Aufkommen der Kognitions- und Neurowissenschaften wurde die menschliche *Informationsverarbeitung* zunehmend durch das Vokabular der (Neuro-) Biologie beschrieben. Die hierbei auftretenden Simplifizierungen oder Reduktionismen („Liebe und Zuneigung ist nichts anderes wie die Ausschüttung von Oxytocin“), die leider auch von prominenten und medienaffinen Neurowissenschaftlern wie Gerhard Roth oder Wolf Singer befeuert wurden, haben zu einer starken – und für den hiesigen Kontext aufschlussreichen – Abwehr auf Seiten vieler Geisteswissenschaftler inkl. geisteswissenschaftlich ausgerichteter Psychologen geführt. Z.B. hat Uwe Laucken (2002, 2003) kritisiert, dass die bedeutungstragenden Inhalte unseres Bewusstseins nicht in der Formalsprache der Naturwissenschaften, wie Größe, Gewicht oder Zeit, erfasst werden können, denn was sollte eine neurowissenschaftlich gewendete Psychologie z.B. zum Thema „Dialog der Kulturen“ (ebd., 2002) aussagen können. Aber selbst wenn man es für möglich erachten würde, „die Geistigkeits-Eigenschaft in der ‚Cm, g, sek‘-Sprache auszudrücken, dann könnte man sich die Geistigkeit als eigenständige Geistigkeit sparen, denn dann hätte sie sich in den Eigenschaften, denen die ‚Cm, g, sek‘-Sprache gemäß ist, verflüchtigt. Es machte dann gar keinen Sinn mehr von einer besonderen Eigenschaft der Geistigkeit zu sprechen“ (ebd.).

Ohne auf die Argumente an dieser Stelle inhaltlich näher einzugehen, zeigt sich hier eindrucksvoll, dass die verschiedenen sprachlich-konzeptuellen Systeme, die wir zur Beschreibung geistiger versus physischer

Phänomene entwickelt haben, eine bedeutsame Ursache für den Umstand sein könnten, dass wir auch zwischen den Bezugsgegenständen dieser Systeme eine ontologische Kluft auszumachen meinen. Eine Möglichkeit wäre demnach, dass der divergierende sprachlich-deskriptive Zugriff als Ursache unseres intuitiven Dualismus fungiert. Wir haben unterschiedliche Beschreibungen ein und desselben Phänomens und erhalten hierdurch den Eindruck, dass durch die unterschiedlichen Bezeichnungs- oder Beschreibungsweisen ein Unterschied in der Sache selbst gekennzeichnet wird.

Trotz einer faktisch vorhandenen psychophysischen Einheit oder Identität haben wir in unserer Alltags- und Wissenschaftssprache verschiedene Ausdrücke entwickelt, um uns einerseits auf Vorgänge in unserem Gehirn zu beziehen und andererseits Bezug auf Geschehnisse in unserem Geist zu machen. Eine solche Idee zeigt sich beim – als Initialzündung der Formulierung ihrer Theorie dienenden – Verweis der Identitätstheoretiker der 1950er Jahre auf Freges Unterscheidung zwischen der Intension (Bedeutung, Konnotation) und der Extension (Referenz, Denotation) eines Ausdrucks, was auf die in der philosophischen Semantik Gottlob Freges getroffene Unterscheidung von Sinn und Bedeutung³¹ zurückgeht (vgl. Place, 1956; Feigl, 1958/1967 bzw. Kap. 6.1 dieser Arbeit).

Frege (1892) hat versucht aufzuzeigen, wie sprachliche Ausdrücke ihre Bedeutung erhalten und wie es möglich ist, dass sich zwei von ihrem Sinn her verschiedene Terme trotzdem auf denselben Teil der Wirklichkeit beziehen können. Zur Illustration kann hier das Beispiel der Identität von

³¹ Die ursprüngliche terminologische Unterscheidung von Sinn und Bedeutung ist irrtumsanfällig, da beide Ausdrücke im heutigen Sprachgebrauch weitestgehend synonym verwendet werden.

Morgenstern und Abendstern dienen: Morgenstern \equiv Abendstern. Da Morgen- und Abendstern an andere Wahrnehmungssituationen gebunden sind und lange Zeit als zwei distinkte Himmelskörper betrachtet wurden, kann man beiden Ausdrücken durchaus eine unterschiedliche Intension/Bedeutung zuordnen, obgleich die Astronomie uns gezeigt hat, dass die Ausdrücke bezugsgleich sind, da sich „Morgenstern“ und „Abendstern“ auf denselben Himmelskörper beziehen, den Planeten Venus. Und so wird ebenso durch den Umstand, dass die Bedeutung psychologischer Prädikate eine ganz andere sein mag als die von neuronalen Prädikaten, nicht ausgeschlossen, dass sie *extensionsgleich* sind, weil sie sich auf ein und denselben Wirklichkeitsausschnitt beziehen.

Wenn das Verhältnis zwischen Gehirn und Geist von den Identitätstheoretikern mit dem Verhältnis von Abendstern und Morgenstern verglichen wird, indem in beiden Fällen von einer Extensionsgleichheit bei einer Intensionsverschiedenheit der verwendeten *Ausdrücke* gesprochen wird (vgl. Kap. 3.3), zeigt sich, dass die Ursache der Verschiedenheit von Gehirn und Geist bzw. unserem intuitiven Dualismus an dieser Stelle in einem divergierenden *sprachlich-konzeptuellen* Zugriff auf ein und denselben Bezugsgegenstand ausgegangen wird. Dies wird insbesondere beim Identitätstheoretiker Herbert Feigl deutlich, wenn er sagt, dass wir es im Fall von neuronalen und psychischen Zuständen nicht mit zwei verschiedenen ontologischen Reichen zu tun haben, sondern „only one reality which is represented in two different conceptual systems – on the one hand, that of physics and on the other hand, where applicable (in my opinion only to an extremely small part of the world) that of phenomenological psychology“ (S. 41; Hervorh.: M.L.).

In ganz ähnlicher Weise wird im Rahmen der Entwicklungspsychologie eine bestimmte Klasse von so genannten *Perspektivenproblemen* besprochen, die darauf zurückgehen sollen, dass wir ein und denselben Gegen-

stand mit unterschiedlichen Worten beschreiben oder aus dem Strom der Ereignisse „herausgreifen“ können (vgl. Perner, 2000, Perner, Brandl & Garnham, 2003). Kinder lernen in diesem Sinne erst allmählich, dass sie sich mit unterschiedlichen Ausdrücken auf ein und denselben Gegenstand beziehen können. Z.B. ließe sich ein bestimmter Gegenstand sowohl als „Orange“ oder „Apfelsine“ (strikte Synonymie) bezeichnen oder ein anderer sowohl als „Hase“ als auch „Tier“ (Wechsel der Kategorie bzw. taxonomischen Ebene). Kinder unter vier Jahren haben hierbei Schwierigkeiten und sagen beispielsweise, dass es sich bei dem Objekt um einen Hasen handelt, womit es nicht zugleich ein Tier sein könne; ein Phänomen, das man als *wechselseitigen Ausschluss* bezeichnet (Carey & Bartlett, 1978; Dockrell & Campbell, 1986).

In diesem Zusammenhang beruft man sich häufig auf den aus der (Sprach-)Philosophie stammenden Ausdruck *Sortals*³² (vgl. Perner, 2000). Sortals sind Wörter, durch deren Anwendung determiniert wird, was für eine Art von Gegenstand (von was für einer „Sorte“) ein Bezugsgegenstand ist. Es ist gleichsam das Kriterium, nach dem wir Dinge abzählen bzw. *sortieren* können, sodass man im Deutschen auch etwas ungelenc von „Zählsubstantiven“ (Von Kutschera, 1999, S. 11) spricht. Je nach Sortal kann ein und derselbe Gegenstand, über den wir sprechen, auf völlig verschiedene Weise (sprachlich) individuiert werden (vgl. Hirsch, 1982) oder wie Von Kutschera (1999) sagt, kann man von „Objekten einer bestimmten Art und von einer bestimmten Anzahl von Objekten dieser Art [...] immer nur relativ zu einer Analyse oder Beschreibung reden“ (S. 16). Demnach bestimmt unser sprachlich-deskriptiver Bezugsrahmen,

³² Eingeführt wurde der Terminus von John Locke (1690/2000), während die aktuelle Debatte und die sprachphilosophische Orientierung vor allem mit Gottlob Frege (1884) und Strawson (1959) verbunden ist.

auf welche Art und Weise wir Objekte individuieren bzw. dass wir ein und dasselbe Objekt auf verschiedene Art und Weise individuieren können.

Im Rahmen der Entwicklungspsychologie durchgeführte Experimente zeigen, dass jüngere Kinder insbesondere Schwierigkeiten mit Aufgaben der alternativen Bezeichnung ein und desselben Gegenstandes haben, wenn man Sortals verschiedener sprachlich-konzeptueller Ebenen (wie Hase vs. Tier vs. Lebewesen) verwendet und weniger bei (striker) Synonymie (Perner, Stummer, Sprung & Doherty, 2002). Überträgt man diese Annahmen wiederum auf die Gehirn-Geist-Relation könnte man sagen, dass sich das Problem des wechselseitigen Ausschlusses entgegen seiner ansonsten relativ frühen Auflösung im Falle von Gehirn und Geist über das gesamte Leben aufrechterhält. Irgendwie scheint das, worauf wir uns mit neurowissenschaftlichen Ausdrücken beziehen, nicht dasselbe sein zu können wie das, worauf wir uns mit genuin psychologischen Ausdrücken beziehen. Die von uns erfasste Differenz ginge dabei auf die Anwendung von Sortals aus unterschiedlichen sprachlich-konzeptuellen Kontexten zurück. Machen wir uns das sprachlich-deskriptive System der Geisteswissenschaften oder (phänomenologischen) Alltagspsychologie zu Eigen, werden wir von *geistigen* Eigenschaften oder Prozessen sprechen, die unserem Bezugsgegenstand kennzeichnen. Individuieren wir ein und denselben Bezugsgegenstand nach Maßgabe neurowissenschaftlicher Termini, werden wir ihn dagegen als *Gehirn* individuieren. Interessant ist, dass „das“, worauf wir uns – der dargestellten Logik zufolge – mit Sortals beziehen, selbst in seiner Existenz nicht festgelegt zu sein scheint. Der Modus unseres sprachlich-konzeptuellen Zugriffs auf diesen Gegenstand bestimmt, als was dieser Gegenstand für uns (sprachlich) in Erscheinung tritt.

Auch wenn divergierende oder konfligierende sprachlich-konzeptuelle Systeme eine entscheidende Rolle in Bezug auf die Existenz und Vehemenz des Leib-Seele-Problems haben, bin ich überzeugt, dass hier nicht die eigentliche Ursache des Problems liegt. Vielmehr sind diese Systeme nur der Ausdruck von tiefer liegenden psychischen Dynamiken. Denn woher rühren die kartesischen Intuitionen, die sich in unserer Alltags- und wissenschaftlichen Sprache widerspiegeln? Woher kommt es zur Ausbildung des persistierenden Phänomens des wechselseitigen Ausschlusses von Gehirn und Geist? Offenbar scheint es hier bestimmte vor- bzw. subsprachliche Determinanten unserer Sprachgewohnheiten zu geben, die sich im Zusammenhang körperlicher versus geistiger Aspekte herausgebildet haben.

3.4. Die Gehirn-Geist-Differenz als Wahrnehmungsphänomen

Was häufig übersehen wird, ist, dass wir uns und unsere Umwelt nicht nur über sprachlich-konzeptuelle Beschreibungen und Erklärungen repräsentieren, sondern – viel elementarer – schon in unserer Wahrnehmung. Wir repräsentieren also einen Gegenstand durch den über unsere Sinne vermittelten Zugriff auf unsere Umwelt auf verschiedene Art und Weise und verfügen daher – je nach Bezugspunkt, Auflösungsgrad unserer spezifischen Wahrnehmungsmodalität oder auch der verwendeten wissenschaftlichen Technik – über verschiedene *Wahrnehmungs- oder Betrachtungsebenen* oder – wie man diese auch bezeichnen kann – *Perspektiven*. Wir haben demnach niemals einen *absoluten* Zugriff auf ein Objekt, sondern immer nur ein bestimmtes *Repräsentationsformat*.

Natürlich sind Perspektivitätsphänomene im Alltag häufig Phänomene, bei denen es um eine deutliche versus undeutliche Wahrnehmung oder Auffassung eines Gegenstands geht. Wenn wir ein Objekt in der Ferne sehen, sehen wir dieses aus unserer Perspektive z.B. verschwommen und erst bei näherer Betrachtung erkennen wir (vermeintlich), was es „wirklich“ ist (quasi unabhängig von uns selbst, entgegen unserer zuvor nur subjektiv verzerrten *Perspektive*, die unserer Distanz zum fraglichen Gegenstand geschuldet war). Fakt ist, dass wir als Menschen über eine sinnliche Wahrnehmung (und konzeptuell-intellektuelle Funktionsweise) verfügen, die physikalische Merkmale der Außenwelt in eine subjektive Darstellungsweise überführt, die sich nicht oder zumindest nicht vollständig aus den Merkmalen der Gegenstände selbst ergibt (vgl. Breuer, 1977). Dass wir z.B. Farben oder Geräusche so erleben, wie wir sie erleben, weist eine gewisse Beliebigkeit auf. Die zugrunde liegenden Wellenlängen reflektierten Lichts oder die über die Luft transportierten Schallwellen haben nichts Farbigen oder Geräuschhaftes an sich. Wenn unser Sinnessystem und/oder unsere neuronal implementierte Informationsverarbeitung anders konstruiert wäre, würden wir dieselben physikalischen Quellen ganz anders erleben (z.B. könnten wir – ähnlich wie ein Synästhetiker – Wellenlängen „hören“ oder Schallwellen „sehen“ oder wir hätten vielleicht phänomenale Eindrücke einer ganz anderen Qualität wie im Beispiel der Fledermaus). Deshalb sind Sinneseindrücke bereits von Locke (1690/2000) als *sekundäre Qualitäten*³³ bezeichnet worden, die mehr mit uns selbst als mit den Gegenständen zu tun haben, an denen wir sie wahrnehmen. Insofern fängt das berühmte Diktum von William James

³³ Primäre Qualitäten, die nach Locke den Gegenständen selbst zukommen sollen, sind dagegen Ausdehnung, Festigkeit und Gestalt.

(1890/1981), das besagt, dass der Mensch ein *Schauplatz geistiger Möglichkeiten* ist, viel Wahrheit ein.

3.4.1. Fechners Zwei-Perspektiven-Theorie

Auch Gustav Theodor Fechner (1851/1922, 1860/1964, 1861, 1879) entscheidet sich für die subjektivistische Auflösung des Spinozistischen Dilemmas. Zugleich grenzt er sich ausdrücklich von Positionen ab, die eine einseitige Abhängigkeit psychischer Prozesse durch ihre körperlichen Grundlagen annehmen. Ihm gehe es gerade nicht um „jene einseitige Bedingtheit des Geistes durch die Materie, bei der der Materialismus stehen bleibt“ (Fechner, 1851/1922, S. 260). Man kann Fechners Ansatz nichtsdestotrotz als *naturalistisch* bezeichnen, weil er das Psychische als ein Naturphänomen betrachtet und er bei seiner Annäherung an das Leib-Seele-Problem eine weitestgehend empirisch-naturwissenschaftliche Vorgehensweise anwendet.

Fechner nimmt bei seiner Theorie und Entscheidung zugunsten einer subjektivistischen Auflösung von Spinozas Dilemma starke Anleihen bei Friedrich Wilhelm Joseph Schelling. Während es bei Spinoza noch strittig war, ob die Verschiedenheit der Attribute auch eine Verschiedenheit der Substanz widerspiegelt, ist bei Schelling der Gegensatz von Körper und Geist (oder „Realem“ und „Idealem“, wie er es nennt) eindeutig ein Oberflächenphänomen einer Wirklichkeit, die (an sich) selbst undifferenziert ist. Nur in unseren Gedanken oder Wahrnehmungen bestehen die vermeintlichen Gegensätze; die Wirklichkeit selbst bleibt indifferent. Unter Schellings Voraussetzungen existiert der dualistische Standpunkt à la Descartes aus diesen Gründen nur im Denken für uns; Reales und Ideales, Körper und Geist, lassen sich nur gedanklich (also im „Idealen“) voneinander ablösen.

Es geht Schelling dabei nicht wie seinem Lehrmeister Immanuel Kant darum, eine *Verstandeskategorie der Einheit* als eine notwendige Bestimmung einzuführen, damit das – wie es im Kantischen Jargon heißt – Mannigfaltige der *Anschauung* zu einer Einheit gebracht werden kann (vgl. Kant, 1787/1966). Schelling hat vielmehr eine *reale* Einheit vor Augen. „Unsere Meinung ist [...] nicht, daß die Gegensätze nur überhaupt in einem allgemeinen Begriff zur Einheit gebracht werden, denn eine solche Einheit wäre wiederum nur formaler Art, sondern daß in allem, was sich ideel entgegengesetzt ist, die Wesenheit eins, und alles nicht bloß durch das äußere Band des Begriffs, sondern der inneren Substanz und gleichsam dem Gehalt selbst nach identisch ist“ (I/4, 411). Daher sei das „Reale und Ideale selbst [...] ein und dasselbe Ding, *nur unter verschiedenen Exponenten betrachtet*“ (I/6, 498f; Hervorh.: M.L.).

Genau diesen Gedanken einer divergierenden Betrachtungsweise, die der Körper-Geist-Differenz zugrunde liegt, greift Fechner auf. Für seine Position, die er selbst *Identitätsansicht* nennt, ist es entscheidend, dass sie aus zwei divergierenden Wahrnehmungsperspektiven entwickelt wird. Der Grundgedanke hierbei ist, dass es in Abhängigkeit von der einem Gegenstand gegenüber eingenommenen Perspektive zu einer qualitativ höchst unterschiedlichen Repräsentation dieses *einen* Gegenstandes kommen kann. Mit einem einfachen Bild zeichnet er nach, dass sich ein und derselbe Vorgang, der sich aus der Sicht eines Neurowissenschaftlers, der heutzutage mit bestimmten Verfahren wie fMRI oder EEG auf das Gehirn zugreift, als Zunahme von Aktivitäten in bestimmten Gehirnregionen präsentiert, für den untersuchten Probanden als z.B. eine sensorische Empfindung wie ein Rot-Erleben darstellen kann:

Vergegenwärtigen wir uns nun einen Menschen, dessen Geist in irgendeiner Weise beschäftigt ist; er mag sehen, hören, denken. Man weiß, daß bei der geistigen Tätig-

keit in seinen Augen, Ohren, Nerven, seinem Gehirn sich gewisse Veränderungen vollziehen. Die körperlichen Erscheinungen würde ein anderer, der von außen in jenen hineinblicken könnte, unter geeigneten Umständen wohl beobachten können; er würde dabei aber nichts von den Empfindungen und Gedanken wahrzunehmen imstande sein, die in der Seele der Ersteren vorhanden sind. Dieser spürt aber umgekehrt nichts von den körperlichen Vorgängen, obwohl sie in ihm selbst vorstatten gehen. Er fühlt nichts von den Bewegungsvorgängen in seinen Nerven, seinem Gehirn; für ihn ist alles nur Sehen, Hören, Denken, d.h. ausschließlich geistiges Erleben. *Wohl empfindet ein jeder seine Nerven, sein Gehirn mit allen darin verlaufenden Vorgängen, aber empfindet sie nicht als Nerven, als Gehirn, sondern als Sinnesempfindung, als Sehen, Hören, kurz als Geist.* (ebd., 1851/1922, S. 112f; Hervorh.: M.L.)

Es wird deutlich, dass Fechner mit dieser Position einen ontologischen Monismus und dennoch eine Art epistemischen Dualismus im Sinne der Dualität der Erscheinung oder Repräsentation einer Grundwesenheit vertritt. Die Verschiedenartigkeit von Gehirn und Geist verdankt sich in den Augen Fechners also vor allem dem „doppelte[n] Standpunkt der Betrachtung“ (ebd., S. 263). Entgegen der substanzdualistischen Tradition gibt es damit nach Ansicht Fechners auch im Falle des Menschen nur ein eingestaltiges Sein, das sich in Abhängigkeit von unserer „arteigenen“ Betrachtungs- bzw. Zugangsweise zum einen als ein geistiges zeigt, wenn man es aus der Perspektive des eigenen Bewusstseins erfährt, und das sich als physisch manifestiert, wenn man sich ihm von außen beobachtend (z.B. als Hirnforscher) nähert. In Fechners eigenen Worten: „Wir erklären uns die Verschiedenheit der Erscheinungsweise damit, daß sie von der Verschiedenheit des Standpunktes herrührt, den der Beobachter dazu einnimmt. Als etwas Geistiges erscheint der Vorgang von einem innerlichen Standpunkte aus, als etwas Leibliches von einem äußerlichen Standpunkte“ (ebd., S. 113f).

Wenn Fechner darauf sagt, dass erst ein anderer kommen muss, „um das, was jedem als geistiges Erleben erscheint, als materiellen Körper und als

materielle Vorgänge darin festzustellen“ (ebd., S. 113), dann zeigt sich darin die fundamentale Bedeutung wahrnehmungsbezogener Aktivitäten für die Divergenz zwischen Gehirn und Geist.

Aber auch jeder Mensch für sich genommen hat eine Art doppelten Zugang zu sich selbst. Was nun aber in der Selbstbezugnahme nicht möglich ist, dass man genau diejenigen physischen Prozesse als physische Prozesse in den Blick bekommt, denen die geistigen Vorgänge entsprechen, denn die betreffende Person fühlt zwar – wie Fechner sagt – ihr gerade „aktives Nervensystem“, nur nimmt sie es nicht als solches wahr, sondern als Gedanken, Wahrnehmungen, Stimmungen usw. Wir fühlen oder erfassen in unserer Erlebnisperspektive nicht die Aktivitäten in unserem Nervensystem als solche, sondern für uns stellt sich eben-diese Aktivität immer nur als geistige Aktivität, als Sehen, Hören, Denken etc., dar. In der Philosophie spricht man in diesem Zusammenhang von der *phänomenalen Transparenz* mentaler Zustände (vgl. Metzinger, 2006), einer Art phänomenologischer „Durchsichtigkeit“ im Hinblick auf ihre strukturellen oder physischen Grundlagen.

Der Fechnersche Ansatz basiert somit auf zwei grundlegenden Prinzipien: Perspektivität und phänomenale Transparenz³⁴. Einerseits ist es der Hinweis auf unterschiedliche Bezugspunkte oder Perspektiven, die man einem Gegenstand gegenüber einnehmen kann und zu einer divergierenden Repräsentation dieses Gegenstandes führen; andererseits wird verdeutlicht, dass es unmöglich zu sein scheint, aus der Perspektive seines eigenen Erlebens die (physischen oder neuronalen) Strukturen in den

³⁴ Die Begriffe „Transparenz“ und „Perspektivität“ werden also von Fechner selbst nicht verwendet, eignen sich aber m.E. vorzüglich, um die zentralen Bestandteile seiner Theorie zu bezeichnen.

Blick zu bekommen, die ebendiesem Erleben zugrunde liegen (vgl. Lüdmann, 2012, 2014a).

Die Idee Fechners besteht also darin, den Fakt unterschiedlicher Wahrnehmungsebenen, die sich ein und demselben Objekt gegenüber ergeben können, auf das Verhältnis von Gehirn und Geist selbst anzuwenden. So wie sich ein Kreis bei Veränderung des Blickwinkels als Ellipse oder – bei vollständiger Rotation in die Ebene – gar als Linie darstellen kann, kann sich auch das, was wir als unseren Geist bezeichnen, einmal als ebensolcher manifestieren, aber – aus einer anderen Perspektive – auch als Gehirn bzw. neuronales Aktivitätsmuster.

Die wahrgenommene Verschiedenheit von Gehirn und Geist entspringt demnach dem Geist selbst und geht nicht auf einen *der Sache nach vorhandenen* Unterschied zurück. Die von Fechner ins Feld geführten Ursachen – Perspektivität und phänomenale Transparenz – können als *psychischer* Ursprung des Leib-Seele-Problems angesehen werden, wodurch Spinozas Dilemma auf *subjektivistische* Art und Weise aufgelöst wird. Es ist der Beobachter, der die Differenz von Gehirn und Geist in die Welt trägt. Es ist kein realer, beobachterunabhängiger Fakt.

Diese Strukturmerkmale unserer psychischen Verfasstheit können auf zwei verschiedenen Wegen erklären, wie es trotz der angenommenen Gültigkeit eines ontologischen Monismus zum Problem der phänomenalen Differenz und damit zur großen intuitiven Überzeugungskraft dualistischer Vorstellungen kommt. So kann über den Verweis auf die von der Wahrnehmungsperspektive abhängige Repräsentation einer Gegebenheit die „äußere“ Facette des Problems der phänomenalen Differenz aufgeklärt werden, indem gezeigt wird, warum z.B. ein Rot-Erlebnis – aus einer externen Perspektive heraus repräsentiert – als neuronaler Prozess in Erscheinung tritt. Dagegen kann über den Verweis auf die Transparenz mentaler Zustände die „innere“ Facette des Problems der phänomenalen

Differenz aufgeklärt werden, indem hierdurch ein explanatorischer Ansatz dafür geschaffen wird zu verstehen, warum z.B. ein – in der Innenperspektive gegebenes – Rot-Erlebnis als mentaler Zustand und nicht als neuronaler Zustand repräsentiert wird. Es gibt somit keine plausiblen Gründe dafür anzunehmen, dass allein die – durch die unterschiedliche Wahrnehmungsperspektive bedingte – divergierende Repräsentation einer Entität eine Verschiedenheit der repräsentierten Entität begründen kann.

Es wird dennoch häufig kritisiert, dass Fechners Perspektivitätserklärung „bestenfalls“ beschreibenden Charakter hat (vgl. Pauen, 2007, S. 85). Denn wenn gilt, dass wir eine Theorie brauchen, die ein *echtes* Verständnis des Zusammenhangs von Gehirn und Geist liefert, mache Fechner mit seiner Unterscheidung zwischen einer „Innen- und der Außenperspektive [...] den Versuch, ein solches Verständnis zu erlangen. Es ist jedoch offensichtlich, dass Fechners einfaches Bild vom Wechsel zwischen der Innen- und der Außenperspektive nicht zu der erhofften Lösung führt: Das Bild liefert bestenfalls eine Beschreibung des Problems, nicht jedoch die geforderte Erklärung des Zusammenhangs von Gehirn und Bewusstsein. Fechner zeigt, *dass* hier irgendein Unterschied vorhanden sein muss, er erklärt jedoch nicht, *warum* genau *dieser* Unterschied auftritt, warum bewusste Prozesse also so erfahren werden, wie sie erfahren werden“ (ebd., Hervorh. im Orig.).

Auch wenn ich denke, dass diese Kritik Pauens überzogen ist, denn wie aufgezeigt wurde, liefert der Ansatz durchaus eine Erklärung der wahrgenommenen Gehirn-Geist-Differenz (oder doch zumindest einen wertvollen Anstoß zu einer solchen), sollte es vielversprechend sein, die Komponenten des Fechnerschen Ansatzes näher zu beleuchten, um auszuloten, ob sein explanatorisches Potential nicht noch stärker ausgeschöpft werden kann. Daher sollen im Folgenden die von Fechner (implizit) postu-

lierten Strukturmerkmale des Psychischen bzw. *Grundstrukturen der Erfahrung* (vgl. Breyer, 2012), Perspektivität und phänomenale Transparenz, in einem ersten Schritt näher betrachtet und in einem zweiten Schritt mit Rückgriff auf wahrnehmungs- und entwicklungspsychologische Gesichtspunkte fundiert werden.

3.4.2. Perspektivität und Transparenz als psychische Strukturmerkmale

Setzt man sich mit Perspektiven oder Perspektivität auseinander, muss man zwischen zwei zentralen Bedeutungsfacetten unterscheiden. Die erste besagt, dass wir die Welt und uns selbst immer von einem bestimmten Bezugspunkt, von einer Perspektive aus betrachten, es ist die so genannte *Erste-Person-Perspektive* als Ausgangspunkt oder Ort mentaler Repräsentationen. Die zweite – für diese Arbeit relevantere – Bedeutung von Perspektivität ist der Umstand, dass Objekte in Abhängigkeit von bestimmten räumlichen oder auch „innerlichen“ Bezugspunkten in ihrem Erscheinungsbild variieren, sich also der *Modus der Repräsentation* verändert.

3.4.2.1. Perspektivität als divergierende Personenperspektiven

Als eine Voraussetzung für die Möglichkeit einer verschiedenartigen Repräsentation von etwas muss es natürlich als konstituierend angesehen werden, dass es ein erlebendes System gibt, das einen „phänomenalen Standpunkt („point of view“) besitzt, der für Systeme mit diesem ‚point of view‘ eine spezifische Form von Innerlichkeit und subjektiver Erlebnisperspektive entstehen lässt“ (Heckmann & Esken, 1998, S. 17).

In der Philosophie hat sich vor allem Thomas Metzinger (1993a, 2000a, 2003, 2003a, 2006) im Rahmen seiner Selbst-Modell-Theorie der Subjek-

tivität³⁵ mit Perspektivität im Sinne der Existenz einer Erste-Person-Perspektive auseinandergesetzt. Zentral ist hierbei die Tatsache, dass Menschen³⁶ über die Fähigkeit verfügen, ihre Umwelt zu repräsentieren. Dabei beziehen sich Repräsentationen zunächst – wie bereits Spinoza festgestellt hat – auf die physische Struktur des repräsentierenden Systems, also einen komplexen neuronalen Aktivierungszustand im Gehirn. Die Theorie geht nun der Frage nach, welche Eigenschaften Repräsentationen in einem bestimmten System haben müssen, um zu *phänomenalen* Repräsentationen zu werden, deren Inhalt also zugleich Inhalt eines bewussten Erlebens ist. Nach Metzinger (2006) besteht „[d]ie phänomenale Gegenwart, das *Erscheinen* einer Welt [...], in der Aktivierung eines einzigen, kohärenten und globalen Modells der Realität [...], innerhalb eines virtuellen Gegenwartsfensters [...], das durch das System, welches es in sich erzeugt, introspektiv nicht *als* Modell erkannt werden kann“ (S. 427; Hervorh. im Orig.). Nach Metzingers Ansicht sollte Subjektivität epistemologisch als eine psychologische Eigenschaft komplexer informationsverarbeitender Systeme verstanden werden, die genau dann instantiiert wird, wenn das System in das von ihm aktivierte Modell der Realität ein Selbstmodell integriert. Da der Ansatz Metzingers recht komplex ist und das Thema der Subjektivität (im Sinne der Entstehung eines erlebten Selbstes) für die Zielstellung dieser Arbeit weniger relevant ist, werde ich auf eine weitergehende Darstellung verzichten.

Interessant ist der Ansatz Metzingers im Zusammenhang dieser Arbeit jedoch, weil er im Hinblick auf Perspektivität und phänomenale Transpa-

³⁵ Er stützt sich hierbei sehr stark auf die psychologische *Theorie mentaler Modelle* von Johnson-Laird (1983).

³⁶ Ich möchte mich im Folgenden der Einfachheit halber auf Menschen beziehen, weil sich die bisherige Erörterung auch auf menschliche Beispiele bezog, was nicht heißen soll, dass die Argumentation nicht ebenso für andere Lebewesen gilt.

renz sowohl wichtige begriffliche bzw. begriffsgeschichtliche Aspekte herausstellt als auch deren funktionale bzw. computationale Hintergründe beleuchtet. Bedingt durch sein Ziel der Formulierung einer Theorie zur Generierung von Subjektivität (in Form eines Selbstmodells) bezieht sich Perspektivität bei Metzinger vor allem auf das Phänomen der „Zentriertheit des repräsentationalen Raumes“ (vgl. ebd., 2006, S. 451) bzw. die Existenz der erlebten Erste-Person-Perspektive.

Perspektivität ist für Metzinger, neben Meinigkeit und Selbstheit, die entscheidende Zieleigenschaft einer repräsentationalistischen Analyse der phänomenalen Erste-Person-Perspektive (vgl. ebd., 2000a, S. 288). Sie lässt sich insofern als eine strukturelle Eigenschaft des phänomenalen Raumes als Ganzem bezeichnen und „besteht in der Existenz eines einzigen kohärenten und zeitlich stabilen Modells der Realität, das repräsentational auf ein einziges, kohärentes und zeitlich ausgedehntes phänomenales Subjekt zentriert ist“ (ebd., 2006, S. 450). Die Erfahrung von Perspektivität macht also im Kern aus, dass der uns erscheinende phänomenale Innenraum ein phänomenales Selbst als Zentrum, als Brennpunkt des Erlebens, aufweist. „Allen höherstufigen, sozial und begrifflich vermittelten Formen von Selbstbewusstsein scheint eine primitive und prä-reflexive Form phänomenalen Selbstbewusstseins zugrunde zu liegen [...] und in dieser nichtbegrifflichen Form der Selbstrepräsentation liegt der *Ursprung* der Perspektive der ersten Person“ (ebd., S. 451; Hervorh. im Orig.).

Die Zentriertheit unseres bewussten Modells der Realität sei eine funktionale Eigenschaft, die eine Entsprechung in der Zentriertheit unseres Verhaltensraumes hat. Hierbei handle es sich um eine funktionale Hintergrundbedingung, die aufgrund ihrer Alltäglich- bzw. Selbstverständlichkeit schnell übersehen würde: „Beim Menschen und allen uns bekannten bewussten Systemen sind sensorische und motorische Systeme

me *physisch* in den Körper eines einzigen Organismus integriert. Dies könnte man den ‚*single-embodiment constraint*‘ nennen, oder die *Bedingung der partikularen Realisierung*“ (ebd., Hervorh. im Orig.).

Metzinger nimmt an, dass das menschliche Selbstmodell und die mit ihm einhergehende Perspektive der ersten Person die einzige phänomenale Repräsentation ist, die „durch ein dauerhaftes funktionales Bindeglied direkt mit bestimmten Hirnregionen verbunden ist“ (ebd.). Nach der gegenwärtigen wissenschaftlichen Befundlage seien Kandidaten hierfür die Aktivität des Vestibularorgans, die als räumliche Matrix des Körperschemas fungiert, viszerale Formen der Selbstrepräsentation sowie die Inputs spezifische Kerne im oberen Hirnstamm, die das ‚interne Milieu‘ homöostatisch regulieren (vgl. ebd.; Parvizi & Damasio, 2001; Damasio, 1999, 2000). Durch diese Mechanismen wird ein hohes Maß an Invarianz und Stabilität generiert, indem „sie das System mit einer kontinuierlichen Quelle inneren, im Gehirn selbst erzeugten Inputs versorgen. Diese Inputquelle *verankert* das menschliche Selbstmodell“ (Metzinger, 2006, S. 451; Hervorh. im Orig.), indem dieses dauerhaft an den Organismus und seine elementaren Regulations- und Selbsterhaltungsprozesse geknüpft ist.

In Abgrenzung zur subjektiven Perspektive der ersten Person spricht man auch von einer *Perspektive der dritten Person*. Mit der Dritte-Person-Perspektive ist die Einnahme eines „äußeren“ oder objektiven Bezugspunktes gemeint. Wenn ich über die geistigen Zustände anderer Menschen sinniere oder hierzu wissenschaftliche Forschung mittels einer *empirisch-analytischen Methode* betreibe, sind mir diese Zustände selbst nicht gegeben, sondern nur indirekte Parameter (Verhalten, (Neuro-) Physiologie, theoretische Modelle etc.). Es ist die Diskrepanz zwischen der Perspektive der ersten und dritten Person, die dem Fechnerschen Ansatz der Identitätsansicht zugrunde liegt.

Es sei erwähnt, dass es nicht zuletzt aufgrund der Kritik am Kriterium der Objektivität der Dritte-Person-Perspektive, das einerseits nicht vollkommen erfüllbar und andererseits auch etwas Entscheidendes auszulassen scheint, in den letzten Jahren – vor allem im Zusammenhang der Erforschung der Entwicklung sozialer Kognitionen – zur Prägung des Begriffes der *Perspektive der zweiten Person* gekommen ist (vgl. z.B. Bohman, 2000; Tomasello, Carpenter, Call, Behne & Moll, 2005; Moll, 2007; Pauen, 2012). Gemeint ist hier die Möglichkeit zur Übernahme der Perspektive einer anderen Person, wie sie in der Entwicklungspsychologie schon länger untersucht wird (vgl. Dimitrova & Lüdmann, 2011, 2014). Es ist entgegen der „distanzierten“ Perspektive der dritten Person eine Art *teilnehmende* Beobachtung. Bedient man sich auch in der Wissenschaft dieser Methode, spricht man auch von einem phänomenologisch-hermeneutischen Vorgehen. Für die vorliegende Arbeit ist die Perspektive der zweiten Person allerdings weniger relevant.

3.4.2.2. *Perspektivität als divergierende Modi der Repräsentation*

Die Hauptbedeutung von Perspektivität liegt aber in dem Umstand, dass wir aufgrund der Tatsache, dass wir unwillkürlich stets einen subjektiven Bezugspunkt im Sinne einer Erste-Person-Perspektive einnehmen, keine unmittelbare und eingestaltige Repräsentation unserer Umwelt (und gewissermaßen auch von uns selbst) vornehmen können, sondern dass wir je nach Ausgestaltung dieses Bezugspunktes über ganz unterschiedliche Eindrücke, Erscheinungsbilder oder eben *divergierende Modi der Repräsentation* verfügen.

In der Philosophie hat der Aspekt der Perspektivität vor allem in der Phänomenologie Edmund Husserls eine profunde Bedeutung und zugleich eine sehr anschauliche Beschreibung erfahren:

Ein Ding ist mir wahrnehmungsmäßig gegeben als dauernd daseiendes in einer synthetisch einstimmigen Wahrnehmungskontinuität, in deren Abfluß das Ding kontinuierlich als das eine und selbe bewußt ist; aber in jeder Wahrnehmungsphase habe ich es in einem besonderen Wie seiner subjektiven Gegebenheitsweise, in wechselnder Orientierung des Nah und Fern, des Rechts und Links usw. Mit dem Wechsel der Orientierung geht im Ganzen Hand in Hand ein Wechsel der perspektivischen Erscheinungsweise; das Ding erscheint in immer neuen Aspekten. Immerfort scheidet sich dabei eigentlich Gesehenes und nur Mitgesehenes, relativ Selbsterfaßtes und nur antizipierend Miterfaßtes. (Husserl, 1976, S. 177f; zit. nach Breyer, 2012)

Husserl verdeutlicht einerseits den Umstand, dass Wahrnehmungsgegenstände zwar konstant in ihrer bloßen Existenz sein können, je nach Bezugspunkt aber ein anderes Erscheinungsbild, eine veränderte mentale Repräsentation hervorrufen können und dass andererseits die konkrete Erscheinungsweise immer ein Produkt ist aus dem, was tatsächlich über die Sinne³⁷ erfasst wird, und dem, was wir (aufgrund von Vorerfahrungen) zusätzlich in unseren komplexen Wahrnehmungseindruck „hineinlegen“. Wie sich im Folgenden noch zeigen wird, sind beide Aspekte höchst bedeutsam und im Hinblick auf unsere (intuitiv-dualistische) Auffassung der Gehirn-Geist-Relation folgenschwer.

Eine ins Auge stechende Tatsache ist außerdem die herausragende Bedeutung *visueller* Repräsentationen unserer Umwelt, denn in den Beschreibungen des Phänomens der Perspektivität geht es fast immer um visuelle Eindrücke. Hintergrund dieses Umstandes ist der sicher kaum zu leugnende Fakt, dass der Sehsinn für unsere Orientierung und

³⁷ Weitergehende Prinzipien des Ansatzes von Husserl wie die Methode der so genannten *transzendentalphänomenologischen Reduktion* (keine Unterscheidung zwischen „inneren“ und „außenweltbezogenen“ Wahrnehmungsgegenständen, Reduktion der Analyse auf den „Fluss der Phänomene“) sollen an dieser Stelle unberücksichtigt bleiben.

Anschauung von etwas (gegenüber allen anderen Sinnen) eine dominierende Stellung innehat.

Blickt man in der Philosophie weiter zurück findet sich der Perspektivitätsgedanke vor allem in grundlegenden (mathematisch-geometrischen) Bezügen bei Gottfried Wilhelm Leibniz (1710/1996) und – in einer stärker epistemologisch-wissenschaftstheoretisch bedeutsamen Hinsicht – bei Friedrich Nietzsche wieder. Leibniz hat herausgestellt, dass Perspektiven durchaus keine reinen Zufallsprodukte sind, sondern in vielen Fällen in einen systematischen Zusammenhang gebracht werden können. So gilt für den von der Perspektive des Beobachters abhängigen Zugriff auf geometrische Objekte durchaus ein strenger gesetzmäßiger Zusammenhang zwischen dem Bezugspunkt bzw. Blickwinkel desjenigen und der Repräsentation der dargestellten Figur:

Dass sich der Kreis in eine Ellipse verwandelt, wenn wir ihn vor unseren Augen rotieren lassen, ist eine durch Erfahrung bestätigte Tatsache; die durch Rotation verursachte perspektivische Verschiebung gehorcht den strengen Gesetzmäßigkeiten der Optik, so dass sie sich auch leicht simulieren und quantifizieren lässt. An der geometrischen Ordnung perspektivischer Erscheinungsverläufe liegt es auch, dass es für Leibniz keine Zufälligkeit in der Verknüpfung von Gegenstand und Repräsentation gibt. Die Verbindung beider ist eine strenge und unbedingte. (Breyer, 2012)

Entgegen einer Vielzahl von Autoren, die in der offengelegten Subjektabhängigkeit der Erkenntnis einen Rückschritt in der Epistemologie sehen, ist für Nietzsche jedes Sehen oder Erkennen immer ein solches unter Maßgabe einer bestimmten Perspektive. Für ihn gibt es überhaupt „*nur* ein perspektivisches Sehen, *nur* ein perspektivisches Erkennen; und *je mehr* Affekte wir über eine Sache zu Wort kommen lassen, je mehr Augen, verschiedene Augen wir uns für dieselbe Sache einzusetzen wissen, um so vollständiger wird unser 'Begriff' dieser Sache, unsere 'Objek-

tivität' sein" (Nietzsche 1887, S. 383; Hervorh. im Orig.). Statt einer Gefahr für eine objektive Erkenntnisgewinnung sieht Nietzsche also in der Akzeptanz unserer immer perspektivisch verfassten Zugriffsweise auf die Welt eine große Bereicherung. Statt eine sowieso prinzipiell unerreichbare (wissenschaftliche) Objektivität anzustreben, sollten wir den Fakt einer subjektabhängigen Wahrnehmung und Erkenntnis zunächst akzeptieren, um dann über den Austausch über und die Integration von verschiedene(n) Perspektiven auf ein Phänomen auf eine größtmögliche *Intersubjektivität* hinzuwirken (vgl. Breyer, 2012).

Das, was Nietzsche hier für die Forschungsobjekte von Wissenschaftlern beschreibt, gilt natürlich auch für unsere alltägliche Auseinandersetzung mit unserer Umwelt und auch uns selbst. Wie Husserl in dem obigen Zitat verdeutlicht hat, ist Perspektivität ein Alltagsphänomen. Je nach unserem räumlichen Standpunkt und auch unserer inneren Einstellung und Vorerfahrung werfen wir einen anderen Blick auf unsere Umwelt oder auch uns selbst. Ein Objekt verändert sein Aussehen je nach Blickwinkel, Lichtverhältnissen oder – gerade wenn es sich um einen Menschen handelt – auch in Abhängigkeit vom Grad der Vertrautheit. Hier kommt das „antizipierend Miterfasste“ Husserls zum Tragen, denn wenn ich einen guten Freund oder Verwandten sehe, schwingen in meiner Wahrnehmung zugleich Attribute mit, die weit über den rein über die Sinne vermittelten Eindruck hinausgehen. Und so verhält es sich auch, wenn ich selbst im Fokus meiner Aufmerksamkeit stehe. Auch der Blick auf mich selbst ist insofern immer perspektivisch, als dass auch hier – je nach kontextueller, gedanklicher oder emotionaler Einbettung eines mentalen Zustandes – ein anderer Bezug zu mir selbst hergestellt wird. Mal stehe ich beispielsweise im völligen Einklang zu einem bestimmten Einfall (ich sollte Fallschirm springen), weil es aufgrund des vorherrschenden Kontextes (im Kreis der Freunde, bei denen es viele Fallschirm-

springer gibt; feucht-fröhliche Stimmung in einer Männerrunde...) angemessen erscheint; mal würde ich mich von dem gleichen Einfall innerlich vielleicht schnell wieder distanzieren, weil der situative Kontext ein anderer ist (im Kreise der Familie, Sicherheits- und Kostenfaktoren rücken in den Mittelpunkt). Das Beispiel soll verdeutlichen, dass man sich selbst ganz unterschiedlich wahrnehmen oder konzeptualisieren kann: mal sehr stark und risikofreudig, mal eher verletzbar und sicherheitsorientiert. Es ist eine andere – in der Regel durch äußere Faktoren beeinflusste – Perspektive, Selbstwahrnehmung oder Selbstkonzeption, die wir vornehmen können.

3.4.2.3. Phänomenale Transparenz und der naive Realismus

Wie Metzinger aufzeigt, geht der Begriff der phänomenalen Transparenz wahrscheinlich auf George Edward Moore (1903) zurück:

[T]he fact that when we refer to introspection and try to discover what the sensation of blue is, it is very easy to suppose that we have before us only a single term. The term 'blue' is easy enough to distinguish, but the other element which I have called 'consciousness' – that which a sensation of blue has in common with a sensation of green – is extremely difficult to fix. [...] And in general, that which makes the sensation of blue a mental fact seems to escape us; it seems, if I may use a metaphor, to be *transparent* – we look through it and see nothing but the blue; we may be convinced that there is something, but what it is no philosopher, I think, has yet clearly recognised. (S. 446; zit. nach Metzinger, 2003a, S. 413; Hervorh.: M.L.)

Und genau für die Existenz dieser besonderen Form der Abwesenheit eines bestimmten Wissens (Metzinger spricht auch von „Dunkelheit“) ist das psychische Strukturmerkmal der *phänomenalen Transparenz* verantwortlich. Es bedeutet, „dass etwas Bestimmtes der subjektiven Erfahrung selbst nicht zugänglich ist, nämlich die repräsentationale Natur des In-

halts der bewussten Erfahrung“ (Metzinger, 2006, S. 426). Der transparente Charakter phänomenaler Zustände verwehrt uns einen introspektiven Zugriff auf deren frühe Verarbeitungsstufen im Gehirn. Wir können die Mittel der Repräsentation nicht selbst als solche repräsentieren. Es ist das Gehirn bzw. ein Gehirnprozess, der als Träger unseres phänomenalen Bewusstseins, unserer inneren Welt, fungiert. Wenn ich ein Buch in den Händen halte, wird der hiermit verbundene Gehirnprozess, „der in keinem konkreten Sinn etwas ‚Buchhaftes‘ besitzt, [...] von mir nicht bewusst erlebt, er ist in dem Sinne transparent, dass ich durch ihn hindurch sehe. Das, *worauf* ich blicke, ist der repräsentationale Inhalt, die Existenz eines Buchs, hier und jetzt, so, wie es mir durch meine Sinnesorgane gegeben ist. Dieser Inhalt ist also eine abstrakte Eigenschaft des konkreten repräsentationalen Zustandes in meinem Gehirn“ (ebd., S. 444; Hervorh. im Orig.).

Wir verfügen daher über eine Innenperspektive, für deren repräsentationalen Charakter wir unter normalen Umständen blind sind. Der Mensch ist – zumindest in seinem alltäglichen Agieren – immer ein *naiver Realist*, da er stets so denkt und agiert, als wäre er mit der Realität selbst und nicht ihrer Repräsentation konfrontiert.

Diesen naiven Realismus, der mit transparentem phänomenalen Erleben zwangsläufig einherzugehen scheint, hat bereits Wittgenstein (1953) in § 275 seinen *Philosophischen Untersuchungen* sehr anschaulich beschrieben:

Schau auf das Blau des Himmels, und sag zu dir selbst „Wie blau der Himmel ist!“ – Wenn du es spontan tust – nicht mit philosophischen Absichten – so kommt es dir nicht in den Sinn, diesen Farbeindruck gehöre nur dir. Und du hast kein Bedenken, diesen Ausruf an einen Andern zu richten. Und wenn du bei den Worten auf etwas zeigst, so ist es der Himmel. Ich meine: Du hast nicht das Gefühl des In-dich-selber-Zeigens, das oft das ‚Benennen der Empfindung‘ begleitet, wenn man über die ‚private Sprache‘ nachdenkt. Du denkst auch nicht, du solltest eigent-

lich nicht mit der Hand, sondern nur mit der Aufmerksamkeit auf die Farbe zeigen.
(zit. nach Metzinger, 2003a, S. 413)

Der naive Realismus entsteht, weil sich das System aufgrund der phänomenalen Transparenz zwangsläufig selbst so versteht, als dass es einen unmittelbaren Kontakt zu den Gegenständen oder Inhalten seines Erlebens hat. Dem System bleibt zumeist verborgen, dass jede bewusste Erfahrung in einem *Medium* abläuft (vgl. Metzinger, 2003a, S. 419).

Phänomenale Transparenz ist kein Alles-oder-Nichts-Prinzip. Es gibt durchaus verschiedene Grade der (Un-)Durchsichtigkeit, wobei gilt, dass der Grad der Transparenz eines mentalen Zustandes umgekehrt proportional zum Grad der attentionalen Verfügbarkeit seiner früheren Verarbeitungsstufen ist³⁸ (vgl. ebd., S. 417). Ein Beispiel für mentale Prozesse, die normalerweise vollständig phänomenal transparent sind, sind unsere alltäglichen Sinneswahrnehmungen. Als Beispiele für Zustände, die zumindest in kurzen Phasen in dem Sinne intransparent bzw. *opak* erscheinen, als dass uns selbst der repräsentationale Charakter unseres Erlebens bewusst wird, nennt Metzinger die Übergangsphasen bei bistabilen Wahrnehmungsobjekten wie dem Neckerschen Würfel (vgl. Abb. 6), wenn man von der einen möglichen (räumlichen) Interpretation zur anderen springt (vgl. ebd., S. 420), oder auch die Phänomene, die sich aufgrund einer durch divergierende nicht fusionierbare Eindrücke auf korrespondierenden Netzhautstellen applizierten *binokularen Rivalität* ergeben (vgl. z.B. Alais & Blake, 2005; Blake & Logothetis, 2002).

³⁸ Metzinger möchte sich hiermit von einer häufig verwendeten Bestimmung von Transparenz abgrenzen, die besagt: Nur Gehalteigenschaften, nicht aber die Vehikeleigenschaften des phänomenal repräsentierten Gegenstandes sind für das System erfassbar. Diese Bestimmung hätte implizit Gedankengut des kartesischen Dualismus in sich, da sie eine gewisse (ontologische) Verdinglichung von Inhalt und Vehikel vornimmt (vgl. ebd., 2003a, S. 417).

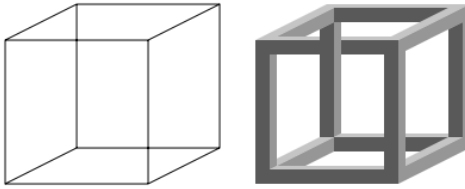


Abb. 6: Neckerscher Würfel

Ein anderes – bei näherer Betrachtung sehr aufschlussreiches – Beispiel sind Pseudohalluzinationen, wie sie unter dem Einfluss halluzinogener Substanzen wie LSD oder Psylocybin auftreten. So berichten fast alle Probanden oder „User“ davon, dass sie zumindest in bestimmten Phasen des drogeninduzierten Erlebens quasi „atmende“ geometrische Muster gesehen haben, die sich in einem vielgestaltigen Farb- und Formenspiel vor ihren Augen bewegt haben (vgl. Siegel & Jarvik, 1975). Für gewöhnlich sind sie sich dabei des lediglich repräsentationalen bzw. halluzinatorischen Charakters ihrer diesbezüglichen visuellen Erfahrung bewusst, sodass wir im Fall solcher Pseudohalluzinationen Zustände vorliegen haben, die einen geringeren Grad an phänomenaler Transparenz aufweisen (vgl. Metzinger, 2003a, S. 421). Aber in diesem Fall ist es nicht nur die bloße Vergegenwärtigung einer konstruierten Wirklichkeit, die sich ereignet, sondern es gibt Hinweise darauf, dass dieses Erleben der sich bewegenden, pulsierenden geometrischen Formen tatsächlich die früheren bzw. tieferen Verarbeitungsstufen selbst sind, für die wir normalerweise „blind“ sind.

So geht Ronald K. Siegel (1992/1998), der wohl weltweit renommierteste Halluzinationsforscher, davon aus, dass diese geometrischen Formen

...unsichtbaren [d.h. für gewöhnlich *phänomenal transparenten*, M.L.] Strukturen innerhalb des visuellen Systems [entstammen], die unter der Einwirkung der Drogen sozusagen illuminiert werden. Tanzende Punkte entstehen beispielsweise,

wenn rote Blutkörperchen durch Netzhautkapillaren fließen und dabei einen Schatten auf die darunterliegenden Stäbchen- und Zapfenzellen werfen. Andere Formen werden erzeugt, wenn Drogen bewirken, daß sich Neutronen in den Sehnerv oder die Netzhaut entladen. Dies ruft Ketten heller Lichter hervor, die Phosphene genannt werden. Phosphene können alle möglichen Formen annehmen: Flecken, konzentrische Kreise, Spiralen, Tunnel, Gitter, sogar Schachbrettmuster. Wieder andere Formen werden vom Sehzentrum des Gehirns erzeugt, wo die Stimulierung organisierter Zellengruppen sich wiederholende Polygone, Mosaik und symmetrische Anordnungen produziert. (S. 23f)

Viele der aufgeführten Zusammenhänge können mittlerweile als gesichert betrachtet werden. So gibt es mathematische Modelle, die die selbstorganisierende Dynamik im visuellen Kortex (vor allem V1) in solchen instabilen Zuständen nachbilden, die Rückschlüsse auf funktionale Architekturprinzipien des visuellen Kortex zulassen (vgl. Bressloff, Cowan, Golubitsky, Thomas & Wiener, 2001). Mit ihrer Hilfe können präzise Voraussagen über phänomenologische Formkonstanten gemacht werden, die von den Probanden in solchen Zuständen berichtet werden. Auch Metzinger ist daher überzeugt, dass durch den Einfluss halluzinogener Substanzen die frühen Verarbeitungsstufen des visuellen Systems zum Gegenstand der introspektiven Aufmerksamkeit werden.

Die atmenden Muster *sind* diese frühen Verarbeitungsstufen. Die funktionale Architektur des Gehirns ist vorübergehend verändert worden. Das pharmakologisch perturbierte System versucht nun verzweifelt wieder in einen stabilen Zustand zu gelangen, und es ist dieser Vorgang der dynamischen Selbstorganisation, den der Halluzinierende erstmals phänomenal – aber *subsymbolisch*, also *nicht-propositional* – repräsentieren kann. [...] In diesen Fällen erleben wir uns als Wesen, die absichtlich abstrakte mentale Repräsentationen konstruieren und mit ihnen operieren. Weil wir erleben, dass sie sich jederzeit als falsch erweisen können, erkennen wir sie *als* Repräsentationen (Metzinger, 2003a, S. 421; Hervorh. im Orig.).

Wie kommt es dann aber überhaupt dazu, dass unsere mentalen Zustände nicht immer einen gewissen opaken Charakter haben? Was ist die Ursache der phänomenalen Transparenz, die die meisten Erlebniszustände kennzeichnet? Bereits Robert van Gulick (1988) hat in einer funktionalistischen Analyse des Begriffes der phänomenalen Transparenz aufgezeigt, dass ihr Aufkommen mit der Zuverlässigkeit, Geschwindigkeit und globalen Interdependenz der repräsentationalen Träger zusammenhängen dürfte:

How can the functionalist account for this subjectively experienced transparency? Possessing information always involves a capacity for appropriate behavior. Being informed about the content of a representation is being able to relate it to other representations and to items in the world in ways appropriate to its content. As long as understanding or being informed is analyzed as a behavioral capacity, even if the relevant behavior is all internal, the functionalist can hope to fit it within his theory. Thus the functionalist should resist any view of phenomenal transparency as a form of nonbehavioral self-luminous understanding. He can undercut the intuitive appeal of that view, by explaining the subjective experience of understanding in terms of smooth and seemingly automatic transitions among representations. The internal component of understanding need involve nothing beyond an ability to interrelate many diverse representations with great speed. For example, understanding the content of a visual representation would require being able to connect it appropriately and rapidly with other visual, non-visual, and nonperceptual representations. [...] *How* this is done is not something to which I have linguistic or introspective access, but there must be powerful processors to produce these seemingly instantaneous transitions. (ebd., S. 178f, Hervorh. im Orig.)

Weil die meisten mentalen Modelle also sehr schnell und zuverlässig aktiviert werden, sodass unser Erleben „stolperfrei“ einfach dahin fließen kann (wie man auch von einem *Bewusstseinsstrom* spricht), bleibt uns der repräsentationale Charakter verborgen, sodass wir durch diese funktionalen bzw. neuronalen Strukturen quasi „hindurchschauen“. Natürlich kann dies aber nur ein erster Ansatz zur Erklärung der Ursache phänomenaler Transparenz sein, denn man hat zwar Bedingungen formuliert,

die es begünstigen dürften, dass wir für gewöhnlich den repräsentationalen Charakter unseres Erlebens ausblenden, gezeigt werden muss aber, wie *genau* der Übergang von der (erlebnismäßig) opaken funktionalen oder neuronalen Ebene zur Ebene des bewussten Erlebens vonstatten geht. Es bleibt also die Frage im Raum, warum wir in unserer Wahrnehmung nicht die Neurone oder die von ihnen realisierten funktionalen Prinzipien vor uns haben, sondern einen repräsentationalen Inhalt in einer bestimmten Gegebenheitsweise.

Die phänomenale Transparenz ist aber nun kein Nachteil, sondern ein klarer Vorteil für die Systeme, die über sie verfügen. Jedes biologische System verfügt nur über begrenzte zeitliche und neuronale Ressourcen und kann durch das Ausblenden von Informationen früherer Verarbeitungsstufen die zu bewältigende computationale Last verringern. Metzinger (2006) führt dies unter Heranziehung einer Computeranalogie in sehr anschaulicher Weise folgendermaßen aus:

Unsere repräsentationale Architektur erlaubt auf der personalen Ebene nur einen sehr begrenzten introspektiven Zugang zu der tatsächlichen Dynamik aus Myriaden von individuellen neuronalen Einzelereignissen, aus denen am Ende unsere phänomenale Welt scheinbar anstrengungslos auftaucht. Phänomenales Bewusstsein funktioniert wie eine evolutionär entstandene Benutzeroberfläche, ein multimodales Interface, das es dem Organismus erlaubt, globale Eigenschaften der eigenen Informationsverarbeitung schnell und zuverlässig zu erfassen. Transparenz ist die Geschlossenheit³⁹ dieser Benutzeroberfläche und sie erzeugt den naiven Realismus der phänomenalen Ebene. (S. 445)

³⁹ Mit Geschlossenheit oder „autoepistemischer Geschlossenheit“ meint Metzinger die epistemische Begrenztheit der attentionalen Verarbeitung unserer eigenen und internen repräsentationalen Dynamik, die keineswegs unter Berücksichtigung des philosophischen Ideals der Selbsterkenntnis konstruiert ist (vgl. Metzinger, 2006, S. 445).

Wäre der repräsentationale Charakter mentaler Zustände dem System selbst stets gegeben, bestünde die Gefahr, dass ein solches System in einem zu starken Maße eine Auseinandersetzung mit seinen internen Mechanismen praktizieren würde, sodass es nicht mehr in einem adäquaten Austausch mit seiner Umwelt stünde (vgl. ebd.). Ein solches System dürfte im evolutionären Überlebenskampf wohl kaum eine Chance haben.

3.4.3. Eine wahrnehmungspsychologische Analyse von Multiperspektivität

Eine weitergehende psychologische Auseinandersetzung mit Perspektivität (sowie zumindest implizit auch mit phänomenaler Transparenz) und ihren Ursprüngen in der konkreten Beschaffenheit unseres Wahrnehmungssystems hat vor allem Rainer Mausfeld (2006, 2007, 2010a) vorgenommen. Diese soll zunächst im Allgemeinen vorgestellt werden, bevor ein dezidierter Rückbezug auf die Gehirn-Geist-Problematik erfolgt (den Mausfeld selbst nicht bzw. kaum vornimmt).

Nach Mausfeld (2006) ist die Fähigkeit zur Multiperspektivität ein „funktionales Designprinzip des menschlichen Geistes“ (S. 1.), das uns so sehr vertraut und mit uns verwoben ist, dass es sich gerade deshalb unserem Blick und Nachsinnen über die Natur unseres Geistes schnell entzieht. Insofern liegen wir mit unserer Alltagskonzeption falsch, dass Wahrnehmung im Sinne einer Widerspiegelung der Außenwelt funktioniert und die Welt selbst im Wesentlichen so aufgebaut ist, wie wir sie wahrnehmen bzw. dass unsere Wahrnehmungskategorien den Kategorien der Außenwelt entsprechen (vgl. ebd., S. 5). Wenn Mausfeld diesbezüglich ausführt, dass in Wirklichkeit unser Gehirn die Funktionsweise unserer Wahrnehmung fast vollständig abschottet und nur ihr Endprodukt in einigen

Aspekten bewusst zugänglich ist, bezieht er sich – obgleich Mausfeld den Begriff nie explizit verwendet – auf den Aspekt einer *phänomenalen Transparenz* gegenüber den funktionalen (und neuronalen) Grundlagen der Wahrnehmung. Es ist vor allem die Diskrepanz zwischen dem Input, der in unseren Organismus eingeht, und dem Output als unserem Wahrnehmungseindruck, der zumeist eine ganze Reihe von bedeutungshaltigen Kategorien mit umfasst und weit über den konkreten Input bzw. die Menge der potentiell verfügbaren Informationen der Außenwelt hinausgeht, der für uns phänomenal nicht einsichtig ist: „Daß wir dieses gewaltige Mißverhältnis von Input und Output in der Regel nicht bemerken und zutiefst überzeugt sind, daß es natürliche Kategorien der Welt sind, die sich im Perzept widerspiegeln – diese Zuschreibung von ‚Realität‘ zu unserem Perzept ist gerade eine der großen *Leistungen* unseres Gehirns“ (ebd., S. 12).

Er verdeutlicht dies anhand einer Vielzahl von Phänomenen aus der Wahrnehmungs- und Gestaltpsychologie (siehe Abb. 7).



Abb. 7: Konzeptform „3D-Objekt“ (Aus: Mausfeld, 2006, S. 8); Kippbild „Vase vs. Gesichter“

Obwohl es in der Abbildung selbst nicht vorhanden ist, sehen wir bei ihrem Anblick die Konzeptform eines dreidimensionalen Objektes, eines weißen Würfels vor dem Hintergrund einiger Kreise, die durch den Würfel teilweise verdeckt werden (linke Seite). Die verschiedene Möglichkeit der Auffassung ein und desselben Gegenstandes wird nach Mausfeld dadurch verdeutlicht, dass wir wie im Falle des auf der rechten Seite in Abbildung 7 dargestellten Kippbildes mühelos „zwischen diesen Ebenen, diesen Schichten von Repräsentationen von ‚Realität‘ geistig hin und herzuwandern [in der Lage sind]“ (ebd., S. 2).

Betrachtet man die breite Spanne an möglichen Phänomenen der Multiperspektivität (die von einer dualen Bildwahrnehmung, über den möglichen „Perspektivenwechsel“ bei der Wahrnehmung und Beurteilung anderer Personen bis hin zur Attribution vs. Nicht-Attribution von Eigenschaften auf unbelebte Gegenstände etc. reicht) gehen diese auf verschiedene Aspekte der funktionalen und computationalen Architektur unseres Geistes zurück, wobei es komplexe Interaktionen zwischen verschiedenen Subsystemen wie dem Wahrnehmungs- und Imaginationssystem oder auch höheren interpretativen bzw. sprachlichen Systemen gibt. Allerdings geht Mausfeld davon aus, dass für eine bestimmte Klasse von multiperspektivischen Phänomenen gilt, dass sie ausschließlich mit strukturellen und funktionalen Eigenschaften des Wahrnehmungssystems selbst zusammenhängen (vgl. ebd., 2010a, S. 97).

Aufgrund der (potentiellen) multiperspektivischen Gegebenheitsweise von Objekten leitet Mausfeld ab, dass ein Wahrnehmungsgegenstand keinesfalls auf seine konkrete Erscheinungsweise zu einem bestimmten Zeitpunkt reduziert werden kann:

Sobald etwas als Objekt im Raume wahrgenommen wird, *muß* es sogleich die Menge seiner potentiellen Erscheinungsweisen unter bestimmten Transformationen,

etwa Änderungen der Betrachter-Objekt-Relation, als intrinsischen Bestandteil seiner internen Codierung mit beinhalten. Insbesondere ist also notwendigerweise in seine Datenstruktur eine ‚optische Multiperspektivität‘ mit eingebaut, weil es andernfalls kein dreidimensionales Wahrnehmungsobjekt im Raume sein könnte. (ebd., S. 3, Hervorh. im Orig.)

Außerdem ist er überzeugt, dass sich in Fällen wie des 3D-Objektes in Abbildung 7 das Phänomen der Multiperspektivität unmittelbar aus der funktionalen Architektur unseres Wahrnehmungssystems ergibt und nicht erst auf höherstufigen interpretativen Ebenen einsetzt. Dass wir den weißen Würfel wahrnehmen, liegt demnach an der Tatsache, dass wir in unserem Wahrnehmungssystem entsprechende konzeptuelle Formen implementiert haben, die durch entsprechende Reize ausgelöst werden „und zwar sogar unter hochgradig reduzierten Bedingungen, bei denen sich ein solches Objekt gleichsam aus dem Nichts konstituiert, nämlich aus dem, was nicht im Reiz vorhanden ist. Dies ist wiederum kein kognitiver Akt⁴⁰, sondern Konsequenz der ‚internen Logik‘ des Wahrnehmungssystems“ (ebd.), bei der der zweidimensionale Reiz eine im Wahrnehmungssystem implementierte Konzeptform auslöst.

Entgegen dem alltagspsychologischen Standardmodell der Wahrnehmung, das nach Ansicht Mausfelds selbst noch in weiten Teilen der Wissenschaft verbreitet ist und – naiv-realistisch – annimmt, dass wir mit unseren Wahrnehmungsinhalten zugleich die Objekte der Außenwelt wahrnehmen, wie sie *an sich* sind, vertritt Mausfeld ein funktionales Modell der Architektur des Wahrnehmungssystems, das auch – entgegen

⁴⁰ Ob man hier von kognitiven Vorgängen spricht oder wie Mausfeld von einer internen Logik des Wahrnehmungssystems selbst, ist im Kontext dieser Arbeit relativ unerheblich. Meines Erachtens spricht nichts dagegen, den Begriff der Kognition so weit zu fassen, dass er auch solche Aspekte mit umfassen kann.

dem „Standardmodell“ – dem Phänomen der Multiperspektivität gerecht werden kann (vgl. ebd., 2010a, S. 101f).

Das in Abbildung 8 dargestellte Modell nimmt an, dass das *sensorische System* mit computationalen Prozessen betraut ist, die in physikalisch-geometrischer Sprache beschrieben werden können (also z.B. mittels Linien, Gradienten, Strukturen, Konturen oder Leuchtkraft/ Helligkeit).

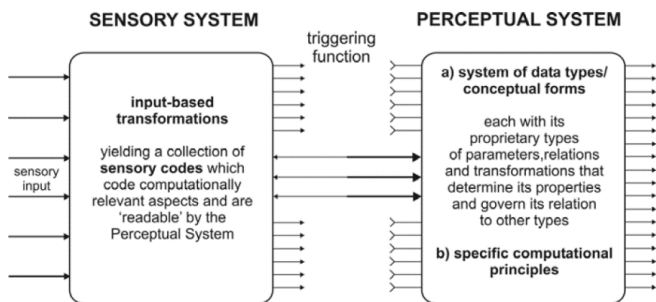


Abb. 8: Funktionale Architektur des Wahrnehmungssystems (Aus: Mausfeld, 2010a, S. 103)

In diesem System findet eine Transduktion physikalischer Energie in einen neuronalen Code statt, der anschließend in ein Format weitertransformiert wird, das den strukturellen und computationalen Anforderungen des perzeptuellen Systems gerecht wird (und so für dieses „lesbar“ wird). Das *perzeptuelle System* beinhaltet dagegen einen umfangreichen Satz an komplexen konzeptuellen Formen, die nicht mittels der Struktureinheiten des sensorischen Systems beschrieben bzw. auf diese reduziert werden können. Hierzu gehören konzeptuelle Formen wie „Oberfläche“, „Essen“, „Feind“, „Gerät“, „Kausalbeziehung“ usw. Sie sind die Grundlage für unsere perzeptuelle Ontologie und konstituieren eine Art „Reich“ perzeptueller Objekte. Erst nachdem die Wahrnehmungsobjekte durch das Raster (bzw. die Anreicherung) der konzeptuellen Formen des per-

zeptuellen Systems gegangen sind, werden die transportierten Inhalte durch höhere interpretative Systeme (wie Systeme für Überzeugungen oder nicht-perzeptuelles Weltwissen) oder Vorstellungssysteme ausgewertet (vgl. ebd., 2010a, S. 104). Mausfeld verweist hierbei auf Chomsky (1981), der solche konzeptuellen Formen als integrierten Komplex in unserem Geist bezeichnet hat, eine Domäne potentieller Fakten zur mentalen Konstruktion unserer Welt. Für Mausfeld liegt diese Domäne aber zu weiten Teilen nicht im Bereich höherer kognitiver Systeme wie dem der Sprache, sondern bereits in der Wahrnehmung selbst. Wahrnehmung in komplexen Systemen basiert in diesem Sinne „on an internal computational organisation in terms of abstract data types or conceptual forms, and that the sensory input serves as a kind of sign for the activation of these biologically given conceptual forms, which constitute the data formats of the computational processes involved“ (Mausfeld, 2010a, S. 103).

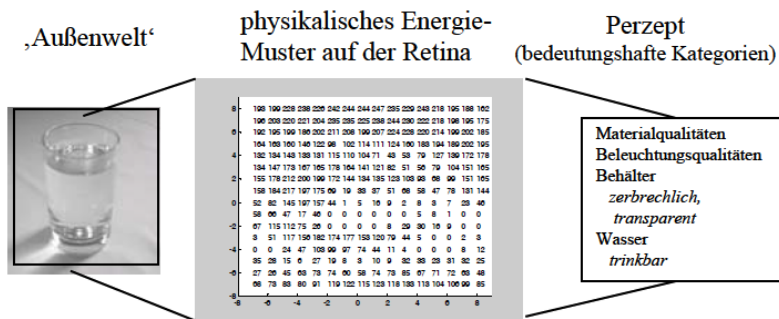


Abb. 9: Explanatorische Lücke der Wahrnehmungsforschung (Aus: Mausfeld, 2006, S. 11)

Wie Abbildung 9 zeigt, löst das Objekt der Außenwelt, ein Glas, ein bestimmtes physikalisches, mathematisch beschreibbares Energiemuster auf unserer Retina aus. Aus diesen Mustern geht dann in höheren Verar-

beitungsstufen und unter Heranziehung der im Wahrnehmungssystem implementierten konzeptuellen Formen das eigentliche Perzept hervor. Dieses beinhaltet neben den durch den tatsächlichen Reizinput bereitgestellten Informationen eine Reihe weiterer bedeutungshaltiger Kategorien wie z.B. die Zerbrechlichkeit des Glases oder den Eindruck der Trinkbarkeit des im Glas enthaltenen Wassers. Wie genau aber nun dieser Übergang zwischen dem bloßen physikalischen Muster hin zur Erschaffung unserer reichhaltigen bewussten Innenwelt funktioniert, ist in weiten Teilen noch unklar (man spricht daher auch von der *explanatorischen Lücke* als Fundamentalproblem der Wahrnehmungsforschung, vgl. ebd., 2006, S. 11f).

Eine evidente Schlussfolgerung ist daher, dass unsere Wahrnehmung „zwar reizvermittelt, aber nicht mehr reizgebunden [ist]“ (ebd., 2006, S. 12; Hervorh. im Orig.), wobei in dieser Dissoziation zwischen dem konkreten Sinnesinput und der bedeutungshaltigen Wahrnehmung ein evolutionärer Vorteil im Sinne eines Zugewinns an Flexibilität liegen dürfte. Die Umwelt liefert nur eine gewisse Anregung zu einem komplexen internen (mental)en Geschehen, welches keineswegs durch die tatsächliche Beschaffenheit des externen Gegenstandes bereits vollkommen festgelegt ist (vgl. ebd.).

Vielmehr funktioniert der Reiz gleichsam als Stichwortgeber für die *Selektion* einer bereits im Gehirn vorhandenen und durch die Evolutionsgeschichte kodeterminierten ‚Außenwelt-Geschichte‘. [...] Wir können gar nicht anders, als die ‚Außenwelt‘ durch die Brille der uns verfügbaren Grundkonzepte wahrzunehmen, die für uns eine *universelle Form der Welterfahrung* darstellen. In unserem metaphorischen Bild können wir also sagen, daß der Reiz lediglich als Stichwortgeber dient für die Aktivierung biologisch vorgegebener Konzeptstrukturen. [...] Der sensorische Input ist – in der schönen Formulierung von Humboldt – *nur ein Reiz für das Hervorbrechen des Innern*. (ebd., S. 12f; Hervorh. im Orig.)

Den evolutionsbiologischen Vorteil erläutert Mausfeld anhand des Vergleichs zu bestimmten Mikroorganismen. So findet man bei den Flagellen von Kolibakterien eine evolutionär äußerst einfache, vielleicht überhaupt die einfachste Form eines Wahrnehmungssystems. Hier sind Rezeptoren und Effektoren durch einen simplen Signalpfad miteinander verbunden, ohne dass es auf diesem Weg dazwischengeschaltete modulatorische (Teil-)Systeme gibt. Damit ist aber auch „der evolutionäre Spielraum für die Rezeptoren durch die Arten der verfügbaren physikalisch-chemischen Energieformen begrenzt“ (ebd., S. 13).

Anders verhält es sich bei höherentwickelten Organismen, die zwischen der Sensorik und Motorik ein neuronales Substrat höherstufiger modulierender Funktionen installieren konnten und damit einen nahezu unerschöpflichen evolutionären Spielraum erhalten haben. Entgegen der eher begrenzten strukturell-funktionalen Veränderbarkeit von Sensoren und Effektoren können sich im dazwischen geschalteten neuronalen Substrat funktional spezialisierte Teilsysteme ausdifferenzieren, die eigene Konzeptstrukturen implementiert haben. „So, wie sich die Wahrnehmungsobjekte von Auge und Ohr unterscheiden, jeder Sinn gleichsam seine eigene Sprache spricht, so hat auch jedes dieser Teilsysteme seine eigene Evolutionsgeschichte und seine idiosynkratische Weise, den sensorischen Input in Kategorien zu segmentieren. Jedes System hat gleichsam seine eigene Konzeptstruktur“ (ebd., S. 15).

Allerdings ergibt sich mit zunehmender Ausdifferenzierung hochspezialisierter computationaler Subsysteme eine zunehmende Anzahl an Schnittstellenproblemen zwischen den bestehenden Teilsystemen, was Mausfeld anhand eines interessanten Beispiels aus der klassischen ethologischen Forschung erläutert. So hat sich aufgrund von Untersuchungen von Tinbergen (1951) gezeigt, dass sich im Laufe der Evolution bei Graugänsen offenbar zwei modulare Teilsysteme in Bezug auf die

Wahrnehmung von Eiern herausgebildet haben, die nicht inferentiell integriert werden können. In System 1 löst jeder annäherungsweise konvexe Gegenstand mit runden Kanten ein Rollverhalten aus (neben Eiern z.B. auch Bierdosen). Es dient der Identifikation von Eiern, die sich außerhalb des Nestes befinden und von der Graugans mit dem Schnabel wieder zurück ins Nest gerollt werden. System 2 hat dagegen zur Aufgabe, alle im Nest befindlichen Objekte, die keine Eier sind, aus dem Nest zu entfernen. Das hier implementierte „Ei-Konzept“ ist dabei viel restriktiver als das in System 1 und „reagiert auf alles, was nicht weitgehend der geometrischen Form und Größe der eigenen Eier entspricht. Beide Systeme operieren gleichsam auf der Grundlage unterschiedlicher Konzeptionen von ‚Ei‘“ (Mausfeld, 2006, S. 14). Aufgrund der für die Gans nicht vorhandenen Möglichkeit zur inferentiellen Integration beider Teilsysteme kommt es mitunter zu eigentümlichen Verhaltensabläufen: Wenn die Gans z.B. eine Bierdose außerhalb des Nestes findet, wird sie diese aufgrund des aktivierten Systems 1 ins Nest rollen, um sie daraufhin durch das Einsetzen der Aktivierung des Systems 2 als „Nicht-Ei“ zu identifizieren und aus dem Nest zu befördern usw. (vgl. ebd.).

Hier zeigt sich also, dass es bei einer Zunahme unabhängiger Teilsysteme, die denselben Input auswerten, ein entscheidender evolutionärer Weiterentwicklungsschritt ist, wenn es dem Organismus gelingt, diese Systeme inferentiell zu integrieren, „gleichsam eine gemeinsame Sprache, ein gemeinsames Datenformat bereitzustellen, in dem diese Systeme miteinander kommunizieren können“ (ebd., S. 15). Wenn allerdings zwei Teilsysteme in ein übergeordnetes System inferentiell integriert werden sollen, muss es für das höhere System möglich sein, die unterschiedlichen Datenformate dieser Teilsysteme zu „lesen“. Insofern bedarf es nach Ansicht Mausfelds eines abstrakteren Codes mit zunehmendem Grad der Integration von modularen Teilsystemen, worin Maus-

feld auch den Ursprung der Möglichkeit zur Ausbildung abstrakter Symbolstrukturen wie unserer Schriftsprache sieht, denn die Folge der skizzierten Entwicklung ist, dass die mit der Modularisierung einhergehende Datenabstraktion „nahezu zwangsläufig eine Tendenz zur Entstehung abstrakter Symbolstrukturen mit sich [bringt]“ (ebd.). Aber nicht nur das. Mausfeld sieht hierin sogar eine „notwendige Voraussetzung für die Möglichkeit eines phänomenalen Bewußtseins“ (ebd., S. 16).

Betrachtet man das perzeptuelle System genauer, sprechen eine Reihe von experimentellen Befunden (z.B. Tse, 1999; Spelke, Breinlinger, Jacobson & Phillips, 1993) dafür, dass dieses durch die konzeptuellen Formen semantisch grundsätzlich nicht vollkommen festgelegt wird (vgl. Mausfeld, 2010a, S. 110). So verbleibt insbesondere in Bezug auf den Output des perzeptuellen Systems an den Schnittstellen zu höheren kognitiven Systemen in Bezug auf die Semantik ein interpretativer Spielraum, der grundlegende computationale Vorteile mit sich bringt:

By postponing disambiguation to higher-order interpretative systems, the *Perceptual System* can increase its global stability with respect to the superordinate ‘interpretations’ provided at its interfaces to subsequent systems. This protects the system from settling, under insufficient or ‘impoverished’ input situations, on some definite ‘interpretation’ that would have to be changed to an entirely different ‘interpretation’ following a small variation in the input. In addition, [...] underspecified conceptual forms boost the potency of generative processes and enhance the conceptual versatility of the *Perceptual System*. (ebd., S. 110f; Hervorh. im Orig.)

Entscheidend ist, dass die Computationen mit den grundsätzlich unter-spezifizierten konzeptuellen Formen auf *verschiedenen Ebenen semantischer Granularität* stattfinden können. Es ist unsere (phylogenetisch) evolvierte und teilweise auch kulturell tradierte Entwicklungsgeschichte, die zur Ausbildung bestimmter konzeptueller Formen geführt hat, die in unser Wahrnehmungssystem implementiert sind und eine vielgestaltige

oder mehrdeutige Repräsentation unserer Umwelt ermöglichen. Entgegen (Alltags-)Theorien, die Multiperspektivität auf sprachlich-konzeptuelle (interpretative) Systeme höherer Ordnung zurückführen, ist diese vielmehr eine intrinsische Eigenschaft des Wahrnehmungssystems selbst (vgl. ebd., S. 111).

Da Wahrnehmung immer durch das Raster (bzw. die Anreicherung) der konzeptuellen Formen verläuft, besteht nach Ansicht Mausfelds bezüglich des grundsätzlichen Operationsmodus des perzeptuellen Systems kein Unterschied in der Wahrnehmung materieller Eigenschaften und der Wahrnehmung mentaler Zustände anderer Menschen. In beiden Fällen werden die sensorischen Informationen über die konzeptuellen Formen des Wahrnehmungssystems (sowie höherer interpretativer Systeme) an unsere arteigene innere Semantik angeglichen und ausgewertet. So lesen wir in beiden Fällen aus den visuellen Eindrücken, die in unseren Organismus eingehen, verborgene Merkmale der Wahrnehmungsobjekte heraus und sagen darüber zukünftige Zustände und Vorgänge dieser voraus (vgl. ebd., S. 106; Chomsky, 1975).

Allerdings müsse man feststellen, dass es eine Besonderheit bei einer Auseinandersetzung mit dem menschlichen Geist ist, dass Erkenntnisgegenstand und Erkenntnisinstrument in diesem Fall eng miteinander verflochten sind. Was Mausfeld (2006) – ganz ähnlich wie Fechner – erläutert: „Machen wir unseren Geist selbst zu einem Erkenntnisobjekt, so verfügen wir durch die Art unserer biologischen Ausstattung über zwei ganz unterschiedliche Zugangsweisen: die Möglichkeit, durch einführendes Nachvollziehen etwas über die Gesetzmäßigkeiten des Psychischen herauszufinden, und die Möglichkeit, die tiefer liegenden verantwortlichen Prinzipien aus einer naturwissenschaftlichen Außenperspektive zu ergründen“ (S. 36). In diesem Sinne differenziert er zwischen einem System, das er „*science forming faculty*“ nennt und uns erlaubt, „unabhän-

gig von einem spezifischen Gegenstandsbereich in hochgradig abstrakter Weise Schlussfolgerungen zu ziehen und somit Mathematik und Naturwissenschaft zu betreiben“ (ebd.) und dem „*theory of mind module*“ (Alltagskonzeption des Geistes, Zuschreibung mentaler Zustände, empathisches Nachvollziehen der Handlungen anderer etc.), auf dem seines Erachtens „unsere Alltagspsychologie und mit ihr Literatur, Kunst, Musik und die gesamten Geisteswissenschaften [beruhen]“ (ebd.). Auch wenn diese Begriffswahl ihre Gründe haben mag, scheint sie mir im Hinblick auf die „*science forming faculty*“ nicht ganz treffend bzw. für den hier verfolgten Zweck zu sehr wissenschaftsbezogen zu sein. Statt einer „die Wissenschaft formenden Fähigkeit“ sollte man eher von einem „*theory of matter module*“ (oder einem Modul für unsere Alltagstheorien über materielle Dinge) sprechen. Man mag – wie eben Mausfeld (2007, 2010) – monieren, dass der Materiebegriff der modernen Physik doch gewissermaßen abhandengekommen ist, sodass man sich nicht mehr auf ihn berufen könnte und den Umweg über die Formulierung „das, worauf sich die Naturwissenschaften (oder deren Basiswissenschaft: die Physik) beziehen“ gehen müsse, aber in unserer Alltagskonzeption von Wirklichkeit, um die es bei dem hier betrachteten psychologischen Ursprung des Leib-Seele-Problems geht, ist die Einteilung in dinglich-materielle und nicht-dingliche, immaterielle Phänomene weiterhin gegeben und maßgeblich.

3.4.4. Entwicklungspsychologische Einbettung bzw. Adaptation

Es muss an dieser Stelle festgehalten werden, dass zwischen Mausfelds allgemeinem Ansatz zur Erklärung von Multiperspektivität und seinen Ausführungen zu der Divergenz zweier höherstufiger Module, die – vereinfacht gesagt – bei der Wahrnehmung körperlicher vs. geistiger

Gegebenheiten aktiviert werden, eine gewisse Lücke klafft. Es wird nicht klar, ob Mausfeld selbst annimmt, dass diese höherstufigen Module bereits auf der Ebene des perzeptuellen Systems selbst greifen (wie er dies bei den dargestellten multiperspektivischen Phänomenen explizit vertritt), ob also die wahrgenommene Divergenz von Gehirn und Geist unter seinen Voraussetzungen für Multiperspektivität auch auf im Laufe der Evolution entstandene und in unser Wahrnehmungssystem implementierte konzeptuelle Formen zurückgehen soll oder ob sie sich ontogenetisch im Zuge bestimmter (kultureller, bildungsbezogener...) Erfahrungen als eine Art alltagspsychologische Überzeugung herausbildet und wir einen „konventionellen“ Fall einer konzeptgesteuerten Top-Down-Wahrnehmung vorliegen haben.

Das Bestehen dieser Kluft dürfte insbesondere damit zu tun haben, dass Mausfeld selbst mit seinem Ansatz nicht unbedingt einen Beitrag zur Auflösung des Leib-Seele-Problems bzw. zur Erklärung der psychologischen Ursprünge der Gehirn-Geist-Differenz leisten will. Wie sich in Kapitel 2.1.2. dieser Arbeit gezeigt hat, ist Mausfeld selbst gerade sehr skeptisch gegenüber psychophysischen Theorien, die den Geist zu „nah“ an seine physischen Grundlagen heranrücken oder gar reduktive/reduktionistische Strategien verfolgen (vgl. ebd., 2003, 2010). Insofern verwundert es nicht, dass er selbst – obwohl dies sehr naheliegend erscheint – das explanatorische Potential seines Ansatzes nicht weitergehend für eine Erklärung der wahrgenommenen Gehirn-Geist-Differenz heranzieht. Denn für Mausfeld scheint die Akzeptanz einer dualistischen Antwort auf das Leib-Seele-Problem keineswegs ausgeschlossen zu sein⁴¹.

⁴¹ Obgleich Mausfeld mit seiner Kritik an der häufig überheblich daherkommenden Neurowissenschaft durchaus recht haben mag (insbesondere wenn er darlegt, dass es absurd wäre anzunehmen, dass sich zukünftig psychologische Phänomene abschließend auf neuronaler Ebene erklären lassen), scheint er nichtsdestotrotz den Fehlschluss zu

Aus diesem Grund stellt sich die hier aufgeworfene Frage, wie es trotz der offensichtlichen Fehlgeleitetheit eines ontologischen Dualismus zur Gehirn-Geist-Differenz im Sinne eines *intuitiven* Dualismus kommt, in dieser konkreten Form für ihn erst gar nicht.

Wenn man nun dennoch diesen Weg einschlägt, könnte man aus Mautfelds Ansatz mit Bezug auf die Gehirn-Geist-Problematik ableiten, dass es in unserem Wahrnehmungssystem unterschiedliche implementierte Konzeptformen gibt, die sich im Laufe unserer phylogenetischen Entwicklung herausdifferenziert haben und in Verbindung mit höherstufigen konzeptuellen Systemen (theory-of-mind-module vs. theory-of-matter-module) stehen oder sogar Teilkomponenten dieser Systeme sind. Sie werden aktiv, wenn es um die Wahrnehmung und (Alltags-)Konzeptualisierung psychischer und physischer Tatbestände geht. So wird nun allerdings die Frage aufgeworfen, ob diese höherstufigen wahrnehmungsbezogenen oder wahrnehmungsimmanenten Module selbst angeboren sind.

Insofern bedarf es an dieser Stelle einer gewissen Präzisierung oder zumindest Problematisierung, denn eine phylogenetische Verankerung und auch das Primat wahrnehmungsbezogener Prozesse bei der Konzeptentwicklung gilt im Rahmen der Erforschung der Entwicklung unserer ontologischen Kategorien bzw. innerhalb des Theory-of-Mind-Ansatzes keineswegs als ausgemacht.

machen, dass nur eine Ablehnung *jeglicher* nicht-dualistischer Ansätze, die mit psychophysischen Identitätsannahmen bzw. der Möglichkeit einer irgendwie gearteten Reduktion aufwarten, die Psychologie davor bewahrt, sich den Vereinnahmungsstrategien solcher Naturwissenschaftler zu widersetzen. Dass diese Befürchtung unberechtigt und bestimmte Identitäts- und Reduktionsannahmen nicht nur harmlos sind, sondern im Gegenteil gerade die Ablehnung einer Identitätsannahme für den Status des Psychischen bzw. der Psychologie hochproblematisch wäre, werde ich an späterer Stelle dezidiert aufzeigen (vgl. Kap. 8 und 9).

So besteht in der Entwicklungspsychologie immer noch Uneinigkeit darüber, ob wir über bestimmte angeborene (phylogenetisch/evolutionär entstandene) Kategorien zur Klassifikation von biologischen vs. nicht-biologischen Arten (und damit auch zwischen Dingen, die über Geist verfügen, und solchen, die dies nicht tun) verfügen. Jemand der ansatzweise von solchen angeborenen Kategorien ausgeht, ist Keil (1992). Er postuliert so genannte KADs (*knowledge acquisition devices*), also Wissenserwerbsstrukturen für abstraktes domänenspezifisches Wissen, die Kinder prädisponieren Fakten zu lernen, die es ihnen ermöglichen, zwischen belebten und unbelebten Dingen zu differenzieren. Zu diesen Fakten gehören nach Ansicht Keils unter anderem die Einsicht, dass etwas den biologischen Arten Intrinsisches die meisten ihrer phänomenalen Eigenschaften hervorbringt oder – etwas spezifischer – dass phänomenale Eigenschaften für gewöhnlich Ausdruck von zugrunde liegenden nicht-phänomenalen Eigenschaften sind. Dies würde mit der Annahme übereinstimmen, dass wir über angeborene (perzeptuelle oder zumindest die Wahrnehmung leitende) Strukturen verfügen, die eine Rolle bei der Entstehung der Gehirn-Geist-Differenz spielen.

Wie Sabina Pauen (1999) aufzeigt, weist der Ansatz von Keil jedoch das große Manko auf, dass er keine Erklärung für typische Fehlleistungen der (früh-)kindlichen Konzeptbildung anbietet. So erklärt die Theorie z.B. nicht, warum jüngere Kinder oft „Leben“ mit der Fähigkeit gleichsetzen, sich (ohne äußeren Anstoß) bewegen zu können, auch wenn es sich dabei um Spielzeuge oder andere unbelebte Objekte handelt: „If our cognitive system were actually tuned to detect specific relational structures in the real world that distinguish living from non-living kinds on the basis of biological features, why then should children use inappropriate distinctions such as the distinction between self-starters and non-self-starters?“ (ebd., S. 20). Insofern verwundert es nicht, dass Keil in der Entwicklungspsycholo-

logie eher eine Randposition einnimmt. Die meisten Forscher gehen von einem relativ offenen Wissenserwerb aus, der durch unsere biologische Ausstattung weniger präformiert ist, als Keil dies annimmt. Dabei bedienen sich die Kinder bei ihren (meist implizit ablaufenden) ontologischen Kategorisierungen in den verschiedenen Entwicklungsphasen unterschiedlicher Quellen und Heuristiken (Carey, 1985; Inagaki, 1990; Gelman & Coley, 1990), die jeweils in gewisser Hinsicht irrtumsanfällig sind (wie eben das (Selbst-)Bewegungskriterium, perzeptuelle Ähnlichkeit zur eigenen Art, Wissen über den „kausalen Hintergrund“ eines Objektes o.Ä.).

Damit wären wir zugleich beim zweiten Aspekt angelangt, also der Frage, ob es ein Primat wahrnehmungsbezogener Prozesse bei der Entwicklung der in diesem Kontext relevanten ontologischen Kategorisierungen gibt. In der Tat gehen einige Forscher davon aus, dass sich der Inhalt von Konzepten im Allgemeinen auf perzeptuelle Strukturen bezieht (z.B. Barsalou, 1993; Jones & Smith, 1993; Smith & Heise, 1992). So gilt es als gesichert, dass in der frühen Kindheit wahrnehmungsbasierte Kategorisierungen dominieren⁴² (vgl. Eimas & Quinn, 1994; Madole & Cohen, 1995) und – was durch-aus auch für Erwachsene gilt – die Zuschreibung von „Leben“ oder „Innerlichkeit“ häufig per Induktionsschluss aufgrund perzeptueller Ähnlichkeit zur eigenen, also menschlichen Spezies erfolgt (Carey, 1985). Je ähnlicher ein fragliches Objekt dem Menschen wahrgenommen wird, desto höher ist demnach die Wahrscheinlichkeit, dass es als „belebt“ eingestuft wird.

⁴² Natürlich könnte man hier auch wieder die Frage stellen, woher die typische Orientierung an Farben, Formen, Bewegung und Größe kommt, nach denen junge Kinder Objekte differenzieren, wenn es sich hierbei nicht um angeborene und damit evolutionär geformte Orientierungen handelt.

Jedoch gibt es auch eine Reihe an Befunden, die suggerieren, dass unser konzeptuelles Wissen eine Menge an untereinander vernetzten *kausalen Überzeugungen* über die Natur eines bestimmten Objektes beinhaltet, die sich nicht auf Wahrnehmungstatbestände reduzieren lassen (z.B. Gelman & Markman, 1987; Keil, 1989). So befähigen kausale Überzeugungen Kinder wie Erwachsene dazu, von irrelevanten Wahrnehmungsaspekten abzusehen und stärker die relevanten *strukturellen* Eigenschaften von Objekten zu berücksichtigen (vgl. Pauen, 1999). So hat z.B. Inagaki (1990) gezeigt, dass Kinder, wenn sie einen Goldfisch einmal als ein Lebewesen erkannt haben, diese Erfahrung nutzen, um potentielle Verhaltensweisen und Eigenschaften anderer im Wasser lebender Tiere zu bestimmen, obwohl sie diesem keineswegs ähnlich sind. Dabei bevorzugen schon 2-4-jährige Kinder bei ihren Entscheidungen das Wissen um eine Kategoriezuordnung gegenüber dem Eindruck perzeptueller Ähnlichkeit, wenn kategoriebasiertes Wissen vorhanden ist bzw. zur Verfügung gestellt wird (Gelman & Coley, 1990; Gelman & Markman, 1987; vgl. Pauen, 1999).

Wahrscheinlich ist, dass sich auch hier die Wahrheit in der Schnittmenge des theoretischen Spektrums befinden wird, denn es ist plausibel und wird durch die aktuelle empirische Befundlage gestützt, dass die Ausbildung und der Wandel unserer ontologischen Kategorien sowohl durch das Lernen von wahrnehmungsbezogenen als auch kausalen Aspekten über Gegenstände vorangetrieben wird (vgl. Keil, 1991; Pauen, 1999). Auch wenn also perzeptuelle Strukturen und auch – wie zuvor beschrieben – sprachlich-konzeptuelle Aspekte für Klassifikationsprozesse (belebt vs. unbelebt; mental vs. nicht-mental) wichtig sind, sind diese dennoch nicht hinreichend dafür, um die Ausbildung und den Wandel ontologischer Kategorisierungen zu erklären (wie die von Mausfeld beschriebenen divergierenden Module für Geistiges und Materielles). Denn wie

Pauen (1999) sehr treffend sagt: „Although perceptual similarities and word use are both known to be important for processes of categorization, these aspects alone cannot explain why humans make certain basic cuts in the world and not others“ (S. 31).

Deshalb schlägt sie ein Modell vor, nach dem Kinder sich bei ihren Kategorisierungen an den (mutmaßlichen) Ursachen und Wirkungen von Veränderungen orientieren, wobei für diese Einschätzung drei Dimensionen möglicher Informationen herangezogen werden sollen: (1) Kausalität, (2) Funktionalität und (3) Vorhersagbarkeit (vgl. ebd., S. 21ff). Z.B. würde zur Dimension der Kausalität gehören, ob ein Objekt selbstinitiierte Veränderungen seiner externalen und/oder internalen Charakteristika oder auch seines Standortes vornehmen kann, zur Dimension der Funktionalität, ob man – wie Keil (1992) sagt – Gegenstände bzw. deren Verhalten nach Maßgabe *teleologisch-funktionaler* Gesichtspunkte erfassen kann (was dann stark mit der Zuschreibung von „Leben“ und „Innerlichkeit“ assoziiert ist) und zur Dimension der Vorhersagbarkeit, ob die Veränderung, die ein Gegenstand bzgl. seiner Bewegung, Form, Geräusche etc. macht, absehbar sind (wobei gelte, dass solche Veränderung bei lebenden Objekten aufgrund ihrer „internen Steuerungsprozesse“ viel weniger vorhersehbar sind als bei nicht-lebenden Objekten).

Dass wir Gegenstände bestimmten Domänen zuordnen, resultiere nun daraus, dass Informationen bzgl. Kausalität, Funktionalität und Vorhersagbarkeit in der Natur für gewöhnlich korrelieren: „some objects behave in mostly unpredictable ways: they perform changes in the absence of external agents, and such changes occur independently of human intentions. Other objects perform changes that are highly dependent on human intentions, follow well-known rules and are comparably easy to predict“ (Pauen, 1999, S. 26). Gemäß diesem Modell ist also das *individuelle* Hintergrundwissen entscheidend dafür, ob und wie man die Unterschei-

dung zwischen lebendig vs. nicht-lebendig (bzw. über geistige Zustände verfügend vs. hierüber nicht verfügend) anwendet. Dieses Modell, das die vermeintlich widersprüchliche Befundlage zur Konzeptbildung relativ gut integrieren kann, spricht also eher gegen den Erwerb ontologischer Kategorien in einer festgefügt (biologisch fixierten) Sequenz. „How much a given child knows about causes of specific changes, the functionality of an object’s activities and properties, or about the predictability of its behaviour depends on his/her particular learning experiences“ (ebd., S. 28).

Wenn also die Frage zu beantworten war, warum es ganz bestimmte Aspekte meines Wahrnehmungsfeldes sind, die bei der Kategorisierung der Wirklichkeit nach Maßgabe bestimmter Module (mind vs. matter) besonders hervorspringen, ob dies mit dem Wahrnehmungssystem selbst zu tun hat und ob dies biologisch bzw. phylogenetisch verankert ist, kann man aufgrund des gegenwärtigen Forschungsstandes der Entwicklungspsychologie sagen, dass die Ausbildung und Erweiterung unserer konzeptuellen Wissensstrukturen hierfür maßgeblich ist. Es mag zwar bestimmte phylogenetische Präformationen geben, die ein Spektrum möglicher Kategorisierungen abstecken und vielleicht auch die Entwicklung in eine bestimmte Richtung lenken, dennoch dürfte es sich hierbei um einen individuell variablen Vorgang handeln, der stark von den ontogenetischen Erfahrungen eines Menschen abhängt. „Concept development and conceptual changes result from the acquisition of new knowledge concerning relevant dimensions and attempts to integrate this knowledge into existing structures. This process may sometimes lead to shifts in the salience of specific aspects that are part of a child’s object presentation“ (Pauen, 1999, S. 31).

3.5. Fazit: Ein Modell der geistigen Entzweigung von Gehirn und Geist

Inwiefern kann die Fechnersche Zwei-Perspektiven-Theorie in Bezug auf die Ursache des Leib-Seele-Problems nun auf der Basis der durchgeführten wahrnehmungspsychologischen Analyse und deren entwicklungspsychologischer Erweiterung fundiert werden?

Obwohl Gehirn und Geist keine substantiell verschiedenen Wirklichkeitsbereiche konstituieren, nehmen wir sie aufgrund der von Mausfeld beschriebenen Mechanismen zwangsläufig völlig unterschiedlich wahr. Betrachte ich physische Dinge wie ein Gehirn, gehen über die Sinne Informationen in unsere internale computationale Informationsverarbeitung ein. Dieser sensorische Input dient dann als ein Hinweisreiz zur Aktivierung konzeptueller Formen, die das Datenformat der nachfolgenden computationalen Prozesse determinieren. Neben dem tatsächlich über die Sinne transportierten Input werden durch die (evtl. zum Teil phylogenetisch verankerten) konzeptuellen Formen des Wahrnehmungssystems bzw. höherstufiger konzeptueller Systeme eine Reihe weiterer bedeutungshaltiger Kategorien mit aktiviert und in unser Perzept integriert. Diese speisen sich in dem Fall aus dem Theory-of-Matter-Modul. So wie wir die Zerbrechlichkeit des Glases beim Anblick eines solchen Trinkgefäßes mit wahrnehmen, nehmen wir beim Anblick eines Gehirns die von Descartes genannten Eigenschaften von physischen Objekten (implizit) mit wahr, die Ausdehnung, Lokalisierbarkeit im Raum und den nicht-denkenden bzw. nicht-intentionalen Charakter. Es werden noch weitere Eigenschaften hinzukommen, die wir individuell-ontogenetisch im Hinblick auf den Gegenstand „Gehirn“ erlernt haben. Es sind Fakto-

ren, die sich gemäß Pauen (1999) der kausalen, funktionalen und prädiktiven Wissensdimension zuordnen lassen werden.

Wie Mausfeld sagt, wird im Zuge dieses Vorgangs eine Art *perzeptueller Ontologie* geschaffen. In der Wahrnehmung selbst liegt gewissermaßen schon der Keim der „dinglich-materiellen“ Auffassung von äußeren Wahrnehmungsgegenständen wie dem Gehirn, da jedes konzeptuelle Teilsystem unseres Wahrnehmungsapparates aufgrund seiner phylogenetischen und ontogenetischen Geschichte eine Segmentierung des sensorischen Inputs in bestimmte Kategorien vornimmt.

Wenn wir unsere Aufmerksamkeit dagegen auf unsere eigenen geistigen Vorgänge richten, die zwar mit neuronalen Vorgängen in unserem Gehirn *identisch* sein mögen (was genau dies heißt, muss noch geklärt werden), erleben wir diese aus zwei zentralen Ursachen zwangsläufig als etwas vermeintlich völlig Andersgeartetes:

Die erste zentrale Ursache dafür, dass uns die Phänomene, die wir in der Perspektive der ersten Person erleben, nun so sehr von ihren neuronalen Korrelaten verschieden erscheinen, die wir in der Perspektive der dritten Person erfassen können, liegt am beschriebenen psychischen Strukturmerkmal der phänomenalen Transparenz. Wir haben aus unserer Erlebnisperspektive für gewöhnlich keinen Zugriff auf die (mit den mentalen Prozessen assoziierten) tieferen Verarbeitungsstufen im Gehirn, was Gerhard Roth (1994) im Zusammenhang der Gehirn-Geist-Problematik sehr schön beschrieben hat:

Durch Introspektion allein kämen wir in der Tat nicht zu der Ansicht, daß Wahrnehmungen, Denken, Vorstellen, Erinnern irgend etwas mit unserem Gehirn zu tun haben könnten; denn wir fühlen unser Gehirn nicht, und wir haben nicht den Eindruck, daß wir *mittels* unseres Gehirns die Umwelt wahrnehmen, über etwas nachdenken, uns an etwas erinnern oder uns etwas vorstellen. Vielmehr ist uns in unserer Wahrnehmung die Umwelt direkt gegeben. Bei allem, was wir hören oder sehen, ist kein Gehirn ‚dazwischengeschaltet‘. Auf dieselbe *unmittelbare* Weise

erleben wir unsere Gedanken, Erinnerungen und Vorstellungen. (S. 81, Hervorh. im Orig.)

Hierzu lässt sich eine interessante Analogie aufführen. In Abbildung 10 ist ein Würfel dargestellt, der sich über vier verschiedene räumliche Dimensionen erstreckt.

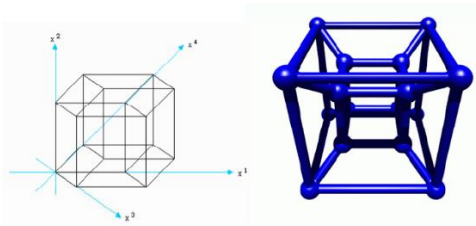


Abb. 10: „4D-Objekt“ (Aus: Mausfeld, 2006, S. 8)

Trotz der im Reiz vorhandenen, mathematisch belegbaren Vierdimensionalität können wir diese auf der Ebene unseres Erlebens nicht adäquat darstellen. Mausfeld (2006) begründet dies damit, dass für uns keine passende Konzeptform zur Wahrnehmung von räumlicher Vierdimensionalität verfügbar ist. Es ist nicht möglich, dass wir etwas wahrnehmen, „für das unser Wahrnehmungssystem konzeptuell nicht ausgestattet ist – selbst dann nicht, wenn es gleichsam im Reiz vorhanden ist“ (S. 8). Und in ganz ähnlicher Weise ist unser Wahrnehmungssystem auch nicht dafür ausgestattet, die funktionalen oder neuronalen Mechanismen, die dem Wahrnehmungsvorgang selbst zugrunde liegen, phänomenal zu repräsentieren. Die Folge ist, wie Roth an späterer Stelle weiter ausführt, dass „[d]em rein reflektierenden Geist die eigenen neuronalen Grundlagen völlig verborgen [sind]. Man könnte meinen, der Anschein der völligen Andersartigkeit der Gehirnprozesse im Vergleich zur phänomenalen Welt, die ja aus diesen hervorgeht, sei vom Gehirn zur Irreführung des

Geistes hervorgebracht“ (S. 87). Wie die vorher getätigte Analyse des funktionalen Hintergrundes der phänomenalen Transparenz gezeigt hat (vgl. Kap. 3.4.2.3), ist diese Irreführung des Geistes in der Tat notwendig, damit wir adäquat auf die Anforderungen der Umwelt eingehen können. Da rein selbstreferentielle Systeme keine evolutionäre Überlebenschance haben dürften, ist die „Strategie“ zur Entwicklung von Systemen, die über eine Durchsichtigkeit der repräsentationalen Grundlagen ihrer mentalen Zustände und die erlebte Unmittelbarkeit der Repräsentationsobjekte im Sinne eines naiven Realismus verfügen, eine fast logische Konsequenz. Die zweite Ursache dürfte darin bestehen, dass wenn wir nun – entgegen dem phänomenal transparenten und naiv-realistischen Alltagserleben – versuchen, die Wahrnehmung mentaler Prozesse *als* mentaler Prozesse in den Blick zu nehmen, dieser Vorgang ganz anders abläuft als die normale sinnliche „selbstvergessene“ Wahrnehmung im Alltag. Es stellt sich aber zunächst die Frage, ob man bei einem solchen „introspektiven“ Vorgang überhaupt von Wahrnehmung sprechen sollte⁴³. Gemäß dem empirisch plausiblen Modell der Wahrnehmung von Mausfeld muss man zwischen dem sensorischen System und dem perzeptuellen, dem eigentlichen *Wahrnehmungssystem* unterscheiden. Wahrnehmen ist eben viel mehr als nur die Darstellung der über die Sinne erhaltenen (und transformierten) Informationen, sondern es ist ein höherstufiger Prozess, der zwar für gewöhnlich zu großen Teilen auf den sensorischen Daten des Sinnessystems basiert, aber keineswegs auf diese allein reduzierbar ist. Insofern kann man auch im Falle der Erfassung unserer

⁴³ Wenn man wie z.B. Pauen (1999) sagt, dass der Wahrnehmungsapparat nicht an der Introspektion beteiligt ist, scheint man vielmehr zu meinen, dass es sich bei der Introspektion nicht um eine sinnliche Wahrnehmung von etwas handelt. Auch wenn es sicherlich richtig ist, dass wir über kein Vehikel wie einen (kantischen) „inneren Sinn“ verfügen, dürfte es sich hierbei durchaus um einen Wahrnehmungsvorgang handeln.

mentalen Prozesse *als* mentaler Prozesse von einem Wahrnehmungsvorgang sprechen, für den allerdings eine Vernachlässigung des konkreten sensorischen Inputs charakteristisch ist. Statt unsere Aufmerksamkeit auf die Objekte zu richten, die über unser Sinnessystem in unseren aktuellen Geisteszustand eingehen, stellt sich uns die schwierige Aufgabe, den Prozess selbst – ungeachtet der konkret in ihm repräsentierten äußeren Wahrnehmungsobjekte – zu erfassen.

Was aber bleibt, wenn man versucht, von den konkreten (äußeren) Wahrnehmungsgegenständen in einem Wahrnehmungsakt zu abstrahieren? Es sind bekanntermaßen die qualitativen und kognitiven Eindrücke, die unseren gegenwärtigen mentalen Zustand kennzeichnen: Wie fühlt es sich an, in diesem Zustand zu sein und welche Gedanken gehen mir durch den Kopf, was verbinde ich mit dieser Situation, welche diesbezüglichen Überzeugungen habe ich? Es sind allesamt Charakteristika, die – da sie sich nunmehr nicht auf konkrete Wahrnehmungsgegenstände (der Außenwelt) beziehen – offenbar selbst „gegenstandslos“ oder eben „immateriell“ sind. Von den Gegenständen meiner sinnlichen Wahrnehmung kann ich sagen, wo sich diese befinden, wie groß, schwer etc. sie sind. Meine Empfindungen und Gedanken *selbst* scheinen dagegen nicht „verortbar“ etc. zu sein.

Durch die gewisse Entkopplung von den sensorischen Inputs und die veränderte Ausrichtung der Aufmerksamkeit kommen im Falle der Wahrnehmung mentaler Prozesse *als* mentaler Prozesse notwendigerweise also ganz andere konzeptuelle Formen zum Einsatz wie bei der Wahrnehmung eines Gehirns (oder anderer physischer Objekte). Statt einer Segmentierung des sensorischen Inputs in bestimmte physische Kategorien, erfolgt eine Art Metarepräsentation des phänomenalen oder intentionalen Gehalts eines mentalen Ereignisses. Wir vergegenwärtigen uns hierbei also, dass unsere mentalen (repräsentationalen) Prozesse

mentale (repräsentationale) Prozesse sind. Dass im Zuge einer solchen Selbstbezugnahme andere mentale Mechanismen greifen werden als im Falle der Wahrnehmung physischer Gegebenheiten in unserer Umwelt ist offensichtlich.

So sagt auch Dörner (1979) ganz in diesem Sinne, dass der Ursprung des Leib-Seele-Problems in der menschlichen Möglichkeit zur Selbstreflexion liegt: „Die menschliche Fähigkeit, sich selbst, d.h. den eigenen Körper und auch z.T. die psychischen Prozesse zum Gegenstand der Betrachtung zu machen, ist wohl eine Ursache für die Unterscheidung von *Leib* und *Seele*, also die Ursache für das Leib-Seele-Problem. Die menschliche Fähigkeit, sich gewissermaßen zweizuteilen in ein beobachtetes Objekt und ein beobachtendes Subjekt, legt eine Aufteilung in zwei verschiedene Instanzen, nämlich Leib und Seele nahe“ (S. 123; Hervorh. im Orig.).

Aus dem hieraus resultierenden Spannungsverhältnis zweier fundamentaler Perspektiven bei der Betrachtung des menschlichen Geistes ergibt sich die von Mausfeld beschriebene (simultane) Aktivierung zweier unterschiedlicher kognitiver Teilsysteme, wenn wir uns mit Geistigem beschäftigen, und zwar die des Theory-of-Matter- und Theory-of-Mind-Moduls. Offenbar besteht zwischen ihnen eine gewisse Entkopplung, wie sie Mausfeld im Hinblick auf die Beobachtungen Tinbergens zur Ei-Wahrnehmung der Graugänse schildert. So wie es die Gans nicht fertigbringt, die beiden perzeptuellen Teilsysteme zur Detektion von Eiern inferentiell zu integrieren und daher potentiell in einer Endlosschleife eine Bierdose mal als Ei erkennt und mal nicht, scheint es im Verlauf der Menschheitsentwicklung bisher nicht gelungen zu sein, eine gemeinsame Sprache bzw. ein gemeinsames Datenformat, wie Mausfeld (2006) es ausdrückt, bereitzustellen, in dem die kognitiven Systeme zur Wahrnehmung von bzw. Reflexion über Gehirn und Geist miteinander „kommunizieren“ können. Wir haben also bisher kein höherstufiges kognitives

Modul entwickeln können, mittels dessen eine „Aussöhnung“ unserer Wahrnehmungseindrücke und konzeptuellen Vorstellungen stattfinden könnte.

Eine Folge dieses Mangels ist daher, dass bei der Wahrnehmung psychischer Prozesse (als psychischer Prozesse) eine Aktivierung höherstufiger sprachlich-konzeptueller Systeme begünstigt wird, die deutlich verschieden ist von derjenigen, die bei der Wahrnehmung eines Gehirns (oder anderer physischer Objekte) aktiviert werden. Hier ist also die Ursache für die beschriebenen sprachlich-konzeptuellen Divergenzen in Bezug auf den Gehirn-Geist-Komplex gelegt. Es sind bestimmte kognitive Systeme, die höchstwahrscheinlich zum Teil phylogenetisch verankert sind und sowohl unsere Wahrnehmung und Aufmerksamkeitsausrichtung als auch unseren Sprachgebrauch beeinflussen oder gar determinieren.

Insofern dürfte Brian Loar (1997/2007) mit seiner Diagnose recht haben, wenn er sagt, dass ein wesentlicher Faktor für die Persistenz des Leib-Seele-Problems darin liegt, dass wir sehr unterschiedliche Begriffe in Bezug auf phänomenale und physische Zustände und Prozesse ausgebildet haben, zwischen denen eine solche Kluft liegt, dass wir meinen, dass auch die Bezugsgegenstände dieser Begriffe ontologisch nicht zusammenfallen können:

Wenn wir [...] phänomenale und physisch-theoretische Begriffe in unseren philosophischen Überlegungen zusammenbringen, sind diese kognitiven Zustände so verschieden, dass die Illusion entstehen kann, dass auch ihre Referenten verschieden sein müssen. Es sieht so aus, als ob die antiphysikalistischen Intuitionen auf einer Ähnlichkeitstheorie mentaler Repräsentationen beruhen würden, als ob wir vom Mangel an Ähnlichkeit bei unseren phänomenalen und physisch-funktionalen Begriffen auf einen Mangel an Gleichheit der Eigenschaften schließen würden, auf die sie sich beziehen. (ebd., S. 41)

Für Loar sind phänomenale Begriffe eine besondere Untergruppe einer Begriffsklasse, die er als Wiedererkennungsbegriffe bezeichnet. Sie sind verankert „in Dispositionen zur Klassifizierung bestimmter Objekte, Gegenstände und Situationen anhand von wahrnehmungsmäßigen Unterscheidungen“ (S. 34). Dass uns phänomenale und physisch-funktionale Gegebenheiten als unabhängig voneinander erscheinen, liegt daran, weil Wiedererkennungsbegriffe und theoretische Begriffe im Allgemeinen voneinander unabhängig sind. Ganz im Sinne des Spinozistischen Diktums „Der Geist als Idee des Körpers“ nimmt auch Loar an, dass Wiedererkennungsvorstellungen von phänomenalen Qualitäten nicht über kontingente Gegebenheitsweisen vermittelt sind, sondern unvermittelt bewusst werden. Hierin unterscheiden sie sich fundamental von den theoretisch-abstrakten Begriffen, die wir – neben einer evtl. vorhandenen phylogenetischen Bahnung – aufgrund unserer Wahrnehmungen über physisch-funktionale Gegebenheiten der Außenwelt bilden und zusammengekommen das Theory-of-Matter-Modul (bzw. die hiermit zusammenhängende *science forming faculty*) konstituieren. „Wenn phänomenale Begriffe die basale Wiedererkennung innerer physisch-funktionaler Zustände widerspiegeln, dann sollten sie von theoretischen physisch-funktionalen Beschreibungen begrifflich unabhängig sein“ (ebd., S. 37). Zusammenfassend kann man festhalten, dass nach Maßgabe der „subjektivistischen“ Auflösung von Spinozas Dilemma die *phänomenale Differenz* zwischen mentalen und neuronalen Aspekten demnach nicht an einem Unterschied in der wahrgenommenen Sache selbst, sondern nur an einem Unterschied in dem repräsentationalen Zugriff auf ebendiese liegt. Für Michael Pauen (1999), einem der gegenwärtigen Hauptvertreter einer solchen Ansicht, gilt es daher als ausgemacht,

daß Differenzen, die sich allein aus dem Verhältnis eines Gegenstandes zum Betrachter und dem von ihm gewählten Referenzsystem ergeben, nicht zur Zuschreibung neuer Eigenschaften führen können: Ein Gegenstand verändert seine Größe nicht dadurch, daß der Betrachter sich von ihm entfernt, er verändert seine Eigenschaften auch nicht dadurch, daß der Betrachter bei der Beschreibung auf ein anderes Bezugssystem zurückgreift. Vielmehr ist es möglich, daß sich unterschiedliche Beschreibungen auf eine Eigenschaft beziehen bzw. umgekehrt Eigenschaften identisch sind, die mit unterschiedlichen Ausdrücken bezeichnet werden. Diese Möglichkeit besteht auch bei neuronalen und mentalen Eigenschaften: Die Verwendung unterschiedlicher Prädikate [oder *Attribute*; M.L.] impliziert nicht, daß hier ein realer Unterschied bestehen muß; es könnte vielmehr auch Fälle geben, in denen beide Ausdrücke nur auf *eine* Eigenschaft verweisen. (S. 34f, Hervorh. im Orig.)

Dass eine psychophysische Identität möglich ist, liegt demnach an der Tatsache, dass Beschreibungen und Wahrnehmungen durch die Entitäten, auf die sie sich beziehen, *unterdeterminiert* sind (vgl. ebd., S. 106). „In den Wissenschaften können wir unterschiedliche Methoden und Beschreibungssprachen verwenden, im Alltag kommen neben sprachlichen Differenzen auch Unterschiede in der Wahrnehmungsperspektive in Betracht. Dies hat zur Konsequenz, daß sich unterschiedliche Beschreibungen oder Wahrnehmungen prinzipiell auf *eine* Entität beziehen können“ (ebd.; Hervorh. im Orig.).

Ist mit der Aufdeckung des psychischen oder psychologisch erklärbaren Ursprungs des Leib-Seele-Problems zugleich auch der Ursprung des Problems *insgesamt* geklärt? Erschöpft sich die Gehirn-Geist-Differenz also *ausschließlich* in einem psychologischen Phänomen? Eine solche Festlegung erscheint zumindest verfrüht, solange nicht auch die zweite Möglichkeit der Auflösung des Spinozistischen Dilemmas einer eingehenden Untersuchung unterzogen wurde. Spinoza haderte mit der unter seinen begrifflichen Voraussetzungen durchaus „eleganten“ Lösungsmöglichkeit, dass es sich bei der phänomenalen Differenz zwischen

Gehirn und Geist *ausschließlich* um ein anthropomorphes Phänomen handelt. Obgleich wir – in seiner Terminologie ausgedrückt – die Substanz (bzw. ihre Modi) immer nur über ihre Attribute und niemals an sich selbst wahrnehmen können, bedeutet dies eben keineswegs, dass das, was sich über diese Attribute präsentiert, nicht zugleich etwas widerspiegelt, was auch *der Sache nach* vorhanden ist und nicht nur in unseren Köpfen entsteht.

Auch in der Wissenschaftswelt und philosophischen Debatte wird in ebendiesem Sinne vielerorts daran gezweifelt, dass die (phänomenale) Gehirn-Geist-Differenz nur ein Artefakt unterschiedlicher Repräsentationsformate ein und desselben Wirklichkeitsausschnitts ist. Eine psychophysische Identitätstheorie, die diese Differenz nivelliert, indem sie ausschließlich als ein rein subjektives (beobachterabhängiges) Phänomen betrachtet wird, wirkt unplausibel. Z.B. sagt Christof Koch (2005):

An diesem Punkt bin ich nicht sicher, ob auch das NCC [Neural Correlate of Consciousness; M.L.] und das damit verknüpfte Perzept in diesem Maße identisch sind. Sind sie wirklich ein und dasselbe, aus unterschiedlichen Blickwinkeln gesehen? Das Wesen von Hirnzuständen und phänomenalen Zuständen scheint zu unterschiedlich, um sich aufeinander reduzieren zu lassen. Ich vermute ihre Beziehung ist komplexer als bislang allgemein angenommen. Momentan ist es wohl am besten, sich nicht festzulegen und sich darauf zu konzentrieren, die Korrelate des Bewusstseins im Gehirn zu identifizieren. (S. 21)

Auch wenn ich den Schluss Kochs bzgl. der vorübergehenden oder gar dauerhaften Akzeptanz eines „ontologischen Schwebezustandes“ nicht teile, stellt sich dennoch die grundsätzliche Frage, welchen Status *wissenschaftliche* bzw. über die Wissenschaften gewonnene Perspektiven auf ein und dasselbe Objekt (den menschlichen Geist) grundsätzlich haben. In der Tat gibt es Gründe daran zu zweifeln, dass es sich hierbei um bloße Perspektivitätsphänomene handelt, die sich – wie Pauen (1999) sagt –

allein aus dem Verhältnis eines Objektes zum Betrachter und dem von ihm gewählten Referenzsystem ergeben. Daher werden im folgenden Kapitel insbesondere psychologische bzw. kognitionswissenschaftliche Modelle mentaler Prozesse vorgestellt, die mit verschiedenen Konzeptualisierungen ihres Untersuchungsgegenstandes operieren. Es wird sich zeigen, dass hier der Begriff der *Ebene* eine entscheidende Rolle bei der Frage nach dem Status der verschiedenen wissenschaftlichen Perspektiven auf den menschlichen Geist spielt. Kognitive und phänomenale Prozesse sind demnach etwas, das sich auf unterschiedlichen wissenschaftlichen *Analyseebenen* vergegenständlichen lässt. Der „Wirklichkeitsstatus“ der psychologischen Analyseebenen und die ontologische Qualität der Beziehung zwischen den postulierten Ebenen bleiben hierbei aber oft noch etwas fraglich, sodass mit der (in der Tat buchstäblich zu nehmenden) Verlagerung der Perspektive auf Ebenenmodelle des Geistes, wie sie in der Psychologie bestehen, zwar ein wichtiger Schritt unternommen worden sein wird, aber hiermit noch keine abschließende Antwort auf die Frage nach dem womöglich objektiven/beobachterunabhängigen Status der Spinozistischen Attribute „Gehirn“ und „Geist“ erzielt werden kann. Im darauffolgenden Kapitel (Kapitel 5) wird daher – zunächst im Allgemeinen – besprochen werden, inwiefern bzw. unter welchen Voraussetzungen Ebenen eine *natürliche* Strukturgliederung eines Objektes widerspiegeln. Um dies zu klären, muss man sich zunächst mit dem allgemeinen Status von perspektivenabhängigen Repräsentationen bzw. Analyseebenen eines Objektes auseinandersetzen. Hierzu müssen verschiedene Individuationsbedingungen natürlicher Organisationsebenen betrachtet und bewertet werden, um dann die Frage nach der Möglichkeit einer „objektivistischen“ Auflösung von Spinozas Dilemma – nunmehr als Frage nach der Möglichkeit der Individuation des Psychischen als einer natürlichen Organisationsebene – beantworten zu können.

4. Psychologische Ebenenmodelle mentaler Prozesse

Wenn aufgrund der Analyse zum Ursprung des Leib-Seele-Problems die Frage im Raum steht, als was man die verschiedenen Perspektiven bzw. sprachlich-konzeptuellen Beschreibungen verstehen kann, die in Bezug auf ein und dasselbe Objekt von den Wissenschaften gewonnen werden und ob es sich hierbei um Aspekte handelt, die dem betrachteten Gegenstand selbst zukommen, ist es unabhkmmlich, sich intensiver mit dem Begriff der *Ebene* (respektive *Stufen* oder *Schichten*) auseinanderzusetzen, wie er in der aktuellen Forschungsliteratur allgegenwrtig ist.

In vielen Modellen der Biologie, Psychologie, Sozialwissenschaften (aber durchaus auch der Chemie oder Physik) wird mit groer Selbstverstndlichkeit von – zumindest im metaphorischen Sinne – „aufeinander liegenden“ oder „aufeinander aufbauenden“ Ebenen ausgegangen, die ein bestimmtes Objekt kennzeichnen sollen. Gerade fr die Gegenstandsbestimmung der immer strker miteinander verschmelzenden Einzeldisziplinen hat der Begriff der Ebene eine zentrale Bedeutung. So findet man in der interdisziplinr agierenden kognitiven Neurowissenschaft, die aus einem Verbund von Neurobiologie, Psychologie, Informatik und der Philosophie des Geistes besteht, verbreitet die Vorstellung, dass die verschiedenen Forschungsgegenstnde wie Synapsen, Hormone, Neuronen, neuronale Netzwerke, kognitive Prozesse usw. verschiedene Ebenen bilden, auf denen man sich der menschlichen Informationsverarbeitung nhern kann.

Nach einer allgemeinen Besprechung des Ebenenbegriffs, seines historischen Hintergrunds sowie gewisser extensionaler und intensionaler Begriffsbestimmungen wird sich im Folgenden zeigen, dass die in der Psychologie und kognitiven Neurowissenschaft verwendeten Ebenenmo-

delle für kognitive und phänomenale Zustände und Prozesse als eine wissenschaftliche oder systematische Ausprägungsform des zuvor entwickelten „Perspektivitätsansatzes“ verstanden werden können. Es gibt demnach – zumindest in erster Linie – verschiedene (funktional konzipierte) *Abstraktionsebenen*, die das, was in unseren Köpfen geschieht, abbilden können. Ein und dasselbe System lässt sich also auf verschiedenen Ebenen mit einem unterschiedlichen Granularitätsniveau oder eben Abstraktionsniveau analysieren. In diesem Kontext wird der menschliche Geist typischerweise als ein informationsverarbeitendes System wie ein Computer aufgefasst. So gilt, dass ein Computer als ein *funktionales*, informationsverarbeitendes System verstanden werden kann, das im Hintergrund der präsentierten Benutzeroberfläche nach Maßgabe der installierten *Software* bestimmte Berechnungen vornimmt. Zugleich ist ein Computer aber auch ein materieller oder physikalischer Gegenstand. Alle zuvor als abstrakte informationsverarbeitende Mechanismen konzipierten Prozesse haben also zugleich eine physikalische Wirklichkeit in der *Hardware* des Computers. Auch die menschliche Informationsverarbeitung lässt sich nun auf verschiedenen Ebenen analysieren. Sie kann auch als ein funktionales Geschehen verstanden werden, bei dem – mehr oder weniger bewusst – bestimmte Berechnungsschritte zur Lösung von Problemen getätigt werden. Sie kann aber auch als ein neuronales Geschehen betrachtet werden, denn auch hier bedarf jeder abstrakte Informationsverarbeitungsprozess offenbar einer bestimmten physischen Grundlage. Die psychologischen Ansätze funktionaler und konnektionistischer Modelle weisen hierbei für gewöhnlich eine Hierarchisierung der postulierten Ebenen auf, während sich die „Ebenen“ bei den ableitbaren Modellen für phänomenale Prozesse wie Emotionen in keine eindeutige Hierarchiefolge einpassen lassen.

Auch wenn es sich bei diesen Ebenenmodellen oder auch (kognitiven) Architekturen des Geistes um einen wichtigen Bezugspunkt in dieser Arbeit handelt, wird sich zeigen, dass die in der Forschungsliteratur konzipierten Architekturen relativ unbestimmt bleiben, wenn es um die Klärung der Frage nach dem Status des Psychischen selbst geht. So bleibt es unklar, welchen *ontologischen* Status die einzelnen Ebenen haben und in welcher Beziehung sie genau untereinander stehen sollen. Es fehlt ein integrativer Ansatz, der diese Ebenen (vor allem die von der Psychologie konzipierten vs. die von der Neurobiologie betrachteten) in eine Beziehung setzt und aufzeigt, wie genau diese Beziehung gestaltet ist und wie genau der mentale Prozess (der kognitionswissenschaftlich beschrieben wird) mit den natürlichen Organisationsebenen des Gehirns (wie sie vor allem die Neurobiologie beschreibt) zusammenhängt bzw. ob das Psychische selbst eine solche natürliche Organisationsebene bildet. Mittels des bloßen Rekurses auf Ebenen der Forschung, Theoriebildung oder der Analyse eines Gegenstandes lässt sich eine profunde Auseinandersetzung mit dem Leib-Seele-Problem jedenfalls nicht umgehen, wenn das Ziel sein soll, eine begrifflich-konzeptuelle Fundierung der Psychologie vorzunehmen.

4.1. Der Begriff der Ebene

Bei Ebenen- oder auch Schichtenmodellen (engl. *levels* resp. *layers*) handelt es sich um eine fächerübergreifende Vorstellung, die sowohl in den empirischen Wissenschaften als auch in der Philosophie von zentraler Bedeutung ist. Es ist die verbreitete Ansicht, dass die Welt bzw. die in ihr enthaltenen Objekte aus verschiedenen Schichten oder Ebenen aufgebaut

sind und dass diese somit auch durch die Wissenschaften auf verschiedenen Ebenen beschrieben, analysiert oder erklärt werden können.

In Anlehnung an das Bild einer „Hierarchie der Wissenschaften“ geht man davon aus, dass auch die Gegenstände, mit denen sich die Wissenschaften beschäftigen, als Elemente eines hierarchischen Systems verstanden werden können. Die fundamentalste Ebene wird demnach durch Entitäten konstituiert, die durch die Mikrophysik erforscht werden. Wenn das heutige „Standardmodell der Materie“ Recht behält, haben wir hier eine bestimmte Anzahl von nicht weiter zerlegbaren Elementarteilchen wie Quarks oder Leptonen. Wenn wir in einem solchen Ebenenmodell weiter „nach oben schreiten“, stoßen wir auf Entitäten wie Atome und Moleküle, die nunmehr vor allem zum Gegenstandsbereich der Chemie gehören. Wenn wir darauf die biologische Ebene betreten, kommen beispielsweise Zellstrukturen, Organe oder auch ganze Organismen mit bestimmten Verhaltensweisen zum Vorschein. Im Fall des Menschen (und anderer Lebewesen) kommt außerdem noch die mentale oder psychische Ebene hinzu, auf der eine Vielzahl phänomenaler und kognitiver Zustände vorzufinden sind. Wenn die soziale Ebene als ein eigener Wirklichkeitsbereich anerkannt wird, könnten auch soziale Gebilde wie Gruppen oder Gesellschaften, wie sie von der Soziologie oder Politologie untersucht werden, eine – in dem Fall sehr hochstufige – Ebene des hierarchischen Modells konstituieren.

Auf den verschiedenen Ebenen lassen sich wiederum spezifische Eigenschaften und Relationen zwischen den hier angesiedelten Entitäten ausmachen. So weisen Elementarteilchen bestimmte Ladungseigenschaften auf; auf der chemischen Ebene haben wir Eigenschaften wie Brennbarkeit oder Viskosität; auf der zellulären (biologischen) Ebene kämen Aktivitäten und Funktionen wie Stoffwechsel und Reproduktion hinzu, während mentale Eigenschaften wie das Vorhandensein von Bewusstsein auf der

(psychischen) Ebene der höheren Organismen zum Vorschein kommen. Häufig spricht man in diesem Zusammenhang von Mikro- versus Makroeigenschaften eines Systems, wobei die Grenze zwischen den Mikro- und Makrobezeichnungen nicht klar gezogen ist. Zumindest im philosophischen Diskurs verläuft die Trennlinie zwischen der Zuordnung der Mikro- versus Makro-Kennzeichnung typischerweise in etwa da, wo man klassischerweise auch zwischen Naturwissenschaften und Geistes- bzw. Sozialwissenschaften trennt. Die Phänomene, die von der Physik, Chemie und Biologie untersucht werden, betreffen Mikroebenen (bzw. Mikroeigenschaften), die Phänomene, die die Geistes- bzw. Sozialwissenschaften (in diesem Falle auch die Psychologie) untersuchen, betreffen dagegen Makroebenen (bzw. Makroeigenschaften).

Was macht nun also den Begriff der Ebene aus? Welche zentralen Bedeutungsfacetten beinhaltet er? Den Minimalbegriff einer Ebene dürfte die bloße Möglichkeit zur Ordnung von Entitäten in eine bestimmte (hierarchische) Abfolge ausmachen. Z.B sagt Craver (2007): „Its application requires only a set of items and a way of ordering them as higher or lower“ (S. 163). Und auch Kim (2002) betont als zentrales Charakteristikum des Ebenenbegriffs die Möglichkeit einer Ordnung von „Höher und Niedriger“ und fügt aber noch einen weiteren Gedanken hinzu: „The idea of ‘higher’ and ‘lower’ is essential to the levels picture: the levels are ordered from lower to higher, and as noted it is often thought that there is a bottom level, a level than which there is no lower“ (S. 3). Es handelt sich hierbei um die weit verbreitete Annahme, dass es so etwas wie eine *Basisebene* geben muss, von der nicht mehr gesagt werden kann, dass sie *über* einer anderen Ebene angesiedelt ist.

4.1.1. Begriffsgeschichtlicher Hintergrund

Die Auffassung, dass die Welt, ihre Objekte und auch das Psychische einer gewissen Schichtung unterliegen, reicht dabei zurück bis in die Antike (vgl. Hartmann, 1943, 1949). Mit dem Gegenwartsbegriff der Ebene haben diese Vorstellungen allerdings nicht viel gemeinsam, sondern beziehen sich eher auf ein „lockeres Nebeneinander“ bestimmter Strukturen oder (geistiger) Vermögen. Auch die Einführung des Ebenenbegriffs kann in diesem Zusammenhang als Antwort auf das Scheitern des Substanzdualismus angesehen werden (vgl. Kap. 3.1.1). Die Idee ist hier, dass man zwar akzeptiert, dass es keine *substantiell* verschiedenen Wirklichkeitsbereiche gibt, aber dass unsere Welt bzw. ihre Objekte dennoch eine Strukturgliederung in Form von verschiedenen *Ebenen* aufweisen, auf denen sich Entitäten oder Eigenschaften manifestieren. Den Unterschied zwischen dem Weltbild, das durch einen Substanzdualismus à la Descartes nahegelegt wird, und einem solchen, wie es für (die meisten) Ebenenmodelle der Wirklichkeit konstitutiv ist, erläutert Jaegwon Kim (1997) folgendermaßen:

Cartesian substance dualism pictures the world as consisting of two independent spheres, the mental and the material, existing side by side, each with its characteristic defining properties (consciousness and spatial extendedness respectively). There are causal interactions across the two domains, but entities in each domain, being 'substances', are ontologically independent of those of the other, and it is metaphysically possible for one domain to exist even in the absence of the other. What has replaced this picture of a bifurcated world is the familiar multilayered model that views reality as stratified into different 'levels', 'orders', or 'tiers', organized in a hierarchical structure. (S. 276)

Im Hinblick auf die Philosophie des Geistes und die Bestimmung der Gehirn-Geist-Relation ist das Aufkommen des Ebenenbegriffs vor allem mit der Emergenztheorie verbunden. Einer der ersten, der eine solche

Abfolge von Ebenen und verschiedener Eigenschaftsklassen explizit postulierte, war der Emergenztheoretiker Conwy Lloyd Morgan (1923). Er nahm eine Hierarchie an, die aus fünf Ebenen besteht und die gesamte natürliche Ordnung aller Dinge umfassen soll (vgl. Abb. 11).

Aus heutiger Sicht erscheint diese Hierarchie natürlich sehr unvollständig und die Sonderstellung des Menschen gegenüber Tieren etwas bizarr. Noch problematischer ist allerdings, dass es sich um keine Hierarchie handelt, bei der sich die höheren Entitäten *durchgängig* aus solchen der niedrigeren zusammensetzen (von Pflanzen bis zu Atomen ist dies gegeben, bei den höheren Ebenen allerdings nicht). Vielmehr hat Morgan eine Hierarchie einer allgemeinen (und nicht lokalen) Seinsordnung im Sinn, bei denen die Entitäten der höheren Ebenen Eigenschaften aufweisen, die die Entitäten der niedrigeren Ebenen nicht aufweisen und somit die höheren Ebenen durch Objekte bzw. Lebensformen konstituiert werden, die den niedrigeren „überlegen“ sind. Man spricht in diesem Zusammenhang auch von *systemischer Emergenz*, was späterer Stelle noch ausführlicher besprochen wird (vgl. Kap. 6.4.1.1).

Menschen
Tiere
Pflanzen
Moleküle
Atome

Abb. 11: Hierarchie natürlicher Ordnungsebenen nach Morgan (1923)

Ein ähnliches sechsstufiges Modell, das ebenfalls dem Anspruch genügen möchte, die natürliche Ordnung der Dinge in ihrer Gesamtheit zu umfassen, findet sich in dem einflussreichen Text von Paul Oppenheim

und Hilary Putnam (1958) „The Unity of Science as a Working Hypothesis“ (vgl. Abb. 12).

Soziale Gruppen
(Vielzellige) Lebewesen
Zellen
Moleküle
Atome
Elementarpartikel

Abb. 12: Hierarchie von Reduktionsebenen nach Oppenheim und Putnam (1958)

Anhand der Erweiterung des Modells durch die Elementarpartikel als niedrigster Ebene zeigt sich bereits der Fortgang der wissenschaftlichen Entdeckungen im 20. Jahrhundert. Der Mensch verliert außerdem seine Sonderstellung des Morganschen Systems und wird der Gruppe der (vielzelligen) Lebewesen zugeordnet. Außerdem fällt auf, dass Oppenheim und Putnam auch soziale Gruppen als eine eigenständige Ebene auffassen, während Morgan noch Individuen fokussiert. Was an dieser Stelle aber noch viel relevanter ist, ist, dass vor allem in der Interpretation der postulierten Ebenen große Unterschiede bestehen (vgl. Kim, 2002). Während Morgan als antireduktionistischer Emergentist annahm, dass auf den höheren Ebenen neuartige Charakteristika auftreten, die aufgrund der Beschaffenheiten der niedrigeren Ebenen *prinzipiell nicht vorhersehbar* sind, sind die von Oppenheim und Putnam genannten Ebenen als *Reduktionsebenen* aufzufassen, d.h., dass die spezifische Beschaffenheit als auch kausale Wirksamkeit der Entitäten und Theorien auf den höheren Ebenen auf die Entitäten und Theorien auf den niedrigeren

Ebenen (und letztendlich auf die Elementarpartikel) zurückgeführt werden müssen. Sie vertreten also bereits zu diesem Zeitpunkt einen so genannten *Mikroreduktionismus* und führen wahrscheinlich diesen Begriff überhaupt als Erste in die wissenschaftstheoretische Diskussion ein.

Einen prominenten psychologischen Ansatz, der – ähnlich wie bei der Annahme einer materiellen Hierarchie von Ebenen – von einer Schichtung des menschlichen *Geisteslebens* ausgeht, hat Erich Rothacker (1938/1966) entwickelt. Er geht von einer sukzessiven Entfaltung unseres geistigen Daseins aus, das mit einer relativ primitiven vital-emotionalen Ebene beginnt, worauf im Zuge der weiteren Entwicklung eine zunehmende Schichtung stattfindet, die in der Ausbildung unserer Person(schicht), die auch unser Denken und unser Gedächtnis beinhaltet, mündet. Einen ähnlichen Ansatz vertritt Nicolai Hartmann (1949). Während er einerseits bezüglich des „realen Seins“ zwischen aufsteigenden Schichten von Unorganischem, Leben, Seele und Geist differenziert, unterteilt er die Schicht des Geistes selbst in einen personalen, objektiven und den objektivierten Geist. Dabei umfasst der personale Geist alle individuellen Bewusstseinsvorgänge, der objektive Geist dient der Verankerung des personalen Geistes in historisch wirksamen Strukturen wie in Erzählungen, Sitten, Wissenschaften oder Recht. Mit objektiviertem Geist ist eine Ebene angesprochen, bei der ein geistiger Gehalt an ein Realgebilde gebunden ist (z.B. an ein bestimmtes Kunstwerk).

4.1.2. Wissenschaftstheoretische Begriffsbestimmung

Über solche eher extensional verfassten Bestimmungen hinaus hat sich insbesondere William Wimsatt (1976, 2007) an einer reichhaltigeren Bestimmung des Ebenenbegriffs versucht. Er ist überhaupt einer der wenigen Wissenschaftstheoretiker, die sich explizit mit dem Ebenenbegriff

bzw. der wissenschaftstheoretischen Bedeutsamkeit von Ebenenmodellen auseinandergesetzt haben.

In einer ersten Annäherung definiert Wimsatt (1976) Ebenen als „features of an abstract (property) space which did not interact with the systems which move through this space and fill it up in different ways“ (S. 239). Aus dieser eher metaphysischen Bestimmung geht zunächst hervor, dass Ebenen *Charakteristika eines abstrakten Eigenschaftsraumes* sein sollen. Gemeint ist hiermit, dass man die Wirklichkeit als einen Raum begreifen kann, der eine Vielfalt von Eigenschaften beinhaltet. Diese Eigenschaften lassen sich nun wiederum gruppieren und die Gruppen wiederum in eine bestimmte Reihenfolge bringen. Außerdem bringt Wimsatt mit dieser Definition zum Ausdruck, dass die Ebenen nicht mit den Systemen interagieren, die durch diesen Eigenschaftsraum verlaufen und ihn auf verschiedene Weisen ausfüllen. Die Ebenen selbst sind also keine kausal wirksamen Größen. Sie sind nur eine Art „Platzhalter“ für Ansammlungen von Eigenschaftsgruppen.

An späterer Stelle scheint sich Wimsatt (1976) seiner vorherigen metaphysischen Begriffsbestimmung einer Ebene entledigen zu wollen. Es sei demnach auch möglich, eine Ebene über die Entitäten und deren Interaktion zu bestimmen, ohne auf einen abstrakten Eigenschaftsraum zu rekurrieren:

If organisms and units of other sizes and properties [...] tend to be good detectors of organization, and thus 'congregate' there, and if the location of the stable, regular and predictable regions is itself fruitfully viewed as a product of what systems are in the neighborhood, then it seems reasonable to drop the introductory fiction of separate levels and their occupants, and to say that levels of organization just are those places (in space of properties) where the greatest densities of types of systems are to be found. This is also where one would expect to find the greatest density of types of interactions among these systems. (S. 239ff)

Aus dieser zweiten Annäherung an die Bestimmung des Gehalts des Ebenenbegriffs geht hervor, dass es möglich und auch vorzuziehen sei „to describe the levels as if they themselves are functions of what kinds of system are in the space and interacting with each other“ (S. 239). Gegenüber der ersten Begriffsbestimmung stellt Wimsatt die Bedeutung der Interaktion zwischen den Systemen oder Individuen in den Mittelpunkt: Die Entitäten, die sich auf einer Ebene befinden und diese hiermit überhaupt erst konstituieren, seien diejenigen, die am stärksten und häufigsten miteinander interagieren⁴⁴ (vgl. S. 215). Gemeinsam ist diesen Bestimmungsversuchen, dass sie Ebenen als *objektive* bzw. *natürliche* Organisationsschemata der Welt auffassen. Natürliche Organisationsebenen haben die strukturelle Beschaffenheit oder kausalen Abläufe von Objekten oder Lebewesen zum Gegenstand. Sie beziehen sich auf bestimmte Aspekte einer *natürlichen* Ordnung der Welt. Wie der Terminus suggeriert, sollen hiermit reale Strukturierungen oder natürliche Gliederungen gemeint sein, die wir in unserer Alltagswelt, aber auch durch verschiedene wissenschaftliche Methoden und Techniken *entdecken* können.

Man kann den Ebenenbegriff aber auch als einen *repräsentationalen* Ordnungsbegriff auffassen, d.h., dass wir mittels unserer Vorstellungskraft bzw. unseres Begriffsvermögens Relationen zwischen Entitäten oder Eigenschaften bestimmen können und so eine Hierarchie von analytischen

⁴⁴ Allerdings kann man den Punkt infrage stellen, dass die Interaktionen und Regelmäßigkeiten zwischen den Dingen verschiedener Größenordnungen weniger ausgeprägt sein sollen als die Interaktionen und Regelmäßigkeiten zwischen Dingen derselben Größenordnung (vgl. Craver, 2007). Wenn man die Beziehung zwischen den Ebenen als kompositorisch begreift, gibt es ganz klare Regelmäßigkeiten zwischen dem Ganzen und seinen Teilen: „The behavior of the whole is dependent on the behavior of the components in such a way that interventions to change the components can change the behavior of the whole and vice versa“ (ebd., S. 183). Hierzu an späterer Stelle mehr (vgl. Kap. 6.3.1 und 7.4).

oder Beschreibungsebenen erhalten. Diese sind nach Maßgabe der Dimension *konkret-abstrakt* geordnet. Entgegen natürlichen Organisationsebenen bedarf es hier lediglich der *gedanklichen* oder *konzeptuellen* Möglichkeit zur Ordnung von Entitäten in ein System des „Höher und Niedriger“. Churchland und Sejnowski (1994/1997) sprechen in diesem Zusammenhang von *analytischen* Ebenen, wenn bestimmte Konzepte ein Phänomen auf eine solche Art und Weise analysieren, dass eine (hierarchische) Abfolge im Sinne eines Systems von konkret zu abstrakt entsteht. In ähnlicher Weise könnte man auch von *theoretischen* Ebenen sprechen.

Analytische Ebenen kann man als Resultat wissenschaftlicher oder philosophischer Annäherungen an bestimmte semantische oder physische Gegenstände verstehen. Wenn sich ein Wissenschaftler oder Philosoph mit einem Gegenstand auseinandersetzt, muss er Begriffe bilden, mittels derer er sich auf diesen bezieht oder derer er sich bedient, um den Aufbau oder die Funktionsweise des Gegenstandes zu analysieren. Häufig sind hierbei verschiedene Formen der Theoriebildung möglich, die so beschaffen sind, dass sich die verschiedenen theoretischen Annäherungen an den zu untersuchenden Gegenstand in eine (hierarchische) Abfolge bringen lassen, sodass sich ein (wissenschaftliches) Ordnungssystem von Ebenen ergibt. Man spricht in diesem Zusammenhang insbesondere von Beschreibungs- oder Erklärungsebenen. Sie tragen dem Umstand Rechnung, dass man bestimmte Inhalte oder Gegenstände auf verschiedenen Granularitätsstufen deskriptiv erfassen kann.

Ein Beispiel wären die in Abbildung 13 dargestellten Beschreibungsebenen der Grammatik, die nach Maßgabe des linguistischen Strukturalismus als hierarchisches System (zur Analyse von Sätzen) verstanden werden können (Bloomfield, 1933/2000; Harris, 1951).

Bedeutungsebene
Satzebene
Wortebene
Lautliche Ebene

Abb. 13: Hierarchie an grammatikalischen Beschreibungsebenen

Auf der untersten, der *lautlichen Ebene*, kann ein Satz in seine kleinsten lautlichen Einheiten (Phoneme) zerlegt werden, während auf der *Wortebene* eine solche Zergliederung ein und desselben Satzes bezüglich seiner kleinsten bedeutungstragenden Einheiten (Morpheme) stattfinden kann. Auf der nächsthöheren, grobkörnigeren *Satzebene* kann eine Analyse der syntaktischen Struktur des Satzes erfolgen sowie auf der höchsten Ebene, der *Bedeutungsebene*, die seiner semantischen Form. Da es nur verschiedene Ebenen der *Analyse* von Sätzen sind, stehen diese Ebenen selbst nicht unbedingt in einer Teil-Ganzes-Beziehung (zumindest im Hinblick auf die Bedeutungsebene und Satzebene liegt keine Relation vor, bei der sich der Bezugsgegenstand der höheren Ebene aus dem Bezugsgegenstand bzw. den Bezugsgegenständen der niedrigeren Ebene zusammensetzt).

Es gilt dabei offenbar, dass einige Eigenschaften eines Objektes, die auf einer bestimmten Beschreibungsebene gegeben sind, auf anderen Beschreibungsebenen dieses Objektes fehlen. So hat Anscombe (1979) darauf hingewiesen, dass ein und derselbe Vorgang unter Maßgabe einer bestimmten (funktionalen) Beschreibungsebene als intentionale Handlung aufgefasst werden kann, während die Eigenschaft der Intentionalität auf anderen (z.B. physikalisch-chemischen) Beschreibungsebenen desselben Systems nicht vorhanden zu sein scheint.

4.2. Psychologische und kognitionswissenschaftliche Ebenenmodelle

Diese Konzeption mehrerer Ebenen der wissenschaftlichen Erklärung gehört zu den großen Neuerungen der Kognitionswissenschaft. (Varela, Thompson & Rosch, 1992, S. 67)

Auch wenn die Kognitionswissenschaft mit vielen anderen Neuerungen verbunden ist und der Ebenenbegriff auch schon zuvor verbreitet war, ist es – was das obige Zitat aufzeigen soll – gerade für die Kognitionswissenschaft in einem besonderen Maße charakteristisch, dass die von ihr untersuchten Modelle mentaler Zustände und Prozesse als etwas konzipiert sind, das mittels verschiedener funktionaler, computationaler oder konnektionistischer Ebenen vergegenständlicht bzw. auf ebensolchen untersucht werden kann. Eine Anbindung der – im Sinne analytischer Ebenen – formalen Modelle an die physischen Grundlagen der menschlichen Informationsverarbeitung hat dabei erst vergleichsweise spät stattgefunden.

4.2.1. Ebenenmodelle für kognitive Zustände und Prozesse

Im Folgenden wird die übliche Einteilung kognitionswissenschaftlicher Modelle übernommen, bei der zwischen informationsverarbeitenden und konnektionistischen Ansätzen unterschieden wird (vgl. z.B. Anderson, 2001). Die meisten wissenschaftlichen Modelle kognitiver bzw. funktionaler Architekturen, insbesondere diejenigen aus der Kognitionsforschung im Rahmen des Informationsverarbeitungsparadigmas, fallen (unmittelbar) in den Bereich analytischer Ebenen.

4.2.1.1. Informationsverarbeitende Modelle

Während nach der vehementen Abkehr von der introspektiven Psychologie in der ersten Hälfte des 20. Jahrhunderts einige Jahrzehnte lang nur noch Untersuchungen manifesten Verhaltens als wissenschaftlich respektabel angesehen wurden und Verhaltensbeobachtungen die einzige und damit „uniforme“ Analyseeinheit darstellten, sind mit dem Aufkommen der Computerwissenschaft und der Künstlichen-Intelligenz-Forschung im zunehmenden Maße informationsverarbeitende Modelle des Geistes in den Mittelpunkt der (kognitiven) Psychologie gerückt (z.B. Dörner, Lutz & Meurer, 1967). Die revolutionär neue Idee war, den Geist als etwas zu verstehen, das ähnliche Beschaffenheiten aufweist wie ein Computer. Also ein System, das in der Lage ist, bestimmte Symbole zu manipulieren, denen beliebige Bedeutungen zugewiesen werden können und das durch den Vorgang des Programmierens in die Lage versetzt werden kann, sich so zu verhalten, dass man semantische Inhalte in dieses System projizieren kann. Nach Ansicht von Allen Newell und Herbert A. Simon (1976) besteht die Essenz des menschlichen Denkens und der Intelligenz daher in der Fähigkeit, wie ein computationales System Symbole manipulieren zu können. Die „Blackbox“ des Behaviorismus wird also durch die Annahme interner Informationsverarbeitungsprozesse gefüllt.

Dies lässt sich gut am Beispiel des kognitionspsychologischen Modells des Gedächtnisses erläutern. Hier werden für gewöhnlich drei verschiedene Speichersysteme unterschieden: die sensorischen Register, das Kurzzeit- bzw. Arbeitsgedächtnis und das Langzeitgedächtnis. Das Arbeitsgedächtnis kann wiederum nach Baddeley (1986) in drei Teilsysteme zergliedert werden: die zentrale Exekutive und die zwei Hilfssysteme des räumlich-visuellen Notizblocks (zum Memorieren von Bildern) und der phonologischen Schleife (Struktur, in der so viel Information

gehalten werden kann, wie in einer bestimmten Zeiteinheit memoriert werden kann). Die zentrale Exekutive (bei der man altmodisch wohl auch von so etwas wie dem „Ich“ sprechen könnte) kontrolliert den Einsatz der beiden Hilfssysteme beim Memorieren. Sie kann Informationen in die Hilfssysteme überführen oder auch solche aus diesen abrufen. Ein solches Modell hat eine große Augenschein-Validität und kann sehr gute Vorhersagen über Gedächtnisleistungen machen, sodass es sich in der psychologischen Forschung und Praxis bewährt hat.

Die durch den Informationsverarbeitungsansatz geprägte Kognitionswissenschaft hat – zumindest in weiten Teilen – eine genauere Anbindung ihrer computationalen Modelle an die neurobiologischen Grundlagen für unnötig gehalten. Dies wird deutlich, wenn man ein klassisches Experiment der Kognitionswissenschaft von Saul Sternberg heranzieht, das mittlerweile aufgrund seines repräsentativen Charakters als *Sternberg-Paradigma* bezeichnet wird (Sternberg, 1966, 1969, vgl. Anderson, 2001, S. 12). Den Probanden wurde in einem ersten Schritt eine Menge von ein bis sechs Ziffern gezeigt, also z.B. „4, 8, 6“, die sie sich merken sollten. Darauf bekamen sie in einem zweiten Durchgang eine Zahl präsentiert (z.B. die „6“), bezüglich der sie schnellstmöglich entscheiden sollten, ob sie sich unter der zu merkenden Zahlenmenge befand. Sternberg konnte hierbei feststellen, dass eine annähernd lineare Beziehung zwischen dem Umfang der zu bearbeitenden Zahlenmenge und der Beurteilungszeit besteht (für jede weitere Ziffer verlängerte sich die Reaktionszeit im Mittel um ca. 38 Millisekunden). Entscheidend ist nun der von Sternberg vertretene Erklärungsansatz seiner Beobachtungen: Er nahm an, dass die Probanden zur Bewältigung der Aufgabe eine bestimmte Anzahl von Informationsverarbeitungsschritten, wie sie in Abbildung 14 dargestellt sind, durchlaufen mussten (Enkodierung der Testzahl, Vergleich mit den vorgegebenen Zahlen, Entscheidungsfindung, Generierung einer Ant-

wort) und das hierfür eben eine Zeit von ca. 38 Millisekunden pro hinzukommender Ziffer vonnöten ist.

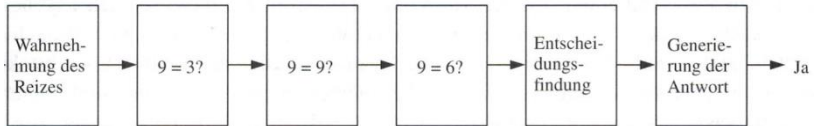


Abb. 14: Abfolge von Informationsverarbeitungsschritten nach Sternberg (Aus: Anderson, 2001, S. 13)

Hier zeigt sich die für die frühe Kognitionswissenschaft charakteristische Erklärungsweise ganz deutlich: Abstrakte Informationsverarbeitungsprozesse werden zur Erklärung der Beobachtungen und Vorhersage zukünftigen Verhaltens postuliert, ohne dass ein Zusammenhang zu Vorgängen im Gehirn hergestellt wird. Viele Forscher waren überzeugt, dass es der Herstellung eines solchen Zusammenhangs grundsätzlich nicht bedarf, da – und dies wurde als entscheidend angesehen – die kognitionswissenschaftlichen Modelle sehr erfolgreich bei der Vorhersage von Verhalten sind.

Ein weiteres Beispiel aus der differentiellen Kognitionsforschung ist das *Drei-Schichten-Modell der Intelligenz* von Carroll (1993). Carroll integriert verschiedene Intelligenzkonzepte zu einem hierarchischen Gesamtmodell. Die allgemeine Intelligenz g (nach Spearman, 1927) beeinflusst eine Reihe von Fähigkeiten auf mittlerer Ebene, zu denen u.a. die kristalline und fluide Intelligenz (Cattell, 1987) gehören. Die weiteren Fähigkeiten auf dieser Ebene können als Abstraktionen der sieben Primärfaktoren von Thurstone (1938) angesehen werden, also: Wortflüssigkeit, Sprachverständnis, schlussfolgerndes Denken, räumliches Vorstellungsvermögen, Rechenfertigkeit, Merkfähigkeit und Wahrnehmungsgeschwindigkeit.

keit. Sie umfassen bei Caroll u.a. Lernen und Gedächtnis sowie die visuelle und auditive Wahrnehmung. Die Faktoren mittlerer Ordnung beeinflussen wiederum jeweils eine Vielzahl von elementaren kognitiven Prozessen, wie z.B. die Gedächtnisspanne (bei „Lernen und Gedächtnis“) oder die Diskriminationsfähigkeit sprachlicher Laute (bei „Auditive Wahrnehmung“). Abbildung 15 zeigt einen Ausschnitt dieses Modells.

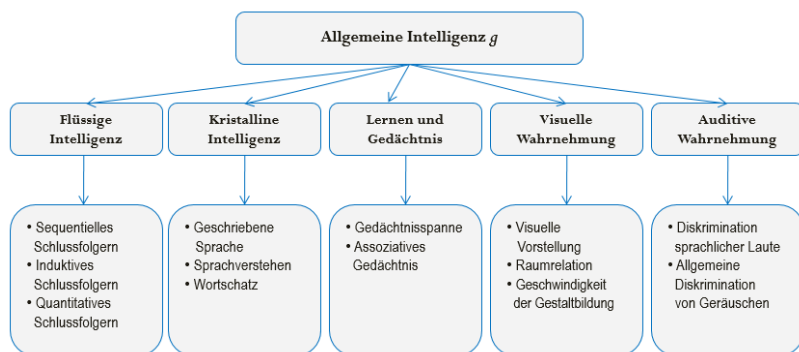


Abb. 15: Ausschnitt aus dem Drei-Schichten-Modell der Intelligenz nach Caroll (1993)

Zwischen den Ebenen wird hier allerdings keine strikte mereologische Relation angenommen, sondern nur so etwas wie eine lockere *Top-down-Beziehung*, die besagt, dass z.B. die allgemeine Intelligenz die fluide Intelligenz beeinflusst, die wiederum mit Mikroaspekten wie sequentielles, induktives oder quantitatives Schlussfolgern in Verbindung gebracht werden kann. Das Modell tätigt dabei keine Aussage zur neuronalen Implementierung der postulierten Intelligenzfacetten bzw. Informationsverarbeitungsprozesse.

Es lässt sich festhalten, dass die (meisten) Modelle des Informationsverarbeitungsansatzes der Kognition die neuronalen Basiseigenschaften nicht direkt in die Theoriebildung einbeziehen. Man könnte sagen, es kommt

zu keiner „wechselseitigen Durchdringung“ der postulierten kognitiven Modelle und der neuronalen Wirklichkeit. Wenn dem so ist, wie lässt sich die Beziehung dann aber näher bestimmen, die zwischen den neuronalen Strukturen und Prozessen, die unserem Denken zugrunde liegen und den kognitiven Modellen besteht? Bereits in der frühen kognitionswissenschaftlichen Literatur stößt man hierbei auf die Vorstellung, dass es sich hierbei um verschiedene *analytische Ebenen* eines Systems handelt: So wie man wie bei einem Computer zwischen Hardware vs. Software unterscheiden kann, sei es auch im Falle des menschlichen Geistes möglich, verschiedene Beschreibungsebenen heranzuziehen. Oder wie Cole (1990) es ausdrückt: „If the computer could be fruitfully investigated and described on two levels – one physical, the other functional – why not the mind itself?“ (S. 3). Bereits Herbert Simon (1969/1996) geht bei seiner Analyse sozialer, biologisch-physikalischer und auch funktionaler Systeme von hierarchischen Systemen aus, die sich aus verschiedenen aufeinander bezogenen Subsystemen zusammensetzen: „By a hierarchic system, or hierarchy, I mean a system that is composed of interrelated subsystems, each of the latter being in turn hierarchic in structure until we reach some lowest level of elementary subsystem“ (ebd., S. 184).

Ein bedeutendes computationales Modell stammt von Marr (1982), bei dem er zur Erklärung der visuellen Wahrnehmung explizit auf ein hierarchisches System von Ebenen zurückgreift (vgl. S. 25). In Anlehnung an Modelle aus der Informatik umfasst es drei Ebenen (vgl. Abb. 16).

Computationale Ebene
Algorithmische Ebene
Physische Implementationsebene

Abb. 16: Ebenen des visuellen Prozessierens nach Marr (1982)

Auf der obersten, *computationalen* Ebene findet eine abstrakte Problemanalyse statt, bei der eine zu bewältigende Wahrnehmungsaufgabe in ihre Hauptbestandteile zerlegt wird. Ein Beispiel wäre die (rechnerische) Bestimmung der dreidimensionalen Tiefe von Objekten aufgrund des zweidimensionalen Musters auf der Retina. Auf der mittleren, *algorithmischen* Ebene wird eine zur Durchführung der Aufgabe formale Prozedur spezifiziert, die es ermöglicht, dass auf eine konkrete Eingabe in das System auch eine konkrete Ausgabe erfolgt. Es ist in diesem Fall die formale Umsetzung der für die Generierung der dreidimensionalen Tiefenstruktur benötigten Berechnungen. Auf der untersten Ebene haben wir schließlich die konkrete *physische Implementation* der Prozesse der höheren Ebenen.

Man kann sagen, dass diese Ebenen eines kognitiven Systems nach Ansicht Marrs in einer Art *Realisierungsbeziehung* zueinander stehen sollen (vgl. Kap. 6.5). Die höhere Ebene wird durch die jeweils niedrigere Ebene auf eine bestimmte Art und Weise konkretisiert. So können die auf der *computationalen* Ebene ablaufenden Berechnungen durch verschiedene formale Algorithmen realisiert werden, wie auch die formale Prozedur zur Aufgabendurchführung der *algorithmischen* Ebene auf verschiedene Weise physisch implementiert werden kann. Dies entspricht dem Vorgehen eines IT-Experten, der eine bestimmte PC-Anwendung (z.B. ein Musikerkenntnisprogramm) programmieren möchte. Er muss sich zunächst fragen, welche Aufgabe die Software genau vollbringen können soll und muss diese in möglichst viele Teilbereiche zerlegen (es bedarf in dem Fall wohl eines Features zur Tonfrequenzanalyse und einer *computationalen* Abbildung der Tonfrequenzcharakteristika eines bestimmten Fundus an Musiktiteln). Daraufhin muss er sich fragen, wie die für den Programmzweck nötigen Berechnungen formal realisiert werden können. Nach welchen Grundprinzipien sollen also die Tonfrequenzanalyse-

technik und die abgespeicherten Charakteristika von Musikstücken in Verbindung gebracht werden? Wenn der Programmierer dieses Problem gelöst hat, bedarf es natürlich noch einer konkreten technischen Grundlage, also einer Hardware, wie die eines Smartphones oder eines PCs, auf der die neue Software dann auch tatsächlich ausgeführt werden kann.

Marr betrachtet die Ebenen insgesamt als relativ unabhängig voneinander. So nimmt er an, dass die Analyse von Problemen auf den höheren Ebenen weitestgehend unabhängig von der Einbeziehung der darunter liegenden Ebenen erfolgen kann. Die computationalen Aspekte der abstrakten Problemanalyse (Ebene 3) können also auch ohne eine Entschlüsselung des Algorithmus (Ebene 2) untersucht werden, der die Berechnungsschritte durchführt. Ebenso sei es möglich, dass man die Frage nach einem bestimmten Algorithmus (Ebene 2) löst, ohne dass man die physische Implementationsebene (Ebene 1) einbezieht. Auch hier findet sich also die Doktrin des Funktionalismus wieder, dass die menschliche Kognition auch ohne den Einbezug neuronaler Basisstrukturen und -prozesse untersucht werden kann bzw. dass es sich hierbei nicht einmal um eine lohnenswerte Forschungsstrategie handeln würde. Wenn Clark (1989/2007) in Folge dessen sagt, dass dies „das Wesen des Denkens auf einer Ebene an[siedelt], die unabhängig von dem physischen Stoff ist, durch das das denkende System konstituiert ist“ (S. 402), wird er hiermit nicht meinen, dass die Position des Funktionalismus oder Kognitivismus einen (Substanz-)Dualismus von Gehirn und Geist mit sich bringen würde. Er stellt vielmehr einfach fest, dass die konkrete *Art* der stofflichen Grundlagen nach Maßgabe des Funktionalismus keine wichtige Rolle spielt. Ob es sich um ein Gehirn, eine Ansammlung von Siliziumchips oder eine völlig andere physische Basis handelt, die mentale Eigenschaften realisiert, ist demnach nicht entscheidend, sondern nur, dass jedes dieser Systeme eine bestimmte *funktionale* Struktur aufweist.

Man spricht in diesem Zusammenhang auch von der *multiplen Realisierbarkeit* mentaler Eigenschaften, was an späterer Stelle im Zusammenhang der hier entwickelten identitätstheoretischen Position ein wichtiger Diskussionsgegenstand wird (vgl. Kap. 6.1.4.2 und 7.5).

4.2.1.2. Konnektionistische Modelle

Die Doktrin der (forschungsbezogenen) Unabhängigkeit psychologischer Modellierungen von den neurobiologischen Grundlagen mentaler Phänomene, wie für das „Sternberg-Paradigma“ kennzeichnend ist, wurde zunehmend kritisiert (vgl. z.B. James Anderson, 1973), da die Verbindung zwischen den abstrakten computationalen Modellen und den konkreten physischen Strukturen, die nun mal in enger Verbindung zu mentalen Prozessen stehen, im Unklaren bleibt. Wenn die physische Architektur des Menschen bei der Etablierung einer kognitiven Architektur dauerhaft nicht berücksichtigt wird, kann dies die Validität der postulierten Modelle bzw. Konstrukte in Zweifel ziehen. „Während das Sternberg-Modell die Daten adäquat voraussagt, stellt sich die Frage, ob das Gehirn tatsächlich auf *diese* Art und Weise eine Aufgabe bewältigt“ (John Anderson, 2001, S. 15; Hervorh.: M.L.). Hierbei handelt es sich um einen Umstand, aus dem Sejnowski und Churchland (1991) eine grundsätzliche Abwegigkeit der Ausblendung der neurobiologischen Forschungsebene ableiten: „It would be convenient if we would understand the nature of cognition without understanding the nature of the brain itself“ (S. 343). Offensichtlich stellen die konkreten neuronalen Strukturen doch stärkere *Constraints*, also einschränkende Vorannahmen, für die Gültigkeit von solchen Modellen dar, als es von den klassischen Informationsverarbeitungstheoretikern angenommen wurde. Wenn man Churchland und Sejnowski (1994/1997) Glauben schenkt, wurde Marr ironischerwei-

se selbst in seiner Arbeit in hohem Maße durch neurobiologische Fakten beeinflusst. „Fakten, die durch die physikalische Ebene vorgegeben waren, wirkten sich auf seine Auswahl an Problemen aus und lieferten die Grundlagen für seine Erkenntnisse auf der Spezifikations- [= Computations-, M.L.] und der algorithmischen Ebene“ (S. 25).

Vor dem Hintergrund der sich in den 1990er Jahren etablierenden kognitiven Neurowissenschaft ist daher zunehmend das *konnektionistische* Paradigma in den Mittelpunkt der Aufmerksamkeit gerückt, das mit Modellen künstlicher *neuronaler Netzwerke* arbeitet. Anstatt von eher linear verlaufenden Informationsverarbeitungsschritten auszugehen (wie im Sternberg-Modell) werden hier Modelle konzipiert, die mit einer dem Gehirn nachempfundenen ausgedehnten *Parallelverarbeitung* distribuerter Informationsverarbeitung operieren (vgl. Rumelhart & McClelland, 1986; Rumelhart, 1989/ 1991).

When we look at essentially different architecture, we see that the architecture makes a good deal of difference. It is the architecture that determines which kinds of algorithms are most easily carried out on the machine in question. It is the architecture of the machine that determines the essentially nature of the program itself. [...] The basic strategy of the connectionist approach is to take as its fundamental processing unit something close to an abstract neuron. [...] The operations in our models then can best be characterized as ‘neurally inspired’. (Rumelhart, 1989/ 1991, S. 134f)

Den von Rumelhart angesprochenen Ansatz einer *Architektur des Geistes* greifen auch Simon und Kaplan (1989/1991) auf und entwickeln den hierarchischen Grundgedanken Simons zu einem hierarchischen Ebenenmodell weiter:

The fundamental design specifications of an information-processing system are called its *architecture*. The components of the architecture represent the underlying physical structures but only abstractly. For example, an architecture for modeling

the human brain might contain neurons as components, but the neurons might be characterized quite grossly as binary on-off elements with certain switching speeds. Another architecture might characterize the brain even more aggregately with units such as long-term memory, short-term memory, sensory organs, and so on. (Simon & Kaplan, 1989/1991, S. 7; Hervorh. im Orig.)

Eine solche Architektur umfasst demnach eine bestimmte Anzahl an *Abstraktionsebenen*, d.h., dass die konkrete Ebene der neuronalen Strukturen und Prozesse auf eine abstrakte Weise modelliert werden soll. Die einzelnen Modelle weisen dabei einen unterschiedlichen Grad an Abstraktion von der neuronalen Wirklichkeit auf und konstituieren so eine Abfolge an Abstraktionsebenen (von On-Off-Elementen mit verschiedenen Schaltgeschwindigkeiten bis hin zu symbolischen Strukturen wie den drei Elementen des Arbeitsgedächtnisses).

Der Konnektionismus ist allerdings keine „Erfindung“ der 1980er oder 1990er Jahre. Er ist ein Teilgebiet der bereits deutlich früher entstandenen *Kybernetik* und hat das Verhalten von Systemen, die durch eine Vernetzung künstlicher Informationsverarbeitungseinheiten gekennzeichnet sind, zum Gegenstand. Das Verhalten des Gesamtsystems ergibt sich unter diesen Voraussetzungen immer aus dem Verhalten einer Vielzahl von Systemkomponenten, die miteinander interagieren. Bereits in den 1940er Jahren haben Warren McCulloch und Walter Pitts (1943, 1947) wegweisende Arbeiten zu neuronalen Netzen publiziert, mittels derer sie zu erklären versuchten, nach welchen kybernetischen Prinzipien das menschliche Gehirn funktioniert. Sie konnten dabei nachweisen, dass Programme, die durch eine so genannte Turing-Maschine⁴⁵

⁴⁵ Bei der Turing-Maschine handelt es sich um eine abstrakte Rechenmaschine, die von dem britischen Mathematiker und Logiker Alan M. Turing eingeführt wurde und mittels derer durch eine Abfolge von Operationen (mathematische) Berechnungen vorgenommen werden können. Eine gute Darstellung findet sich bei Kim (1996/1998, S. 89ff), der sich wiederum auf Davis (1958) bezieht.

berechenbar sind, auch durch ein endliches Netzwerk künstlicher Neuronen berechnet werden können. Lange Zeit vor Rumelhart hat McCulloch (1965/2000) dabei das Neuron als die zentrale logische Einheit der Informationsverarbeitung in unseren Gehirnen herausgestellt. Des Weiteren konnte Dörner (1974) neuronale Netze über Computersimulationen gewissermaßen zum „Denken“ bringen.

Neuronale Netzwerke sind aus Einheiten zusammengesetzt, die sich zu einem komplexen Geflecht an Eingabe-, Ausgabe- oder vermittelnden Schichten verdichten, zwischen denen wiederum modulierende Verbindungen bestehen. In Abbildung 17 finden sich zwei Beispiele für Netzwerke mit parallelverteilter Informationsverarbeitung. Generell gilt, dass Input-Einheiten entweder direkt oder indirekt mit Output-Einheiten verknüpft sind. In geschichteten Netzwerken (linke Seite) gibt es keine Rückkopplung von einer „späteren“ zu einer „früheren“ Ebene, während in rekurrenten Netzwerken (rechte Seite) einige Einheiten des Schaltkreises mit früheren Einheiten verbunden sind (vgl. Kupfermann, 1996, S. 367).

Die Stärke jeder Verbindung kann hierbei unterschiedlich sein. In einem Versuch bekommen die Input-Einheiten jeweils ein Aktivierungsniveau (einen Wert zwischen 0 und 1). Die Aktivierung jeder Output-Einheit ist dann durch die Aktivitätssumme der Input-Einheiten festgelegt. Der Grad der Aktivierung, der von einer Input-Einheit geliefert wird, ist eine Funktion aus dem Produkt des Aktivierungsniveaus der Input-Einheit und dem Maß der Verbindungsstärke zur Output-Einheit. Ein geschichtetes Netzwerk besteht mindestens aus einer Eingabe- und einer Ausgabeschicht. Ein geschichtetes Netzwerk kann auch Schichten aus „versteckten Einheiten“ (*hidden units*) enthalten [Abbildung auf der linken Seite; M.L.]. Die Eingabeschicht ist in diesem Beispiel eine Fläche wie die Retina. Sie wurde hier mit dem Buchstaben C (Anordnung der dunklen Einheiten) stimuliert. Die Ausgabeschicht hat vier Möglichkeiten: a, b, c oder d. Durch Trainieren des Netzwerks kann die Verbindungsstärke zwischen den Einheiten so justiert werden, daß die Präsentation eines Buchstabens hauptsächlich die entsprechende Output-Einheit aktiviert. [...] In einem rekurrenten Netzwerk [Abbildung auf der rechten

Seite; M.L.] sind einige Einheiten des Schaltkreises mit früheren Einheiten rückgekoppelt (dargestellt durch die gebogenen Pfeile). (ebd.)

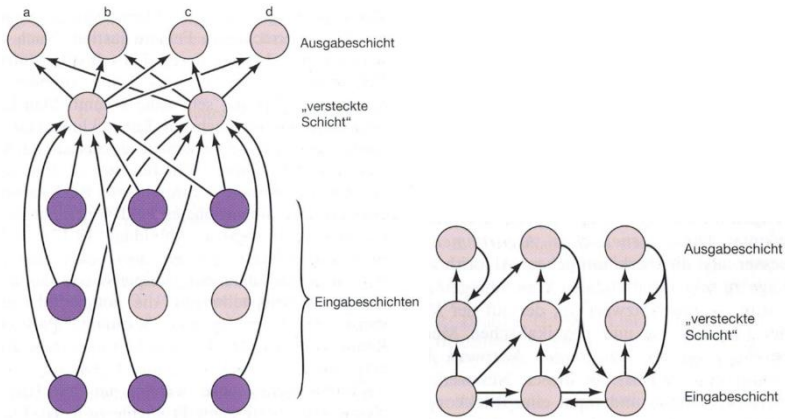


Abb. 17: Geschichtetes Netzwerk mit versteckten Einheiten (links); rekurrentes Netzwerk (rechts) (Aus: Kupfermann, 1996, S. 367)

Wir haben hier also Verhältnisse, die dem Aufbau unseres Nervensystems stärker nachempfunden sind (obgleich sie natürlich keine exakte Abbildung dessen sind, was tatsächlich in unseren Gehirnen geschieht). So lassen sich die Eingabeschicht mit den Sinneszellen, die Ausgabeschicht mit den Motoneuronen und die vermittelnden Schichten mit allen zwischengeschalteten Nervenzellen in Verbindung bringen. Die modulierenden Pfade zwischen den Einheiten sollen in erster Linie den durch Transmittersubstanzen gesteuerten Neuromodulationsvorgängen im synaptischen Spalt zwischen zwei Neuronen entsprechen. Entgegen den computationalen Modellen, bei denen es lediglich auf eine gute Vorhersagbarkeit von Verhalten ankam, wird nun auch die physische Implementationsebene des Gehirns als Constraint in die Theoriebildung mit

einbezogen. So hat James Anderson (1973) für das von Sternberg beschriebene Modell des Lern- bzw. Gedächtnisvorganges beim Lernen einer Ziffernliste herausgestellt, dass das, was hier tatsächlich in unseren Köpfen geschieht, besser durch ein neuronales Netzwerkmodell beschrieben werden kann:

When reaction times can be interpreted as showing an "exhaustive scan" of the learned list, the form taken by the filter is assumed to be an "interactive memory." An interactive memory is the vector sum of vectors of high dimensionality representing the items on the learned list. It is assumed that attentional mechanisms program a set of elements (perhaps neurons or groups of neurons) with the interactive memory, and when a test item is presented, the interactive memory and the input to the set of elements are multiplied together at the elemental level. (S. 417)

Auch für die konnektionistisch gefärbte Kognitionswissenschaft ist der Rekurs auf Ebenenmodelle kognitiver Zustände charakteristisch. Die spezifischen kognitiven Architekturen sollen dabei die tatsächlichen zugrunde liegenden physischen Strukturen weiterhin in einer abstrakten Form widerspiegeln, wenn auch die neuronalen Netzwerkmodelle den tatsächlichen physischen Vorgängen stärker entsprechen werden als die Modelle auf der Basis des klassischen computationalen Ansatzes:

The connectionist models considered are massively parallel numerical computational systems that are a kind of continuous dynamical system. The numerical variables in the system correspond semantically to fine-grained features below the level of the concepts consciously used to describe the task domain. The level of analysis is intermediate between those of symbolic cognitive models and neural models. (Smolensky, 1988, S. 11)

Es ergibt sich demnach wiederum eine dreigliedrige Hierarchie, in der das basale (niederstufige) menschliche Nervensystem auf höherstufigen Ebenen modelliert wird; die konnektionistischen Modelle lassen sich da-

bei auf einer Ebene ansiedeln, die sich zwischen den neuronalen Strukturen und Prozessen und der Ebene der „symbolischen“ Informationsverarbeitung befindet. Sie erfüllen hier also die Funktion, die zuvor bei Marr durch einen relativ abstrakten Algorithmus erfüllt worden ist. Es kann also das folgende hierarchische Modell zur Beschreibung kognitiver Zustände abgeleitet werden (vgl. Abb. 18).

Computationale Ebene
Konnektionistische Ebene
Physische Implementationsebene

Abb. 18: Computational-konnektionistisches Modell

Auf der physischen Implementationsebene haben wir die neuronalen Strukturen, die an der Realisierung von Gedächtnisleistungen beteiligt sind (also z.B. den Hippocampus oder den frontalen Kortex). Auf der konnektionistischen Ebene, befindet sich ein mehrschichtiges neuronales Netzwerkmodell, das bei einem bestimmten Eingabemuster (also der Einspeisung von Informationen in das Netzwerk), ein Ausgabemuster oder Verhalten zeigt, das unseren Erwartungen und Vorhersagen zu einem durch Gedächtnisprozesse beeinflussten System entspricht. Es soll die Vorgänge auf der neuronalen Ebene bestmöglich abbilden. Auf der computationalen Ebene finden wir z.B. das Arbeitsgedächtnis mit seinen drei Teilsystemen vor. Als ein Kriterium für die Gültigkeit der psychologischen Konzeptionen auf den höherstufigen Ebenen kann nun – neben der guten Vorhersagbarkeit des Systemverhaltens – auch berücksichtigt werden, inwieweit sie die konkrete Beschaffenheit der niederstufigen Ebenen (also der neuronalen Constraints bzw. der diese abbildenden konnektionistischen Modelle) widerspiegeln.

4.2.2. Ebenenmodelle für phänomenale Zustände und Prozesse

Auch in Bezug auf phänomenale Zustände, also mentale Ereignisse mit einer spezifischen Qualität des Erlebens wie Schmerz- oder Farbbeimpfindungen, Emotionen usw., finden sich in der aktuellen psychologischen Fachliteratur viele Beispiele für Modelle, die explizit auf Ebenen zurückgreifen. Dies liegt zum einen daran, dass neben der Kognitionspsychologie auch die anderen Bereiche der (Allgemeinen) Psychologie eine zunehmend stärkere Anbindung an biologische oder neurologische Daten suchen und daher kaum umhin kommen, mehrere Forschungszugänge umspannende Theorien zu formulieren. Zum anderen dürften aber auch die zuvor skizzierten Modellvorstellungen für die kognitiven Zustände, die mit Ebenen operieren und in der Psychologie großes Gehör fanden, eine gewisse Vorbildfunktion für die Analyse von phänomenalen Zuständen gehabt haben.

Beispielsweise definieren Pauli und Birbaumer (2000) *Emotionen* mit Bezug auf Lang (1993) als „Reaktionsmuster auf positiv verstärkende oder aversive körperexterne oder -interne Reize, die auf den drei *Ebenen* des Verhaltens ablaufen [...]: der motorischen, der physiologischen und der subjektiv-psychologischen Verhaltensebene“ und führen darüber hinaus aus, „dass psychophysiologische Verhaltenstheorien – historisch betrachtet – geprägt [sind] von gegensätzlichen Auffassungen über die Wechselwirkung zwischen diesen Verhaltensebenen⁴⁶“ (S. 75, Hervorh.: M.L.). Bei Schandry (2006) heißt es ganz ähnlich: „Eine Emotion ist i.Allg. gekennzeichnet durch drei Elemente: (1) einen bestimmten Funktionszustand des Gehirns; (2) eine typische subjektive Befindlichkeit; (3) charak-

⁴⁶ Gemeint ist hiermit die langwierige Debatte darüber, ob das Auftreten eines Gefühls die Ursache (alltagspsychologische Vorstellung) oder die Folge (James-Lange-Theorie) von bestimmten physiologischen Vorgängen ist oder ob Gefühl und physiologische Reaktionen gleichzeitig erfolgen (Cannon-Bard-Theorie) (vgl. z.B. Pinel, 1997, S. 460).

teristische Vorgänge auf der körperlichen *Ebene*. Sowohl auf der subjektiven als auch auf der körperlichen *Ebene* ist das Geschehen in der Regel spezifisch für die Art der Emotion“ (S. 493; Hervorh.: M.L.). Weiterhin beschreibt er, dass man Emotionen auf vier verschiedenen Ebenen untersuchen kann, auf der Ebene der Gehirnprozesse, auf der Ebene der vegetativen und hormonellen Prozesse, auf der Ebene der motorischen Reaktionen sowie auf der Ebene der eigentlichen subjektiv erlebten Gefühle.

Wenn man die Unterschiede zwischen den einzelnen Ansätzen an dieser Stelle außer Acht lässt, zeichnet sich ein Bild von Emotionen ab, das zumindest drei Ebenen umfasst (vgl. Abb. 19).

Subjektiv-psychologische Ebene
(Funktional-)neuronale Ebene
Physiologische Ebene

Abb. 19: Ebenenmodell der Emotionen

Auf der *physiologischen Ebene* laufen im Emotionsprozess verschiedene vegetative, kardiovaskuläre, muskuläre und hormonelle Vorgänge ab. Wenn auch die physiologische Erregung insgesamt eher diffus und unspezifisch ist, konnte dennoch gezeigt werden, dass es für die verschiedenen Basisemotionen durchaus unterschiedliche autonome Reaktionsmuster gibt (Ekman, Levenson & Friesen, 1983). Bezüglich der (*funktional-)neuronalen Ebene* hat die Forschung gezeigt, dass vor allem verschiedene Teile des limbischen Systems (wie die Amygdala, der Hippocampus, der cinguläre Kortex und die Insula) und des präfrontalen Kortex am emotionalen Geschehen beteiligt sind (vgl. Schandry, 2006). Schließlich haben wir auf der *subjektiv-psychologischen Ebene* zum einen die erlebten

Gefühle, die eine ganz spezifische phänomenale Qualität aufweisen, d.h. sich auf eine ganz bestimmte Art und Weise „anfühlen“. So hat das Erleben von Freude eine völlig andere phänomenale Qualität als z.B. das Erleben von Ärger. Zum anderen können aber auch kognitive Bewertungsprozesse für den Emotionsprozess relevant sein, insofern sie an der Auslösung und Regulation der Emotion beteiligt sind. Allerdings haben sie offenbar nicht den Stellenwert einer *notwendigen* Emotionskomponente, den ihr die frühen Emotionstheorien (z.B. die Zwei-Faktoren-Theorie nach Schachter & Singer, 1962) zugeordnet haben, denn gemäß der aktuellen Forschungslage können Emotionen auch „ohne Attributionsbedürfnisse, ohne bewusste Klassifikationsprozesse und sogar gegen die bewusste Einstellung und Ursachen entstehen“ (Pauli & Birbaumer, 2000, S. 77).

Ähnliche, verschiedene Ebenen umspannende Vorstellungen wie zu Emotionen lassen sich auch im Hinblick auf die anderen von der Psychologie untersuchten phänomenalen Zustände (wie Schmerzen oder sensorische Empfindungen) vorfinden.

Es muss an dieser Stelle aber problematisiert werden, dass die „Emotionsebenen“ keinen eindeutig hierarchischen Charakter haben, so wie er bei den Modellen für kognitive Zustände angenommen wird. Je nach spezifischem Einzelfall und konsultierter Theorie kann es ein „Übereinander“ oder auch „Nebeneinander“ der (funktional-)neuronalen und der physiologischen Ebene geben. So gibt es Ansätze, die sich damit befassen, aufzuzeigen, welchen Einfluss physiologische Prozesse auf die neuronale Informationsverarbeitung haben, die wiederum unseren phänomenalen Zuständen zugrunde liegt. In diesem Fall wäre es möglich, ein hierarchisches (lokal bestimmtes) Verhältnis anzunehmen. Es gibt aber ebenso Modelle, die davon ausgehen, dass ein phänomenaler Zustand sowohl durch die mit ihm assoziierten neuronalen Vorgänge als auch durch ein

bestimmtes Ausmaß an physiologischer Erregung bestimmt wird. Es handelt sich dann nicht um eine Hierarchie von Ebenen im Hinblick auf ein lokales Ereignis, sondern vielmehr um so etwas wie (mehr oder weniger gleichrangige) *Einflussfaktoren* auf die Hervorbringung bzw. Ausprägung von phänomenalen Prozessen.

4.2.3. Fazit

Man kann also festhalten, dass sowohl die Arbeiten von Simon zu hierarchisch verfassten (informationsverarbeitenden) Systemen, Marrs Modell der Analyseebenen zur visuellen Wahrnehmung als auch die neueren konnektionistischen Modellvorstellungen einen nachhaltigen Einfluss auf die (zumindest implizit vorhandenen) ontologischen und epistemologischen Grundlagen der Kognitionsforschung ausgeübt haben. Das Ebenenmodell kann als eine allgemein hin akzeptierte Hintergrundannahme in der heutigen kognitiven Neurowissenschaft und in weiten Teilen der Psychologie aufgefasst werden und findet sich in einer Vielzahl von Forschungsarbeiten und insbesondere Arbeiten wieder, die sich mit einer biologischen Fundierung der Kognitionswissenschaften auseinandersetzen (z.B. Lycan, 1987/ 2007; Clark, 1989/2007; Pylyshyn, 1989/1991; Sejnowski & Churchland, 1991). Aber auch im Zusammenhang der Konzeptualisierung phänomenaler Eigenschaften wie Emotionen kommt er vielfach zur Anwendung, obgleich er hier häufig etwas „zweckentfremdet“ gebraucht wird.

Auch wenn so wichtige Grundgedanken zu einer *Architektur des Psychischen* geliefert werden, bleiben dennoch viele Fragen ungeklärt. Zunächst wird in Modellen grundsätzlich offen gelassen, welchen *ontologischen* Status die einzelnen Ebenen haben und in welcher Beziehung sie genau untereinander stehen sollen. Sollen die postulierten Analyseebenen nur

bestimmte Perspektiven (bzw. Betrachtungs-ebenen) oder sprachlich-konzeptuelle Beschreibungsformen sein, wie sie im letzten Kapitel beschrieben wurden, die mehr mit dem Beobachter selbst als dem Beobachteten zu tun haben? Marr deutet für sein Modell eine Art Realisierungsbeziehung an, aber was genau bedeutet es, dass höhere Ebenen durch niedrige *realisiert* werden? Ist es nur eine Art *theoretische* Realisierung, in dem Sinne, dass eine Theorie in eine andere umgewandelt wird, ist es eine Art *mathematische* Realisierung, bei der „höherstufige“ mathematische Modelle mit relativ vielen Freiheitsgraden in niederstufige Modelle mit weniger Freiheitsgraden umgewandelt und damit konkretisiert werden? Haben die psychologischen Beschreibungen (wie z.B. die verschiedenen Komponenten des Arbeitsgedächtnisses) eine physische bzw. biologische Entsprechung oder sind es „lediglich“ theoretische Konstrukte? Wie verhält es sich des Weiteren mit psychischen Prozessen oder Zuständen, die eine bestimmte Erlebnisqualität haben? Kann diese phänomenale Qualität auf die Vorgänge oder die Verfasstheit der neuronalen Ebene zurückgeführt werden?

Des Weiteren kann – vor allem in Bezug auf den „klassischen“ Informationsverarbeitungsansatz – kritisiert werden, dass hier zumeist der Eindruck erweckt wird, dass es nur zwei Ebenen geben würde, deren Beziehung zueinander geklärt werden müsste: *die* strukturelle (oder physische) und *die* funktionale Ebene⁴⁷. Lycan (1987/2007) spricht in diesem Zusammenhang daher von einem *Zwei-Ebenen-Dogma* („Two-Levelism“, S. 449). In Wirklichkeit gibt es aber eine Vielzahl von Organisationsebenen, die bei der Konzeption von Modellen psychischer Eigenschaften potentiell

⁴⁷ Dass die funktionale Ebene selbst hierbei wie bei Marr (1982) oder auch Caroll (1993) nochmals unterteilt werden mag, ändert nichts an dieser grundsätzlichen Gegenüberstellungsweise nichts.

auf gewinnbringende Art und Weise mit einbezogen werden könnten (der Organismus insgesamt, Moleküle, Synapsen, Nervenzellen, neuronale Netzwerke, Gewebeschichten usw.). So stellt sich die Frage, ob es auf diesen verschiedenen Organisationsebenen relevante Mechanismen gibt, die für die Hervorbringung mentaler Eigenschaften und Prozesse relevant sind. Es fehlt ein integrativer Ansatz, der diese Ebenen in eine Beziehung setzt und aufzeigt, wie genau diese Beziehung gestaltet ist und wie genau die mentale Eigenschaft oder der mentale Prozess (der kognitions-wissenschaftlich beschrieben wird) mit diesen natürlichen Organisations-ebenen zusammenhängt bzw. ob das Psychische sogar selbst eine solche natürliche Organisationsebene bildet.

5. Die natürliche Organisation des Psychischen

I am suggesting, roughly, that there are special sciences not because of the nature of our epistemic relation to the world, but because of the way the world is put together. (Fodor, 1974, S. 113)

Die bisherige Leitfrage nach dem *Ursprung des Leib-Seele-Problems* wird in diesem Kapitel zu einem Abschluss gebracht. Mit der Einführung des Begriffs der Ebene, der Betrachtung von kognitionswissenschaftlichen Ebenenmodellen oder Architekturvorstellungen des Geistes ist eine wichtige konzeptuelle Weiterentwicklung getätigt worden, die selbst zwar noch keine abschließende Antwort auf die Frage nach der Möglichkeit einer „objektivistischen“ Auflösung von Spinozas Dilemma geben konnte, aber – wie sich im Folgenden zeigen wird – als äußerst fruchtbare begriffliche Grundlage für die Klärung ebendieser Frage abgibt.

Wir befinden uns in folgender Lage: Über ein bestimmtes Objekt liegen verschiedene theoretische Beschreibungen unterschiedlicher Granularität vor. Wenn nun die Frage aufgeworfen wird, inwieweit es sich hierbei um eine Widerspiegelung natürlicher Verhältnisse handelt, muss berücksichtigt werden, dass es sich – wie der philosophiegeschichtliche Abriss im Hinblick auf das Verhältnis von psychischen und physischen Aspekten gezeigt hat (vgl. Kap. 3.1) – nicht um Unterschiede handeln wird, die „substantieller“ Art sind. Wenn der Begriff der Ebene daher die Funktion haben soll, zu erklären, wie wir dennoch eine Gliederung der Welt bzw. ihrer Gegenstände selbst haben können, stellt sich die Frage, inwieweit dies tatsächlich verständlich gemacht werden kann. Woran liegt es, dass wohl kaum ein Zweifel daran angemeldet werden kann, dass es sich bei den Ebenen von Atomen, Molekülen, Zellen, Zellverbänden etc. nicht nur um Aspekte handelt, die auf unsere Auffassung der Welt zurückgehen,

sondern um eine *natürliche* Ordnung oder Organisation der Welt bzw. eines betrachteten Gegenstandes selbst? Mittels welcher Prinzipien ist also in Bezug auf Perspektiven im Allgemeinen eine berechtigte Möglichkeit zur Individuation⁷⁵ natürlicher Organisationsebenen gegeben und – was hier entscheidend ist – kann auch das Psychische nach ebendiesen Prinzipien als etwas verstanden werden, das eine natürliche Organisationsebene eines Organismus konstituiert?

Im Hinblick auf die hier angesprochene natürliche Organisation von Objekten ist es vor allem die *Teil-Ganzes-Beziehung* zwischen den verschiedenen strukturellen Ebenen eines Objektes oder Organismus, die verständlich machen kann, dass ein Gegenstand divergierende Erscheinungsweisen hat, die nicht nur ein „Artefakt“ unseres beobachterabhängigen Zugriffs auf diesen Gegenstand sind. Nach einer Besprechung der hiermit zusammenhängenden mereologischen und mechanistischen Individuationskriterien natürlicher Organisationsebenen wird nachgewiesen werden, dass sich auch das Psychische unter Heranziehung ebendieser Individuationskriterien als natürliche Organisationsebene verstehen lässt und somit die Möglichkeit einer „objektivistischen“ Auflösungs-möglichkeit von Spinozas Dilemma gegeben ist.

Abschließend wird sich darüber hinaus zeigen, dass beide Stränge, die „subjektivistische“ und „objektivistische“ Antwort auf das Dilemma, durchaus miteinander versöhnt werden können und so nicht nur ein sehr

⁷⁵ Der philosophische Begriff der *Individuation* eines Gegenstandes meint, dass man über bestimmte Prinzipien zu der berechtigten Annahme gelangt, dass es sich bei diesem Gegenstand tatsächlich um ein *Individuum* handelt, nicht unbedingt ein „Unteilbares“, wie im ursprünglichen Sinne des Wortes „In-dividuum“, aber eine echte konzeptuelle oder *natürliche* Einheit. In diesem Sinne definiert auch Prechtel (2004) eine Individuation als „Verfahren, mit dessen Hilfe etwas in einem Gegenstandsbereich als Individuum konstituiert oder aufgefaßt wird. Das Individuationsprinzip hat anzugeben, mit welchen Mitteln die Gliederung eines Gegenstandsbereichs in einzelne identifizierbare Einheiten möglich ist“ (S. 102).

kohärenter Erklärungsansatz zum Ursprung des Leib-Seele-Problems, sondern auch eine tragfähige Basis für die – im weiteren Verlauf der Arbeit verfolgte – Formulierung einer Theorie zur begrifflich-konzeptuellen Grundlegung der Psychologie gegeben ist.

5.1. Natürliche Organisationsebenen: Eine Systematisierung

Welchen Status haben also nun im Allgemeinen die verschiedenen Gegebenheitsweisen eines Objektes, die sich z.B. in Abhängigkeit vom Auflösungsgrad der methodisch-apparativen Zugangsweise ergeben? Beispielsweise kann man den Menschen als (Neuro-)Biologe mittels verschiedener Instrumentarien vergegenständlichen und hat z.B. eine Repräsentation des gesamten Organismus oder auch von bestimmten Organen wie dem Gehirn, von Nerven- und Gliazellen oder Zellbestandteilen (durch Mikroskope) oder bestimmter neuronaler Aktivitäten (fMRI, PET, EEG etc.) gegeben. Durch die verschiedenen Zugriffsformen auf den Forschungsgegenstand liegen unterschiedliche Repräsentationen dieses Gegenstandes vor, die offenkundig nicht nur Ausdrucksformen der veränderten Beobachterperspektive oder überhaupt Artefakte unserer Subjektivität sind. Was aber sind sie dann? Für gewöhnlich spricht man in diesem Kontext von *natürlichen Organisationsebenen* eines Organismus. Der Begriff natürlicher Organisationsebenen ist allerdings vieldeutiger, als es vielleicht auf den ersten Blick scheinen mag. Es bedarf daher einer grundlegenden Klärung, was genau unter dem Begriff einer solchen überhaupt verstanden werden sollte. Hierzu ist es notwendig, die zentralen Bedeutungsvarianten zu besprechen und aufzuzeigen, inwieweit diese die hier gesuchte Funktion erfüllen. Während bloße quantitative

Prinzipien als Individuationskriterien nicht hinreichend erscheinen, ist es der kompositorische Ebenenbegriff, der über mereologische bzw. mechanistische Prinzipien (wobei letztere eine Sonderform ersterer sind) eine plausible Grundlage für die Zuschreibung natürlicher Organisationsebenen abliefern. Perspektiven können daher immer dann als natürliche Organisationsebenen angesehen werden, wenn sie kompositorisch aufeinander abgebildet werden können und sich die jeweiligen Entitäten ohne überkreuzende Überlappungen in einer Hierarchie anordnen lassen.

5.1.1. Der quantitative Begriff natürlicher Organisationsebenen

Das gemeinsame Merkmal aller natürlichen Organisationsebenen oder auch Ordnungsebenen ist, dass sie beanspruchen, grundlegende „objektive“ Aspekte der Wirklichkeit einzufangen (also solche, die nicht nur *Artefakte* unserer subjektiven Auffassung oder Vergegenständlichung der Welt sind). Wenn man z.B. das menschliche Nervensystem mittels unterschiedlicher Methoden untersucht, kann man dabei feststellen, dass sich je nach Auflösungsgrad der herangezogenen Methode unterschiedliche organisierte Strukturen offenbaren. Sie reichen in Bezug auf die im Rahmen der Neurowissenschaften herangezogenen Methoden von Molekülen über Synapsen, Nervenzellen, neuronale Netzwerke, Schichten, Karten bis hin zu komplexeren (Teil-)Systemen und dem zentralen Nervensystem selbst (siehe Abb. 20).

Wenn man von der Anbindung an die neurowissenschaftliche Methodologie absehen würde, könnte man natürlich noch niederstufigere Ebenen (wie die der Atome und Elementarteilchen) und höherstufigere Ebenen (wie die des menschlichen Organismus oder vielleicht auch die der sozia-

len Gruppe) hinzunehmen, so wie es bei dem Modell von Oppenheim und Putnam (1958) der Fall war.

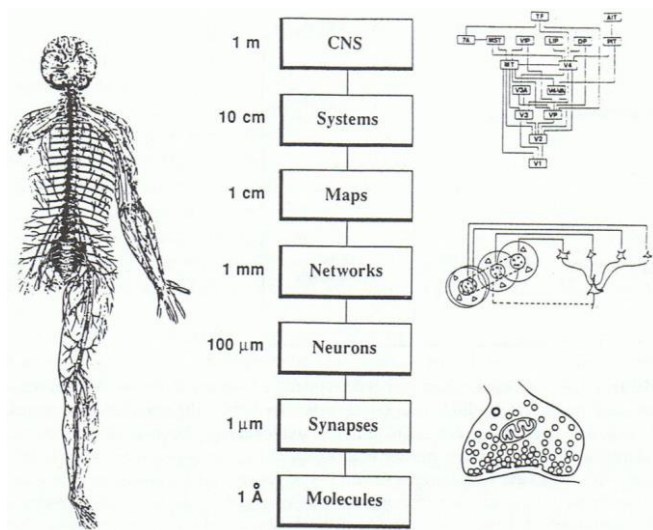


Abb. 20: Organisationsebenen des Nervensystems (Aus: Churchland & Sejnowski, 1994/1997, S. 15)

Wie kommt man nun allerdings dazu anzunehmen, dass solche Organisationsebenen bzw. die sie konstituierenden Entitäten, aus denen z.B. unser Gehirn aufgebaut ist, nicht nur theoretische Konstrukte oder „Artefakte“ unserer Wahrnehmung sind? Welche Möglichkeiten gibt es also, aus berechtigten Gründen trotz der substantiellen Einheit (also einem Substanzmonismus) der Welt auf eine *natürliche* Gliederung der Welt in bestimmte Ebenen zu schließen? In der Wissenschaftstheorie werden verschiedene Möglichkeiten angegeben, natürliche Organisationsebenen zu individuieren. Sie sind danach zu unterscheiden, ob man nach Größenunterschieden der zu ordnenden Entitäten, nach ihrer spezifi-

schen Zusammensetzung (Komposition) oder nach einem Ordnungsgefüge kausaler Prozesse fragt. Diese verschiedenen Begriffsvarianten sollen im Folgenden beleuchtet und es soll geklärt werden, inwiefern sie dazu geeignet sind, eine natürliche Gliederung der Welt in Ebenen verständlich zu machen.

Ausgangspunkt für eine Bestimmung verschiedener Bedeutungsvarianten des Ebenenbegriffs als Ausdruck verschiedener natürlicher Organisationsebenen ist die Frage, wie die angesprochene Gruppierung von Entitäten (und Eigenschaften) in ein System des Höher und Niedriger zustande kommt. Intuitiv einleuchtend scheint zu sein, dass die *Größenunterschiede* zwischen den Entitäten hierfür von zentraler Bedeutung sind. Die „klassischen“ Ebenenmodelle von Morgan (1923) als auch von Oppenheim und Putnam (1958) sind so verfasst, dass die Größe der aufgeführten Entitäten mit höherer Position in der Hierarchie – vom Grundsatz her – zunimmt. Entsprechendes gilt für die Organisationsebenen des Nervensystems bei Churchland und Sejnowski (1994/1997).

Auch wenn die quantitative Differenz zwischen Entitäten *prima facie* ein brauchbares und in der wissenschaftlichen Praxis auch ein oft herangezogenes Kriterium abgeben mag, gibt es im Einzelfall doch Probleme. Die Größe und die dynamischen Charakteristika einer Entität können nämlich unabhängig voneinander variieren (vgl. Wimsatt, 1976, S. 237). Dies kann man an der Gegenüberstellung von Schwarzen Löchern und Bakterien sehen. Beide können die Größe eines Staubkörnchens haben, obwohl sie ein völlig unterschiedliches Verhalten in ähnlichen Situationen aufweisen werden. Da ein schwarzes Loch eine gewaltige Masse haben kann, wird es im Gegensatz zum Bakterium keine Brownsche Bewegung aufweisen (vgl. ebd., 2007, S. 208f). Beide anhand ihrer Größe auf ein und

derselben Ebene verorten zu wollen, erscheint abwegig⁸³. Ebenso müsste man in Kauf nehmen, dass sich z.B. einzellige Lebewesen auf einer Ebene mit dem genannten Staubkörnchen oder anderen unbelebten materiellen Teilchen befinden, wenn sie von gleicher Größe sind.

Wimsatt (1976) unterbreitet deshalb den Vorschlag, das Kriterium der Größe von Entitäten mit einem weiteren Kriterium zu verbinden. So könne man eine Ebene in Abhängigkeit von der Größenrelation auch als ein *lokales Maximum an Vorhersagbarkeit und Regelhaftigkeit* betrachten (S. 238), was er mittels einer Grafik verständlich zu machen versucht (Abbildung 21).

In der Abbildung sind Vorhersagbarkeit und Regelhaftigkeit als eine Funktion der Größe der Entitäten abgetragen. Beim Weltmodell a („regular periodic“) haben wir – abgetragen anhand der Größe der Entitäten – eine vollkommen geordnete Welt mit klar bestimmten Peaks an Vorhersagbarkeit und Regelhaftigkeit. Beim darunter liegenden Weltmodell b („random fluctuation“) liegen nahezu gegenteilige Verhältnisse vor. Es gibt keine klar umrissenen Ebenen⁸⁴. In Abhängigkeit von der Größenordnung der Entitäten sind keine geordneten Veränderungen der Regelhaftigkeit und Vorhersagbarkeit ausfindig zu machen, sondern nur Zufallsschwankungen.

⁸³ Allerdings stellt sich trotzdem die Frage, ob die Größenunterschiede nicht trotzdem im Bereich der Lebewesen ein Kriterium für die Verortung von Entitäten und Eigenschaften zu bestimmten Ebenen sein können. Zumindest ist mir kein gleich oder ähnlich gelagertes Gegenbeispiel aus diesem Bereich ersichtlich.

⁸⁴ Die Interpretation von Craver (2007), die besagt, dass hier gar keine Ebenen vorliegen, dürfte den Intentionen Wimsatts widersprechen. Der *Verlauf* des Ausmaßes an Regelhaftigkeit und Vorhersagbarkeit ist mit zunehmender Größe der Entitäten hier sicherlich diffus. Nichtsdestotrotz sind diese Ordnungsmaße aber – im Vergleich zu Welt e – vorhanden und damit könnten einzelne Peaks auch als Ebenen bezeichnet werden, obgleich diese nicht so klar umrissen sind. Vielleicht ist es aber angezeigt, in diesen Fällen von *Perspektiven* zu sprechen.

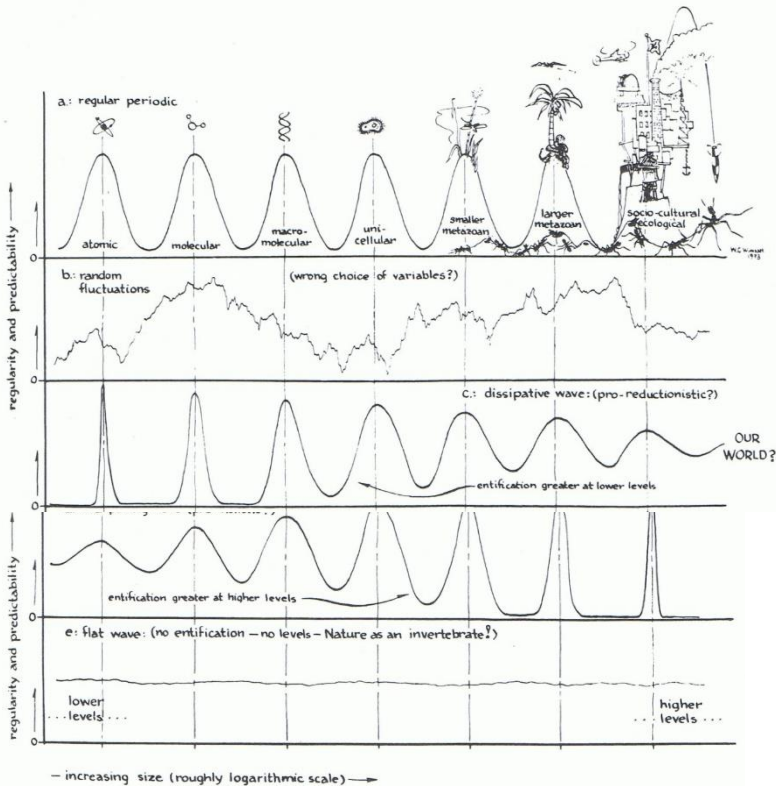


Abb. 21: Ebenen als lokale Maxima an Regelmäßigkeit und Vorhersagbarkeit (Aus: Wimsatt, 1976, S. 240)

Wimsatt nimmt an, dass das Weltmodell c („dissipative wave“), das hinsichtlich seines Ordnungsgrades irgendwo zwischen den beiden vorherigen Modellen angesiedelt ist – am ehesten unserer Welt entspricht. In diesem – eher reduktiven Vorstellungen entsprechenden – Modell haben wir das größte Maß an Ordnung auf den niedrigeren Ebenen und zu den höheren Ebenen hin eine Abnahme der Vorhersagbarkeit und Regelmäßigkeit.

tigkeit. Weltmodell d („sharpening wave“) entspricht dem vorherigen Modell, allerdings haben wir eine Umkehrung der Richtung der zunehmenden Ordnung, d.h., nach diesem – eher holistischen oder emergentistischen – Modell gibt es eine Zunahme an Vorhersagbarkeit und Regelmäßigkeit zu den höheren Ebenen hin, während auf den unteren Ebenen ein geringes Maß an Ordnung herrscht. Weltmodell e („flat wave“) stellt zum Schluss noch eine Welt vor, in der es keinerlei Regelmäßigkeit und Vorhersagbarkeit gibt und damit auch keine Ebenen.

Auch wenn so über quantitative Aspekte eine Etablierung *verschiedener* Organisationsebenen möglich sein mag, können sie kaum zur Begründung dafür herangezogen werden, dass es überhaupt eine *natürliche* Ordnung der Welt gibt. Außerdem bleibt im Hinblick auf die Grafik Wimsatts fraglich, inwieweit sich die postulierte oberste Ebene, auf der sich sozial-kulturell-ökologische Phänomene befinden sollen, überhaupt sinnvoll mittels des von Wimsatt verwendeten Kriteriums integrieren lassen. Zwar könnte man auch im Hinblick auf kulturelle oder soziale Phänomene von bestimmten „Maxima an Vorhersagbarkeit und Regelmäßigkeit“ sprechen, inwieweit aber die hier vorgenommene Einordnung nach Größeneigenschaften berechtigt ist, bleibt (mir) unklar. Kann man von kulturellen Phänomenen in irgendeiner Hinsicht sinnvoll sagen, dass sie größer sind als die vielzelligen Organismen der darunter liegenden Ebene? Vielleicht nur insofern, als dass diese vornehmlich ganze Gesellschaften und nicht nur einzelne Personen betreffen.

Noch nebulöser erscheint die Lage, wenn man nach der Einordnung von mentalen Eigenschaften in eine solche Ordnung fragt. Sollen diese zur sozial-kulturell-ökologischen Ebene gehören? Wenn dem so ist, warum sollen mentale Eigenschaften dann in *quantitativer* Hinsicht physikalischen bzw. biologischen Entitäten oder Eigenschaften übergeordnet sein? Dies scheint sich nicht verständlich machen zu lassen.

Es kann konstatiert werden, dass der quantitative Ebenenbegriff für ein weites Feld von physikalischen Entitäten ein nützliches Individuationskriterium bereitstellt. Probleme scheint es dann zu geben, wenn man ihn wie Wimsatt als eine Art *universelles* Ordnungsschema betrachtet, mittels dessen man jegliche räumlich-zeitlich unverbundenen Gegenstände (wie ein Staubkörnchen und ein schwarzes Loch), die nicht selbst Teile *ein und desselben* Systems sind, in Verbindung zu setzen versucht. Oft ist nicht klar, ob man bei zwei Dingen sagen soll, sie sind auf einer Ebene oder dass das eine höher oder niedriger als das andere einzustufen ist. Hierauf haben vor allem Craver (2007), aber auch bereits Kim (2002) hingewiesen:

It does not seem to me that the question whether sophisticated industrial robots or powerful supercomputers, say, are higher or lower in comparison with sea slugs or birch trees makes sense or, even if it did, what possible significance the answer would have. [...] On what possible ground would an answer to this question be based, and what possible use could the answers have? [...] I think that attempts to construct an overarching levels ontology for the whole of the natural world in which every object has its 'appropriate' place are rather pointless if not hopeless. (S. 16)

Es ist daher angezeigt, dass man Ebenen nur *lokal* bestimmt, d.h., im Hinblick auf einzelne Systeme wie den Menschen oder das menschliche Gehirn nach einer Ordnung von Elementen sucht.

5.1.2. Der kompositorische Begriff natürlicher Organisationsebenen

Anstatt auf bloße Größenunterschiede zu rekurrieren, wird im Fall kompositorischer Ebenen zum Kriterium für die Individuation von Ebenen erhoben, wie genau bestimmte Entitäten *zusammengesetzt* sind und in welchem Verhältnis eine bestimmte Zusammensetzung oder *Komposition* zu den Kompositionen steht, die sich auf höheren und niedrigeren Ebenen desselben Objekts finden.

Für eine Individuation von Ebenen über kompositorische Verhältnisse scheint es – nicht zuletzt aufgrund der obigen Ausführungen zum kaum einlösbaren Anspruch einer Einführung „universeller“ Ebenen – sinnvoll anzunehmen, dass sie von nur *lokaler* Natur sind, d.h., dass die Einführung einer natürlichen Ebene auf ganz konkrete Wirklichkeitsausschnitte oder Vorkommnisse, wie auf ein Lebewesen bzw. seine Organe wie das Gehirn, beschränkt ist⁸⁵. Die drei zentralen Formen des kompositorischen Begriffs natürlicher Organisationsebenen sind mereologische, mechanistische und aggregative Verhältnisse.

5.1.2.1. Mereologische Ebenen

Bei mereologischen Ebenen geht es um *Teil-Ganzes-Beziehungen*. Im Gegensatz zu dem Modell von Wimsatt, bei dem Ebenen (in Abhängigkeit von ihrer Größe) über die Interaktion von den Individuen eines bestimmten Organisationsgrades bestimmt werden, spielen die Beziehungen zwischen den Individuen auf einer Ebene in der Mereologie keine ausgezeichnete Rolle. Jede komplexe Entität wird als ein Aggregat angesehen (vgl. Craver, 2007, S. 184). Ebenen werden hier nicht über die Größe der Entitäten bestimmt (wenn diese sicherlich auch hier zu den niedrigeren Ebenen hin abnimmt) oder die Interaktion zwischen den Entitäten einer Ebene, sondern darüber, ob man zwischen zwei Gegebenheiten ein *kompositorisches* Verhältnis ausfindig machen kann. Aufgrund dieser Bestim-

⁸⁵ Die Möglichkeit, alle kompositorischen Ebenen als universell zu betrachten, besteht nichtsdestotrotz prinzipiell in dem Sinne, dass eine Aussage bzgl. der Zusammensetzung eines Objektes induktiv verallgemeinert werden kann, wie z.B., dass sich Moleküle generell aus Atomen zusammensetzen etc. Inwieweit die Betrachtung solcher einzelnen Vorkommnisse (Token) zu einer Möglichkeit der Formulierung allgemeiner Verhältnisse (Typen) führen kann, wird an späterer Stelle diskutiert (siehe Kap. 7.5.2).

mungen muss der Umfang der Elemente, die eine Ebene konstituieren, von höheren Ebenen zu niedrigeren Ebenen stetig zunehmen.

Eine anschauliche Beschreibung dieses Prinzips liefert Lycan (1987/2007):

Zellen [...] setzen sich aus kooperierenden Gruppen kleinerer Einheiten zusammen, zu denen u.a. die Membran, der Zellkern und die Mitochondrien gehören: Diese Einheiten selbst sind *Systeme* noch kleinerer, ebenfalls kooperierender Bestandteile. Wenn wir schon mal dabei sind, gibt es darunter noch zahlreiche und deutlich unterscheidbare Ordnungsebenen der Natur: die chemische, die molekulare, die atomare sowie die (traditionelle) subatomare oder mikrophysikalische Ebene. Diese Ebenen bestehen aus Geflechtem interessanter gesetzesartiger Verallgemeinerungen und werden nach den Typen der betreffenden Generalisierungen individuiert. Verfolgt man die Hierarchie jedoch wieder aufwärts, gruppieren sich die Zellen zu Geweben, aus denen sich Organe zusammensetzen, die sich wiederum zu Organismen anordnen, die – auf wunderbare Weise – zusammenwirken, um ganze *Organismen* wie etwa menschliche Lebewesen zu bilden. (S. 433f; Hervorh. im Orig.)

Wenn man die über die Wissenschaften i.d.R. relativ gut gesicherten Erkenntnisse über Entitäten und ihre Zusammensetzungen als Tatsachen akzeptiert, dann dürften mereologische Aspekte ein sehr valides Kriterium zur Individuation natürlicher Organisationsebenen abliefern. Immer, wenn wir eine solche Teil-Ganzes-Beziehung ausfindig machen können, haben wir einen hinreichenden Grund, von zwei natürlichen Ebenen zu sprechen. Es ist daher nicht verwunderlich, dass dies auch das gebräuchlichste Kriterium für die Etablierung von Ebenen darstellen dürfte⁸⁶.

⁸⁶ Was nicht heißen soll, dass *alles*, was existiert, sich mittels einer Teil-Ganzes-Beziehung individuiert lässt (man denke z.B. an Kräfte wie Magnetismus, Gravitation oder quantenphysikalische Aspekte). Die Teil-Ganzes-Hierarchie ist speziell im Hinblick auf den *Ebenenbegriff* die schlüssigste Interpretation, wenn es um ein Verständnis natürlicher Ordnungsbeziehungen zwischen Entitäten oder Prozessen eines Objektes geht und damit kein *univer-*

Mit Verweis auf mereologische Ebenen kann man also durchaus Substanzmonist sein, ohne die offen-kundige Vielfalt⁸⁷ der Welt und ihrer Objekte leugnen bzw. nur auf unsere subjektive Auffassung des betreffenden Gegenstandes zurückführen zu müssen, da über eine Teil-Ganzes-Beziehung, die zwischen Entitäten besteht, eine natürliche Gliederung von Objekten angenommen werden kann. Wenn sich also ein Gehirn auf der Ebene neuronaler Netzwerke, auf der zellulären oder vielleicht auch auf der molekularen Ebene untersuchen lässt, dann können diese Ebenen als natürliche Gliederungen des Gehirns angesehen werden, da sie in einem mereologischen Verhältnis zueinander stehen. Eine relevante Implikation mereologischer Beziehungen ist, dass zwischen dem Ganzen und seinen Teilen ein Verhältnis besteht, was man am besten wohl als *Koextensivität* bezeichnen kann, d.h., dass beide das gleiche räumlich-zeitliche „Feld“ einnehmen. Dies ergibt sich bereits aus der Lokalität kompositorischer Prinzipien zur Individuation von Ebenen. Es ist *ein und derselbe* Gegenstand (Lebewesen, Organ...), der durch eine bestimmte Anzahl von mereologischen Ebenen gekennzeichnet ist und im Hinblick auf diese „zerlegt“ werden kann. Ein Herz und die Gefäße, Zellen, Moleküle oder Atome, aus denen dieses Organ besteht, nehmen daher mit Notwendigkeit den gleichen Platz in der Welt ein. Die Annahme eines Teils, das außerhalb des räumlichen Feldes des Ganzen vorkommt, ist kein kohärenter Gedanke. Ebenso wäre es unsinnig anzunehmen, dass etwas Teil eines Ganzen sein kann, obwohl es nicht zur gleichen Zeit wie das Ganze existiert. Es ist im Zusammenhang der Argu-

selles Individuationskriterium.

⁸⁷ Eine andere Möglichkeit mag darin bestehen, die Diversität der Entitäten der Welt auf unterschiedliche Eigenschaften von Objekten zurückzuführen (Eigenschaftsdualismus bzw. -pluralismus). Solche Positionen sind gegenüber Substanzdualismen oder -pluralismen allerdings kaum ein Fortschritt.

mentation der vorliegenden Arbeit nicht relevant, die logischen Beziehungen zwischen Teilen und Ganzen in all ihren vielfältigen Facetten zu analysieren, wie es in der mereologischen Forschung beispielsweise Simons (1987) getan hat. Fragen derart, ob die Gesamtheit der Teile mit dem Ganzen identisch ist, ob ein Teil für bestimmte Eigenschaften des Ganzen konstitutiv ist oder nicht und dergleichen, bleiben hier (zunächst) unberührt.

Obleich dies eine triviale Feststellung sein mag, sei an dieser Stelle explizit darauf hingewiesen, dass sich ein Teil natürlich nicht über das *gesamte* räumlich-zeitliche Feld des Ganzen erstreckt. Mit Koextensivität in dem hier intendierten Sinne ist vielmehr gemeint, dass ein Teil, um Teil eines Ganzen zu sein, notwendigerweise einen *Teilbereich* der räumlich-zeitlichen Ausdehnung des Ganzen besetzt (ist dies nicht gegeben, wird die Annahme einer Teil-Ganzes-Beziehung hinfällig). Worauf es hier ankommt, ist, dass ein Teil keinen anderen räumlich-zeitlichen Bereich einnehmen kann als einen Teilbereich des Ganzen, dessen Teil er ist, und dass das Ganze und seine Teile in diesem Sinne koextensiv sind.

5.1.2.2. Aggregative Ebenen

Auch aggregative Ebenen gehen auf ein spezifisches mereologisches Individuationskriterium zurück, bei dem nunmehr *Eigenschaften* fokussiert werden. Aggregative Ebenen ergeben sich aus der Beziehung zwischen den Eigenschaften des Ganzen und den Eigenschaften der Teile, wobei gilt, dass die Eigenschaften der höheren Ebenen als die Summe der Eigenschaften der niedrigeren Ebenen aufzufassen sind. Ein Beispiel hierfür liefert Craver (2007): „The mass of a pile of sand is an aggregate of the masses of the individual grains” (S. 186).

Das bedeutet auch, dass sich das Ganze in Abhängigkeit davon verändert, ob einzelne Teile hinzugefügt oder entfernt werden. Dagegen macht allerdings ein Austausch von einzelnen Teilen (durch funktional äquivalente Teile) keinen Unterschied für die fokussierte Eigenschaft des Ganzen: „The pile gets heavier continuously as one adds new grains of sand, and moving them about has no effect on the weight. Replacing individual grains with equally weighted replicas has no effect on the weight of the pile, and the grains do not interact with one another in ways that influence the weight of the pile” (ebd.).

Allerdings kritisiert Craver (2007) zu Recht, dass es viel häufiger Fälle von Ebenen geben wird, die die Bestimmungen zur Aggregativität verletzen als Fälle, die diese erfüllen. Es ist eine sehr anspruchsvolle Beziehung zwischen Eigenschaftsklassen, die in der Natur nur sehr selten zu finden ist. Oftmals ist nämlich eine bestimmte Eigenschaft des Ganzen nicht nur von ebendieser Eigenschaft der Teile abhängig, sondern auch von einer Reihe anderer Eigenschaften der Teile. So sei z.B. das Volumen von Gin nicht nur auf das Volumen der Moleküle, aus denen er besteht, zurückzuführen, sondern auch auf die mittlere kinetische Energie sowie die elektromagnetische Energie der Moleküle (vgl. ebd., S. 186f).

Insofern kann man von Aggregativität streng genommen nur dann sprechen, wenn die *organisationale* Beschaffenheit der Ebene der Teile keine Rolle spielt: „if the mode of organization of the parts *didn't* matter at all, the system property would seem to collapse as nothing more than an aggregate of parts' properties“ (Wimsatt, 2007, S. 174; Hervorh. im Orig). Nach Ansicht Wimsatts ist die (implizite) Annahme von Aggregativität zwischen Ebenen ein zentraler Grund für die negative Einstellung zum Thema Reduktion (vgl. Kap. 2.1.2), denn sie suggeriert, dass die höherstufigen Eigenschaften eines Systems *nichts anderes als* eine Ansammlung niederstufiger Eigenschaften dieses Systems sind: „With aggregativity,

there's no tendency to separate system property from sum of parts properties; they are simply identified“ (ebd., S. 375). Dass die Eigenschaften des Ganzen (z.B. mentale Eigenschaften) sich vollkommen auf die Eigenschaften ihrer Teile (neuronalen Netzwerke...) reduzieren lassen sollen, erscheint abwegig. Dies ist insbesondere der Fall, da ein so verstandener Reduktionismus häufig als eine Form des *eliminativen Materialismus* verstanden wird, der mentalen Eigenschaften ihre Existenz abspricht (vgl. Kap. 3.1.3). Außerdem wirkt die Annahme unplausibel, dass die organisationale Beschaffenheit der neuronalen Ebene (wie sie der mechanistische Ansatz fokussiert) irrelevant für die Beschaffenheit eines mentalen Vorgangs sein soll.

5.1.2.3. Mechanistische Ebenen

Mechanistische Ebenen finden sich im mechanistischen Erklärungsansatz wieder, der gerade in der Gegenwartsphilosophie stark diskutiert wird. Dieser Ansatz geht vor allem auf Arbeiten von Bechtel und Richardson (1993), Machamer, Darden und Craver (2000), Bechtel (2001, 2007, 2008, 2009), Craver (2002, 2007) bzw. Craver und Bechtel (2007) zurück. Gemeinsam ist allen Autoren, dass sie mechanistische Ebenen als natürliche Organisationsebenen verstanden wissen möchten.

Ein Mechanismus kann nach Craver (2007) verstanden werden als „a set of activities organized such that they exhibit the phenomenon to be explained“ (S. 5). Der mechanistische Ansatz individuiert ein Phänomen darüber, welche Aufgaben oder kausale Rollen es für das betreffende System innehat. Wenn man also den Mechanismus ergründet hat, der einem Phänomen zugrunde liegt, hat man auch das Phänomen selbst erklärt. „Understanding the relevant mechanism consists in decomposing the system into parts, specifying the properties and causal roles of the parts,

and understanding the spatial and temporal organisation of the parts” (Fazekas & Kertész, 2011).

Abbildung 22 zeigt die allgemeine formale Struktur mechanistischer Ebenen. An der Spitze des Mechanismus (S) befindet sich das zu erklärende Phänomen, das Craver für gewöhnlich als ψ bezeichnet. Durch die beigesteuerte Endung „-ing“ und die eingezeichneten Verlaufspfeile soll die Prozesshaftigkeit mechanistischer Ebenen ausgedrückt werden. Das Phänomen ψ kann nun wiederum in agierende Teilkomponenten zerlegt werden, wobei Craver hier X als Bezeichnung für die als Komponente von Ψ fungierenden Entitäten und φ als Bezeichnung für deren Aktivitätsmuster verwendet. Wiederum lässt sich jede Komponente dieser Ebene auf eine noch niedrigere Ebene des Mechanismus herunterbrechen, indem hier wiederum die aus einer Entität (P) und einer Aktivität (ρ) bestehenden Teilkomponenten vorzufinden sind.

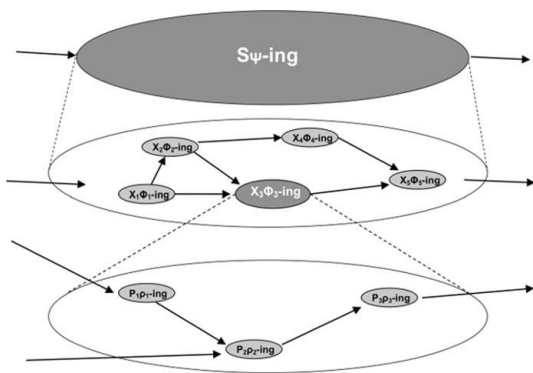


Abb. 22: Die formale Struktur mechanistischer Ebenen (Aus: Craver, 2007, S. 189)

Auch in Bezug auf die agierenden Entitäten der verschiedenen Ebenen eines Mechanismus gilt das *Koextensivitätsprinzip*, sodass es auch in einer solchen mechanistisch begründeten Hierarchie ein höherstufiges Ganzes gibt, das dasselbe räumlich-zeitliche Feld einnimmt wie seine Komponenten, die wiederum ein koextensives Ganzes gegenüber ihren eigenen Komponenten darstellen usw. Da sich mechanistische Ebenen auf kompositorische bzw. Teil-Ganzes-Verhältnisse beziehen, stellen sie ebenso wie mereologische Ebenen eine hierarchische Gliederung in *lokalen* Einzelfällen dar und kein (metaphysisches) Charakteristikum der ganzen Welt, keine „monolithischen“ Strukturgliederungen der Welt, wie sie von Oppenheim und Putnam oder (teilweise) von Wimsatt angenommen wurden:

However, levels of mechanisms are far more *local* than the monolithic image suggests. They are defined only within a given compositional hierarchy. Different levels of mechanisms are found in the spatial memory system, the circulatory system, the osmoregulatory system, and the visual system. How many levels there are, and which levels are included, are questions to be answered on a case-by-case basis by discovering which components at which size scales are explanatory relevant for a given phenomenon. They cannot be read off a menu of levels in advance. (Craver, 2007, S. 191; Hervorh.: M.L.)

Mechanistische Ebenen können als ein prozess- oder kausalitätsbezogener Spezialfall von mereologischen Ebenen verstanden werden. Der Unterschied zu mereologischen Ebenen ist, dass es sich bei mechanistischen Ebenen nicht um Ebenen von (statischen) Objekten handelt, sondern um „levels of behaving components“ (ebd., S. 190). „Levels of mechanisms are levels of composition, but the composition relation is not, at base, spatial or material. In levels of mechanisms, the relata are behaving mechanisms at higher levels and their components at lower levels. These

relata are properly conceived neither as entities nor as activities; rather, they should be understood as *acting entities*” (ebd., S. 189; Hervorh.: M.L.). Während also mittels des rein mereologischen Individuationskriteriums *strukturelle* Ebenen von kompositorisch aufeinander bezogenen *Objekten* herausgestellt wurden, wird hier eine Individuation im Hinblick auf *Aktivitäten* oder *Prozesse* vorgenommen. Auch wenn somit die „Prozesshaftigkeit“ und nicht die „Dinghaftigkeit“ in den Blick genommen wird, gilt trotzdem, dass es sich um einen Prozess handelt, durch den ein bestimmtes Objekt gekennzeichnet ist; es gibt also immer *etwas*, das eine bestimmte Aktivität aufweist. In diesem Sinne handelt es sich um *agierende Entitäten*. Agierende Entitäten sind des Weiteren durch eine bestimmte Organisation gekennzeichnet. So haben die Komponenten des Mechanismus bestimmte räumliche (Position, Größe, Form und Bewegung), zeitliche (Reihenfolge, Frequenz und Dauer) und aktivitätsbezogene (z.B. Rückkopplungseffekte) Beziehungen untereinander, um ein Phänomen hervorzubringen.

Craver (2007) erläutert den Ansatz mechanistischer Ebenen am Beispiel des räumlichen Gedächtnisses⁸⁸ (siehe Abb. 23). Auf der obersten Ebene, der psychischen bzw. psychologisch beschreibbaren Ebene, haben wir eine Maus, die mittels ihres räumlichen Gedächtnisses versucht, einen Ausweg aus einem Labyrinth zu finden, in das sie gesetzt wurde. Um verschiedene Formen des räumlichen Gedächtnisses zu testen, können die Forscher eine Reihe von Variablen manipulieren (z.B. den Ort, wo sie die Maus ins Labyrinth setzen, Umgebungsvariablen usw.) (vgl. ebd., S. 167). Auf der neuronalen Ebene findet die neuronal realisierte Informa-

⁸⁸ Natürlich handelt es sich hierbei um eine starke Vereinfachung, denn an der räumlichen Orientierung sind deutlich mehr Prozesse beteiligt, wie bestimmte Aktivierungen im so genannten dorsalen Pfad (vgl. Stiles, 2007).

tionsverarbeitung statt, bei der u.a. im Hippocampus eine Art räumliche Karte generiert wird (vgl. ebd.).

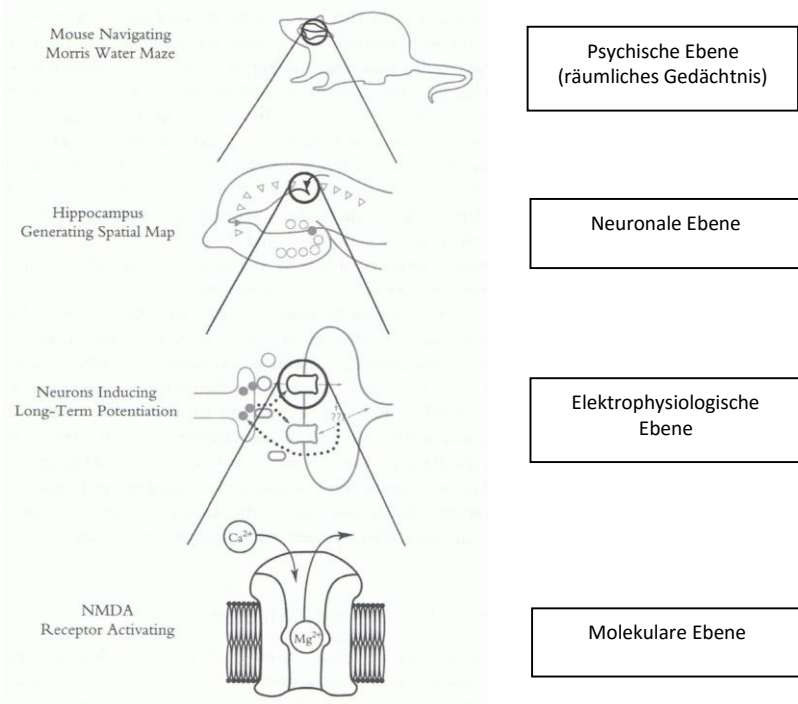


Abb. 23: Mechanistische Ebenen des räumlichen Gedächtnisses (Aus: Craver, 2007, S. 166)

Wenn man sich nun fragt, wie dies der Hippocampus bewerkstelligt, muss man die Antwort auf der nächsten Ebene, der elektrophysiologischen Ebene, suchen. Ein entscheidender Prozess dürfte hierbei die sogenannte Langzeitpotenzierung (LTP) sein, bei der es aufgrund bestimmter (wiederholter) Erfahrungen, zu einer dauerhaften Verstärkung der postsynaptischen Antwort bei einem konstanten präsynaptischen Signal kommt (vgl. ebd.). Und schließlich gibt es noch die „molekulare Ebene“

ne“, die aus den molekularen Mechanismen besteht, die wiederum die elektrochemischen Aktivitäten der nächsthöheren Ebene überhaupt erst ermöglichen. Ein Beispiel wäre die hier dargestellte Blockade eines NMDA-Rezeptors durch Magnesium-Ionen. Bei einer Depolarisation der postsynaptischen Zelle werden diese aus dem Kanal gepresst und Calcium-Ionen können in die Zelle einströmen (vgl. ebd., S. 169).

Zu beachten ist, dass bei den mechanistischen Ebenen, wie sie Craver im Hinblick auf das räumliche Gedächtnis vorschlägt, nicht unbedingt eine *vollständige* Zerlegung des höherstufigen Prozesses in „darunterliegende“ Teilprozesse der *gleichen* Art erfolgt. Wie in Abbildung 23 aufgezeigt, wird hier nur ein Teilbereich herausgegriffen, dessen mechanistische Struktur bzw. Aktivität dann auf einer niedrigeren Stufe veranschaulicht wird, während andere Teilbereiche unbeachtet bleiben. Allerdings dürfte dies wohl vor allem unserem derzeitig noch sehr unvollständigen Wissen über die neuronalen Grundlagen von Gedächtnisleistungen geschuldet sein. Eine vollständige Zerlegung der „acting entities“ der höherstufigen Ebenen des Mechanismus in solche der niedrigeren Ebenen sollte jedoch zumindest *prinzipiell* möglich sein. Es dürfte sich aber um eine Vielzahl relativ unterschiedlicher prozessualer Teilkomponenten handeln, die den kausalen Hintergrund eines mentalen Zustandes oder komplexen Verhaltensablaufes (wie das durch das Labyrinth Navigieren der Maus) ausmachen.

Des Weiteren muss problematisiert werden, dass Craver (2007) zwar von einer Ebene des räumlichen Gedächtnisses als zu erklärendes Zielphänomen und somit (zumindest indirekt) von einer *psychischen* Ebene spricht, jedoch keine weitergehende Klärung herbeiführt, wie genau sich diese Ebene des Psychischen in sein Schema natürlicher Organisationsebenen einfügt. Er bestimmt diese Ebene nur verhaltensbezogen (räumliche Orientierung als das Verhalten der Ratte in dem Labyrinth) bzw. über die

Zuordnung zu Forschungsdisziplinen („[t]his level is typically associated with scientific fields, such as experimental psychology and ethology“, S. 165), ohne ihren konkreten (ontologischen) Status näher zu klären. In Bezug auf die Fragen, ob es sich beim Psychischen also tatsächlich um eine eigene natürliche Organisationsebene handelt und in welcher genauen Beziehung es zu den Prozessen steht, die durch solche Modelle mechanistischer Hierarchien beschrieben werden können, liefert Craver also keine Antwort. Wenn Psychisches überhaupt nicht über seine *psychische* Struktur erfasst wird, sondern nur verhaltensbezogen (bzw. die disziplinäre Zuordnung), dann besteht im Hinblick auf phänomenale Eigenschaften (im Gegensatz zu der hier thematisierten kognitiven Größe) sogar eine noch größere Schwierigkeit zur Einbindung, da sich diese Eigenschaften häufig nicht eindeutig über behaviorale Parameter erfassen lassen. Im Folgenden wird allerdings aufgezeigt werden, dass dieses Modell durchaus ein großes Potential hat, diese hier noch ausgeblendeten Fragen zu klären.

Insgesamt lässt sich festhalten, dass auch im Falle der Anwendung mechanistischer Prinzipien ein hinreichendes Kriterium zur Individuation von natürlichen Organisationsebenen eines Systems vorliegt. Allerdings könnte es sein, dass mechanistische Verhältnisse gegenüber einfachen mereologischen Verhältnissen nur für einen kleineren Phänomenbereich zutreffen. Es sind eben speziell lokale Beschreibungen von *Kausalprozessen* (im Sinne des Agierens – von in einen Mechanismus eingefassten – Entitäten verschiedener Komplexitätsebenen eines Systems), die wir in „vertikaler“ Hinsicht von Makro- zu Mikroebenen (oder auch andersherum) untersuchen können.

5.1.3. Perspektiven/analytische Ebenen versus natürliche Ebenen

Wenn also kompositorische Verhältnisse in einer mereologischen oder mechanistischen Form das gesuchte valide Individuationskriterium für natürliche Organisationsebenen liefern, muss noch geklärt werden, wie genau die Beziehung zwischen Perspektiven und Ebenen zu verstehen ist.

Eine aufschlussreiche Bestimmung liefert wiederum Wimsatt (2007). Er geht dabei von einem (zumindest quasi-)subjektiven Verständnis von Perspektiven aus. Sie sind seines Erachtens „intriguingly quasi-subjective (or at least observer, technique, or technology-relative) cuts on the phenomena characteristic of a system, which needn't be bound to given levels" (ebd., S. 222). An anderer Stelle heißt es: "What I call perspectives is a diverse range of things that nonetheless appear to have at least some of the properties of being 'from a point of view' or to have a subjective or quasi-subjective character" (ebd., S. 227).

Für Perspektiven in diesem Sinne stellt Wimsatt nun auch eine Vielzahl von Eigenschaften heraus, von denen ich nur diejenigen aufführen möchte, die mir in diesem Zusammenhang relevant erscheinen: Perspektiven beinhalten *erstens* ein Set von Variablen, die gebraucht werden, um ein System zu charakterisieren oder Objekte in Teile aufzuteilen, was zusammen einen systematischen Ansatz einer *Phänomendomäne* ergibt und auf eine charakteristische Art und Weise für einen Beobachter oder eine Gruppe von Beobachtern hervorstechend ist und zwar aufgrund der kausalen Interaktion der Beobachter mit dem fraglichen System (vgl. ebd.). Für ihn sind Perspektiven die divergierenden „Phänomendomenen“ (insofern phänomenale Perspektiven), die sich durch die verschiedenen theoretischen oder methodisch-apparativen Verfahren ergeben, die für eine Wissenschaft kennzeichnend sind.

Zweitens können einfache Systeme genauso gut wie komplexe Systeme von einer Vielzahl von Perspektiven aus beschrieben werden, aber sie werden sich in dem Grad unterscheiden, in dem sie „transperspektivische“ Probleme haben (also solche, deren Lösungen Informationen von mehr als einer Perspektive benötigen) (vgl. ebd., S. 228).

Drittens variiert die Komplexität von transperspektivischen Problemen ebenso von einfachen zu komplexeren Systemen damit, ob sie Systeme auf eine Weise zerlegen, die...

- a. räumlich koinzident ist (in dem Fall sind die verschiedenen Perspektiven entweder auf derselben Ebene oder sie „spannen denselben Ebenenbereich auf“);
- b. die hierarchisch erfasst werden können und zwar relativ zu anderen, sodass sich Teile einer Perspektive in einer anderen Perspektive als ganze Systeme darstellen können;
- c. sich in beliebigen Weisen überlappen (vgl. ebd., S. 228f).

Was lässt sich nun aus den von Wimsatt getroffenen Charakterisierungen in Bezug auf die Beziehung zwischen Perspektiven und natürlichen Organisationsebenen entnehmen? Natürliche Organisationsebenen können nach Maßgabe der Wissenschaftstheorie als ein Spezialfall von Perspektiven angesehen werden – *eine Klasse von Perspektiven, die kompositorisch aufeinander abgebildet werden kann und deren Entitäten sich ohne überkreuzende Überlappungen in einer Hierarchie anordnen lassen*. Für einen Teilbereich von Perspektiven bzw. repräsentationalen Gegebenheitsweisen eines Gegenstands gilt also, dass er – vor allem aufgrund der Implementation einer Teil-Ganzes-Beziehung – strukturiert und aufgrund dieser Ordenbarkeit angenommen werden kann, dass es sich hierbei um eine natürliche Gliederung des betrachteten Gegenstands selbst handelt.

Des Weiteren sollte man – ganz im Sinne Nietzsches (1887) – die divergierenden Repräsentationsformate eines zu erforschenden Gegenstandes durchaus als etwas Positives für unsere Erkenntnisgewinnung ansehen, da sie uns letztendlich mehr über ein Objekt sagen. Es mag zwar am Anfang einer Untersuchung darum gehen, ein hohes Maß an Übereinstimmungen der verschiedenen Arten der Detektion, Messung, Ableitung etc. zu erhalten; ab einem bestimmten Stadium kann man die beobachteten Unterschiede, die sich durch die verschiedenen Zugänge zu einem Objekt ergeben, aber dazu nutzen, mehr über dieses Objekt zu erfahren (vgl. Wimsatt, 2007, S. 197).

5.2. Das Psychische als mereologisch-mechanistisch individuierte Organisationsebene

Wenn es um das Verhältnis zwischen Gehirn und Geist geht, fühlen sich neokantianisch ausgerichtete Philosophen und Psychologen wie z.B. Bitbol (2012) im Angesicht einer „subjektivistischen“ oder – wie man früher gesagt hätte – „idealistischen“ Auflösung des Leib-Seele-Problems bzw. seines Ursprungs, wie sie in Kapitel 3 dargestellt wurde, zu Schlussfolgerungen folgender Art ermutigt: „No metaphysical distinction between the higher and basic levels of organization is supposed, but only a duality of modes of access“ (S. 233). Ist dem aber wirklich so? Könnte nicht hier in der Anwendung des Ebenenmodells natürlicher Organisationsebenen ein gewisses Potential liegen, Spinozas Dilemma „objektivistisch“ aufzulösen, d.h. nachzuweisen, dass die Differenz zwischen Gehirn und Geist mehr als ein „Artefakt“ unserer eigenen Auffassungsweise ist?

Wenn bisher von natürlichen Organisationsebenen die Rede war, wurde hierbei in der Tat auf ausschließlich klassisch „materielle“ oder – vielleicht besser – nicht genuin psychische Entitäten Bezug genommen. So wurde gesagt, dass Zellen aus Molekülen, diese wiederum aus Atomen und diese wiederum aus Elementarteilchen bestehen und dass jede Möglichkeit der Zerlegung einer höherstufigen (agierenden) Entität in (agierende) Entitäten niedrigeren Stufen jeweils eine Individuation natürlicher Strukturebenen erlaubt. In diesem Kapitel soll nun geklärt werden, ob es sich auch beim Psychischen, also in Bezug auf mentale Zustände und Prozesse, um eine natürliche Organisationsebene handelt. Wenn die (beobachterunabhängige) Vielgestaltigkeit eines Objektes im Allgemeinen darauf zurückgeht, dass dieses Objekt über eine natürliche Strukturgliederung in Form der Teil-Ganzes-Beziehung verfügt, dann stellt sich die Frage, ob nicht auch das Psychische nach denselben Individuationskriterien als natürliche Organisationsebene des menschlichen Organismus aufgefasst werden kann.

Im Folgenden wird sich zeigen, dass es durchaus *möglich* ist anzunehmen, dass das Psychische eine eigene höherstufige Organisationsebene bildet. Trotz der vielzähligen Aspekte, die verdeutlichen werden, dass es keine prinzipiellen Gründe gibt, die eine solche (kontraintuitive) These ausschließen würden, muss ich einräumen, dass zunächst jedoch wenig geleistet werden wird, um sie in *positiver* Hinsicht argumentativ abzusichern. Es sei daher angemerkt, dass die hier erfolgende Argumentation diesbezüglich nur einen vorläufigen Charakter hat und nach der Besprechung einiger hierfür entscheidender Zusammenhänge eine weitere Fundierung und Erklärung in Kapitel 7 erfahren wird.

5.2.1. Die Gehirn-Geist-Differenz als divergierende mereologische Ebenen

Zunächst muss festgestellt werden, dass für die Frage nach der Möglichkeit einer Individuation des Psychischen als einer natürlichen Organisationsebene der Rekurs auf quantitative Prinzipien offensichtlich wenig hilfreich ist. Eine Möglichkeit, einer solchen Einordnung des Mentalen – zumindest teilweise – Sinn abzugewinnen, könnte lediglich darin bestehen, dass man, wenn man Ebenen als etwas ansieht, das sich über quantitative Aspekte begründet, nicht auf die bloße Größe von Entitäten rekurriert, sondern auf so etwas wie *Komplexität*. In diesem Fall wäre nicht mehr die bloße physikalische Ausdehnung des Gegenstandes entscheidend (von der man bei mentalen Phänomenen bekanntermaßen nicht sinnvoll sprechen kann), sondern das quantitative Ausmaß seiner strukturellen Verfasstheit. Die höheren Ebenen weisen grobkörnigere (coarser-grained) Eigenschaften auf, während die niedrigeren Ebenen feinkörnigere (finer-grained) Eigenschaften besitzen (vgl. z.B. McClamrock, 1991). Dies trifft nach Ansicht vieler Autoren auch auf das Verhältnis mentaler und physikalischer Eigenschaften zu. So sagt z.B. Lockwood (1993),

...the phenomenal objects of introspective awareness are far less finely structured than are any plausible physiological correlates. Consider, for example, a phenomenally flawless auditory experience, of a note, say, on a violin. Its physiological substrate, presumably, is a highly structured, not to say messy, concatenation of changes in electrical potential within billions of neurons in the auditory cortex, mediated by the migration of sodium and potassium ions within the chemical soup at the synapses⁹⁰. (S. 274)

⁹⁰ Allerdings sind mit diesen Unterschieden zwischen der physischen und der mentalen Ebene eigene Schwierigkeiten verbunden, wenn es um das Leib-Seele-Problem geht. Es stellt sich die Frage, inwieweit physikalistische Ansätze, die in irgendeiner Form eine Identität dieser Ebenen behaupten, wahr sein können, wenn wir doch so unterschiedliche Eigenschaften (z.B. der „Körnung“) auf ihnen vorfinden.

Allerdings hätte man hiermit aber den Bereich rein quantitativer Beziehungen zwischen Gegenständen oder Eigenschaften verlassen, da bereits bestimmte *kompositorische* Verhältnisse angenommen werden. Außerdem könnte kritisiert werden, dass Komplexität im Hinblick auf mentale Zustände nicht unbedingt dasselbe meint wie im Hinblick auf neuronale Zustände.

Wie gezeigt wurde, sind es vor allem *kompositorische* Prinzipien, die eine plausible Möglichkeit zur Individuation *natürlicher* Ordnungs- oder Organisationsebenen bereitstellen. Trotz der Annahme eines Substanzmonismus‘ ermöglicht die Einteilung in Ganzes und Teile eine einleuchtende *natürliche* Gliederung von Objekten. So bestehen ein Lebewesen oder auch ein Organ wie das Gehirn aus einer Anzahl von Zellen, die ihrerseits aus Molekülen zusammengesetzt sind und von denen sich wiederum jedes einzelne Molekül in Atome zerlegen lässt. Eben solche Zergliederungen sind im Hinblick auf kausale Abläufe (z.B. in Form von Mechanismen) möglich. Mereologische Beziehungen können verständlich machen, wie die Vielheit an unterschiedlichen Repräsentationen eines Objektes trotz dessen substantieller Einheit zustande kommt. Zusätzlich erweitern mechanistische Ebenen, wie sie vor allem von Bechtel oder Craver propagiert werden, das mereologische Prinzip um die prozess- bzw. kausalitätsbezogene Dimension. Im Gegensatz zu „reinen“ mereologischen (Objekt-) Ebenen, die sich auf eher statische Objekte beziehen, beziehen sich mechanistische Ebenen auf *kausal agierende Entitäten*.

Auf den ersten Blick mag es abwegig erscheinen, die Existenz einer natürlichen Ebene des Psychischen über kompositorische Prinzipien begründen zu wollen. Denn kann man mentale Phänomene allen Ernstes als etwas verstehen, das aus physischen Gegebenheiten in irgendeiner Form *zusammengesetzt* ist? Auch wenn dies in der Tat gewissen Intuitionen widersprechen mag, scheint mir eine solche Beziehung nicht nur mög-

lich, sondern fast eine zwingende Konsequenz der gegenwärtigen wissenschaftlichen Befundlage zu sein, was ich anhand einiger Beispiele aufzeigen möchte.

Zunächst zu einem Beispiel aus der visuellen Wahrnehmung: Es ist bekannt, dass Farbinformationen im Kortex durch so genannte doppelte Gegenfarbenzellen in den „Blob-Regionen“ des visuellen Kortex im Okzipitallappen verarbeitet werden (vgl. Gegenfurtner, 2003, S. 16ff ; Tessier-Lavigne & Gouras, 1996, S. 468ff). Wir haben hier ein funktionales System der Informationsverarbeitung als neuronales Korrelat der Farbwahrnehmung⁹¹ vorliegen, was aus Gründen der Einfachheit als V_{CORF} bezeichnet werden soll. Im Kern macht dieses die Blob-Regionen des Parvo-Blob-System⁹² aus, das Fasern von der okzipitalen Region V1 in das Areal V2 und von dort in den ventralen Okzipitalkortex V4 entsendet. Insbesondere V4 zeigt hierbei eine farbsensitive Aktivierung (vgl. Zeki, 1990; Engel, Zhang & Wandell, 1997). Aber auf die empirischen Details muss es an dieser Stelle nicht so sehr ankommen.

Es liegt auf der Hand, dass die Funktionseinheit V_{CORF} als ein Ganzes gegenüber den Funktionseinheiten einer niederen Stufe, in erster Linie also den Neuronen und synaptischen Verbindungen, aus denen der Kor-

⁹¹ Natürlich könnte man auch einen realistischeren komplexeren phänomenalen oder kognitiven Zustand heranziehen, der aus verschiedensten phänomenalen Elementen wie akustischen, olfaktorischen Eindrücken, Gedanken, Überzeugungen etc. zusammengesetzt ist und an dem verschiedenste neuronale Netzwerke beteiligt sind. Die Gültigkeit der hier dargestellten *prinzipiellen* Erwägungen bleibt hiervon unberührt.

⁹² Das Parvo-Blob-System stellt eine abgegrenzte Bahn für die Verarbeitung von Farbinformationen dar. „Die Zellen in den Blobs haben runde rezeptive Felder und besitzen deshalb auch keine Orientierungspräferenz“ (Tessier-Lavigne & Gouras, 1996, S. 470). Im Gegensatz dazu werden von den orientierungselektiven Zellen in den Interblob-Regionen des visuellen Cortex, die ihren Ausgangspunkt in den parvozellulären Schichten des Corpus geniculatum laterale haben („Parvo-Interblob-Bahn“), Informationen über Formen verarbeitet (vgl. ebd., S. 401).

tex in diesem Bereich bzw. Netzwerk besteht, angesehen werden muss. Es kann hier also insofern von einem mereologischen Verhältnis gesprochen werden, als dass V_{CORF} (bzw. im Allgemeinen: die neuronale Funktionseinheit zur Hervorbringung eines mentalen Zustandes) die Ebene des Ganzen konstituiert und die hier angesiedelten Neurone die Teile dieses Ganzen sind. Damit wäre zunächst einmal eine Individuation von zwei natürlichen Ebenen gegeben: V_{CORF} lässt sich als höhere Ebene und die sie konstituierenden Zellen als niedrigere Ebene ansehen.

Wenn nun gelten soll, dass die Farbempfindung selbst in einem kompositorischen Verhältnis zu ihren neuronalen Grundlagen steht, müssen die Farbempfindung und das V_{COR} -konstituierende neuronale Netzwerk denselben räumlich-zeitlichen Bereich in der Welt einnehmen, sie müssen also *koextensiv* sein. Wenn ein Ganzes in Teile zerlegt werden kann, dann muss für die Teile gelten, dass sie zur gleichen Zeit das gleiche räumliche Feld einnehmen wie das Ganze. Bevor dies eine weitere Begründung erfährt, soll zunächst das bereits thematisierte Beispiel des räumlichen Gedächtnisses als eines kognitiven Prozesses wieder aufgegriffen werden.

In Bezug auf die hier vorliegenden mechanistischen Ebenen, wie sie Craver (2007) herausgestellt hat (vgl. Kap. 5.1.2.3), bedeutet dies, dass der Hippocampus, der eine räumliche Kartierung *generiert*, die eine LTP *induzierenden* Neuronen und die NMDA-Rezeptor-Aktivierung gemäß der in Abbildung 23 dargestellten Struktur des Mechanismus in einem kompositorischen und damit koextensiven Verhältnis stehen. Im gleichen Sinne gilt dies dann aber auch für das „Zielphänomen“ des räumlichen Gedächtnisses der Maus bei der Labyrinth-Aufgabe. Dieser kognitive Prozess müsste – wenn er selbst eine Ebene in einer Hierarchie mereologisch-mechanistischer Ebenen darstellen soll – in einem kompositorischen und damit koextensiven Verhältnis zu seinen neuronalen (pro-

zessualen) Grundlagen stehen, was in dem Fall ein komplex-distribuiertes neuronales Netzwerk sein wird, in dem der Hippocampus-Formation – wie von Craver vorgeschlagen – eine zentrale Stellung („constitutive relevance“) zukommt. Allerdings liegt das Augenmerk hier nicht auf mentalen *Zuständen*, sondern auf mentalen *Prozessen*, die wiederum eine aktivitätsbezogene neuronale Grundlage aufweisen, was der aktuellen neurowissenschaftlichen Forschungslage bzgl. des neuronalen Korrelats des Psychischen gerechter werden dürfte als eine eher statische mereologische Beziehung.

So weit, so gut. Gibt es aber noch weitere Bedingungen, die erfüllt sein müssen, damit man mit Berechtigung von einem Ganzen sprechen kann? Nach der plausiblen Analyse von Minimalanforderungen an ein Ganzes von Rescher und Oppenheim (1955) gilt:

1. Das Ganze muss ein besonderes Merkmal oder eine Charakteristik aufweisen, mittels dessen es als ein Ganzes verstanden werden kann.
2. Die Teile des Ganzen müssen in einer besonderen und charakteristischen Abhängigkeitsbeziehung zueinander stehen; sie müssen eine spezifische Beschaffenheit aufweisen, mittels derer sie ihren Status als Teil erfüllen.
3. Das Ganze muss eine gewisse *Struktur* besitzen, kraft derer es spezifische Strukturcharakteristika besitzt (vgl. ebd., S. 90).

Auch wenn manche Autoren diese Kriterien für relativ trivial halten mögen (vgl. Simons, 1987), sind sie doch – vielleicht gerade aufgrund ihrer analytisch-apriorischen „Trivialität“ – die minimalen (metaphysisch unspektakulären) Anforderungen an ein Ganzes.

Die erste Bedingung kann als klar erfüllt betrachtet werden. Denn die neuronale Funktionseinheit wird genau genommen erst zu einer *Einheit*,

indem sie durch eine bestimmte *Funktion* oder Aufgabe gekennzeichnet ist. Erst durch die Hervorbringung einer z.B. phänomenalen Röte (vs. der räumlichen Orientierung) kann man im Hinblick auf das relativ undifferenzierte Geflecht zellulärer Gegebenheiten, die unser Gehirn ausmachen, ein Teilnetzwerk oder „Bereich“ herausgreifen und als funktionale Einheit verstehen⁹³. Demnach begründet die Farbempfindung überhaupt erst den Funktionsbereich *V_{CORF}*. Dieser wird also erst durch die Hervorbringung von Farbempfindungen zu einem echten „Ganzen“. Ebenso gilt im Hinblick auf den Prozess des räumlichen Gedächtnisses, dass erst durch die Annahme seiner Existenz aus der „blinden“ Feuerungsrate von Neuronen im Hippocampus etc. eine funktionale Einheit geschaffen wird, die sich in verschiedene mechanistische Ebenen zerlegen lässt. Dies deckt sich mit den Bestimmungen Wimsatts (2007): „Ontologically, one could take the primary working matter of the world to be causal relationships, which are connected to one another in a variety of ways – and together make up patterns of causal networks. [...] Under some conditions, these networks are organized into larger patterns that comprise *levels of organization* [...]”. (S. 200; Hervorh. im Orig.).

In ähnlicher Weise wird auch die zweite Bedingung erfüllt. Denn die Teile des Ganzen stehen in beiden Beispielen insofern in einer besonderen und charakteristischen Abhängigkeitsbeziehung zueinander, als dass sie ein neuronales Netzwerk mit einer speziellen Funktion mit kausalen

⁹³ Natürlich könnte man an dieser Stelle mit epistemologischen Einwänden kommen und kritisieren, dass es gerade dies ist, was zu zeigen wäre. Aber es soll an dieser Stelle nicht um die Herleitung konkreter psychophysischer Verhältnisse gehen, sondern um die Möglichkeit einer Individuation mentaler Zustände über mereologische Prinzipien. Die wissenschaftliche Forschungslage, die besagt, dass Farben in diesem Bereich verarbeitet werden, wird dabei als Tatsache akzeptiert (wenn sich auch in der zukünftigen Forschung herausstellen sollte, dass es doch eine andere kortikale Region ist, die Farben verarbeitet, bleibt das hier diskutierte grundsätzliche *Prinzip* unberührt).

Verflechtungen zwischen den Elementen dieses Netzwerkes bilden. Jedes Neuron in diesem Netzwerk ist durch seine spezifische funktionale Rolle an der Hervorbringung der Rot-Empfindung bzw. der räumlichen Orientierung beteiligt.

Der dritte Punkt ist nicht ganz so eindeutig. Gemeint ist hiermit, dass es eine zentrale Bedeutungsfacette eines Ganzen ist, dass ein Ganzes mit einer strukturierten Organisation von Elementen zusammenhängt. „A structured whole in this sense involves three things:

- (1) its *parts*,
- (2) a *domain of 'positions'* which these parts 'occupy' (this need not necessarily be spatial or temporal, but may have any kind of topological structure whatever), and
- (3) an *assignment* specifying which part occupies each of the positions of the domain” (Rescher & Oppenheim, 1955, S. 100; Hervorh. im Orig.).

Ein klassisches Beispiel hierfür ist nach Ansicht der Autoren die Darbietung einer Musikkomposition: „the time-interval of the performance serves as the *domain of positions*, the various tones (as characterised by their pitch, volume, and duration) which are played throughout the performance are the *parts*, and the score which fixes the distribution of these notes throughout the performance is the *assignment*” (ebd., S. 100, Hervorh. im Orig.).

Lässt sich dies auf die Beispiele der Rot-Empfindung und des räumlichen Gedächtnisses übertragen? (1) und (2) dürfte man nunmehr relativ problemlos zustimmen können. Wie oben aufgezeigt, ist es prinzipiell möglich, dass eine Rot-Empfindung etwas ist, das sich aufgrund der Koextensivitäts- oder Identitätsannahme aus Teilen (einem neuronalen

Netzwerk V_{COR}) zusammensetzt. V_{COR} wäre die in (2) geforderte topologische „domain of positions“. Entsprechendes gilt für das räumliche Gedächtnis, bei dem eine solche Gliederung durch die Struktur des Mechanismus noch evidentester ist.

Schwieriger könnte der dritte Aspekt sein, denn ist es wirklich so, dass man von den Teilen von V_{COR} oder des Hippocampus (sprich: den Neuronen und wiederum deren Komponenten) sagen kann, dass sie eine ganz bestimmte Position in dem Funktionsbereich einnehmen? Es könnte aufgrund der neuronalen Plastizität zumindest schwierig sein, eine solche Aussage bezüglich des mentalen *Typs* oder der mentalen *Eigenschaft* von Rot-Empfindungen oder auch räumlicher Gedächtnisleistungen treffen zu wollen. Im Hinblick auf lokale Einzelfälle (Token), wie sie hier im Hinblick auf kompositorische Ebenen betrachtet werden sollen, sollte es dagegen schon möglich sein, für jeden Teil des neuronalen Netzes eine konkrete Position zu bestimmen. Letztendlich ist dies aber eine Frage, die die empirische Forschung *en détail* beantworten muss, für die es *prima facie* keine prinzipiellen Hinderungsgründe geben sollte.

5.2.2. Kategorienfehler und der „Mythos des Gegebenen“

Inwieweit kann man aber aus berechtigten Gründen behaupten, dass eine Farbempfindung (resp. räumliche Gedächtnisleistung) gegenüber ihren neuronalen Grundlagen ein *Ganzes* darstellt bzw. dass die Neuronen eines bestimmten neuronalen Netzwerkes *Teile* einer Farbempfindung (resp. Gedächtnisfunktion) sind? Ist eine Farbempfindung überhaupt etwas, von dem man sinnvoll sagen kann, dass es aus Teilen *zusammengesetzt* ist? Ist nicht bereits an dieser Stelle eine nicht zu akzeptierende Prämisse eingestreut worden? Genauer: ist es nicht ein *Kategorienfehler* in dem Sinne, dass man zwar von materiellen Dingen sagen kann, dass sie

aus Teilen bestehen, während wir bei mentalen Zuständen und hier gerade bei solchen phänomenaler Art wie Farbempfindungen nicht-aufspaltbare Einheiten haben (im Sinne von so etwas wie phänomenalen „A-tomen“, also nicht-teilbaren Entitäten). Natürlich wurde und wird von vielen Philosophen und Wissenschaftlern in genau diesem Sinne argumentiert (z.B. Laucken, 2002), was ich für nicht statthaft halte. Erstens hat man hiermit von vornherein einen problematischen (substantiellen) Dualismus festzementiert. Denn genau solche Intuitionen waren es, die Descartes (1641/1986) dazu bewogen haben, von einer klaren Trennung einer körperlichen und geistigen Substanz auszugehen. Hier zeigt sich ein großes Manko der *ordinary language philosophy* (vgl. Kap. 3.3). Wenn man etwas von einem Gegenstand nicht vernünftig *sagen* kann, dann handelt es sich nach Maßgabe dieses Ansatzes um einen Kategorienfehler. Problematisch ist hieran, dass unsere sprachlichen Gewohnheiten, die durch den kartesischen Dualismus infiltriert sind, so zum unangreifbaren Maßstab für die Beurteilung des Realitätsgehaltes einer Aussage erhoben werden. Zweitens scheint es allgemein keine sinnvolle Strategie zu sein, von der Art des Erlebens darauf zu schließen, was dieses Erleben genau ist bzw. was es nicht sein kann. Sellars (1963/1971) nennt dies den *Mythos des Gegebenen*. Wir haben den Eindruck, dass uns unsere mentalen Zustände unmittelbar gegeben sind, sodass wir, wenn wir sie genau beobachten, unmittelbar wissen, *was* sie sind. Dabei hat die wissenschaftliche Erforschung des Geistes immer wieder gezeigt, dass sich unsere gewöhnlichen Meinungen über die Natur mentaler Phänomene als falsch herausstellen können. So muss die Modularität bei der Verarbeitung und „(Re-)Konstruktion“ kognitiver, aber auch phänomenaler Zustände als neurowissenschaftlich gesicherter Fakt angesehen werden, auch wenn dies der von Phänomenologen gepriesenen erlebten „Einheit der Wahrnehmung“ oder „Einheit des Ichs“ vermeintlich entgegenstehen mag. So haben wir

den Eindruck, einer gewissen Einheitlichkeit des Erlebens und empfinden uns als „ein psychisch integrales Ganzes und nicht etwa als ein Sammelsurium einzelner Komponenten“ (Mausfeld, 2006, S. 4f), obwohl dies der faktisch bestehenden extremen Selektivität bei der Verarbeitung unserer Wahrnehmungsinhalte diametral entgegensteht.

Auch wenn eine Zerlegbarkeit eines mentalen Zustandes (bedingt durch die Zerlegbarkeit der mit ihm koextensiven physischen Struktur) also kontraintuitiv sein mag, heißt es also nicht, dass eine solche unmöglich ist. Des Weiteren zeugen eine Vielzahl neuropsychologischer Störungsbilder wie Agnosien oder Neglekte von der Möglichkeit eines hoch selektiven Ausfalls bestimmter phänomenaler oder kognitiver Elemente bei komplexeren mentalen Zuständen (also solchen, die selbst aus verschiedenen mentalen Teilkomponenten wie z.B. Farb-, Geruchs-, Sprach- oder räumlichen Aspekten zusammengesetzt sind). So können z.B. selektiv Bewegungs- oder Schmerzinformationen im Rahmen bestimmter Störungsbilder ausfallen.

Tatsächlich lässt sich bei einer Läsion in der V4-Region eine *zerebrale Achromasie*, eine Farbenblindheit trotz ansonsten normaler visueller Funktionstüchtigkeit, beobachten (vgl. Gegenfurtner, 2003). Vermutlich nimmt die Empfindungsstärke dabei mit dem Ausfall von Teilkomponenten des Systems ab bis hin zu einer nicht mehr kompensierbaren Störung des Gesamtsystems. In Bezug auf das von Craver herangezogene räumliche Gedächtnis stellt Kerkhoff (2003) eine Gruppe von *räumlich-topographischen Störungen* heraus als „räumliche Navigationsdefizite im vorgestellten oder realen dreidimensionalen Raum“ (S. 181), die mit unilateralen rechts- oder linksseitigen (para)hippocampalen Läsionen verknüpft sind. Mit dem durch die Läsion bedingten Ausfall einzelner Komponenten dieses mentalen Prozesses kommt es also offensichtlich dazu, dass der Genauigkeits-, Detailliertheitsgrad o.Ä. der räumlichen Repräsentation

tion der Umgebung abnimmt bis hin – wenn der Ausfall der Komponente nicht kompensiert werden kann – zu einer empfindlichen Störung des Gesamtsystems und damit zu einem „Totalausfall“ der Funktion.

Darüber hinaus findet sich in der *neuronalen Synchronisation* (vgl. Singer, 1999) ein weiteres Beispiel, das den „kompositorischen Charakter“ mentaler Prozesse verdeutlicht. In vielen Fällen komplexerer mentaler Prozesse, bei denen eine Integration von verschiedenen repräsentationalen/ phänomenalen Einzelaspekten erforderlich ist, zeigt die wissenschaftliche Befundlage, dass ihre *Bindung* über eine synchrone oszillatorische Aktivität von hochgradig verteilten neuronalen Antworten in kortikalen Netzwerken realisiert wird. Auch hier kann der mentale Zustand als die höherstufige Einheit bestimmt werden, in die die verschiedenen neuronalen Zentren mit den gleichen Schwingungsmustern eingehen, wodurch im Umkehrschluss wiederum von einer Zerlegbarkeit dieses mentalen Prozesses in seine kortikal verteilten Einzelkomponenten ausgegangen werden muss. Dass es sich so verhält, wird auch hier durch pathologische, relativ selektive Ausfallerscheinungen kognitiver Funktionen z.B. im Rahmen der Schizophrenie gestützt, die bei einer fehlerhaften neuronalen Synchronisation verschiedener Gehirnteile auftreten (vgl. Uhlhaas & Singer, 2006).

Man kann also festhalten, dass die phänomenale Nicht-Erfassbarkeit einer durch Teile konstituierten Organisationsstruktur einer Farbempfindung kein Hinderungsgrund dafür sein sollte, eine solche Organisationsstruktur als *möglich* zu betrachten. Auch andere Teil-Ganzes-Beziehungen wie die zwischen Atomen und Elementarteilchen sind phänomenal wenig einsichtig (obgleich man sich hier bestimmter Annäherungen/ Analogien bedienen mag).

Wenn wir uns nun also wieder der Frage nach der Ursache des Leib-Seele-Problems bzw. der Gehirn-Geist-Differenz zuwenden, dann wäre die Ant-

wort, die die objektivistische Auflösung von Spinozas Dilemma hierauf gibt, dass es sich bei Gehirn und Geist um unterschiedliche Ebenen einer natürlichen Strukturgliederung handelt. Psychische Prozesse sind auf einer höheren Strukturebene angesiedelt als neuronale Prozesse. Neurone sind gewissermaßen die „Bausteine des Geistes“ (Dörner, 1998/2008, S. 61). Des Weiteren ist hier schön zu sehen, wie einseitig die Betrachtung lediglich neuronaler vs. mentaler Prozesse eigentlich ist. Denn auch neuronale Prozesse bilden eine relativ hochstufige Strukturebene, die sich in mannigfaltigste niedrigere Ebenen aufspalten lässt, auf denen ebenso kausale relevante (agierende) Entitäten vorzufinden sind wie auf der neuronalen Ebene.

Jedenfalls kann aber bereits an dieser Stelle konstatiert werden, dass sowohl für eine Individuation des Psychischen über mereologische als auch mechanistische Ebenen gilt, dass sie eine *Koextensivität* psychischer Zustände und Prozesse mit ihren neuronalen Grundlagen implizieren. Wenn ein psychischer Zustand wie eine Rot-Empfindung ein Ganzes und ein neuronales Netzwerk, das vor allem verschiedene Areale des okzipitalen Kortex' umfasst, dessen Teile ausmacht, dann müssen sie den gleichen räumlich-zeitlichen Bereich in der Welt einnehmen, wie es für alle kompositorischen Verhältnisse gilt. Man könnte also sagen, dass die *Koextensivität* des Geistes mit seinen neuronalen Grundlagen der „Preis“ ist, den man dafür zahlen muss, damit ein mentaler Zustand über kompositorische Prinzipien als natürliche Organisationsebene individuiert werden kann.

5.2.3. Mereologisch-mechanistische versus aggregative Individuation

Es bleibt zu klären, ob mit Rekurs auf *aggregative* Verhältnisse eine Individuation der mentalen Ebene vorgenommen werden kann. Natürlich

ließe sich, wie in der gerade diskutierten Form behaupten, dass mentale Eigenschaften letztendlich als eine Summe niederstufiger Eigenschaften angesehen werden können. Dies wäre aber eben nur eine bloße *Behauptung*, die offenbar nicht in irgendeiner explanatorischen oder deskriptiven Form abgeleitet werden kann, sondern sie ist gerade das, was zu zeigen wäre. Inwiefern sind mentale Eigenschaften denn höherstufige Eigenschaften und neuronale Eigenschaften niederstufige Eigenschaften? Unter Heranziehung mereologischer Prinzipien ließ sich über die nicht unplausible Annahme einer Koextensivität mentaler Ereignisse und ihrer neuronalen Grundlagen eine Teil-Ganzes-Beziehung zwischen diesen annehmen und darüber eine natürliche Ebene des Mentalen individuierten. Eine solche Ableitung scheint hier – zumindest ohne den zusätzlichen Rekurs auf mereologische Prinzipien – nicht gegeben zu sein. Dennoch könnten aggregative Verhältnisse zwischen mentalen und neuronalen Eigenschaften natürlich *möglich* sein. Gerade unter der aus mereologischen Prinzipien abgeleiteten Konsequenz, dass eine psychoneuronale Koextensivität vorliegt, scheint der Weg zu der Annahme nicht weit zu sein, dass auch eine Koextensivität mentaler und neuronaler *Eigenschaften* besteht. Demnach wäre eine mentale Eigenschaft (z.B. das Haben von Rot-Empfindungen) koextensiv mit einer Summe neuronaler Eigenschaften (Entladungsrate des zugrundeliegenden Neuronenverbandes, Neurotransmitterkonzentrationen etc.). Es bleibt allerdings unklar, inwieweit eine solche Koextensivität von Eigenschaften überhaupt eine inhaltlich kohärente Annahme ist und ob mentale Eigenschaften tatsächlich als ein bloßes Aggregat, also eine *Ansammlung* neuronaler Eigenschaften angesehen werden können. Der letzte Punkt wäre eine stärkere These als die

einer mereologischen Gehirn-Geist-Beziehung. Wie sich zeigen an späterer Stelle wird, ist eine solche Annahme aber tatsächlich abwegig (vgl. Kap. 7).

5.3. Spinozas fehlendes Puzzlestück: Gehirn und Geist als Trigger divergierender konzeptueller Formen

Zurück zur Leitfrage, wie sie in Kapitel 3 bei der Frage nach dem Ursprung des Leib-Seele-Problems mit Bezug auf Spinoza aufgeworfen wurde: Handelt es sich also bei der wahrgenommenen Differenz zwischen Gehirn und Geist um ein Phänomen, das nur durch uns selbst in die Welt getragen wird, ist der Ursprung des Leib-Seele-Problems also im Psychischen selbst zu suchen, oder handelt es sich bei dieser Differenz doch um etwas, das eine objektive, beobachterunabhängige Basis aufweist? Wenn die Achillesferse des Spinozistischen Systems also der strittige Status der Attribute der „all-einen“ Substanz ist (und damit eine Art *Knotenpunkt des Weltknotens*⁹⁵ markiert), dann kann dieses Dilemma aufgrund der hier getätigten Analysen bzw. unter Heranziehung wissenschaftlicher Theorien und Befunde nun sehr elegant aufgelöst werden. Die Teil-Ganzes-Beziehung ist gewissermaßen das fehlende Puzzlestück Spinozas. Wenn die Substanz über verschiedene Modi „an sich selbst“ gegliedert ist und wir dies über deren Attribute erfassen können, dann sind diese Strukturgliederungen mereologische Verhältnisse. Die Diskrepanz zwischen der Einheit und nichtdestotrotz Vielgestaltigkeit der

⁹⁵ Arthur Schopenhauer (1813/2007) hat das Leib-Seele-Problem in seiner Dissertationsschrift als *Weltknoten* bezeichnet und dürfte mit dieser emphatischen Begriffsprägung so falsch nicht liegen.

Welt geht (zumindest in einem großen Ausmaß) auf das Vorhandensein einer natürlichen Strukturgliederung der Welt zurück (die Modi der Substanz), die wir mittels verschiedener perspektivengebundener Zugriffe (Attribute) erfassen können. Oder mehr an die Gegenwartsterminologie angepasst und weniger metaphysisch pompös aufgeladen, kann man sagen, dass Gehirn und Geist koextensive Strukturen in einer mereologisch-mechanistischen Hierarchie sind, wie sie für lebende Systeme wie den Menschen charakteristisch ist, und wir diese verschiedenen natürlichen Strukturebenen mittels bestimmter Repräsentationsformate voneinander unterscheiden können. Gehirn und Geist sind gewissermaßen *Trigger divergierender konzeptueller Formen*.

Wenn wir eine Verschiedenheit von Gehirn und Geist wahrnehmen, hat dies demnach nicht nur mit unserem Wahrnehmungs- und Kognitionssystem selbst zu tun, sondern geht darauf zurück, dass es sich hier um einen *faktisch vorhandenen* Unterschied handelt. Es ist allerdings keine problematische substantielle Verschiedenheit von neuronalen und psychischen Prozessen, sondern nur eine solche der Divergenz der Organisationsebene. Sie gehören gemäß Wimsatts Bestimmungen zu der Subklasse von Perspektiven, die kompositorisch aufeinander abgebildet werden können und deren Entitäten sich ohne überkreuzende Überlappungen in einer Hierarchie anordnen lassen. Es sind daher nicht nur Perspektiven, die auf eine subjektiv sprachliche, perzeptuelle oder kognitive Zugriffsweise zurückgehen oder – wie Pauen (1999) sagt – auf die Entfernung zum fraglichen Gegenstand. Psychische Prozesse sind auf einer höheren Strukturebene angesiedelt als neuronale Prozesse (Erklärung der Vielheit). Dennoch gilt wie für alle (agierenden) Entitäten mereologischer bzw. mechanistischer Ebenen, dass die Strukturkomponenten der verschiedenen Ebenen in einem *koextensiven* Verhältnis stehen (Erklärung der Einheit). Psychische Prozesse nehmen mit Notwendigkeit das

gleiche räumlich-zeitliche Feld in der Welt ein, wie die (agierenden) Entitäten aller ihnen zugrunde liegenden mereologisch-mechanischen Ebenen. Aufgrund der vorgestellten Wahrnehmungsmechanismen löst die Betrachtung physischer Phänomene und psychischer (als psychischer) Phänomene allerdings eine ganz andere Aktivierung von konzeptuellen Formen der Wahrnehmungs- und höherstufigen kognitiven Systeme aus, die mit der Verarbeitung materieller vs. psychischer Gegebenheiten betraut sind⁹⁶. Gehirn und Geist triggern die Einnahme bestimmter Perspektiven, die wiederum aufgrund unserer (phylogenetischen und ontogenetischen) Dispositionen inkommensurabel erscheinen. Die Gehirn-Geist-Differenz ergibt sich folglich aus einem *Zusammenspiel von objektiven und subjektiven Faktoren*. Wenn an früherer Stelle von Sortals die Rede war, die determinieren, wie wir einen Gegenstand sprachlich-konzeptuell individuieren, so können wir auch diesen Aspekt nun noch genauer fassen. So können wir einen Gegenstand als Statue ansehen, aber auch als einen Klumpen Lehm oder – um noch einen Schritt weiterzugehen – als Ansammlung von Atomen. Als was wir den Gegenstand wahrnehmen, hängt von unserer Perspektive im Sinne einer divergierenden Art und Weise der Repräsentation ab. Typischerweise würden wir sagen, die Statue ist aus Lehm geformt und dieser wird letztendlich durch eine Ansammlung von Atomen konstituiert. Der Bezugspunkt ist hier also die Statue, über die wir Aussagen treffen. Aber warum die Statue und nicht der Lehm oder die Ansammlung von Atomen? Die Antwort ist in dem

⁹⁶ Man mag hier einen bestimmten argumentativen Zirkel ausfindig machen zu meinen. Solchen Zirkeln unterliegt man bei der Heranziehung mentaler Charakteristika zur Erklärung bestimmter Phänomene aber mit Notwendigkeit, denn jedes zu erklärende Phänomen ist ein solches, das wir *mental* erfassen müssen. Dennoch ist bereits aus der psychologischen Forschung klar, dass psychologische Mechanismen (wie Attributionen, kognitive Tendenzen und dergleichen) ein großes explanatorisches Potential haben.

Umstand zu suchen, dass wir mit Statuen in unserer – immer nach bedeutungshaltigen Mustern Ausschau haltenden – Alltagskonzeption der Wirklichkeit bzw. aufgrund der Beschaffenheit unseres sensorischen und perzeptuellen Systems (die nur einen ganz bestimmten *Auflösungsgrad* zulässt) eher konfrontiert sind als mit genuin stofflich-materiellen oder (mikro-)physikalischen Größen und Eigenschaften. Insofern ist die Statue unser Bezugspunkt, wenn wir Aussagen über diesen Wirklichkeitsausschnitt vornehmen, obgleich dieser potentiell auf völlig verschiedene Arten und Weisen individuiert werden könnte.

Man muss sich also die Konsequenz vor Augen führen, dass es genau genommen gar keinen festen Bezugspunkt gibt, von dem aus man etwas über die Beschaffenheit eines Gegenstandes aussagen kann. Was wir als ein Objekt individuierten, hängt immer von bestimmten Vorannahmen, Gewohnheiten etc. ab. Denn inwiefern sollten die (agierenden) Entitäten irgendeiner beliebigen Ebene dessen eigentliche Realität ausmachen können? Unsere eigene physisch-psychische Beschaffenheit als Maßstab heranzuziehen, ist genau genommen eine anthropomorphe Vereinseitigung. Bereits Spinoza (1677/1977) sagte deshalb, dass der Substanz unendlich viele Attribute zukommen müssen (vgl. ebd., S. 5). Nur der Umstand, dass wir daraus nur wenige (oder doch zumindest nicht alle) erblicken, ist unserer spezifisch menschlichen Verfasstheit zuzurechnen, nicht aber die Tatsache der essentiellen „strukturellen Gliederung“ der Substanz selbst.

An dieser Stelle lohnt es sich, das von Nagel aufgeworfene Problem der Subjektivität aufzugreifen. Wie bereits in Kapitel 2.1.4 aufgezeigt wurde, würden Ansätze, die mentale Phänomene auf irgendeine Art mit physikalischen Phänomenen identifizieren, nach Ansicht Nagels vor dem Problem stehen, dass sie der subjektiven Erlebnisqualität des menschlichen Geistes nicht gerecht werden können. Jeglicher Objektivierungsversuch

bzgl. psychischer Prozesse mit qualitativen Eigenschaften sei daher zum Scheitern verurteilt, weil eine hierzu herangezogene Theorie immer einen entscheidenden Aspekt auslassen müsse: das *Wie-es-ist* in solchen Zuständen zu sein.

Zur Erinnerung: Nagel monierte abschließend, dass es – wenn zwei Dinge phänomenal so verschieden sind wie Gehirn und Geist – unverständlich bliebe, wie sie miteinander identifiziert werden könnten. „Wir mögen dann nicht einmal eine vage Idee davon haben, wie oder in was die beiden referentiellen Pfade konvergieren könnten. Ein theoretischer Rahmen mag hinzugefügt werden müssen, damit wir dies verstehen können“ (ebd., 1974/1997, S. 270). Oder – wie Bieri (1981a/1997) sagt – sei die Aufgabe, ein Verständnis unserer Innenperspektive zu etablieren, das mit der „materialistischen Einsicht in Übereinstimmung gebracht werden kann“ (S. 207).

Der geforderte theoretische Rahmen, um diese Aufgabe zu erfüllen, ist hiermit etabliert worden. Zunächst muss man die Prämisse Nagels infrage stellen, dass die objektive Natur eines Phänomens zu erkennen, bedeute herauszufinden, was dieses Phänomen *unabhängig* von allen subjektiven Perspektiven ist (vgl. Bieri, 1981/1997, S. 305). Wie aufgezeigt wurde, gibt es keine Perspektive, die die eigentliche Wirklichkeit eines Gegenstandes ausmachen könnte. Die einzige Differenz mit ontologischen Implikationen, die man in Bezug auf Perspektiven ausfindig machen kann, ist, dass nur für einen bestimmten Teilbereich von Perspektiven bzw. perspektivengebundenen Repräsentationsformen gilt, dass sie eine objektive (beobachterunabhängige) Entsprechung haben (entgegen einer „verschwommenen“ versus „klaren“ Perspektive oder dergleichen). Und für diese nach Maßgabe kompositorischer Prinzipien feststellbaren natürlichen Organisationsebenen gilt, dass keine von ihnen einen „höheren“ Realitätsstatus hat als eine andere. Man könnte ganz ähnlich wie

Mausfeld oder Nietzsche sagen, dass – entgegen Nagel – nur die (potentielle) Gesamtheit aller Perspektiven das eigentliche Phänomen darstellt und zwar solcher Perspektiven, die eine beobachterunabhängige Entscheidung in Form einer natürlichen Organisationsebene haben.

Wie gezeigt wurde, spricht sehr viel dafür, dass es sich bei den Perspektiven „Gehirn“ und „Geist“ um solche natürlichen Organisationsebenen handelt, die in einem kompositorischen Verhältnis stehen. Unser Erleben ist nichts, was der materiellen Seinsform in irgendeiner Weise außen vor ist, dennoch erschöpft es sich nicht in seinen physischen Grundlagen. Es ist eine eigene höherstufige Komplexitätsebene, die in einem koextensiven Verhältnis zu ihren physischen Grundlagen (also den agierenden Entitäten aller niedrigeren Ebenen) steht. Es findet also durchaus eine gewisse *Identifizierung* des Psychischen mit seinen physischen Grundlagen statt, denn es handelt sich bei psychischen Prozessen um Vorgänge, die koextensiv sind mit den neuronalen (zellulären, molekularen...) agierenden Entitäten der niedrigeren Organisationsebenen in einer mereologisch-mechanistisch konstituierten Hierarchie. Die von Nagel geäußerte Befürchtung, dass die subjektive Perspektive ihre Gültigkeit oder Bedeutung verlieren könnte, wenn sie mit physischen Zuständen identifiziert wird, ist trotzdem offensichtlich völlig unberechtigt.

Was bleibt dann noch vom Subjektivitätsproblem, das Nagel beschrieben hat? Wie Pauen (1999) sagt, bleibt die Schwierigkeit bestehen, dass Zustände oder Prozesse, die über die Dritte-Person-Perspektive (was nunmehr korrekterweise eigentlich heißen müsste: über eine niedrige Organisationsebene als die psychische selbst) beschrieben werden, „aus einer Perspektive beschrieben [werden], die nicht charakteristisch für die Tatsache ist, daß sie Bewußtsein instantiieren. Solche Beschreibungen aus der Perspektive der dritten Person können außerdem wertvolle Hinweise auf die Präsenz von Bewußtsein geben, keinesfalls vermögen sie jedoch die

Bedeutung des Begriffes ‚Bewusstsein‘ zu erschöpfen“ (S. 28). Insofern zeigen das Leibnizsche Mühlengleichnis oder Bieris „Rundgang durch das Gehirn“ (vgl. ebd., 1994) weiterhin eine gewisse *epistemische* Kluft zwischen Gehirn und Geist auf. Was hieraus aber auf keinen Fall abgeleitet werden kann, ist, dass das Psychische – nur weil es vielleicht auf basaleren Organisationsebenen nicht abschließend beschrieben werden kann – irgendetwas ist, dass sich nicht in die natürliche Strukturierung der Welt eingliedern würde. Der einzige Unterschied ist, dass es manche Gegenstände gibt, die eine Hierarchie natürlicher Ebenen aufspannen, die bis zur psychischen Ebene „hinauf“ reicht, und manche, die dies nicht tun. Die Ursache dieser Differenz liegt aber nun nicht am Hinzukommen einer mysteriösen Seelensubstanz oder Eigenschaft, sondern einfach daran, dass nur bestimmte Systeme wie der Mensch die strukturelle oder funktionale Komplexität besitzen, die hierfür notwendig ist (hierzu an späterer Stelle mehr, vgl. Kapitel 7).

Es muss nun aber noch geklärt werden, was genau ein solches koextensives Verhältnis des Mentalen zu seinen neuronalen Grundlagen bedeutet. Ich möchte aufzeigen, dass man aus einer *Koextensivität*, die sich aufgrund einer kompositorischen Beziehung zwischen mentalen Zuständen als Ganzheiten und ihren neuronalen Grundlagen als Teilen auf eine bestimmte Form psychophysischer *Identität* schließen lässt. Welche Auswirkungen hat dies dann aber für den Status des Psychischen und die wissenschaftliche Disziplin der Psychologie? Könnte hierdurch der Wirklichkeitsstatus des Psychischen gemindert oder auch die Autonomie der Psychologie infrage gestellt werden?

Wenn eine (über Teil-Ganzes-Verhältnisse begründete) psychophysische Identität vorliegt, inwieweit kann man dann noch von einer Kausalbeziehung zwischen den Ebenen sprechen (Inter-level-Causation) oder überhaupt die Möglichkeit einer mentalen Verursachung physischer Ereignisse-

se verständlich machen? Gibt es insbesondere so etwas wie abwärts gerichtete Kausalität, also eine Top-Down-Einwirkungsmöglichkeit? Muss es überhaupt eine solche geben, um die Eigenständigkeit (der Entitäten) der höherstufigen Wissenschaften bewahren zu können? Eine weitere wichtige Frage, die es gilt zu beantworten, ist, inwieweit es Unterschiede zwischen dem Status von kognitiven und phänomenalen Zuständen gibt. Lässt sich die Individuation mentaler Zustände und Prozesse für beide Cluster genauso gut vornehmen?

Bevor all diese wichtigen Fragen beantwortet werden können, müssen die für die weitere Besprechung erforderlichen begrifflich-konzeptuellen Grundlagen noch klarer herausgearbeitet werden. Es wird daher im folgenden Kapitel zunächst das begriffliche Gerüst der gegenwärtigen Philosophie des Geistes näher beleuchtet, um dann in den darauffolgenden Kapiteln die genannten Fragen systematisch beantworten zu können.

6. Licht und Schatten: Psychophysische Konzepte der gegenwärtigen Philosophie des Geistes

Nachdem nun bereits eine schlüssige Antwort auf die Frage nach dem Ursprung des Leib-Seele-Problems gegeben und damit auch ein entscheidender Schritt zu seiner (Auf-)Lösung unternommen wurde, könnte die Frage aufkommen, warum es an dieser Stelle überhaupt notwendig sein soll, sich intensiver mit den Konzepten der psychophysischen Relation – wie sie in der gegenwärtigen Philosophie des Geistes diskutiert werden – zu beschäftigen. Warum kann nicht direkt auf der Basis der bisher gewonnenen Erkenntnisse die eigentliche Ausformulierung eines allgemeinen Modells einer Architektur des Geistes bzw. die Bestimmung des Gegenstands der Psychologie erfolgen? Für dieses Vorgehen gibt es eine Reihe wichtiger Gründe:

Wenn in den vorherigen Kapiteln von einer Gehirn-Geist-Einheit oder einer psychophysischen Identität die Rede war, ist dies primär in Abgrenzung zu (substanz-)dualistischen Theorien des Geistes geschehen, ohne dass eine solche Identitätsannahme selbst weitergehend analysiert und auf möglicherweise vorhandene Schwierigkeiten hin untersucht worden ist. Dabei ist die so genannte *Identitätstheorie* in der Philosophie des Geistes aber bereits ein festgefügtter Begriff für einen Lösungsvorschlag zum Leib-Seele-Problem, der auf ganz bestimmten (nicht unproblematischen) Vorannahmen basiert, sodass der Begriff einer psychophysischen Identität hierdurch stark geprägt ist. Es muss daher gezeigt werden, welche genauen Aussagen die Identitätstheorie trifft, inwieweit diese mit dem hier besprochenen Ansatz übereinstimmen oder auch von diesem abweichen. Es soll hierbei aufgezeigt werden, dass die (Typen-)Identitätstheorie vor allem durch die Ausformulierung bzw. Verteidigung durch Jack

Smart eine höchstproblematische Interpretationsweise erfahren hat, der keineswegs Notwendigkeit zukommt. Diese Identitätstheorie erhebt den Anspruch haben aufzudecken, was das tatsächliche „Wesen“ mentaler Prozesse ist, nämlich dass sie letztendlich *nichts anderes als* Vorgänge im Gehirn sind. Sie wurde in Folge dessen häufig gleichgesetzt mit der Position eines recht kruden *reduktiven* bzw. *reduktionistischem Physikalismus* oder *Materialismus*. Trotzdem wird durch die Besprechung der Identitätstheorie auch der Blick für Problemstellungen freigelegt, die im nächsten Kapitel dezidiert entfaltete Konzeption einer psychophysischen Identität von großer Bedeutung sind. Vor allem handelt es sich hierbei um das so genannte Argument der *multiplen Realisierbarkeit*. Aber auch die so genannte *Erklärungslücke* und das Argument der *starren Designatoren* von Saul Kripke sollen vorgestellt werden.

Auch die zweite Form einer Identitätstheorie, die (Token-)Identitätstheorie nach Donald Davidson, weist Probleme auf, vor allem weil diese an dessen These eines *anormalen Monismus* geknüpft worden ist. Die Folge ist, dass mit der Ablehnung von Davidsons „metaphysischem Rahmen“ auch die an sich brauchbare These einer psychophysischen Identität auf Token-Niveau vorschnell abgelehnt wird, obwohl auch für diese konzeptuelle Verschränkung der Sache nach kein Anlass besteht. Das Konzept einer psychophysischen Identität muss daher gewissermaßen aus seinen „begriffsgeschichtlichen Fesseln“ befreit werden, um im nächsten Kapitel eine gewisse Rehabilitation (auf der Basis einer mereologisch begründeten Koextensivität) erfahren zu können.

Des Weiteren muss der Frage nachgegangen werden, ob es denn überhaupt nötig ist, eine ontologisch so „starke“ These, wie sie eine Identitätsannahme darstellt, aus der Nicht-Dualität von Gehirn und Geist abzuleiten. Ist die Eigenständigkeit der Psychologie nicht gerade umso gefährdeter, je näher die Entitäten und Prozesse, wie sie von der Psychologie

postuliert und untersucht werden, an diejenigen „herangerückt“ werden, wie sie von der Biologie postuliert und untersucht werden? Ist dann eine psychophysische Identität nicht der partout zu vermeidende Extremfall von „ontologischer Nähe“, der die Möglichkeit zu einer Reduktion der Psychologie (bzw. ihrer Entitäten) auf die Biologie (bzw. ihrer Entitäten) eröffnen würde? Schenkt man dieser nicht unplausiblen Intuition Glauben (dass sie nicht greift bzw. selbst wenn sie greifen würde, für die Psychologie unproblematisch ist, wird an späterer Stelle zu zeigen sein; vgl. Kapitel 8 und 9), kann gefordert werden, dass geklärt werden muss, ob es nicht andere Möglichkeiten der näheren Bestimmung der psychophysischen Koextensivität gibt, die mit einer geringeren „ontologischen Nähe“ des Psychischen zu seinen neuronalen Grundlagen verbunden sind. Kann also nicht doch bereits auf der Basis von – gegenüber einer Identitätsannahme – ontologisch „schwächeren“ Konzepten eine gehaltvolle psychophysische Theorie begründet werden? Neben den Konzepten der *Emergenz* und *Supervenienz* muss hierzu vor allem die derzeit in der philosophischen Debatte dominierende *Realisierungstheorie* genauer untersucht werden. Grundsätzlich wird sich zeigen, dass der Einbezug dieser Konzepte der psychophysischen Relation – vor allem im Sinne der Etablierung einer größeren begrifflichen Klarheit – durchaus gewinnbringend ist. So nimmt z.B. das Konzept der *Emergenz* explizit Bezug auf die Teil-Ganzes-Beziehung und stellt – primär im Hinblick auf die so genannte *systemische Emergenz* – Charakteristika dieser Relation heraus, die auch für die vorliegende Arbeit von großer Bedeutung sind. Ebenso ist der Begriff der *mereologischen Supervenienz* sehr nützlich, weil er dem wichtigen Grundsatz, dass es keine Änderung auf der Ebene des Ganzen geben kann, ohne dass eine Änderung auf Ebene der Teile dieses Systems stattgefunden hat, Ausdruck verleiht. Die physikalische Realisierungstheorie fungiert schließlich zumeist als philosophische Antwort auf die

Frage nach der Interpretation der konkreten (ontologischen) Beziehung zwischen den in Kapitel 4 vorgestellten kognitionswissenschaftlichen Ebenen des Geistes. Abstrakte funktionale Eigenschaften werden demnach – vermittelt über ihr kausales Profil – durch bestimmte physische (neuronale) Eigenschaften bzw. Strukturen *realisiert*.

Auch wenn allen drei besprochenen Ansätzen also durchaus einiges abzugewinnen ist, wird sich zeigen, dass sie keine eigenständige Theorie des Geistes begründen können, die als echte Alternative zu einer psychophysischen Identitätsthese infrage kommt: Verweigern sie sich prinzipiell, mit Identitätsannahmen zu operieren, weisen sie innere Widersprüchlichkeiten oder zumindest doch eine wenig hilfreiche Beliebigkeit in der Charakterisierung der Gehirn-Geist-Beziehung auf. Lassen sie hingegen Identitätsannahmen zu, verschwinden diese Schwierigkeiten zwar, sie können dann jedoch nicht mehr als eine echte Alternative – in der hier verfolgten Hinsicht – zu einer identitätsbasierten Gehirn-Geist-Theorie angesehen werden.

6.1. Die psychophysische Identitätstheorie (Type-Identity-Theory)

Ein unglücklicher Umstand bei den meisten gegenwärtigen Darstellungen zur Philosophie des Geistes ist, dass sie den Eindruck erwecken, dass die aktuelle Besprechung der Probleme quasi aus dem Nichts entstanden wäre bzw. dass sie zumindest mit ihrer Art Lösung des Leib-Seele-Problems einen fundamentalen Bruch mit den „primitiven“ Lösungsversuchen der Vergangenheit vollzogen hätten. Beispielsweise meint Kim (1998), dass alle früheren Arbeiten entweder weit unter dem Niveau der neuen Entwicklung in der Philosophie des Geistes geblieben wären oder

sich nur mit der Logik mentaler Ausdrücke beschäftigt hätten (wie Ryle oder Wittgenstein) und somit das eigentliche Thema verfehlten. Angeblich sei die gesamte philosophische Leib-Seele-Debatte tief in der kartesischen Unklarheit und Verwirrung verfangen gewesen, bis die Identitätstheoretiker der 1950er Jahre auf den Plan traten, um diesem unsäglichem Zustand ein Ende zu bereiten (vgl. Heidelberger, 2002). So sagt Kim (1998) anerkennend, dass „the brain state theory [= Identitätstheorie; M.L.] helped set the basic parameters and constraints for the debates [...] – a set of broadly physicalist assumptions and aspirations that still guide and constrain our thinking today“ (S. 2). Tatsächlich reicht der Gedanke einer psychophysischen Einheit oder Identität viel weiter zurück und hat auch bereits hier durchaus tiefgründige und empirisch gehaltvolle Ausformulierungen erlebt (Lüdmann, unveröffentlichtes Manuskript; Heidelberger, 2002). Bereits in der Antike sind bei Parmenides, den Atomisten Leukipp und Demokrit und auch bei der älteren Stoa identitätstheoretische Elemente vorzufinden. Die Seele wird hier meist als eine spezifische Art von Stoff begriffen. Ebenso finden sich bei Aristoteles wichtige Ideen zur Etablierung einer nicht-dualistischen Theorie des Geistes, obgleich seine Ansicht nicht unbedingt als Propagierung einer psychophysischen *Identität* angesehen werden kann (hierzu im nächsten Kapitel mehr).

Es soll Erwähnung finden, dass es meines Wissens der stark durch Spinoza beeinflusste Friedrich Wilhelm Joseph von Schelling (1985, 1802/2005) war, der als erster den Begriff einer Identität im Hinblick auf die Beziehung von Körper und Geist (bzw. das „Reale“ und „Ideale“) verwandt hat. Die Bezeichnung *Identitätsphilosophie* hat sich zur Charakterisierung von Schellings Frühwerk deshalb eingebürgert, weil er bestrebt war aufzuzeigen, dass sich der Unterschied zwischen Freiheit und Determinismus, Subjekt und Objekt und damit auch zwischen Körper

und Geist auf eine hinter den Gegensätzen liegende Identität zurückführen lässt. In den Augen Schellings weist die materielle Natur immer schon ein geistiges Moment auf, während es wiederum keinen völlig unabhängig von der Materie existierenden Geist geben kann. Wie bereits besprochen wurde, entschied sich Schelling für eine extreme Auslegung der „subjektivistischen“ Auflösung von Spinozas Dilemma, indem jede Verschiedenheit der „all-einen“ Substanz und damit auch der Gegensatz von Körper und Geist auf unseren gedanklichen oder perzeptuellen Zugriff auf die Welt zurückgeht. Die Substanz selbst bleibt indifferent bzw. sie ist „an sich selbst“ vollkommen undifferenziert.

Mit den unabhängig voneinander entstandenen Texten der britisch-australischen Philosophen Ullin T. Place und John J. C. Smart auf der einen und dem österreichisch-amerikanischen Philosophen Herbert Feigl auf der anderen Seite²⁰⁵ ist in den 1950er Jahren – nach einigen Jahren ihrer Verbannung aus dem philosophischen Tagesgeschäft – die Besprechung der Leib-Seele-Frage wieder in den Mittelpunkt der Diskussion gerückt. In ihren Arbeiten plädieren die Autoren auf ähnliche Art und Weise für eine Identifikation von mentalen Zuständen mit bestimmten physischen Zuständen des Gehirns. Sie sind der zentrale Bezugspunkt, wenn in der heutigen Philosophie des Geistes von einer psychophysischen Identität die Rede ist, sodass eine Besprechung ihrer Arbeiten unabkömmlich ist, wenn man die derzeitige philosophische Fachdebatte und die immer noch diskutierten Probleme psychophysischer Identitätsaussagen (bzw. ihren „schlechten Ruf“) verstehen möchte.

Anlass für die Formulierung der Identitätstheorie war insbesondere die Frage, wie es möglich sein soll, dass psychische Gegebenheiten die Ursa-

²⁰⁵ Eine aussagekräftige Gegenüberstellung der „australischen“ gegenüber der „österreichischen“ Version der Identitätstheorie findet sich in Stubenberg (1997).

che von physischen Ereignissen sein können, obwohl die physische Welt ein in sich kausal abgeschlossenes System darstellt, bei dem jedes physische Ereignis bereits hinreichend durch vorhergehende physische Ereignisse determiniert wird (und damit auch jeder Vorgang in unseren Gehirnen bereits hinlänglich durch vorhergehende neuronale, hormonelle oder andere physische oder physiologische Ereignisse bedingt wird; vgl. Kap. 3.1.1). Also lag es durchaus nahe anzunehmen, dass die psychischen Prozesse selbst nichts anderes sind als Vorgänge in unserem Gehirn. Schließlich finden sich zu mehr oder weniger allen psychologischen Größen bestimmte neuronale Korrelate, die einen engen Zusammenhang unseres Erlebens und Denkens mit den Vorgängen in unseren Gehirnen belegen.

6.1.1. Ullin T. Place: Ist Bewusstsein ein Gehirnzustand?

In seinem Aufsatz *Is consciousness a brain process?* bezieht sich Place (1956) bei der Argumentation für eine psychophysische Identität vor allem auf die begriffliche Unterscheidung zwischen der Intension (Bedeutung, Konnotation) und der Extension (Referenz, Denotation) eines Ausdrucks, was auf die in der philosophischen Semantik Gottlob Freges getroffene Unterscheidung von Sinn und Bedeutung²⁰⁶ zurückgeht (vgl. Kap. 3.3). Die Identität von Gehirn und Geist verhalte sich daher wie diejenige von Morgenstern und Abendstern; wir mögen mit den Ausdrücken inhaltlich etwas anderes meinen, beziehen uns aber *de facto* auf denselben Himmelskörper, den Planeten Venus.

²⁰⁶ Die ursprüngliche terminologische Unterscheidung von Sinn und Bedeutung ist irrtumsanfällig, da beide Ausdrücke im heutigen Sprachgebrauch weitestgehend synonym verwendet werden.

Dass es uns dennoch so schwer falle, auch nur die Möglichkeit zu akzeptieren, dass jedes psychische Prädikat durch eine physikalische Eigenschaft ausgedrückt wird, liege nach Place vor allem an einem *phänomenologischen Fehlschluss* (vgl. S. 48). Diesem Fehlschluss sitzen wir dann auf, wenn wir phänomenale Inhalte (wie Geruch, Farbe, Geräusche) unseres Bewusstseins verdinglichen, indem wir von ihnen – wörtlich genommen – wie von Eigenschaften von Dingen oder Ereignissen sprechen. Durch diesen Vorgang kommt es nach Place erst zur (ungerechtfertigten) Annahme nicht-physischer Gegebenheiten, dem *intuitiven Dualismus*.

Im Gegensatz zu früheren Theorien, die eine solche Identifikation zum Gegenstand hatten, soll die von ihm behauptete Identität von Bewusstseins- und Gehirnzuständen keine sein, die von analytischer bzw. rein logischer Qualität wäre. Wie Place betont, kommt der psychophysischen Identität keine logisch-begriffliche Notwendigkeit zu; es handele sich vielmehr um eine lediglich *kontingente* Identität. Die Gültigkeit der Identitätsbehauptung ergibt sich deshalb nicht aus der Bedeutung der eingesetzten Prädikate²⁰⁷, sondern muss an der Erfahrung überprüft werden. So wie auch die Behauptung, dass sich die Ausdrücke „Morgenstern“ und „Abendstern“ auf denselben Himmelskörper beziehen, den Status einer empirischen Hypothese hat, die sich als unzutreffend hätte herausstellen können, so hat auch die Identitätsthese den Status einer empirischen Hypothese, die *a posteriori* durch wissenschaftliche Forschung widerlegt oder bestätigt werden kann.

Wenn es bei der Identität von mentalen und neuronalen Eigenschaften um eine empirisch zu ermittelnde bzw. zu begründende Annahme han-

²⁰⁷ An dem Versuch, mentale Ausdrücke in physikalischer Sprache zu definieren und auf diesem Wege das Leib-Seele-Problem auszuräumen, ist zuvor der vor allem von Carnap (1932) propagierte so genannte *Logische Behaviorismus* bzw. *Semantische Physikalismus* gescheitert.

delt, stellt sich die Frage, wann aufgrund von Beobachtungen zweier Größen auf deren Identität geschlossen werden sollte und wann es sich um zwei eigenständige Größen handelt, zwischen denen eine kausale Beziehung besteht. Nach Place sei es ein sinnvolles Kriterium, dass wir zwei Beobachtungsreihen in denjenigen Fällen als Beobachtungen eines Ereignisses ansehen können, in denen die wissenschaftlichen Beobachtungen oder Modelle im Kontext eines entsprechenden wissenschaftlich-theoretischen Gebäudes eine unmittelbare Erklärung der Alltagsbeobachtungen dieses Ereignisses liefern (vgl. ebd., S. 48). Das Kriterium sei im Fall von Blitzen und der Bewegung elektrischer Ladungen erfüllt. Die wissenschaftliche Beobachtung bzw. das wissenschaftliche Modell der Bewegung elektrischer Entladungen kann eine unmittelbare Erklärung unserer Alltagsbeobachtungen von Blitzen gewährleisten. Die Annahme „Blitze sind Bewegungen elektrischer Ladungen“ ist daher gerechtfertigt.

Natürlich lässt sich kritisieren, dass sich durchaus Fälle denken lassen, in denen dieses Kriterium erfüllt wird, aber nichtsdestotrotz für gewöhnlich kein Identitätsverhältnis angenommen wird. Eine rein empirische oder gar experimentelle Entscheidung zwischen einer Dualität und Identität zweier Größen scheint schwerlich möglich zu sein. So weist auch Smart (1959) darauf hin, dass es zwar eine empirische Frage sei, ob eine Identifikation des Geistes mit dem Gehirn oder doch eher mit dem Herzen, der Leber oder der Niere vorgenommen sollte, eine Entscheidung zwischen z.B. Epiphänomenalismus und Materialismus könne aber nicht auf Grundlage eines Experimentes erfolgen, da kein solches denkbar ist (vgl. ebd., S. 155).

Bei Place zeigt sich – wie auch bei Smart und Feigl – die Tendenz, ausschließlich phänomenale Zustände in den Blick zu nehmen. Dies mag daran liegen, dass für kognitive Zustände noch die Möglichkeit der Durchführung des damals vorherrschenden behavioristischen Pro-

gramms angenommen wird. So heißt es bei Place (1956): „In the case of cognitive concepts like ‘knowing’, ‘believing’, ‘understanding’, ‘remembering’ and volitional concepts like ‘wanting’ and ‘intending’, there can be little doubt, I think, that an analysis in terms of dispositions to behave [...] is fundamentally sound” (S. 44).

6.1.2. Herbert Feigl: Zwischen neutralem und idealistischem Monismus

Der österreichisch-amerikanische Philosoph Herbert Feigl (1958/1967, 1960/1970) hat in etwa zur gleichen Zeit wie Place eine eigene Variante der Identitätstheorie entwickelt. Wie Place beruft sich Feigl bei seiner Formulierung einer psychophysischen Identitätstheorie auf die von Frege (1892) getroffene Unterscheidung zwischen Intension (Bedeutung, Konnotation) und der Extension (Referenz, Denotation): „Certain neurophysiological terms denote (refer to) the very same events that are also denoted (referred to) by certain phenomenal terms. [...] [N]europhysiological terms and the corresponding phenomenal terms, though widely differing in *sense*, and hence in the modes of confirmation of statements containing them, do have identical *referents*” (Feigl, 1960/1970, S. 38; Hervorh. im Orig.).

Gemeinsam ist beiden Formen weiterhin die Ansicht, dass sich mentale Zustände gänzlich auf komplexe Eigenschaften von neuronal-biochemischen Systemen zurückführen lassen und dass es bei der Formulierung der Identitätsbeziehung um eine kontingente Aussage handeln soll, deren Gültigkeit in Bezug auf die Empirie (zumindest teilweise) festgestellt werden kann und muss. Während allerdings für Place (und auch Smart) das eigentliche Problem an der Leib-Seele-Beziehung das Mentale ist, das sie durch die Annahme der psychophysischen Identität in das physikalische Weltbild integrieren wollen, ist für Feigl – genau entgegengesetzt

– das eigentliche Problem auf der physischen Seite zu suchen²⁰⁸. In Bezug auf den Begriff des Physischen bzw. Physikalischen gäbe es nämlich ein weit verbreitetes Missverständnis, das auf die Verwechslung von Evidenz und Referenz zurückgeht. Wenn wir uns mit neuronalen Zuständen auseinandersetzen, dann verwechseln wir die bildhaften Vorstellungen, die wir uns hierzu machen, schnell mit den tatsächlichen neuronalen Zuständen (die wir uns versucht haben vorzustellen). So könnte man z.B. aufgrund seiner bildhaften Vorstellungen zu der Aussage kommen, dass das Gehirn eine graue, glitschige Masse ist (vgl. McGinn, 1989), obwohl dem realen Gehirn keines dieser Prädikate zukommen dürfte. Feigl betont deshalb, dass man sich darüber im Klaren sein müsse, dass physikalische (oder auch neurophysiologische) Begriffe abstrakt und deshalb unanschaulich sind. Für den Bereich der einfachen Sinnesempfindungen (also phänomenale Zustände) fallen nun allerdings Evidenz und Referenz zusammen. Feigl bezeichnet sie als „die Wirklichkeiten selbst“, auf die sich die neurophysiologischen Beschreibungssysteme erst in einem zweiten Schritt beziehen. Es gibt hier also „genau eine grundlegende Wirklichkeit, die den Gegenstandsbereich sowohl introspektiver Aussagen als auch bestimmter physikalischer oder neurowissenschaftlicher Theorien bildet“ (Metzinger, 2007b, S. 130). Daraus schlussfolgert Feigl, dass es denkbar und sogar plausibel ist, eine bestimmte Teilmenge der Bezugsgegenstände physikalischer Begriffe mit dem zu identifizieren, was uns in unserem subjektiven Erleben direkt gegeben ist.

Ähnlich wie Fechner, durch den er zumindest mittelbar beeinflusst war (vgl. Heidelberger, 2002), geht also auch Feigl (1960/1970) davon aus,

²⁰⁸ Feigl übernimmt hier im Wesentlichen einen Argumentationsansatz, der sich bereits in der *Erkenntnislehre* seines Mitstreiters aus dem Wiener Kreis, Moritz Schlick (1918/1979), findet.

dass es im Falle der Entitäten, auf die sich Psychologie und Hirnforschung bzw. Biologie beziehen, nur eine zugrunde liegende Wirklichkeit gibt, deren Verschiedenheit dadurch zustande kommt, dass wir diese unterschiedlich repräsentieren: „Instead of conceiving of two realms or two concomitant types of events, we have only *one reality which is represented in two different conceptual systems* – on the one hand, that of physics and on the other hand, where applicable (in my opinion only to a extremely small part of the world) that of phenomenological psychology“ (S. 41; Hervorh.: M.L.). In der zuletzt genannten Einschränkung auf einen wahrscheinlich nur sehr kleinen Realitätsausschnitt, in dem die psychophysische Identität gilt, möchte er seine Position allerdings klar von einem Panpsychismus abgrenzen, der von Fechner (1879) in seinen späteren Arbeiten vertreten wurde.

Feigl bekennt sich zum Physikalismus und seiner Grundthese, dass die grundlegenden Gesetze des Universums physikalischer Natur sind. Aber im Gegensatz zur „australischen“ Version der Identitätstheorie nach Place (und Smart) möchte sich Feigl einer Aussage darüber enthalten, welchen ontologischen Status die der Identitätsannahme zugrunde liegende Wirklichkeit hat. So sagt er, dass die Anerkennung des Physikalismus „does not commit me in the least as to the nature of the *reality* whose regularities are formulated in the physical laws. This reality is known to us by acquaintance only in the case of our direct experience which [...] is the referent also of certain neurophysiological concepts“ (ebd., 1960/1970, S. 40; Hervorh. im Orig.). Auch wenn Feigls Ansatz aufgrund dieser Einschätzung häufig als *neutraler Monismus* (vgl. z.B. Metzinger, 2007b) bezeichnet wird, da er sich programmatisch über den „wahren Bezugsgegenstand“ der von ihm getroffenen Aussagen enthalten möchte, findet sich doch an einigen Stellen eine nähere Bestimmung des bezugsgleichen Referenten der neuronalen und phänomenalen Ausdrücke. Diesen

bezeichnet Feigl als unmittelbare Erlebnisqualitäten (was er auch *raw feels* nennt): „I take these referents to be the immediately experienced qualities, or their configurations in the phenomenal fields” (Feigl, 1960/1970, S. 38). Insofern ist es sicherlich nicht unberechtigt, wenn gesagt wird, dass es sich im Gegensatz zur materialistischen Version nach Place (und noch viel verstärkter nach Smart) im Falle Feigls um eine *idealistische* oder *mentalistische* Version der Identitätstheorie²⁰⁹ handelt (vgl. z.B. Borst, 1970, S. 19).

Soweit die Konzeptionen einer Identitätstheorie nach Place und Feigl. Hätten sie sich durchgesetzt, wäre die Diskussion in den darauffolgenden Jahrzehnten sicher ganz anders und im Hinblick auf den Stellenwert und Anklang einer psychophysischen Identitätsannahme viel harmloser verlaufen. Aber es war ein anderer Text, der sich zum eigentlichen Bezugs-

²⁰⁹ Der Hintergrund dieser Interpretation der Identitätsrelation dürfte in seiner Beeinflussung durch den Phänomenalismus und Positivismus Ernst Machs zu suchen sein. Mach (1886/1903) war davon überzeugt, dass alle Wissenschaften, wie auch die Physik, die Tatsache anerkennen müssen, dass sie keinen direkten Zugriff auf eine unabhängig vom menschlichen Bewusstsein existierende Welt haben, sondern dass alle Erkenntnis auf unseren spezifischen Erfahrungsqualitäten basiert bzw. durch diese vermittelt wird. Entgegen einer Vielzahl anderer Physiker der damaligen Zeit trat er also dafür ein, dass Empfindungen nicht aus dem wissenschaftlichen Weltbild zu eliminieren sind, sondern gerade als Ausgangspunkt aller Wissenschaften anerkannt werden müssen. Die sich uns anbietenden Empfindungen sind der eigentliche Forschungsgegenstand; eine Setzung von etwas, das über diese hinausgeht, ist metaphysischer und damit unwissenschaftlicher Natur (vgl. Heidelberger, 2000; Lüdmann, unveröffentlichtes Manuskript). Feigls Bestimmung des gemeinsamen Referenten als „raw feels“, als das in der Erfahrung Gegebene, dürfte also vor allem dem Umstand geschuldet sein, dass er nicht in einen unlauteren Metaphysikverdacht geraten möchte (was im Wiener Kreis einem Hochverrat gleichkäme), als dass er wirklich eine idealistische Position im klassischen Sinne vertreten möchte. Jedenfalls wird hieran auch noch mal verdeutlicht, warum Feigl das Problematische am Leib-Seele-Problem eher auf der physikalischen als auf der mentalen Seite sieht. Das Mentale ist unmittelbar gegeben und der Ausgangspunkt all unserer Überlegungen. Das Physikalische muss erst – vermittelt über die sinnliche Erfahrung – erschlossen werden, wobei Irrtümer und Fehlbestimmungen auftreten können.

punkt der Debatte entwickeln sollte und lange Zeit einer der meistzitierten philosophischen Texte überhaupt war.

6.1.3. John J. C. Smart: Strikte psychophysische Identität

Es handelt sich hierbei um den Aufsatz *Sensations and brain processes*, in dem Smart (1959) die von Place formulierte These gegen mögliche Einwände zu verteidigen und diese systematisch auszuräumen versucht. Da die einzelnen Einwände und Smarts Er widerungen hierauf für die vorliegende Arbeit unerheblich sind (und in der Gegenwartsdebatte auch keine Rolle mehr spielen), werde ich auf ihre Darstellung verzichten²¹⁰ und nur auf die resultierenden Bestimmungen der Identitätstheorie eingehen.

Smart nimmt eine Verstärkung der Identitätsthese Place' vor, indem er betont, dass Sinnesempfindungen und Gehirnvorgänge in einem *strikten Sinne* identisch sind, was für ihn klar von einer bloßen räumlich-zeitlichen Koextensivität, die für Place' Konzeption noch im interpretativen Spielraum liegt – verschieden ist: „When I say that a sensation is a brain process or that lightning is an electric discharge I do not mean just that the sensation is somehow spatially or temporally continuous with the brain process or that the lightning is just spatially or temporally continuous with the discharge” (ebd., S. 145). Die sensorische Empfindung ist schlicht und einfach *nichts anderes als* ein Gehirnprozess, wie Blitze *nichts anderes als* elektrische Entladungen sind. Dieser explizite Ausschluss einer Etablierung eines Identitätsverhältnisses auf der Basis eines Koextensivitätsverhältnisses durch die Formulierung einer strikten Identität

²¹⁰ Eine dezidierte Darstellung der damaligen Debatte findet sich in Borst (1970).

markiert einen gewissen Wendepunkt in der Debatte und stellt den entscheidenden Unterschied zu dem in dieser Arbeit verfolgten Ansatz dar. In der Philosophie wurde von dieser entscheidenden Wendung allerdings kaum Notiz genommen bzw. es hat den Anschein, als wenn diese „Klarstellung“ Smarts etwas benannt hätte, was einer jeden Identitätstheorie ohnehin zu Eigen sein müsse (sodass man sich hiermit auch gar nicht weiter beschäftigen muss). Vielmehr wurde Smarts Aufsatz für seine begriffliche Schärfe gelobt. Im Gegensatz zu Place, der seinen Aufsatz (für einen Philosophen „unziemlich“) stark an empirischen bzw. forschungsbezogenen Gesichtspunkten ausgerichtet hat, wird Smart allem Anschein nach dafür gelobt, dass er die Besprechung der Leib-Seele-Frage wieder zurück in die „sprachlogische“ Argumentationsweise der Philosophie des Geistes gebracht hat.

So impliziere die Identitätsbehauptung „Sinnesempfindungen sind Gehirnprozesse“ trotz der Annahme einer strikten Identität nicht, dass *Aussagen* über Sinnesempfindungen in *Aussagen* über Gehirnprozesse *übersetzt* werden können. Außerdem werde auch nicht behauptet, dass beide Aussagetypen dieselben logischen Bestimmungen aufweisen, sondern nur „in so far as a sensation statement is a report of something, that something is in fact a brain process. Sensations are *nothing over and above* brain processes“ (ebd.; Hervorh.: M.L.).

Die Identifikation von mentalen Eigenschaften mit Eigenschaften unseres Gehirns ist also so aufzufassen, dass das, worauf wir uns beziehen, wenn wir von Schmerzen sprechen, *in Wirklichkeit* C-Faser-Reizungen sind. Der physischen Seite der Identitätsbehauptung wird gewissermaßen eine ontologische Vorrangstellung zugesprochen. Insofern wäre anstatt „nothing over and above“ (was wortwörtlich genommen nur eine Koextensivität impliziert) die für Smarts Ansicht treffendere Aussage ein „nothing but“.

Der Hintergrund dieser Betonung der physischen Seite der Identitätsgleichung dürfte nicht zuletzt darin begründet sein, dass man der Identitätstheorie vorwarf, mit dem symmetrischen Charakter von Identitätsgleichungen der vermeintlich asymmetrisch verfassten Gehirn-Geist-Relation nicht gerecht zu werden. Wenn eine Identität eine symmetrische Relation ist, folgt nämlich „dann nicht, daß, wenn mentale Phänomene mit bestimmten physischen Phänomenen identisch sind, einige physische Phänomene mental sind?“ (Bieri, 1981/1997, S. 42). Dies eingestehen zu müssen war (und ist) materialistisch ausgerichteten Philosophen ein Dorn im Auge und war einer der Gründe für das Aufkommen des bereits besprochenen *Eliminativen Materialismus* (vgl. Kap. 3.1.3). Denn dieser hat es sich – durch seine Leugnung mentaler Eigenschaften – zur Aufgabe gemacht, dafür Gewähr zu tragen, dass man zwar daran festhalten kann, dass mentale Prozesse *nichts anderes als* gewisse neuronale Prozesse sind, ohne damit sagen zu müssen, dass gewisse neuronale Prozesse mentale Prozesse sind. Auch wenn eine partielle Nähe Smarts zum Eliminativen Materialismus nicht zu leugnen ist²¹¹ (siehe z.B. Smart, 1970), geht er aber nicht so weit, mentalen Eigenschaften die Existenz abzuspochen, noch sie als rein theoretische Elemente einer obsoleten Alltagspsychologie zu betrachten. Wenn wir Schmerzen empfinden, einen rot schimmernden Sonnenuntergang betrachten oder ein Stück von Mozart hören, sitzen wir in den Augen der von Smart propagierten Identitätstheorie keiner Illusion auf. Was die Identitätsthese vermitteln soll, ist vielmehr eine Auskunft darüber, was Schmerzen oder andere mentale Zustände tatsächlich sind, was ihr eigentliches *Wesen*

²¹¹ In der Anfangszeit der neu aufgekommenen Debatte um das Leib-Seele-Problem innerhalb der Philosophie des Geistes gab es auch noch keine klare Unterscheidung zwischen Identitätstheorie und Eliminativem Materialismus.

ausmacht, nämlich Eigenschaften unseres Gehirns zu sein. Wenn dies aber ihr Wesen sein soll, warum erfahren wir dann diese Vorgänge in unserem Gehirn *als* Schmerzen, *als* schimmerndes Rot oder *als* eine anmutige Melodie? Hier bleibt Smart eine Antwort schuldig, obwohl dies für eine Akzeptanz der Identitätstheorie von zentraler Bedeutung wäre. Sowohl Smart als auch Place und Feigl haben – ob wirklich vollkommen intendiert oder eher einer sprachlichen Laxheit geschuldet, bleibt unklar – ihre These so formuliert, dass sie eine Identifizierung von mentalen und neuronalen *Typen* vornimmt: Jeder mentale Typ ist demnach identisch mit einem neuronalen Typ. Wenn uns also die Wissenschaft eine strenge Korrelation von z.B. C-Faser-Reizungen in unserem Nervensystem und dem Auftreten von Schmerzzuständen aufzeigen würde, dann kann dieser Umstand nach Ansicht der Identitätstheoretiker darauf zurückgeführt werden, dass Schmerzen (nichts anderes als) C-Faser-Reizungen *sind*. So ist der mentale Typ ‚Schmerzen‘ identisch mit dem physischen bzw. neuronalen Typ der ‚C-Faser-Reizungen‘²¹². Statt einem expliziten Rekurs auf Typen hat sich in der Philosophie des Geistes allerdings die Besprechung einer Identität von *Eigenschaften* eingebürgert. Es ist dementsprechend die Identität des „Habens einer Schmerzempfindung“ und des „Habens von C-Faser-Reizungen“, die als Diskussionsgrundlage psychophysischer Identitätsaussagen (auf Typen-Niveau) herangezogen wird.

²¹² Dies ist neurophysiologisch betrachtet natürlich eine extreme und sehr willkürliche Vereinfachung. C-Fasern sind eine bestimmte Art von Nervenfasern, die Schmerzinformationen an das Gehirn weiterleiten. An der Entstehung eines phänomenalen Schmerzerlebnisses sind eine Vielzahl anderer Strukturen und Prozesse beteiligt. Da sich dies aber in der Debatte um die Identitätstheorie als zentrales Beispiel behauptet hat, möchte ich es auch hier aufgreifen. Für eine *prinzipielle* Untersuchung der Möglichkeit solch einer (Typen-) Identität ist die empirische Gültigkeit des zur Diskussion verwendeten Beispiels aber zu vernachlässigen.

6.1.4. Stellenwert und Grenzen

Neben einer großen Anzahl an Vorzügen, die die Identitätstheorie aufweist (diese gelten genau genommen für alle Lösungsvorschläge zum Leib-Seele-Problem, die sich psychophysischer Identitätsannahmen bedienen), sind bereits kurz nach ihrer Formulierung durch Place, Smart und Feigl bestimmte Kritikpunkte erhoben worden, die die Identitätstheorie als unhaltbar erscheinen ließen. Des Weiteren soll im Folgenden auch verdeutlicht werden, welche problematische und inhaltlich keineswegs notwendige Interpretation das Konzept einer psychophysischen Identität in der maßgeblichen Auslegung durch Smart erhalten hat.

6.1.4.1. Stärken und Vorzüge der Identitätstheorie

Eine große Stärke identitätstheoretischer Annahmen ist, dass sie auf plausible Weise erklären können, warum es die vielfältigen psychophysischen Korrelationen gibt. Nämlich einfach deshalb, weil es sich um *ein und dieselben* Vorgänge bzw. Eigenschaften handelt. Durch die hiermit einhergehende Vereinheitlichung der Welt, die dem Psychischen keine „metaphysische“ Sonderrolle zuordnet, weist sie eine große Kohärenz zu naturwissenschaftlichen Theorien bzw. eine große (meta-)theoretische Konsistenz auf (vgl. Metzinger, 2007a). Im Gegensatz zu einem interaktionistischen Dualismus widerspricht die Identitätstheorie in keiner Weise der kausalen Geschlossenheitsannahme der physischen Welt und auch nicht dem Energieerhaltungssatz. Es gibt auf der einen Seite keine durch das Einwirken eines immateriellen Geistes geschuldeten Unterbrechungen des Ursache-Wirkungs-Vollzugs der physikalischen bzw. physischen Phänomene. Auf der anderen Seite können psychische Erscheinungen

durchaus durch psychische Vorläuferprozesse erklärt werden. Außerdem besticht diese Konzeption im Gegensatz zu voraussetzungsreicheren dualistischen Vorschlägen durch ihre begriffliche bzw. ontologische Sparsamkeit und kann damit in Hinblick auf „Ockham’s Razor“ Boden für sich gut machen.

Was hiermit einhergeht, ist eine schlüssige Erklärung der Möglichkeit einer mentalen Verursachung, die im Rahmen dualistischer Konzeptionen völlig nebulös bleibt (vgl. Kap. 3.1). Um das vormals verwendete Beispiel wieder aufzugreifen, lässt es sich unter den Voraussetzungen der Identitätstheorie völlig unproblematisch annehmen, dass es mein Wunsch nach einer Cola war, der mich zum Kühlschranks gehen ließ, da mein Wunsch mit einem physischen Prozess in meinem Gehirn identisch ist, der natürlich andere physische Vorgänge anstoßen kann. Eine kausale Überdetermination ist hierbei nicht zu fürchten.

Metzinger (2007a) erwähnt auch, dass die Offenheit der Identitätstheorie für eine Anbindung an Ebenen- oder Schichtenmodelle der Wirklichkeit eine ihrer großen Stärken sei. Auch wenn psychophysische Identitätskonzeptionen in der Tat hierfür *prinzipiell* offen sind, kann doch im Hinblick auf die sich durchsetzende Interpretation der Identitätstheorie im Sinne einer „strikten“ Identität angezweifelt werden, dass diese *faktisch* für Ebenenmodelle offen ist, denn Smart widerspricht explizit der Auffassung, dass eine Identität auf einer (für kompositorische Ebenenmodelle charakteristischen) Koextensivität aufgebaut werden sollte.

Auch wenn das Aufkommen der Identitätstheorie in mehreren Hinsichten eine Art Wendepunkt in der Philosophie des Geistes markiert, wurden bereits kurz nach ihrem Aufkommen einige entscheidende Kritikpunkte formuliert, die dazu führten, dass die These einer psychophysischen Identität schnell für inakzeptabel angesehen wurde und für längere

Zeit wieder aus dem Aufmerksamkeitsfokus²¹³ rückte. Insbesondere – und dies ist für die vorliegende Arbeit von zentraler Bedeutung – hat die Interpretation einer psychophysischen Identität im Zuge der Diskussion eine m.E. problematische Entwicklung erfahren, die den Begriff einer solchen Identität in ein sehr einseitiges und negatives Licht gerückt hat. Beide Aspekte werden im Folgenden systematisch herausgearbeitet.

6.1.4.2. Multiple Realisierbarkeit

Der angesprochene entscheidende Fallstrick der Identitätstheorie war die so genannte Möglichkeit *multipler Realisierungen* und ist vor allem mit Putnam (1967), Block und Fodor (1972) bzw. Fodor (1974) verbunden. Putnam (1967) räumt zunächst ein, dass es nach seiner Einschätzung keine *apriorischen* Gründe für ein Scheitern der Identitätstheorie gibt. Die Annahme einer psychophysischen Identität sei sprachlich und begrifflich durchaus konsistent. Gleichzeitig sei das Zutreffen der „Gehirnzustandstheorie“ aber extrem unwahrscheinlich, da diese eine unangemessene Verallgemeinerung der Beziehung zwischen mentalen und neuronalen Typen beinhalte. Wenn die Identitätstheorie nämlich behauptete, dass mentale Eigenschaften neuronale Eigenschaften sind und damit – wie zumeist als Beispiel verwendet – die mentale Eigenschaft (des Habens von) Schmerzen mit der neuronalen Eigenschaft (des Habens von) C-Faser-Reizungen identisch ist, dann müsste jeder nur denkbare Organismus, der über Schmerzen verfügt, die gleichen neuronalen Basisprozesse, sprich: C-Faser-Reizungen, aufweisen. Jedes über Schmerzzustände verfügende Wesen (inklusive allen potentiell vorhandenen außerirdi-

²¹³ Gegenwärtig gibt es wieder eine wachsende Anzahl von Verfechtern einer Typen-Identität (vgl. z.B. Pauen & Stephan, 2002; Gozzano & Hill, 2012). Meist ist hierbei allerdings Dreh- und Angelpunkt die Frage der *Reduzierbarkeit* des Psychischen (vgl. Kap. 8).

schen Lebens) müsste demnach ein Gehirn von angemessener physikalisch-chemischer Struktur besitzen und sich in einem ganz bestimmten physikalisch-chemischen Zustand befinden. „Thus it is at least possible that parallel evolution, all over the universe, might *always* lead to *one and the same* physical ‘correlate’ of pain” (Putnam, 1967, S. 44; Hervorh. im Orig.). Dass Schmerzen in jedem nur denkbaren Fall durch C-Faser-Reizungen realisiert werden, scheint jedoch extrem unplausibel zu sein. Wahrscheinlicher ist, dass es eine Vielzahl an (physischen) *Realisierungen* von Schmerzen gibt. Im Zuge dieser Überlegungen spricht sich Putnam für eine Alternativhypothese aus, die in den kommenden Jahrzehnten die Philosophie und Psychologie geprägt hat und in weiten Teilen bis heute prägt, nämlich dass mentale Zustände *funktionale* Zustände sind, die durch verschiedenste physische Zustände realisiert werden können. Dem konkreten physischen Substrat eines mentalen Zustandes kommt dabei eine relative Beliebigkeit zu. Der Funktionalismus ist zugleich die philosophische Hintergrundtheorie, auf der die besprochenen computationalen Modelle kognitiver Zustände fußen (vgl. Kap. 4.2.1.1).

Block und Fodor (1972) stellen drei relevante Möglichkeiten multipler Realisierungen heraus. Die *erste* plausible Möglichkeit hängt mit der Plastizität des Gehirns zusammen. Aus der neuropsychologischen Forschung ist bekannt, dass z.B. im Falle von Läsionen des Gehirns andere Bereiche die ausgefallenen Funktionen (in einem bestimmten Ausmaß) übernehmen können (vgl. z.B. Knecht & Ringelstein, 1999). Ebenso gibt es auch zwischen gesunden Menschen individuelle Unterschiede in der genauen Lokation der neuronalen Korrelate ihrer mentalen Zustände (vgl. z.B. Braun & Bogerts, 2001). *Zweitens* sei es mit Bezug auf die Evolutionstheorie plausibel anzunehmen, dass auch Organismen mit einer stark divergierenden stammesgeschichtlichen Entwicklung und einem unterschiedlichen Körperbau ähnliche Verhaltensweisen und Erlebniszustände auf-

weisen werden. „Psychological similarities across species may often reflect convergent environmental selection rather than underlying physiological similarities. For example, we have no particular reason to suppose that the physiology of pain in man must have much in common with the physiology of pain in phylogenetically remote species” (Block & Fodor, 1972, S. 161). *Drittens* sei es auch vorstellbar, dass wir im Rahmen der Forschung zur künstlichen Intelligenz irgendwann soweit sein werden, dass wir Maschinen bauen können, die eine ähnliche Bewusstseinswelt wie wir selbst haben werden, obwohl sie phys(iolog)isch völlig verschieden von uns sind. Obgleich dieser Punkt sicher streitbar ist, könnte man ihn mit dem Blick auf die Errungenschaften der neurowissenschaftlichen Forschung der letzten Jahrzehnte insoweit modifizieren, dass man nach dem Stellenwert von so genannten „Neuroprothesen“, also z.B. bestimmten Gehirn-Implantaten, die den Aktivitätsgrad bestimmter Hirnregionen erhöhen oder eindämmen, für die Erzeugung und Qualität bestimmter psychischer Zustände fragt. Wenn man also in Betracht zieht, dass es – wie es häufig heißt – bald Siliziumchips geben wird, die die Funktion von Nervenzellen übernehmen können, scheint es auch hier einen plausiblen Fall einer multiplen Realisierbarkeit mentaler Zustände zu geben, der einer psychophysischen Identitätstheorie (auf Typen-Niveau) offensichtlich zuwider läuft.

Eine brisante, wirkungsgeschichtlich äußerst bedeutsame Wendung erfuhr das Argument der multiplen Realisierbarkeit durch den Aufsatz von Fodor *Special Sciences* aus dem Jahre 1974. Fodor verknüpft die Möglichkeit zu multiplen Realisierungen mentaler (oder anderer höherstufiger) Eigenschaften nämlich mit der Existenzberechtigung und Nicht-Reduzierbarkeit der höherstufigen Spezialwissenschaften wie der Psychologie. Fodor (1974) sagt zwar ganz ähnlich wie Putnam, dass die von ihm vorgebrachte Kritik den (für ihn mit der Identitätstheorie untrennbar verbun-

denen) Reduktionismus nicht mit Notwendigkeit widerlege, dass dieser aber bei genauerer Analyse extrem unplausibel sei, was er am Beispiel eines höherstufigen Prozesses aus dem ökonomischen Bereich verdeutlicht: „A natural kind like a monetary exchange *could* turn out to be co-extensive with a physical natural kind, but if it did, that would be an accident on a cosmic scale (S. 104; Hervorh. im Orig.). Sind höherstufige Eigenschaften wie mentale Eigenschaften durch verschiedene physikalische Eigenschaften realisierbar, ist nach Fodor keine Reduktion der Psychologie auf die Biologie oder Physik möglich. Die multiple Realisierbarkeit höherstufiger Eigenschaften wird somit zu einem Garant der Existenzberechtigung der „Spezialwissenschaften“. Was im Umkehrschluss aber auch heißt, dass man, wenn man die Gültigkeit der Identitätstheorie, die offenbar keine multiplen Realisierungen zulässt, akzeptieren würde, gleichzeitig unweigerlich einer Reduzierbarkeit der höherstufigen Wissenschaften zustimmen und deren Eigenständigkeit infrage stellen würde: „If psychology is reducible to neurology, then for every psychological natural kind predicate there is a co-extensive neurological natural kind predicate, and the generalization which states this co-extension is a law²¹⁴“ (S. 104). Die verbreitete Annahme, dass eine psychophysische Identität zwangsläufig mit einer Gefahr für den Status oder gar die Existenz der Psychologie einhergeht, dürfte spätestens hiermit festzementiert worden sein. Wie es Shapiro & Polger (2012) etwas überspitzt auf den Punkt bringen:

²¹⁴ Fodor bezieht sich hierbei auf das Reduktionsmodell Nagels (1961), das für die hierfür notwendige Verbindung der Begriffe der zu reduzierenden höherstufigen Theorie mit der niederstufigen Theorie, auf die reduziert werden soll, Brückenprinzipien vorsieht, die zumeist als (Typen-)Identitäten aufgefasst werden. Das Thema Reduktion(ismus) wird an späterer Stelle ausführlich behandelt werden (vgl. Kap. 8).

The concern, as philosophers like Hilary Putnam and Jerry Fodor have expressed it, is that a victory for identity or reduction is a defeat for psychology. For if minds are physical, or if mental events are physical events, then psychologists might as well disassemble their laboratories, making room for the neuroscientists and molecular biologists who are in a better position to explain those phenomena once misdescribed as 'psychological'. (S. 264)

6.1.4.3. Starre Designatoren und der Anschein von Kontingenz

Neben dem Vorwurf, dass die Identitätstheorie multiplen Realisierungen mentaler Eigenschaften nicht gerecht werde, drehte sich ein zweiter zentraler Kritikpunkt um den angenommenen *kontingenten* Charakter der Identität. Die Identitätsthese sollte nach Ansicht der Identitätstheoretiker als eine „quasi-empirische“ Hypothese verstanden werden, die – obgleich es kein geeignetes Argument zu ihrer Überprüfung geben mag – zumindest prinzipiell falsch sein könnte. Da die herausgestellte Identität von z.B. Schmerzen und C-Faser-Reizungen also nicht *notwendigerweise* wahr ist, könnte sich im Fortgang der empirischen Forschung zeigen, dass diese Annahme falsch ist.

Die Identitätstheoretiker gehen somit von einer Konkordanz von Apriorität und Notwendigkeit aus. Jede Aussage, deren Wahrheitsgehalt nicht schon im Begriff selbst enthalten ist, sondern an der Erfahrung überprüft werden muss, hat einen nur kontingenten Charakter, da sie sich als falsch herausstellen könnte. Saul Kripke (1971, 1972) hat mit seinem – sprachlogische und modale Aspekte in den Blick nehmenden – Argument der so genannten *starrten Designatoren* jedoch darlegt, dass es bei den in einer Identitätsbehauptung verwendeten Ausdrücken um solche handelt, die sich in allen möglichen Welten *notwendigerweise* auf immer dieselben Entitäten beziehen müssen. Sie bezeichnen in allen möglichen Welten starr dieselben Gegebenheiten. Entweder gilt damit die behauptete Identität *notwendigerweise* (in allen möglichen Welten) oder sie gilt gar nicht

(also in keiner möglichen Welt und damit auch nicht in unserer). Da man sich nun problemlos eine Welt vorstellen kann, in der es Schmerzen ohne C-Faser-Reizungen gibt, sei die psychophysische Identitätsthese notwendigerweise falsch. Soweit das Argument in aller Kürze. Um den Ansatz Kripkes aber wirklich verstehen zu können, muss man etwas weiter ausholen und sich den sprachlogischen Hintergrund für diese Schlussfolgerungen genauer anschauen.

In der Philosophie galt als lange Zeit als ausgemacht, dass es zwei Formen von Wahrheiten gibt: *notwendige* Wahrheiten *a priori* (auch Begriffswahrheiten genannt) und *kontingente* Wahrheiten *a posteriori* (auch empirische Wahrheiten genannt). So ist es *a priori* und notwendig war, dass Junggesellen unverheiratet sind. Es liegt bereits im Begriff des Junggesellen selbst, dass es sich hierbei um ledige Männer handelt, sodass es keinerlei empirischer Untersuchung bedarf, um die Wahrheit dieses Satzes festzustellen. Die Aussage „Die Wahlbeteiligung in Deutschland ist auf ein Rekordniveau gesunken“ muss dagegen erst an der *Erfahrung* überprüft werden, um festzustellen, ob sie zutrifft. Des Weiteren kann man sich problemlos vorstellen, dass sich die Wahlbeteiligung in Deutschland ganz anders ausnimmt wie in der getroffenen Aussage. Insofern ist sie offensichtlich (wenn sie überhaupt zutrifft) eine kontingente Wahrheit *a posteriori*. Wenn eine Aussage prinzipiell widerlegbar ist, weil sie nicht bereits begrifflich wahr ist, dann kann sie demnach nicht notwendig sein.

Dem widerspricht Kripke. Denn Aussagen wie „Wasser ist H₂O“ oder „Cicero ist Tullius“ seien zwar Aussagen, die an der Erfahrung geprüft werden müssen, denn sie sind keine „Begriffswahrheiten“. Dennoch sind sie – wenn sie wahr sind – *notwendigerweise* wahr. An dieser Stelle kommt nun Kripkes modallogische Argumentationsweise bezüglich „möglicher Welten“ ins Spiel. Eine mögliche Welt kann hierbei als eine vollständige

Art und Weise, wie sich die Dinge verhalten können, verstanden werden (vgl. Beckermann, 1999. S. 129). Sie ist also nicht als ein „Paralleluniversum“ im physikalischen Sinne zu verstehen, sondern mehr als etwas wie ein „Möglichkeitsraum“. Man bezieht sich z.B. auf eine solche „Mögliche-Welten-Argumentation“, wenn man sagt, dass es undenkbar ist, dass es eine Welt gibt, in der die Gesetze der Mathematik oder Logik nicht gelten (eine Welt, in der $1 + 1 = 2$ nicht gilt, ist also *keine* mögliche Welt). Dagegen ist eine Welt, in der unsere phänomenalen Eindrücke von Rot und Grün komplett vertauscht sind, durchaus vorstellbar bzw. widerspruchsfrei denkbar. Auch von einer Welt, in der andere Naturgesetze herrschen als in unserer (z.B. einer, in der es keine Gravitation gibt), wird für gewöhnlich gesagt, dass sie eine mögliche (widerspruchsfrei denkbare) Welt darstellt.

Für Eigennamen gelte nun, dass sie sich nach Ansicht Kripkes in jeder möglichen Welt auf ein und dieselbe Sache beziehen. So referiert z.B. der Ausdruck „Jesus“ in jeder möglichen Welt, die „Jesus“ enthält, auf „Jesus“ (den Sohn Gottes nach christlicher Denkweise). Eigennamen bezeichnen in jeder möglichen Welt „starr“ dieselbe Person oder Entität und sind daher *starre Designatoren*. Das gelte ebenso für alle Ausdrücke für natürliche Arten oder natürliche Phänomene wie z.B. Eisen, Wärme, Helium und auch für phänomenale Gegebenheiten wie Schmerzen. Auch sie sind Kripke zufolge *starre Designatoren*. So ist z.B. „Wasser“ ein *rigider Designator*, weil er unter allen, auch kontrafaktischen, Umständen dieselbe Substanz meint (H_2O). Dagegen ist der Begriff „das Lieblingsgetränk meiner Tante“ kein *rigider Designator*, da damit zwar in unserer Welt ebenfalls Wasser gemeint sein mag, jedoch problemlos anzunehmen ist, dass dies nicht der Fall wäre (sprich: dass es mögliche Welten gibt, in denen meine Tante ein anderes Lieblingsgetränk hat).

Hintergrund dieser Bestimmungen ist, dass sich Kripke von theoretischen Ansätzen abheben möchte, die meinen, dass die Referenz von Ausdrücken ausschließlich eine Angelegenheit sprachlicher Beschreibungen oder von reinen Konventionen über die Bedeutung von Ausdrücken ist. Wie Boyd (1980) herausstellt, hat man in der Philosophie²¹⁵ lange angenommen, dass die Zuweisung von Ausdrücken zu bestimmten Entitäten lediglich eine Sache der (sprachlichen) Konvention ist, die den semantischen Gehalt eines Ausdrucks festlegt, da ein Schluss auf die *tatsächliche* Beschaffenheit von Entitäten, also auf deren *essentielle* Eigenschaften, als unzulässige spekulative Metaphysik angesehen wurde: „Questions regarding the real essence of natural kinds are held to lie beyond the range of possible empirical investigation and thus to involve unscientific and pointless speculation“ (Boyd, 1980, S. 70). Auch wenn eine solche Fixierung auf bestimmte (sprachliche) Konventionen in einer Zeit des Wildwuchses spekulativer philosophischer Welterklärungen ihre Gründe gehabt haben mag, zieht diese Denkart absurde Konsequenzen nach sich. Denn unter diesen Voraussetzungen wäre es möglich, dass wenn sich alle Sprecher einer Sprache darin einig sind, dass Fledermäuse zur Art der Vögel gehören, *de facto* gilt: Fledermäuse *sind* Vögel. Keine empirische Entdeckung zur Beschaffenheit der Fledermäuse *selbst* könnte eine solche Zuordnung ausschließen, solange es nicht zugleich eine Veränderung der sprachlichen Konventionen gibt, die bestimmte Kriterien für die Zuordnung des Ausdrucks „Vogel“ beinhalten (vgl. ebd.). Und so wäre es auch *nur* eine (sprachliche) Konvention, wenn wir meinen, dass Wasser H₂O ist. Wenn wir unsere sprachlichen Konventionen oder Bedeutungs-

²¹⁵ Eine solche Sichtweise findet sich (in Bezug auf Eigennamen) z.B. bei Frege, Russell und Wittgenstein (vgl. Boyd, 1980).

zuweisungen ändern würden, dann könnte sich prinzipiell auch die Gültigkeit dieser Aussage verändern.

Das ist für Kripke nicht haltbar. Es ist für ihn offenkundig, dass der Stellenwert der Sprache oder von Konventionen bei der Frage nach der Referenz von Ausdrücken oder Feststellungen (wie Identitätsaussagen) stark überhöht wird. So ist es völlig unplausibel anzunehmen, dass es sich bei der Feststellung, dass Wasser H₂O ist, um lediglich eine Sache von Bedeutungszuschreibungen oder des allgemeinen Sprachgebrauchs handelt. Nach seinem Ansatz *haben* natürliche Phänomene ihre wesentlichen Eigenschaften unabhängig von jeder linguistischen Konvention oder der Auswahl einer bestimmten Beschreibung. Es ist also eine essentielle Eigenschaft von Wasser *selbst*, dass es aus H₂O besteht, sodass ihm diese Eigenschaft unabhängig von irgendwelchen Modalitäten der Beschreibung und damit *notwendigerweise* zukommt. Akzeptiert man dies, gilt aber auch der Grundsatz nicht mehr, dass jede empirische Behauptung, eine Behauptung, die prinzipiell also zurückweisbar wäre, einen lediglich kontingenten Sachverhalt bezeichnet. Obgleich also die Behauptung, dass Wasser H₂O ist, grundsätzlich widerlegbar ist, kann sie eine notwendige Wahrheit sein. Die Konzepte der Apriorität und Notwendigkeit werden auf diese Weise voneinander getrennt (vgl. ebd.).

Allgemein gelte nun für Identitätsaussagen, bei denen beide Seiten der Aussage starre Designatoren sind, dass sie – wenn sie wahr sind – *notwendig* wahr sind. Dabei ist im Falle der Identität von Wasser und H₂O, der Ausdruck H₂O nicht nur ein starrer Designator, sondern auch ein *kompositionsspezifizierender* Ausdruck (vgl. Boyd, 1980). Daher gilt, dass Wasser in jeder möglichen Welt immer aus Wasserstoff und Sauerstoff bestehen muss. Wenn man nun trotzdem den Eindruck hat, dass die Beziehung zwischen Wasser und H₂O kontingent ist, dann kann man die scheinbare

Kontingenz dadurch ausräumen, dass man sich vergegenwärtigt, dass wir hierbei von bestimmten kontingenten Annahmen zu Wasser/H₂O geleitet sind. Wir mögen z.B. – ganz im Sinne des „Sprach-Konventions-Ansatzes“ – gedacht haben, dass es eine bestimmte eindeutige *Beschreibung* von Wasser als deren Definition gibt, die in jeder möglichen Welt gilt (z.B. Wasser ist die durchsichtige, geruchs- und geschmacklose Flüssigkeit einer bestimmten Konsistenz, die bei 100 Grad kocht und bei 0 Grad gefriert...). So können wir uns Welten vorstellen, die eine auf diese Weise spezifizierte Flüssigkeit enthält, die damit *per Definition* Wasser ist und dieses „Wasser“ aber eine andere Mikrostruktur als H₂O aufweist. Die vermeintliche Kontingenz besteht also in der Diskrepanz zwischen Erscheinung und Realität: Wir mögen zwar gedacht haben (oder denken), dass Wasser etwas anderes als H₂O ist. Wenn wir dies tun, dann *erscheint* es aber nur so; wir verkennen dabei die (unabhängig von unseren Zuschreibungen) bestehenden *realen* Verhältnisse. In modallogischer Terminologie müssen dabei genau genommen die Intuitionen ausgeräumt werden, dass es erstens eine mögliche Welt gibt, die Wasser enthält, aber kein H₂O und zweitens, dass es eine mögliche Welt gibt, in der H₂O existiert, aber kein Wasser. Wenn es uns so scheint, dass Wasser und H₂O nicht identisch sind, liegt es im ersten Fall daran, dass wir uns hier eine Welt vorstellen, die eine Flüssigkeit aufweist, die nicht H₂O ist, aber alle qualitativen Eigenschaften besitzt, die Wasser in dieser Welt besitzt; und in dem zweiten Fall daran, dass wir uns eine Welt vorstellen, die zwar H₂O enthält, dieses aber nicht die qualitativen Eigenschaften besitzt, mittels derer wir Wasser in unserer tatsächlichen Welt erkennen. Obwohl die Möglichkeit solcher Welten tatsächlich besteht, bleibt die Aussage „Wasser ist H₂O“ in jeder dieser möglichen Welten wahr (vgl. Boyd, 1980).

Auch psychophysische Identitätsaussagen wie „Schmerzen sind C-Faser-Reizungen“ beinhalten eine Spezifikation der Komposition und haben den Anschein kontingent zu sein. Denn wir können uns vorstellen, dass es Schmerzen ohne C-Faser-Reizungen und C-Faser-Reizungen ohne Schmerzen gibt. Um den Anschein der Kontingenz auszuräumen, wozu der Identitätstheoretiker nach Kripkes Vorgaben genötigt ist, wenn er an seiner These festhalten möchte, muss er aufzeigen, warum die Annahme einer Welt abwegig ist, in der zwar Schmerzen vorkommen, aber keine Materie. Genauso muss er die Abwegigkeit der Annahme einer Welt aufzeigen, in der es C-Faser-Aktivierungen gibt, ohne dass jemand Schmerzen fühlt. Die auszuräumenden Intuitionen wären diesmal also, dass es erstens eine mögliche Welt gibt, in der es Schmerzen ohne C-Faser-Reizungen gibt, und zweitens eine mögliche Welt, in der es C-Faser-Reizungen ohne Schmerzen gibt (vgl. ebd.).

Versucht man dies nach dem gleichen Prozedere auszuräumen, müsste man sagen, dass die vermeintliche Kontingenz darauf zurückgeht, dass man sich erstens eine Welt vorstellt, in der Entitäten existieren, die die qualitativen Eigenschaften von Schmerzen besitzen ohne dass sie C-Faser-Reizungen und damit Schmerzen sind (es gibt also schmerzartige Nicht-Schmerzen) und dass es zweitens eine Welt gibt, in der es C-Faser-Reizungen und damit Schmerzen gibt, die nicht die qualitativen Eigenschaften besitzen, die Schmerzen in unserer Welt haben (es gibt also C-Faser-Reizungen, die sich nicht anfühlen wie Schmerzen, aber dennoch Schmerzen sind). Diese Ausräumung würde aber nun offensichtlich scheitern, da nach Ansicht Kripkes alles, was sich anfühlt wie Schmerz, in jeder möglichen Welt Schmerz ist und jedes natürliche Phänomen, dass Schmerz ist, muss sich so anfühlen wie Schmerz in dieser Welt. Die skizzierten Welten sind also nach Ansicht Kripkes gerade keine *möglichen* Welten.

Die Kontingenz im Fall von psychophysischen Identitätsannahmen könne grundsätzlich nicht ausgeräumt werden, da Erscheinung und Realität hier zusammenfallen. Schmerzen sind schmerzvoll; alles, was uns wie Schmerz erscheint, das *ist* auch Schmerz. „To be in the same epistemic situation that would obtain if one had a pain *is* to have a pain; to be in same epistemic situation that would obtain in the absence of pain *is* not to have a pain. The apparent contingency of the connection between the mental state and the corresponding brain state thus cannot be explained by some qualitative analogue as in the case of heat” (Kripke, 1972, S. 152; Hervorh. im Orig.). So schließt Kripke, da es – ungeachtet der für unsere Welt womöglich ausnahmslos geltenden psychoneuralen Verbindungen – zumindest *vorstellbar* ist, dass es z.B. zu einem Auftreten von Schmerzen kommt, ohne dass irgendwelche C-Fasern gereizt werden, und wir diese Kontingenz nicht über das epistemische Modell der Erscheinungs-Realitäts-Diskrepanz ausräumen können, dass Schmerzen nicht mit C-Faser-Reizungen identisch sein können. Überhaupt sei die psychophysische Identitätstheorie in den Augen Kripkes damit zum Scheitern verurteilt.

Auch wenn dieser Schluss von einer bloßen Vorstellbarkeit auf die Wirklichkeit selbst in der Philosophie des Geistes kaum noch Zustimmung findet, haben Kripkes Ausführungen dazu geführt, dass man im Hinblick auf psychophysische Identitätsbehauptungen nunmehr von *notwendigen* Beziehungen *a posteriori* ausgeht. *A posteriori* sind diese Relationen, weil Identitäten, wie die von Wasser und H₂O, uns nicht bekannt sein müssen. Erst die moderne Naturwissenschaft hat diese notwendige Identität aufgedeckt, obwohl sie schon immer Bestand hatte und Bestand haben muss.

Durch diese Bestimmungen verschärft sich das Argument der multiplen Realisierbarkeit zusätzlich: „In post-Kripkean world, to say, e.g., that c-

fibers are identical to pain is to say that in any possible world where an organism is experiencing pain, that organism is also in possession of active c-fibers. This entails, obviously, that organisms that differ from human beings in lacking c-fibers cannot feel pain. They *necessarily* cannot feel pain” (Shapiro & Polger, 2012, S. 265f, Hervorh. im Orig.). Im Zuge dessen ist es in der Philosophie des Geistes für jeden Verteidiger einer Identitätsannahme maßgeblich geworden zu erklären, wie die scheinbare Kontingenz von mentalen und neuronalen Gegebenheiten zustande kommt.

6.1.4.4. Die Erklärungslücke

Da die Identitätstheorie in der „Lehrbuchinterpretation“ als eine materialistische Lösung des Leib-Seele-Problems verstanden wird, wird auch ein Argument gegen sie hervorgebracht, das bereits eine längere Tradition in Philosophie und (empirischer) Wissenschaft aufweist. Es beinhaltet den Umstand, dass, auch wenn uns ein vollkommenes Wissen über all die neuronalen Prozesse zugänglich wäre, die einem bestimmten mentalen Erlebnis zugrunde liegen, dieses damit nicht befriedigend zu *erklären* ist. Man kann nur feststellen, dass eben diese Bedingungen vorhanden sein müssen, damit dieses Bewusstseinsgeschehen auftritt – eine Erklärung, warum es genau diese Bedingungen sind und nicht andere, ist daraus nicht zwangsläufig ableitbar.

Gemeint ist, dass es trotz einer an sich schlüssigen Identitätsannahme zwischen C-Faser-Reizungen und Schmerzen fraglich bliebe, warum es gerade C-Faser-Reizungen sind und nicht z.B. A-Faser-Reizungen, die mit Schmerzen identisch sind. Oder andersherum: warum erleben wir gerade Schmerzen und nicht eine andere sensorische Qualität, wenn C-Fasern feuern (vgl. Kim, 1998, S. 95f)? Es scheint der Identitätstheorie an

einer *Erklärung* dafür zu mangeln, warum die von ihr herausgestellten Identitätsverhältnisse gelten sollen, weshalb Levine (1983) hierfür den Begriff der Existenz einer gewissen *Erklärungslücke* („*explanatory gap*“) geprägt hat. Chalmers (1995, 1996) spricht in diesem Kontext vom so genannten *schwierigen Problem* des Bewusstseins („*hard problem of consciousness*“).

In ähnlicher Form findet sich dieses Argument aber bereits bei Thomas Henry Huxley (1866): „But what consciousness is, we know not; and how it is that anything so remarkable as a state of consciousness comes about as the result of irritating nervous tissue is just as unaccountable as the appearance of the Djin when Aladdin rubbed the lamp in the story, or as any other ultimate fact of nature“ (zit. nach Kim, 2005, S. 94).

Im deutschen Sprachraum war es zur fast gleichen Zeit die berühmte *Ignorabimus-Rede* des Berliner Physiologen Emil Du Bois-Reymond, die diesem Gedanken eine große Resonanz verlieh. Auf der 45. Versammlung deutscher Naturforscher und Ärzte im Jahr 1872 äußerte er sich in seinem Vortrag zunächst sehr optimistisch, was das naturalistische Forschungsprogramm der Wissenschaften anbelangt. So sei es prinzipiell möglich, aufgrund der Kenntnis der relevanten Ausgangsbedingungen und Gesetze das zukünftige Verhalten komplexester Systeme genau zu berechnen (vgl. Pauen, 2007, S. 81). Auch daran, dass geistige Vorgänge „Erzeugnisse der materiellen Bedingungen im Gehirn [sind]“ (Du Bois-Reymond, 1880/1974, S. 76), bestehe keinerlei Zweifel. Was allerdings zu bezweifeln sei, ist, dass wir jemals erklären könnten, warum Bewusstsein aus Materie entsteht:

Die astronomische Kenntnis des Gehirns, die höchste, die wir davon erlangen können, enthüllt uns nichts als bewegte Materie. Durch keine zu ersinnende Anordnung oder Bewegung materieller Teilchen aber lässt sich eine Brücke ins Reich des Bewusstseins schlagen. [...] Die neben den materiellen Vorgängen im Gehirn

einhergehenden geistigen Vorgänge entbehren also für unseren Verstand des zureichenden Grundes. Sie stehen außerhalb des Kausalgesetzes, und schon darum sind sie nicht zu verstehen, so wenig, wie ein mobile perpetuum es wäre. Aber auch sonst sind sie unbegreiflich. (ebd., S. 70)

So schloss Du Bois-Reymond seine Rede mit den berühmten Worten: *Ignoramus et ignorabimus* (lateinisch für: „Wir wissen es nicht und wir werden es niemals wissen“). In diesem Sinne sagt nun auch Mausfeld (2010), dass man neurobiologische Befunde keineswegs als Erklärungen mentaler Phänomene ansehen dürfe. Sie würden den Erklärungsbedarf im Gegenteil sogar noch vergrößern, „denn nun müssen wir neben den psychologischen Phänomenen auch noch erklären, *warum* sie gerade mit diesen oder jenen neurobiologischen Vorgängen korrelieren (S. 182; Hervorh. im Orig.).

Zwar enthält die Identitätstheorie als ontologische Position genau genommen „keine direkten Festlegungen darüber, wie sich unser *Wissen über* neuronale Prozesse zu unserem *Wissen über* mentale Prozesse verhält, vielmehr macht sie eine Aussage über das Verhältnis dieser Prozesse selbst“ (Pauen, 2005, S. 181). Dennoch wird ihre Gültigkeit im Zuge der Erklärungslückenproblematik an ein epistemisches Kriterium geknüpft. So würde ein dauerhaftes Scheitern unserer Erklärungsversuche...

...letztlich die Frage aufwerfen, ob die behauptete Identitätsbeziehung wirklich existiert. Denn wenn die Identitätstheorie Recht hat, müssten Theorien über Gehirnprozesse auf Dauer als intuitiv plausible Erklärungen der korrespondierenden mentalen Zustände anerkannt werden. Wenn Identität nämlich die Übertragbarkeit der zugehörigen Erklärungen impliziert, dann müsste ein solcher Fehlschlag zu dem Schluß führen, daß es sich hier doch um zwei unterschiedliche Entitäten handelt. (ebd., 1999, S. 13)

6.1.4.5. Die problematische Interpretation der Identität

Auch wenn man rückblickend konstatieren muss, dass das Argument der multiplen Realisierbarkeit den entscheidenden Fallstrick der Identitätstheorie ausgemacht hat, stellt es für das Vorhaben, eine psychophysische Identität als Forschungsgrundlage der Psychologie heranzuziehen, kein allzu großes Problem mehr dar. So werde ich an späterer Stelle im Zusammenhang mit meiner Neuformulierung einer psychophysischen Identitätstheorie, die auf mereologisch-kompositorischen Prinzipien fußt, aufzeigen, dass das Argument der multiplen Realisierbarkeit trotz seiner augenscheinlichen Validität weniger schlagkräftig ist als es auf den ersten Blick scheint (vgl. Kap. 7.5).

Eine Tragik von ganz anderem Ausmaß haben dagegen bestimmte (metaphysische) Deutungen psychophysischer Identitätsaussagen mit sich gebracht, die sich in der Diskussion der Identitätstheorie herauskristallisiert haben. Ich gebe Shapiro und Polger (2012) zwar durchaus recht, wenn sie kritisieren, dass die „Lehrbuchinterpretation“ den Intentionen der Identitätstheoretiker kaum gerecht wird und ihnen hier zumeist eine allzu abgedroschene These untergeschoben wird. Dennoch ist es für mich aber auch eine Tatsache, dass sich bei allen „klassischen“ Identitätstheoretikern eine Vielzahl problematischer (metaphysischer) Annahmen findet, die als Steilvorlage für die nachfolgende harsche Kritik an der Identitätstheorie gedient und damit zu einer sehr negativen Auffassung jeglicher Ansätze, die sich auf psychophysische Identitätsaussagen stützen, geführt haben.

Erstens: Was an den Identitätskonzeptionen zunächst bemängelt werden muss, ist, dass sie keine empirisch plausible Antwort auf die Frage nach der Ursache der von ihnen behaupteten psychophysischen Identität abliefern. Sie bieten keine genauere Bestimmung des Umstandes an, warum zwei (phänomenal) so verschiedene Dinge wie Gehirn und Geist mitein-

ander identisch sein können. Bei Feigl findet sich zwar andeutungsweise ein interessanter Ansatz, wenn er sagt, dass wir uns mittels verschiedener begrifflicher Systeme auf ein und dieselbe Sache beziehen und dadurch unterschiedliche Eindrücke ebendieser Sache erhalten. Es fehlt hierbei allerdings eine nähere, empirisch plausible Ausbuchstabierung dieser Annahme (wie sie hier in Kapitel 3-5 geleistet wurde).

Zweitens: Wenn das Kriterium für eine Lösung des Leib-Seele-Problems – wie häufig angenommen – darin besteht, dass eine Theorie aufzeigen kann, was mentale Eigenschaften *wirklich* sind, was ihr *Wesen* ausmacht, dann gilt dies auch für die Identitätstheorie. Wenn gelten soll: „Schmerzen sind C-Faser-Reizungen“, muss dies nach dieser Logik also so gedeutet werden, dass das auf der rechten Seite der Identitätsgleichung Stehende die verborgene Natur des auf der linken Seite der Gleichung Stehenden angeben soll. Genau dies entspricht auch der tatsächlichen Stoßrichtung der Identitätstheoretiker. Ihr Anliegen war es, eine *asymmetrische* Beziehung zwischen neuronalen und mentalen Eigenschaften zu formulieren. Bei Place und Smart finden wir eine Deutung der physischen Seite der Identitätsannahme als die wahre Natur von mentalen Zuständen, bei Feigl eine solche, die – zumindest tendenziell – der mentalen Seite den Vorrang gibt. Die Feiglsche Variante der Identitätstheorie hat in der nachfolgenden Diskussion allerdings kaum eine Bedeutung gespielt; *de facto* wurde die Identitätsthese als materialistische oder physikalistische These gedeutet, die der neuronalen Seite der Identitätsgleichung die *eigentliche* Realität zuspricht.

Wie bereits Deutscher (1967) sehr treffend herausgestellt hat, hat die Identitätstheorie weniger den Charakter einer *Identifikation mit* als vielmehr den einer *Identifikation als*. Sie möchte also weniger mentale Eigenschaften *mit* neuronalen Eigenschaften identifizieren; vielmehr geht es

ihr darum, dass mentale Eigenschaften *als* neuronale Eigenschaften herausgestellt werden.

Vor allem wurde diese „strenge“ Auslegung der Identität von Smart vorangetrieben. Bei Place und Feigl findet sich dagegen durchaus an einigen Stellen in Ansätzen eine *kompositorische* Interpretation der Identitätstheorie. So heißt es bei Place (1956), dass das, was die These der Identität von mentalen und neuronalen Eigenschaften bereitstellen soll, eine Offenlegung der *Mikrostruktur* bzw. „composition“ (S. 44) von mentalen Eigenschaften sein soll. Schmerzen sind in diesem Sinne C-Faser-Reizungen, weil der mentale Typ „Schmerz“ die Mikrostruktur „C-Faser-Reizungen“ aufweist. Hier wird also eine höherstufige Eigenschaft auf ihre Mikrostruktur zurückgeführt, ohne dass damit dem höherstufigen Phänomen die Existenz abgesprochen wird.

Eine ähnliche Interpretation der Identitätsbeziehung findet sich auch bei Feigl (1960/1970):

Laws of concomitance in the physical world could usually be accounted for in terms of underlying *identical* structures. Thus, for example, the correspondence of certain optical, electrical, and magnetic properties of various substances, as expressed in simple functional relations between the refraction index, the dielectric constant, and the magnetic permeability, is explainable on the basis of *the atomic structure* of those substances. [...] Similarly, it is at least a partially successful working programme of psycho-physiology to reduce certain correlated macro-behavioural features to *underlying identical neurophysiological structures and processes*. (S. 37f; Hervorh.: M.L.)

Auch wenn bereits Borst (1970) zu Recht darauf hinweist, dass „[t]he examples purport to give the physico-chemical or micro-structure of the things or properties in question, and it is possible to read them as providing *explanations* of familiar phenomena rather than as specifying what things or properties *really are*“ (S. 22; Hervorh. im Orig.), wird hieraus keineswegs eine neutrale oder symmetrische Interpretation der Identität

tätstheese abgeleitet. Im Gegenteil: Es ist die Symmetrie von Identitätsbehauptungen, die den Identitätstheoretikern Sorge bereitet hat, denn es sollte in ihren Augen partout vermieden werden, dass man – wenn mentale Eigenschaften neuronale Eigenschaften sind – auch von bestimmten neuronalen Eigenschaften sagen muss, dass sie mental sind. In diesem Sinne muss aber nun auch der Verweis auf die Mikrostruktur eines Systems gedeutet werden: Die Identität bedeute nicht einfach, dass mentale Zustände und Prozesse denselben räumlich-zeitlichen Bereich einnehmen wie ihre mikrostrukturellen (neuronalen) Grundlagen, also koextensiv sind. Die Identität bestehe vielmehr darin, dass sich in der Mikrostruktur eines Systems die *wahre* Natur seiner Existenz zeigt, sodass mentale Prozesse *nichts anderes als* neuronale Prozesse seien. So wie die wahre Natur eines Blitzes in elektrischen Entladungen besteht (vgl. Place, 1956) oder höherstufige physikalische Phänomene auf ihre atomare Struktur zurückgeführt werden können, seien Schmerzen *in Wirklichkeit* C-Faser-Reizungen.

Insofern trifft die folgende Kritik Nagels (1974/1997) unter diesen Voraussetzungen durchaus zu: „Ergibt es [...] einen Sinn zu fragen, wie meine Erlebnisse *wirklich* sind – im Gegensatz zu der Art und Weise, wie sie mir erscheinen? Wir können kein echtes Verständnis der Hypothese erlangen, daß ihre Natur von einer physikalischen Beschreibung erfaßt wird, ohne die grundsätzlichere Idee zu verstehen, daß sie eine objektive Natur *haben* (oder daß objektive Prozesse eine subjektive Natur haben können)“ (S. 271; Hervorh. im Orig.).

Drittens: Im Hintergrund der Abwehr gegenüber einer symmetrischen Interpretation der Identitätsrelation steht für die Identitätstheoretiker die Befürchtung, dass wenn man sie zulassen würde, als Konsequenz das Vorhandensein irreduzibler mentaler Eigenschaften in Kauf nehmen müsste. Die damals sehr stark an einer sprachlichen und logischen Ana-

lyse orientierte Diskussion suggerierte nämlich, dass man sich, wenn man eine Identität mentaler Zustände mit bestimmten Hirnzuständen behauptet, mit dieser Behauptung – quasi aussagenlogisch vorgeordnet – bereits auf bestimmte mentale Zustände beziehen muss und damit ihre Existenz einräumen würde. Dies wäre nach Maßgabe einer Identitätstheorie, die das wahre Wesen des Psychischen (ob nun Prozesse, Zustände oder Eigenschaften) im neuronalen Seinsmodus sucht, völlig inakzeptabel. Smart (1959) hat daher eine Sprechweise entwickelt, die ohne den Bezug auf mentale Eigenschaften auskommen soll. Statt zu sagen: „Ich sehe ein gelblich-oranges Nachbild“, könne man denselben Sachverhalt *ontologisch neutral* wie folgt wiedergeben: "*There is something going on which is like what is going on when I have my eyes open, am awake, and there is an orange illuminated in good light in front of me, that is, when I really see an orange*" (vgl. S. 149; Hervorh. im Orig.). In den Augen vieler Autoren hat die Identitätstheorie damit den Charakter oder Auftrag, eine „ontologisch-neutrale Analyse mentalistischer Terme“ (Metzinger, 2007a) durchzuführen. Auch wenn dies aus heutiger Sicht – insbesondere der eines theoretischen Psychologen – recht bizarr anmutet, war dies lange Zeit einer der Hauptdiskussionspunkte im Rahmen der „konzeptuell-sprachlogischen Überprüfung“ der Identitätstheorie, sodass auch diese Bestimmung immer noch mitschwingt, wenn von einer psychophysischen Identitätsannahme die Rede ist.

Viertens: Aufgrund der asymmetrischen Interpretation der Identitätstheorie, die der neuronalen Seite die eigentliche Realität zuspricht, ist die Identitätstheorie (und leider damit überhaupt das Konzept einer psychophysischen Identität) mit einem relativ kruden reduktiven Physikalismus gleichgesetzt worden. Als Synonym für die „Identitätstheorie“ hat sich der Ausdruck „Typen-Physikalismus“ eingebürgert und ist die heute sogar gebräuchlichere Redeweise. Ebenso werden die Ausdrücke „Identität“

und „Reduzierbarkeit“ (oder gar „Reduktionismus“) als so etwas wie unzertrennliche Zwillinge betrachtet. Auch wenn mit einer psychophysischen Identitätskonzeption jeglicher Spielart notwendigerweise eine bestimmte (ontologische) Reduktionsmöglichkeit gegeben ist, muss diese – wie ich an späterer Stelle nachweisen werde (vgl. Kap. 8 und 9) – keineswegs so stringent und folgenschwer (für die „reduzierten“ Entitäten) ausfallen, wie der Anschein erweckt wird.

Fünftens: Eine besondere Brisanz liegt in dem Umstand, dass in (zumindest mittelbarer) Folge der Annahme einer Typen-Identität fast in der gesamten philosophischen Debatte bis zum heutigen Tage nur noch um die Frage nach einer Identität, Reduzierbarkeit oder die Möglichkeit der Formulierung einer physikalistisch akzeptablen Theorie mentaler *Eigenschaften* geht²¹⁹. Zusätzlich dürfte man in diesem omnipräsenten Rekurs auf Eigenschaften auch eine willkommene *sprachliche Vereinfachungsmöglichkeit* gesehen haben, die darin begründet ist, dass man alle anderen in der Leib-Seele-Debatte relevanten ontologischen Bezugsgrößen (Prozesse, Zustände, Ereignisse) in Eigenschaftsaussagen umformulieren kann. Der mentale Prozess einer räumlichen Orientierung kann in die Eigenschaft des „Habens räumlicher Orientierung“ umformuliert werden; genauso kann der mentale Zustand einer Motiviertheit durch das „Haben von Motivation“ ausgedrückt werden, wie auch mentale Schmerzereignisse als die mentale Eigenschaft des „Habens von Schmerzen“ aufgefasst werden können. Oder wie Metzinger (2007c) die Beziehung zwischen mentalen Typen, Vorkommnissen (Token) und Eigenschaften auf den Punkt bringt: „Wenn ich jetzt gerade ein visuelles Erlebnis von *Röte*

²¹⁹ Beckermann (1996) nennt die gesamte in der Folgezeit der Identitätstheorie diskutierte Position des Physikalismus daher sicher nicht zu Unrecht „Eigenschafts-Physikalismus“ (S. 3).

habe, dann befinde ich mich jetzt und hier in einem *token* des Zustands-*typs* des Ein-Rot-Erlebnis-Habens“ (S. 168).

Aus dieser scheinbar harmlosen sprachlichen Konvention ergibt sich aber ein handfestes Problem: Man muss sich vor Augen führen, dass Eigenschaften immer nur von einem gesamten Objekt oder System ausgesagt werden können. Ein Mensch hat z.B. bestimmte psychische und bestimmte physische Eigenschaften, deren Verhältnis man im Allgemeinen diskutieren und z.B. Identitätsbehauptungen aufstellen kann. Wenn sich die Diskussion einer Identitätsthese dabei nun ausschließlich – ohne weiteren Einbezug struktureller, mereologischer oder mechanistischer Aspekte – auf Eigenschaftsrelationen bezieht (wie eben in der Aussage, dass die psychischen Eigenschaften eines Organismus nichts anderes als seine neuronalen Eigenschaften sind), kommt es daher mehr oder weniger *begriffsnotwendig* zu der problematischen These, dass es sich hierbei um ein bloßes *aggregatives* Verhältnis und somit eine krude Form von Reduktivität bzw. Reduktionismus handeln muss (vgl. Kap. 5.1.2.2). Mentale Eigenschaften wären demnach nicht anderes als eine *Ansammlung* neuronaler (oder anderer physischer/physiologischer) Eigenschaften und die These, dass es letztendlich der (Neuro-)Biologie zukäme, diese Eigenschaften und die hiermit zusammenhängenden Verhaltensweisen (von Menschen) adäquat erklären zu können, erschiene nur konsequent.

Sechstens – und dies könnte wiederum eine Ursache für einige der anderen genannten Punkte sein – entsteht der Eindruck, dass das Argument der multiplen Realisierbarkeit die Identitätstheorie zu einer Position gemacht hat, auf die sie vor dem Aufkommen dieses Argumentes überhaupt nicht festgelegt war. Mit der Akzeptanz dieses Argumentes ist die Identitätstheorie selbst unbemerkt zu einem Spezialfall einer allgemeineren Theorie physikalischer Realisierung (vgl. Kap. 6.5) geworden. Vormalig bereits bestehende Annahmen, die Identität auf kompositorischen

Prinzipien zu begründen, sind spätestens seit dem Entflammen der Diskussion um die multiple Realisierbarkeit von der Bildfläche verschwunden.

Zu sagen, es besteht eine Identität zwischen mentalen und physischen Eigenschaften, wird auf diese Weise (implizit) gleichgesetzt mit der These, dass auch hier mentale Eigenschaften durch physische Eigenschaften *realisiert* werden, mit der einzigen Einschränkung, dass die klassische Identitätstheorie annimmt, dass es immer ein bestimmter Typ physikalischer Eigenschaften ist, der einen bestimmten Typ mentaler Eigenschaften *realisiert* (was diesen Spezialfall *empirisch* unplausibel macht). Gemäß der Realisierungstheorie werden mentale Eigenschaften dadurch „in die Welt gehoben“, dass sie eine *physische* Existenz haben. Das, was mentale Eigenschaften zu wissenschaftlich respektablen, ontologisch beständigen und kausal wirksamen Größen macht, ist in diesem Sinne ihre Hervorbringung durch (oder lokale Identität mit) physikalische(n) Eigenschaften. Dabei konstituiert das Psychische – entgegen den Annahmen dieser Arbeit (vgl. Kap. 5) – keine eigene natürliche Organisationsebene, sondern ist auf derselben Ebene angesiedelt wie seine neuronalen Realisierer (zumeist werden mentale Zustände hierbei im Sinne des Funktionalismus als *Eigenschaften zweiter Ordnung* aufgefasst). Dabei kommt dieser Interpretation des Konzeptes psychophysischer Identität keineswegs Notwendigkeit zu. Wie in Kapitel 7 aufgezeigt werden wird, kann durchaus auch auf der Grundlage kompositorischer Prinzipien eine Theorie psychophysischer Identität konzipiert werden, die prinzipiell ohne das Realisierungsideom (und ohne funktionalistische Analysen mentaler Zustände) auskommt.

6.2. Die Identität von psychischen und physischen Ereignissen (Token-Identity-Theory)

Entgegen der Identitätstheorie nach Place, Smart oder Feigl kann man die Identitätsthese auch als eine Identität von einzelnen Ereignissen oder Vorkommnissen auffassen, auch wenn in diesem Zusammenhang nur selten explizit von einer *Identitätstheorie* gesprochen wird (was nicht zuletzt an den vielen herausgestellten Kritikpunkten am Identitätskonzept liegen wird). Verbunden ist die Token-Identitätsthese vor allem mit dem Aufsatz *Mental Events* von Donald Davidson (1970/ 2007) und weist in der Philosophie (gerade auch in anderen Bereichen als der Philosophie des Geistes) eine große Anhängerschaft oder doch zumindest einen großen Sympathisantenkreis auf.

6.2.1. Davidsons anomaler Monismus

Der große Unterschied zwischen Davidsons Lösungsansatz zum Leib-Seele-Problem und der „klassischen“ Identitätstheorie nach Place, Smart oder Feigl besteht im Rekurs auf Ereignisse anstatt auf Typen oder Eigenschaften. Die als *Anomaler Monismus* bezeichnete Position versucht nämlich nicht, das Verhältnis von mentalen und neuronalen Eigenschaften insgesamt zu klären, sondern geht nur von einer Identität von einzelnen mentalen und neuronalen *Ereignissen* oder *Vorkommnissen* (Token) aus. Für Davidson haben Eigenschaften keinen akzeptablen ontologischen Status. Ihre Existenz erschöpft sich vielmehr in unseren sprachlichen Zuschreibungen zu Objekten: „Daß ein Gegenstand eine Eigenschaft hat, heißt seiner Meinung nach nichts anderes, als daß ein bestimmtes Prädikat auf diesen Gegenstand zutrifft“ (Beckermann, 1999, S. 185). Daher wird verständlich, warum der Versuch einer Formulierung einer psycho-

physischen Identitätstheorie auf Eigenschaftsniveau für Davidson von vornherein ausscheidet.

Seine These leitet Davidson (1970/2007) aus der scheinbaren Unvereinbarkeit dreier Prinzipien her. Das erste Prinzip besagt, dass zumindest einige mentale Ereignisse mit physischen Ereignissen interagieren. Insbesondere verweist er dabei auf die Bedeutung von Wahrnehmung, Aufmerksamkeit, Beurteilungen, Entscheidungen und Meinungsäußerungen hin (vgl. S. 223), die Körperbewegungen verursachen können, woraus deutlich wird, dass seine Theorie in erster Linie auf *kognitive* Ereignisse bezogen ist. Auch hier besteht also ein Unterschied zur „klassischen“ Identitätstheorie, die sich von ihrer Ausformulierung her maßgeblich auf phänomenale Ereignisse bezieht.

Das zweite Prinzip des „nomologischen Charakters der Kausalität“ (ebd.) sagt aus, dass es, immer wenn es Kausalität gibt, auch ein allgemeines, streng deterministisches Gesetz geben muss, das die beteiligten Ereignisse in eine Ursache-Wirkungs-Relation setzt.

Das dritte Prinzip beinhaltet die für Davidsons Ansatz zentrale *Anomalität des Mentalen* und besagt, dass es keine streng deterministischen Gesetze gibt, mittels derer man mentale Ereignisse voraussagen oder erklären könnte (vgl. ebd., S. 224). Strikte Gesetze sind für Davidson solche, die sich in einem einheitlichen Vokabular formulieren lassen und ausnahmslos und unter allen Bedingungen gelten. Seiner Ansicht nach sind die einzigen strikt deterministischen Gesetze die der Physik. Es mag zwar auch psychophysische Gesetze geben, nur sind diese nicht strikt: „[E]s kann wahre Allgemeinaussagen geben, die das Mentale und das Physikalische aufeinander beziehen, Aussagen, die die logische Form von Gesetzen haben; sie sind jedoch nicht [...] gesetzesartig. Würden wir durch einen absurd unwahrscheinlichen Zufall auf eine nichtstochastische wahre psychophysische Verallgemeinerung stoßen, so hätten wir keinen

Grund anzunehmen, sie sei mehr als eine ungefähre Wahrheit“ (ebd., S. 232). Psychophysische Gesetze haben demnach bestenfalls den Status von „heteronomen *ceteris paribus*-Gesetzen“ (Beckermann, 1999, S. 197). Davidson ist bemüht, diese These über eine diffizile Analyse von Rationalitätsprinzipien bzw. dem so genannten Argument eines *Holismus des Mentalen* nachzuweisen. Da diese primär begriffs- und sprachlogisch orientierte Argumentation für die Zielstellung des Kapitels allerdings eher unerheblich ist bzw. unter anderen Vorzeichen bereits hinlänglich diskutiert wurde, möchte ich auf ihre eingehende Darstellung verzichten. Der zentrale Gedanke ist in diesem Zusammenhang jedenfalls, dass es keine strikten psychophysischen Gesetze geben kann, weil die „Zuschreibungsbedingungen mentaler und physikalischer Begriffe nicht zueinander passen“ (Beckermann, 1999, S. 200).

Für die Stoßrichtung der Formulierung seiner eigenen Identitätstheorie ist es für Davidson (1970/2007) entscheidend nachzuweisen, dass die vermeintliche Unvereinbarkeit der drei genannten Prinzipien in Wirklichkeit gar nicht zutrifft. In der zentralen Passage zur Herleitung einer psychophysischen Identität auf der Ebene von Ereignissen heißt es hierzu:

Nach dem Prinzip der kausalen Bedingtheit des Mentalen nehmen wir an, dass zumindest einige mentale Ereignisse die Ursachen und Wirkungen physikalischer Ereignisse sind; unser Argument betrifft nur sie. Ein zweites Prinzip (das Prinzip des nomologischen Charakters der Kausalität) besagt, dass hinter jeder wahren Kausalaussage ein strenges Gesetz steht, das Ereignisse der Typen verknüpft, zu welchen die als Ursache und Wirkung genannten Ereignisse gehören. [...] Die Physik schließt das Versprechen ein, ein umfassendes geschlossenes System zu liefern, aus dem sich verbürgtermaßen eine genormte, eindeutige Beschreibung eines jeden physikalischen Ereignisses in einem Vokabular ergibt, das für Gesetze geeignet ist. Es ist einfach deshalb unplausibel, dass mentale Begriffe von sich aus ein solches System ergeben können, weil das Mentale kraft unseres ersten Prinzips kein geschlossenes System bildet. Zu viel beeinflusst nun einmal das Mentale, das nicht

selbst ein systematischer Teil seiner ist. Verbinden wir diese Beobachtung jedoch mit der Konklusion, dass keine psychophysische Aussage in ein strenges Gesetz eingepasst werden oder selbst ein solches Gesetz sein kann, so erhalten wir das Prinzip der Anomalie des Mentalen: Es gibt überhaupt keine strengen Gesetze, mit Bezug auf die wir mentale Ereignisse voraussagen und erklären können. Die Identität lässt sich im Anschluss hieran leicht dartun. Nehmen wir an, ein mentales Ereignis *m* habe ein physikalisches Ereignis *p* verursacht; *m* und *p* müssen dann unter irgendeiner Beschreibung ein strenges Gesetz instantiieren. Dieses Gesetz kann nach dem vorigen Abschnitt nur ein physikalisches sein. Fällt *m* jedoch unter ein physikalisches Gesetz, so gibt es von *m* eine physikalische Beschreibung; *m* ist demnach ein physikalisches Ereignis. Ein entsprechendes Argument funktioniert für den Fall, dass ein physikalisches Ereignis ein mentales Ereignis verursacht. Folglich ist jedes mentale Ereignis, das in einer Kausalbeziehung zu einem physikalischen Ereignis steht, ein physikalisches Ereignis. (S. 240f)

Der vermeintliche Widerspruch zwischen den drei grundlegenden Prinzipien Davidsons lässt sich also dadurch auflösen, dass mentale Ereignisse physische Ereignisse sind. Wenn nur physikalische Gesetze Kandidaten für strikt deterministische Gesetze sind und man gleichzeitig an der Gültigkeit einer mentalen Verursachung festhalten möchte, dann kann dieser nur zustande kommen, wenn mentale Ereignisse eben selbst physische Ereignisse sind.

Was in dem Zitat von Davidson ebenso deutlich wird, ist die zentrale Bedeutung der Möglichkeit von *Beschreibungen* für seine Theorie bzw. Ontologie. Beschreibungen sind bei Davidson mit ontologischen Erwägungen stark verschränkt; was mittels des begrifflich-sprachlichen Inventars einer bestimmten Wissenschaft beschrieben werden kann, muss auch (als eine Entität dieser Wissenschaft) existieren. Für ihn sind alle Gesetze letzten Endes *sprachliche* Formulierungen und die Ebene der Sprache sei es ebenso, die das Leib-Seele-Problem erst hervorbringe (vgl. Metzinger, 2007d, S. 220). Aufgrund dessen müsse auch eine Lösung des Leib-Seele-Problems eine *sprachliche* Klärung herbeiführen. Davidsons Ansatz ist hierbei, von unterschiedlichen Beschreibungen ein und dessel-

ben Ereignisses auszugehen (vgl. Kap. 3.3). Da mentale Ereignisse, die wir für gewöhnlich über psychologische Begriffe beschreiben, die Ursache von physikalischen Ereignissen sein können, muss es auch physikalische Beschreibungen mentaler Ereignisse geben. Wovon es wiederum eine physikalische Beschreibung gibt, ist ein physikalisches Ereignis.

6.2.2. Die problematische Identitätskonzeption der Token-Identitätstheorie

An Davidsons Ansatz wurden in der philosophischen Debatte verschiedene Aspekte kritisiert. Zentral war hierbei seine kaum begründete These, dass zwei Ereignisse nur in einer Kausalbeziehung stehen können, wenn es ein strikt deterministisches Gesetz gibt, unter das sie fallen. Es scheint offensichtlich zu sein, dass es eine Kausalbeziehung zwischen zwei Ereignissen geben kann und wir diese in den Blick nehmen können, ohne dass wir uns hierbei zwangsläufig auf Gesetze beziehen müssen.

Starke Kritik hat vor allem Kim (1976, 1998, 2005) an Davidsons Position geübt. Kim kritisiert *erstens*, dass Davidson nur eine negative Bestimmung des psychophysischen Verhältnisses abgibt: „In particular, it famously claims that there are no laws connecting mental kinds or properties with physical ones. This component of the doctrine, the anomalousness of the mental in relation to the physical, is a *negative* thesis: it tells us how the mental is *not* related to the physical, and says nothing about how the two *are* related” (Kim, 1998, S. 4; Hervorh. im Orig.). Allerdings kann man Kim entgegenhalten, dass mit der Bestimmung, dass jedes mentale Ereignis mit einem physischen Ereignis identisch ist, durchaus eine positive ontologische Bestimmung des psychophysischen Verhältnisses vorliegt, die aber aufgrund der Anomalität des Mentalen nicht auf eine epistemisch-nomologische Ebene gehoben werden kann.

Was Kim Davidson *zweitens* ankreidet, ist in der weiteren Rezeption des Anomalen Monismus bzw. der Besprechung des Leib-Seele-Problems überhaupt zu einer annäherungsweise allgemein akzeptierten Doktrin geworden, nämlich, dass die Diskussion der Beziehung zwischen Ereignissen die eigentliche Kernfrage der Leib-Seele-Debatte verfehlen würde. Ereignisse sind für Davidson „nicht-wiederholbare, zeitlich-bestimmte Individuen, wie etwa der Ausbruch des Vulkans, die (erste) Geburt oder der Tod einer Person“ (ebd., 1970/2007, S. 225). Sie sind für ihn also offenbar *ontologisch basale* Einzeldinge, die nicht auf andere ontologisch grundlegendere Entitäten zurückgeführt werden können (vgl. Beckermann, 1999, S. 185). Eine solche Klärung der Beziehung von mentalen und physischen Ereignissen sei aber nun – zumindest in den Augen der Kritiker – nicht hinreichend, um das Leib-Seele-Problem zu lösen. Vielmehr müsse man hierzu das Verhältnis von mentalen und physischen *Eigenschaften* verständlich machen. Kim weist darauf hin, dass Davidsons Begriff von ontologisch basalen (d.h. nicht zusammengesetzten) Ereignissen anzuzweifeln ist. Zumindest sei dieser nicht konkurrenzlos. Seiner Ansicht nach kann jedes Ereignis als etwas angesehen werden, das drei Bestimmungsstücke aufweist: Ein bestimmter *Gegenstand* hat zu einem bestimmten *Zeitpunkt* eine bestimmte *Eigenschaft* (vgl. Kim, 1976). Damit hat man nun auch die gewünschte Verlagerung zu einer Besprechung von Eigenschaften erreicht, die für die derzeitige Debatte in der Philosophie (des Geistes) charakteristisch ist. Analysiert man Davidsons Argument in Form von Eigenschaftsbeziehungen, dann würde es nach Ansicht Kims (1998) nur aussagen, dass jedes Ereignis, das eine bestimmte mentale Eigenschaft hat, ebenso eine bestimmte physische Eigenschaft aufweist, was letztendlich kaum etwas aussagt.

Insbesondere liege dies *drittens* an der Ausblendung von Typen-zu-Typen-Verhältnissen. Nach Kim (1998) ist eine Theorie, die solche Verbindun-

gen nicht integrieren kann, eine solche „that in effect says there is no connection between our mental nature and our physical nature. This means that the monism of anomalous monism is no less a negative thesis than the anomalism of anomalous monism” (S. 5). Davidson wurde im Nachgang daher unter Einbeziehung des für ihn selbst ungebräuchlichen Eigenschaftsbegriffs so interpretiert, dass unter Maßgabe seiner Bestimmungen gelten würde, dass bestimmte Ereignisse, als basale Einzeldinge, *sowohl* physische *als auch* mentale Eigenschaften aufweisen (vgl. z.B. Höhle, 2006). Hiermit sei nach Kim (1998) aber nur eine recht beliebige bzw. uneindeutige Kennzeichnung der Gehirn-Geist-Relation vorgenommen: Denn wenn man es beispielsweise als Tatsache erachtet, dass jedes Objekt, das eine Farbe hat, auch eine bestimmte Form aufweist, dann würde Davidsons Anomaler Monismus ebenso wenig über die Beziehung des Mentalen und Physischen aussagen, wie die Behauptung, dass jedes Objekt, das eine Farbe hat, auch eine Form hat, etwas über die Beziehung zwischen Farbe und Form aussagt (vgl. ebd., S. 5).

Soweit die Kritik an Davidsons Ansatz, wobei sich im Verlauf der vorliegenden Arbeit allerdings zeigen wird, dass die von Kim vorgetragene Kritik an Identitätsaussagen auf Token-Niveau in vielen Punkten nicht stichhaltig ist. Problematisch oder zumindest nicht hinreichend für eine Theorie des Geistes, die als Forschungsgrundlage der Psychologie dienen könnte, erscheint mir die spezielle Token-Identitätstheorie Davidsons aus ganz anderen Gründen.

Seine Theorie ist zunächst durchaus ein sehr interessanter Ansatz zur Versöhnung materialistischer und „mentalistischer“ Intuitionen und könnte als eine Art neutraler Monismus verstanden werden. Davidson wählt damit eine (Auf-)Lösung des Leib-Seele-Problems, die auf unterschiedliche Beschreibungsebenen (als bestimmte Form theoretischer Ebenen) ein und desselben Phänomens hinausläuft. Wie in Kapitel 3

herausgestellt wurde, dürften der Möglichkeit zu solch unterschiedlichen theoretischen Beschreibungen aber zumeist unterschiedliche perzeptuelle *Perspektiven* zugrunde liegen. Daher sind unterschiedliche (sprachlich-konzeptuelle) Beschreibungsformate eines Systems ein sekundäres Phänomen und machen nicht den eigentlichen Ursprung des Problems aus. Selbst wenn man Davidson also zugesteht, dass er einen wichtigen Schritt zur (Auf-)Lösung des psychophysischen Problems tätigt, indem er die Geist-Gehirn-Differenz auf unterschiedliche (deskriptive oder perzeptuelle) Repräsentationsformate ein und desselben Gegenstandes zurückführt, bleibt diese „Analyse“ auf halbem Wege stecken. Es bleibt unklar, warum es diese unterschiedlichen perspektivischen Erscheinungsformen überhaupt gibt bzw. inwieweit es zustande kommt, dass ein und derselbe Gegenstand nach Maßgabe physikalischer und auch psychologischer Begrifflichkeiten beschrieben werden kann.

Entspringen diese Beschreibungen lediglich unserer „subjektiven“ Auffassung eines an sich ungegliederten Gegenstandes (was relativ unplausibel ist) oder gehen sie auf eine natürliche Gliederung, auf eine bestimmte Form natürlicher Organisationsebenen, durch die ein Gegenstand gekennzeichnet ist, zurück? Gerade wenn es Davidson tatsächlich darum geht, eine gewisse Autonomie und Nicht-Reduzierbarkeit mentaler Ereignisse nachzuweisen, und das trotz seines Eingeständnisses, dass mentale Ereignisse physische Ereignisse sein müssen, damit ihre kausale Wirkmächtigkeit gewährleistet werden kann, sollte er auch aufzeigen können, warum mentale Ereignisse dann eben doch nicht *nur* physische Ereignisse sind. Der alleinige Verweis auf die von ihm propagierte „nomologische Inkommensurabilität des Mentalen“ (Metzinger, 2007d, S. 220) erscheint in meinen Augen hierfür nicht ausreichend zu sein, da es sich hier um ein rein epistemisches Argument handelt.

Die Token-Identitätsthese besagt, dass *ein bestimmtes* mentales Ereignis mit *einem bestimmten* physischen Ereignis identisch ist. Die Identität wird dabei einfach indirekt aus der Tatsache abgeleitet, dass mentale Ereignisse, da sie offenkundig kausal wirksam sind, physische Ereignisse sein müssen, da sonst die kausale Geschlossenheit der physikalischen Welt infrage gestellt wäre bzw. eine kausale Überdetermination bestimmter physikalischer Ereignisse vorliegen würde. Was genau meint dann hier aber *Identität*? Es fehlt eine genauere positive Bestimmung der psychophysischen Beziehung, wobei insbesondere die Frage nach der Bedeutung kompositorischer Verhältnisse Berücksichtigung finden müsste.

Die Besprechung der Leib-Seele-Frage ist bei Davidson, wie es bei den meisten *analytischen* Philosophen des Geistes lange Zeit der Fall war (und bei so manchen noch immer ist), primär eine Angelegenheit der *sprachlichen* Analyse. Wir müssen unsere Sprache entwirren, um dann, wenn wir unsere Begriffe auf ein Maximum an Klarheit gebracht haben, mittels einer sauberen logischen Ableitung eine Lösung auf diese schwierige Frage zu finden. Ein Einbezug der Forschungswirklichkeit der Fächer, die sich mit der empirischen Untersuchung des Psychischen beschäftigen (sprich: Psychologie bzw. die Kognitions- und Neurowissenschaften) erscheint unnötig. Insofern verwundert es nicht, dass konkrete kompositorische Verhältnisse (mereologischer oder mechanistischer Art) für die Theorie Davidsons keine Rolle gespielt haben. Es scheint einfach unter seinen Voraussetzungen *logisch* zwingend zu sein, dass mentale Ereignisse physische Ereignisse sind. Eine empirisch plausible Ableitung und exakte Ausbuchstabierung dessen, was mit einer psychophysischen Identität gemeint ist, schien ihm offensichtlich unnötig zu sein.

6.3. Supervenienz und der „Mythos des nicht-reduktiven Materialismus“²²⁰

Das Prinzip der Supervenienz (im heutigen Sinn des Wortes²²¹) stammt ursprünglich aus der Diskussion bestimmter Problemstellungen innerhalb der Ethik. Auch hier gab es eine Debatte um die Art der Beziehung, die zwischen bestimmten normativen Eigenschaften wie Geboten und Verboten auf der einen sowie den physischen Eigenschaften eines Menschen auf der anderen Seite besteht. Es herrscht zwar eine weitestgehende Einigkeit vor, dass sich solche moralischen Prinzipien nicht einfach auf die Physis eines Menschen zurückführen lassen. Dennoch ist ohne physische Grundlagen auch keine Realisierung solcher ethischer Prinzipien denkbar. Darüber hinaus erschien es höchst unplausibel anzunehmen, dass Individuen, die physisch völlig indifferent sind, sich in ihren moralischen Eigenschaften unterscheiden können, denn „mit unseren moralischen Intuitionen scheint die Annahme unvereinbar, daß es zwei physisch völlig gleiche Gegenstände – also z.B. zwei gleich große, gleich rote, gleich schmeckende Äpfel mit der gleichen chemischen Zusammensetzung [...] – geben kann, von denen der eine gut und der andere schlecht ist“ (Beckermann, 1999, S. 204). Inwieweit liegt hier also eine Abhängigkeit des Moralischen von einer physischen Basis vor? Es war George Edward Moore (1922), der hierauf eine Antwort im Sinne des Supervenienzkonzeptes formulierte (ohne allerdings selbst den Begriff explizit zu ge-

²²⁰ Es handelt sich um einen Aufsatztitel von Kim (1993).

²²¹ Nach Kim (1993) stammt die früheste nachgewiesene Verwendung des Adjektivs „supervenient“ aus dem Jahr 1594, während das Nomen „Supervenienz“ wahrscheinlich erstmals 1664 erwähnt wurde. In der Philosophie hatte der Terminus bis ins 20. Jahrhundert hinein zunächst die Bedeutung von „etwas Zusätzliches“ oder „etwas Hinzukommendes“.

brauchen): Die Abhängigkeit ist so gestaltet, dass es keinen Unterschied in den wertenden bzw. moralischen Eigenschaften eines Menschen geben kann, ohne dass es auch Unterschiede in dessen physischen Eigenschaften gibt. Mit dieser Idee versuchte Moore gleichzeitig einen realistischen und nichtsdestotrotz nicht-reduktiven moralphilosophischen Standpunkt einzunehmen. So gibt es demnach zwar irreduzible moralische Tatsachen, diese supervenieren aber über physischen Eigenschaften.

6.3.1. Psychische Supervenienz: Prinzip und Grenzen

Wenn auch hier – wie bei der Identitätsthese – die Formulierung des Konzeptes einer *psycho-physischen* Supervenienz zumeist als Errungenschaft der modernen Philosophie des Geistes gepriesen wird, kann gezeigt werden, dass das Grundprinzip einer psychophysischen Supervenienz spätestens durch Spinoza in wichtigen Zügen beschrieben wurde, obgleich auch ihm wie Moore der Ausdruck „Supervenienz“ noch fremd war (vgl. Lüdman, unveröffentlichtes Manuskript). So heißt es bei ihm: „Je befähigter ein Körper ist, vieles zugleich zu tun oder zu leiden, desto befähigter ist auch sein Geist, vieles zugleich zu erfassen“ (ebd., 1677/1977, S. 143) oder an anderer Stelle sogar noch markanter: „Alles, was das Tätigkeitsvermögen unseres Körpers vermehrt oder vermindert, hemmt oder fördert [...], fördert oder hemmt das Denkvermögen unseres Geistes“ (S. 279). Doch Spinoza geht noch hierüber hinaus, indem er die Vielgestaltigkeit des menschlichen Geistes mit der Tatsache begründet, dass der menschliche Körper aus Verschiedenartigem zusammengesetzt ist (ebd., S. 159). Damit müsste nun auch gelten, dass in dem gleichen Maße, in dem die Komplexität eines menschlichen Gedankens oder überhaupt einer Bewusstseinsqualität zunimmt, auch die körperliche Entsprechung

an Komplexität zugenommen haben muss (was empirisch/neurowissenschaftlich durchaus nicht abwegig ist).

Ist die Supervenienz-Relation dem Spinozistischen System eher implizit inhärent, nimmt Fechner den heutigen Begriff in seiner vollen Bedeutungsweise vorweg. Fechner – als entscheidender Begründer der Psychophysik – hatte bereits ganz klar erkannt, dass geistige Ereignisse auf gesetzesartige Weise von physischen Prozessen im Gehirn begleitet bzw. funktionell abhängig sind. Für diesen Zusammenhang verwendet Fechner verschiedene Ausdrücke. So spricht er z.B. in den *Elementen der Psychophysik* vom „Namen des Functionsprinzipes“ (1860/1964, S. 380) oder in der *Seelenfrage* vom „psychophysischen Grundgesetze“ (1861, S. 203). Dieses besagt, „daß nichts im Geiste bestehen, entstehen, gehen kann, ohne daß etwas im Körper mit besteht, entsteht, geht [...]. Man kann es kurz so ausdrücken, daß alles Geistige seinen Träger oder Ausdruck in etwas Körperlichem und hiedurch seine weiteren Wirkungen und Folgen im Körperlichen hat“ (ebd., S. 211).

In die Gegenwartsdebatte um das Leib-Seele-Problem wurde der Begriff der Supervenienz wiederum von Donald Davidson (1970/2007) eingeführt. Da er von einer Unmöglichkeit der Formulierung strenger psychophysischer Gesetze ausgeht, ist eine Reduktion psychischer Ereignisse auf die korrespondierenden physischen Ereignisse nach seiner Ansicht ausgeschlossen. Der von ihm propagierte *Anomale Monismus* beschränkt sich daher auf die Feststellung, dass es zu jedem Ereignis, zu dem es eine mentale Beschreibung gibt, auch eine physikalische Beschreibung geben muss. Die ontologische Beziehung zwischen mentalen und physikalischen Ereignissen ist damit allerdings nur vage umrissen. Um der kaum zu leugnenden Abhängigkeit des Mentalen vom Vorhandensein bestimmter physischer Entitäten Rechnung zu tragen, führt Davidson daher das Supervenienzprinzip ein:

Obgleich die von mir beschriebene Position die Existenz psychophysischer Gesetze bestreitet, ist sie mit der Auffassung verträglich, dass mentale Eigenschaften gegenüber physikalischen Eigenschaften in irgendeinem Sinne dependent oder supervenient sind. Mit dieser *Supervenienz* könnte gemeint sein, dass sich zwei Ereignisse, die sich in jeder physikalischen Hinsicht gleichen, nicht in einer mentalen Hinsicht unterscheiden können, oder dass sich ein Objekt nicht in einer mentalen Hinsicht verändern kann, ohne sich in einer physikalischen Hinsicht zu verändern. (Davidson, 1970/2007, S. 230; Hervorh. im Orig.)

Die beiden zentralen Bestimmungen, die aus dieser allgemeinen Formulierung des Supervenienzprinzips hervorgehen, sind also, dass erstens eine vollkommene physische Gleichheit mit einer vollkommenen psychischen Gleichheit einhergehen muss und dass zweitens jede Variation im Bereich des Mentalen eine Variation im Bereich des Physischen voraussetzt. Eine zentrale Intention bei der Formulierung des Supervenienzprinzips besteht darin, eine *asymmetrische* Abhängigkeit des Mentalen vom Physischen zu begründen. Die Relation ist nicht einfach umkehrbar. So wird nicht impliziert, dass sich zwei Wesen, die mental ununterscheidbar sind, auch physisch identisch sein müssen oder dass eine Veränderung in der Physis dieses Wesens immer auch mit einer mentalen Variation einhergehen muss. Auf das Verhältnis von Gehirn und Geist angewandt, verlangt eine Supervenienzbeziehung nur, dass es *de facto* keine zwei Merkmalsträger gibt, die sich in ihren mentalen Eigenschaften, aber nicht in ihren neuronalen Eigenschaften unterscheiden, auch wenn es *möglich* wäre, dass es so etwas gibt. Den Kern des Prinzips kann man also wie folgt zusammenfassen: Es kann keinen Unterschied auf der mentalen Ebene geben, ohne dass es einen Unterschied auf der physischen (neuronalen) Ebene gibt.

Eine weitere Ausformulierung und Ausdifferenzierung des Supervenienzprinzips hat vor allem Jaegwon Kim (1993, 1995, 1998, 2005) vorgenommen. Seine ursprüngliche, dann aber doch schnell verworfene Inten-

tion war, auf der Grundlage einer psychophysischen Supervenienzbeziehung eine *nicht-reduktive* Version des Physikalismus zu begründen. Der Gedanke ist, dass unter den Voraussetzungen der Supervenienz einerseits der physikalischen Domäne und ihrer Gesetze ein klarer (ontologischer) Vorrang zugesprochen werden kann, ohne dass andererseits hierdurch gleichzeitig ein physikalischer Reduktionismus impliziert würde, der dem Mentalen seine Realität abspricht. Ein weiterer Vorteil liegt in dem Umstand, dass die multiple Realisierbarkeit mentaler Eigenschaften (vgl. Kap. 6.1.4.2) ermöglicht wird, indem es potenziell verschiedene (physikalische) Grundlagen bzw. Supervenienzbasen für supervenierende Eigenschaften geben kann.

Auch Kim (1998) bringt zum Ausdruck, dass das Verhältnis, was man durch die Supervenienzbeziehung ausdrücken möchte, ein Determinations- oder Abhängigkeitsverhältnis sein soll. Eine solche Beziehung muss offenbar *asymmetrischer* Natur sein: „The relation of dependence, or determination, is asymmetric: if x depends on, or is determined by y , it cannot be that y in turn depends on or is determined by x ” (S. 11). Es scheint eben keine sinnvolle Annahme zu sein, dass mein Wunsch nach einem Schokoladeneis durch mein Gehirn determiniert wurde und gleichzeitig anzunehmen, dass mein Wunsch in diesem Fall mein Gehirn determiniert hat. Damit stellt sich allerdings die Frage, ob solch eine einseitige Abhängigkeit durch die Formulierung dieses Supervenienzprinzips wirklich abgesichert wird. Auch wenn die Intention bei der Formulierung der Supervenienzrelation war, eine asymmetrische Beziehung herauszustellen, wird eine symmetrische Beziehung der herangezogenen Variablen zumindest nicht *prinzipiell* ausgeschlossen. Denn eine Supervenienz von A auf B schließt nicht automatisch aus, dass es auch eine Supervenienz von B auf A gibt. Wir haben hier also letzten Endes nur eine bestimmte Form einer *Kovariation* zweier Variablen gesichert,

die *per se* noch keinerlei faktische Abhängigkeit oder Determination beinhaltet. Die Annahme einer solchen symmetrischen Beziehung kann die Zusammenhänge zwischen dem Mentalen und seinen physischen Grundlagen kaum verständlich machen und bleibt weitestgehend inhaltsleer. Um die gewollte einseitige Abhängigkeit abzusichern, müsste eine symmetrische Beziehung daher genau genommen per Zusatzannahme ausgeschlossen werden. Aber selbst in dem Fall bleibt gänzlich unklar, *warum* eine solche Abhängigkeit der mentalen Ebene von der neuronalen Ebene besteht bzw. *warum* die mentalen Makroeigenschaften durch die neuronalen Mikroeigenschaften determiniert sein sollen. Auch Kim räumt ein, dass die durch die Supervenienz herausgestellte Beziehung zwischen dem Mentalen und seinen physischen Grundlagen rein *deskriptiver* Natur ist. Man kann sagen, dass sie nur einen oberflächlich wahrnehmbaren Zusammenhang zwischen zwei Eigenschaftsklassen beschreibt, ohne eine Erklärung hierfür zu geben. Die womöglich auf einer tieferen Ebene liegende Ursache der Kovariation (oder Determination) der Elemente dieser Klassen bleibt unberührt: „...questions still remain: Is *M* reducible to *P* in some appropriate sense? Can we explain why something has *M* in terms of its having *P*?” (ebd., 1998, S. 18).

Er weist allerdings darauf hin und dies hat für den hier verfolgten Ansatz eine große Relevanz, dass man dem Supervenienzprinzip mehr abgewinnen kann, wenn man eine Verbindung zwischen ihm und der Vorstellung einer *hierarchischen Ordnung* der Welt herstellt:

What has replaced this picture of a dichotomized world [of dualism; M.L.] is the familiar multilayered model that views the world as stratified into different ‘levels’, ‘orders’, or ‘tiers’ organized in a hierarchical structure. The bottom level is usually thought to consist of elementary particles, or whatever our best physics is going tell us are the basic bits of matter out of which all material things are composed. As we go up the ladder, we successively encounter atoms, molecules, cells, larger living

organisms, and so on. The ordering relation that generates the hierarchical structure is the mereological (part-whole) relation. (Kim, 1998, S. 15)

Die Vorstellung einer hierarchischen Ordnung der Wirklichkeit ist für Kim also eindeutig kompositorisch begründet (vgl. Kap. 5.1.2). Entitäten, die zu einer bestimmten Ebene gehören, können in Entitäten fundamentalerer Schichten zerlegt werden. Nur die Entitäten auf der Basisebene (wenn es eine solche denn gibt) können selbst nicht mehr in noch kleinere Teile zerlegt werden. Wenn man nun dieses Modell natürlicher Organisationsebenen mit der Supervenienzthese verbindet, kommt man Kim zufolge zu einer *mereologischen Supervenienz*, die besagt, dass Eigenschaften des Ganzen durch Eigenschaften und Beziehungen seiner Teile bestimmt werden²²². So wird angenommen, dass zwei auf einer bestimmten Ebene befindliche Größen ununterscheidbar sein müssen, wenn die Eigenschaften (und die Beziehungen zwischen diesen), in die sie sich auf einer fundamentaleren Ebene zergliedern lassen, ununterscheidbar sind. Kim (1998) verwendet in diesem Zusammenhang den Ausdruck „microindiscernible“ (S. 17). Die Idee ist, dass sich Gegebenheiten auf einer höherstufigen Ebene nicht unterscheiden können, wenn sie „mikrounterscheidbar“ sind, d.h. auf niederstufigen Ebenen die gleichen Muster an Eigenschaften und Beziehungen zwischen diesen aufweisen müssen. So ist es nicht möglich, dass (eine Person) x die mentale Eigenschaft M aufweist und (eine Person) y , die auf den Mikroebenen ununterscheidbar von x ist, nicht über die mentale Eigenschaft M verfügt.

²²² Genau genommen bezieht sich Kim hierbei allerdings nach den in dieser Arbeit getroffenen Unterscheidungen nicht auf mereologische, sondern *aggregative* Prinzipien, die suggerieren, dass sich die auf einer höheren Ebene manifestierenden *Eigenschaften* eines Ganzen auf die auf einer niedrigeren Ebene vorzufindenden *Eigenschaften* seiner Teile zurückführen bzw. zumindest in Zusammenhang bringen lassen. Psychische Eigenschaften wären demnach eine Art Agglomerat niederstufiger (neuronaler) Eigenschaften.

Wenn Peter und Paul zu einem gegebenen Zeitpunkt in all ihren Mikroeigenschaften übereinstimmen, kann es nicht sein, dass Peter zu diesem Zeitpunkt Schmerzen empfindet und Paul nicht. Man könnte auch sagen, dass sich eine Makroeigenschaft eines Systems (Schmerz), die auf dessen Mikroeigenschaften basiert, nicht ändern kann, ohne dass auch eine Variation der Mikroeigenschaften des Systems stattgefunden hat. So lässt sich die mereologische Supervenienz formal folgendermaßen definieren: Wenn x eine mentale Eigenschaft M aufweist, dann gilt für jedes y , das auf einer (bzw. allen) Mikroebene(n) ununterscheidbar ist von x , dass es ebenfalls M aufweisen muss.

Dennoch bleibt es trotz dieser sinnvollen Erweiterung dabei, dass hiermit keine *Erklärung* mentaler Zustände abgegeben wird, so wie man in der Physik höherstufige Eigenschaften durch niederstufige erklärt, was auch Kim (1998) einräumt: „But supervenience or determination is one thing, explanation quite another. We may know that B determines A [...] without having any idea why this is so – why A should arise from B, not C, or why A, rather than D, arises from B“ (S. 18). Somit kann konstatiert werden, dass die Körper-Geist-Supervenienz für sich genommen keine eigenständige Leib-Seele-Theorie ist. Sie liefert keine *Erklärung* der von ihr herausgestellten Korrelationen bzw. Abhängigkeitsverhältnisse. Es ergibt sich aus diesem Ansatz selbst also keine Antwort auf die Frage, *warum* das Mentale über dem Physischen superveniert. Auch wird in Folge dessen keine Aussage über den ontologischen Status von mentalen Eigenschaften getroffen. Diese treten zwar nur in Abhängigkeit von physischen Eigenschaften auf, aber heißt das nun, dass sie selbst physischer Natur sind oder dass sie etwas sind, das über physische Eigenschaften selbst hinausgeht? Die Supervenienzthese schweigt hierzu. Aufgrund dieses explanatorischen Defizits ergibt sich die Möglichkeit, die Lücke mittels verschiedener Leib-Seele-Theorien zu schließen (vgl. Hüttemann, 2004,

S. 31). So ist eine bloße psychophysische Supervenienz neben dem Physikalismus bzw. einer psychophysischen Identitätsannahme ebenso vereinbar mit Positionen wie dem noch zu behandelnden Emergentismus oder auch dem Epiphänomenalismus. „If mind-body supervenience is a commitment of each of these diverse, and conflicting, approaches to the mind-body-problems, it cannot itself be a position on this problem that can be set alongside these classic alternatives. What this shows is that the mere claim of mind-body supervenience leaves unaddressed the question what *grounds* or *accounts* for it” (Kim, 1998, S. 12f; Hervorh. im Orig.). Kim insistiert nichtsdestotrotz darauf, dass man Supervenienz als einen gemeinsamen Nenner aller grundsätzlich *physikalistischen* Positionen ansehen könne, als eine Art Trennlinie zwischen Physikalismus und Substanzdualismus à la Descartes (vgl. ebd., S. 14f). Ein so definierter minimaler Physikalismus umfasst dann auch den Emergentismus und den Epiphänomenalismus²²³, was allerdings eine streitbare Ausweitung des Physikalismus-Begriffs sein dürfte.

6.3.2. Problematische Konsequenzen der Supervenienzkonzeption

Mit der Supervenienzrelation sind aber sogar Probleme verbunden, bei denen es nicht nur um Aspekte der fehlenden explanatorischen Tiefe dieses Konzeptes geht, sondern um weitgehende Konsequenzen für die

²²³ Wenn man dem Diktum Alexanders (1920) folgt, kann letztendlich nur von dem, was kausale Wirkmächtigkeit besitzt, gesagt werden, dass es existiert. Der Epiphänomenalismus mit seinen – um Feigl's Ausdruck zu gebrauchen – *nomological danglers* würde in dem Fall keinen eigenständigen Bereich des Mentalen postulieren und könnte daher auch als physikalistische Position betrachtet werden. Schwieriger sieht es mit dem Emergentismus aus. Auch seine Vertreter nehmen zwar eine starke Abhängigkeit des Mentalen von seiner physischen Realisierungsgrundlage an, jedoch werden mentale Eigenschaften als etwas angesehen, was in seinem Auftreten und seiner Qualität prinzipiell nicht aufgrund der korrespondierenden physischen Eigenschaften vorhergesagt werden kann.

Kausalität auf den höheren Ebenen bzw. die Möglichkeit mentaler Verursachung. Ein Problem ist, dass es unter den Voraussetzungen der Supervenienzrelation offenbar zu einem Ausschluss von Kausalität auf den höheren Ebenen und damit auch auf der mentalen Ebene kommt (*causal exclusion*).

Angenommen ein mentales Ereignis m^* ist die Ursache eines nachfolgenden mentalen Ereignisses m (ich erkenne in der Menschenmenge ein bekanntes Gesicht [m^*] und freue mich [m]) oder ein mentales Ereignis m^* ist die Ursache eines physischen Ereignisses p (ich erkenne in der Menschenmenge einen alten Freund [m] und strecke ihm meine Hand zur Begrüßung entgegen [p]). Gemäß der Supervenienzrelation superveniert jede mentale Eigenschaft über einer physischen Basiseigenschaft. Auf der Seite der Ursachen in der oben skizzierten Situation wird also die mentale Eigenschaft m^* durch eine physische Eigenschaft p^* realisiert, auf der Seite der Folgen wird die mentale Eigenschaft m durch eine physische Eigenschaft p realisiert. Das Problem besteht nun darin, dass man gemäß der Geschlossenheitsannahme der physischen Welt für jedes physische Ereignis p bereits eine hinreichende physische Ursache p^* ausfindig machen kann, womit ein kausaler Ausschluss oder Epiphänomenalismus des Mentalen droht: „Given that p has a physical cause p^* , what causal work is left for m to contribute? The physical cause therefore threatens to exclude, and preempt, the mental cause” (Kim, 1998, S. 37). Es ergibt sich die in Abbildung 24 aufgezeigte Konsequenz, dass das Auftreten von p bereits durch p^* verursacht wird, was ebenso für m gilt, das über p superveniert. Möchte man das nicht akzeptieren und nimmt an, dass m^* dann eben einen *zusätzlichen* Einfluss auf p (bzw. m) ausgeübt hat, führt auch das nicht weiter, da man sich in dem Fall eine problematische kausale *Überdetermination* einhandeln würde, da das Auftreten von p (bzw. m) bereits *hinreichend* durch p^* kausal determiniert ist. Auch eine

solche kausale Überdetermination ist mit dem Prinzip der kausalen Geschlossenheit der physischen Welt unvereinbar und damit grundsätzlich inakzeptabel.

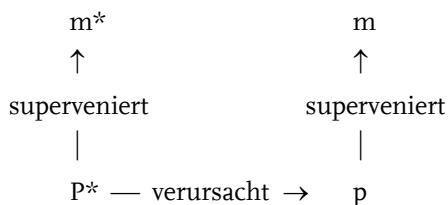


Abb. 24: Das kausale Ausschlussprinzip (in Anlehnung an Kim, 2003, S. 159)

Dieser Eindruck wird dadurch noch verstärkt, dass unter den Voraussetzungen der Supervenienzrelation nicht nur die Gefahr eines kausalen Ausschlusses (des Mentalen oder anderer höherstufiger Eigenschaften) droht, sondern offenbar sogar die eines allgemeinen Abdriftens der Kausalkräfte zu den immer niedrigeren Ebenen (*causal drainage*; Block, 1990, 2003). Denn wenn man die Supervenienz psychischer Eigenschaften über neuronalen Eigenschaften so versteht, dass die eigentliche Kausalität auf der neuronalen Ebene stattfindet, scheint sich die Kausalität im Allgemeinen auf die Ebene der Supervenienzbasen zu verlagern (kausaler Drift). Wenn dem so ist, warum soll dann aber hier bereits das Ende des kausalen Drifts erreicht sein? Wenn die neuronalen Kausalkräfte über denen der molekularen und diese über denen der atomaren Ebene etc. supervenieren, dann muss doch offenbar auch die „eigentliche Kausalität“ auf der jeweils niedrigeren Ebene und damit letztendlich auf der untersten (mikrophysikalischen) Ebene zu finden sein. Wenn wir mentale Eigenschaften als kausal irrelevant betrachten, müssten wir dies demnach auch bei chemischen oder biologischen Eigenschaften machen und

alle Kausalität (den Entitäten oder Eigenschaften) der elementarsten Ebene der Mikrophysik zuschreiben. Was ist aber, wenn es gar keine solche elementare mikrophysikalistische Ebene gibt (wie sie heute vom so genannten „Standardmodell“ der 18 Elementarpartikel postuliert wird) und sich Materie bis ins Unendliche zerteilen lässt²²⁴? Sind Ursache-Wirkungs-Zusammenhänge dann – vielleicht ganz im Sinne Humes – eine Illusion?

Allgemein scheint die Supervenienzrelation unserem allgemeinen Verständnis von Kausalität zuwider zu laufen. Um es noch einmal in Erinnerung zu rufen: Was Supervenienz (asymmetrisch verstanden) im Kern zu implizieren scheint, ist eine *Mikro-Makro-Determination*: Wenn es keine Veränderungen auf der Makroebene geben kann ohne zeitgleiche Veränderung auf der Mikroebene, dann gilt, dass die Makroeigenschaften eines Systems auf bestimmte Art und Weise durch die synchrone (also zeitgleich vorhandene) Mikrostruktur eines Systems determiniert werden. Man kann in diesem Zusammenhang auch von *vertikaler* (synchroner) Determination sprechen, die der herkömmlichen *horizontalen* (diachronen) Kausalität in gewisser Hinsicht entgegensteht. Denn welchen Stellenwert kann die horizontale Kausalität überhaupt noch haben, wenn es gemäß der Mikro-Makro-Determination bereits eine Festlegung der Makroeigenschaften des Systems durch dessen Mikrokonfiguration gibt? Wenn z.B. die Eigenschaft eines Goldklumpens, von goldgelber Farbe zu sein, bereits durch dessen Mikrostruktur zu einem Zeitpunkt t festgelegt ist, welchen Einfluss soll dann noch das haben, was *vor* dem Zeitpunkt t passiert ist (vgl. Kim, 2005, S. 37)? Verflüchtigt sich also im Zuge der Akzeptanz der Supervenienzthese unsere herkömmliche, horizontal verlaufende Kausalität etwa gänzlich aus der Welt?

²²⁴ Block (2003) beruft sich hier auf den Nobelpreisträger Hans Dehmelt (1989).

Kim meint, dass trotz dieser gravierenden Folgen keine generelle Zurückweisung des Supervenienz-Argumentes folgen muss, sondern dass dies nur bedeute, dass man sich von dem Vorhaben, mittels der Supervenienzrelation einen *nicht-reduktiven* Physikalismus begründen zu wollen, verabschieden müsse. Eine zentrale Bedingung für das Aufkommen der zuvor vorgestellten Probleme sei nämlich unsere Intuition oder auch unser Wunsch, dass alle Verursachung auf den höheren Ebenen irreduzibel gegenüber der physikalischen Basisverursachung sein müsse. Um den Wirklichkeitsstatus höherstufiger Entitäten und den Eigenständigkeitsstatus höherstufiger Wissenschaften wie der Psychologie abzusichern, erscheint es vielen als notwendig anzunehmen, dass die kausalen Beziehungen auf den höheren Ebenen etwas sind, was zur unterliegenden physikalischen Verursachung *hinzukommt* (vgl. ebd., 2005, S. 54). Bei genauerer Betrachtung dürfte aber genau das Gegenteil der Fall sein – nur eine Identitätsannahme zwischen höherstufiger und niederstufiger Kausalität bzw. die Akzeptanz eines reduktiven Verhältnisses zwischen diesen Größen scheint den kausalen Ausschluss der höheren Ebenen oder das Abdriften von Kausalität verhindern zu können. Denn wenn man sich als Physikalist nicht mit den in Abbildung 24 dargestellten Verhältnissen abfinden möchte, dann müsse man sich nach Ansicht Kims notwendigerweise mit der in Abbildung 25 veranschaulichten Lage arrangieren.

Eine solche Reduktion bedeute nun aber nicht, dass die höherstufige Kausalität eine Illusion darstellt, sie sei nur nichts, was zu der basaleren physikalischen Kausalität *hinzukommt*: „Indeed, the supervenience argument may be generalizable, but all that would show is that if there is biological causation, biological properties are, or are reducible to, physical or physico-chemical properties; it does not show that biological causation does not exist” (S. 54f).

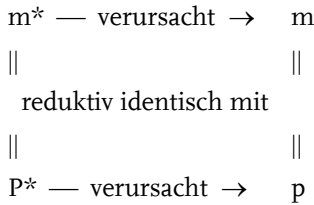


Abb. 25: Identitätsbasierte Reduktion (in Anlehnung an Kim, 2003, S. 166)

Eine weitere Ausbuchstabierung dessen, was mit Reduktion und Identität hier gemeint ist (und was nicht), nimmt Kim im Zusammenhang der Konzeption einer Realisierungstheorie vor (vgl. Kap. 6.5), die er als die notwendige explanatorische Erweiterung der Supervenienzrelation auffasst. Fakt ist und bleibt allerdings, dass die Supervenienzrelation selbst keine Theorie des Geistes begründen kann, die das Verhältnis zwischen der mentalen und neuronalen Ebene verständlich macht.

Die „Rettung“ des Supervenienzprinzips über eine solche identitätsbasierte Reduktion lässt das Konzept der Supervenienz selbst – über die bloße Beschreibungsfunktion einer Mikro-Makro-Abhängigkeit hinaus – überflüssig erscheinen. Denn es stellt sich unweigerlich die Frage, welche Bedeutung es überhaupt haben kann, wenn es um identitätstheoretische oder reduktive Annahmen erweitert werden muss, um selbst ein explanatorisch gehaltvolles und inhaltlich kohärentes Konzept zu sein. Wenn es zwangsläufig erforderlich ist, die Ursache einer supervenienten Beziehung zweier Eigenschaftsklassen zu bestimmen, die z.B. in einer Identitätsbeziehung liegt, verliert die Betrachtung dieser Eigenschaftsklassen unter Rückgriff auf das Supervenienzprinzip selbst offensichtlich jeden Anreiz, da die Identitätsrelation in dem Fall die *exaktere* Beschreibung der Beziehung der betrachteten Variablen liefert.

6.4. Emergenz und das unerklärliche Auftauchen des Geistes

Das Konzept der Emergenz hat in der Philosophie eine weit zurückreichende Geschichte. Bereits in der Antike finden sich seine Ursprünge. So heißt es z.B. im siebten Buch der *Metaphysik* bei Aristoteles (1995): „Dasjenige, was so zusammengesetzt ist, daß das Ganze eines ist, nicht wie ein Haufen, sondern wie die Silbe, *ist nicht nur seine Elemente*. Die Silbe nämlich ist nicht einerlei mit ihren Elementen (Buchstaben), das *ba* nicht einerlei mit *b* und *a* [...]. Also ist die Silbe etwas außer diesen, nicht bloß nämlich die Sprachelemente, Vokale und Konsonanten, sondern auch noch etwas anderes“ (S. 168/1041b; Hervorh.: M.L). Hieraus ist später die populäre (aber streitbare) These „Das Ganze ist mehr als die Summe seiner Teile“ abgeleitet worden, die heute vor allem mit der Gestaltpsychologie in Verbindung gebracht wird.

Im Kern bedeutet Emergenz, dass es aufgrund der Aggregation und Organisation von (niedrigstufigen) Elementen eines Systems zu einer spontanen Ausprägung neuer (höherstufiger) Eigenschaften oder Strukturen dieses Systems kommt. Diese als *emergent* bezeichneten Eigenschaften lassen sich dabei nicht auf die Eigenschaften der *einzelnen* Bestandteile des Systems zurückführen. Im Gegensatz zur „klassischen“ Identitätstheorie spielt die Beziehung zwischen dem Ganzen und seinen Teilen im Rahmen des emergenztheoretischen Ansatzes also eine konstitutive Rolle.

Im Zusammenhang mit der Erklärung von Bewusstsein wurde der Begriff Emergenz zum ersten Mal explizit von George Henry Lewes (1875) verwendet und hat daraufhin in der ersten Hälfte des 20. Jahrhunderts vor allem in der philosophischen Debatte um die Erklärung des

Aufkommens von Leben in der Welt eine Hochzeit erlebt. Weil es vor allem britische Philosophen (wie Samuel Alexander, Conwy Lloyd Morgan und Charlie Dunbar Broad) waren, die den Emergenzbegriff ins Feld führten, hat sich die Bezeichnung *Britischer Emergentismus* eingebürgert. Die Annahme emergenter Eigenschaften sollte die in der damaligen Diskussion konkurrierenden mechanistischen²²⁵ und vitalistischen Ansätze miteinander versöhnen. Während erstere annehmen, dass sich die charakteristischen Eigenschaften von Lebewesen auf die Eigenschaften ihrer Organe zurückführen lassen, wie das Verhalten einer Uhr auf das Zusammenspiel der Zahnräder, Gewichte etc. zurückgeführt werden kann, nehmen letztere an, dass es für die Erklärung von Leben einer besonderen Substanz – einer Entelechie oder schöpferischen Lebenskraft – bedürfe (vgl. Beckermann, 1999, S. 216). Der Grundgedanke der Emergenz sagt dagegen, dass es trotz der Akzeptanz eines Monismus höherstufige Eigenschaften (wie die des Lebens oder eben Bewusstseins) gibt, die einen *neuartigen* Charakter haben und vor ihrem ersten Auftreten *prinzipiell unvorhersehbar* waren.

Nach ihrem Aufblühen verschwand die Emergenztheorie daraufhin allerdings ziemlich schnell aus dem Fokus des Interesses, hat dann aber in den letzten 30 Jahren im Zuge der so genannten Qualia- bzw. Bewusstseinsdebatte in der Philosophie des Geistes ein gewaltiges „Comeback“ erfahren. Vielfach wird nun behauptet, dass das Psychische (insbesondere phänomenale Eigenschaften) am besten dadurch zu charakterisieren

²²⁵ Trotz einer sicherlich großen Schnittmenge ist der damalige mechanistische Ansatz nicht deckungsgleich mit dem heutigen Ansatz mechanistischer Erklärungen, wie ihn Bechtel oder Craver vertreten (vgl. Kap. 5.1.2.3). Vor allem halten sich die heutigen Vertreter dieser Denkfigur bzgl. ontologischer Deutungen mechanistisch beschreibbarer Systeme viel stärker zurück als ihre damaligen Gesinnungsgenossen. Während die damaligen Mechanisten eher reduktionistisch orientiert waren, verfolgen die meisten heutigen Protagonisten des Ansatzes kaum reduktionistische Strategien (vgl. z.B. Craver, 2007, S. 107ff).

ist, dass es *emergent* ist. So könne vermeintlich durch das Prinzip der *Emergenz* eine Körper-Geist-Theorie etabliert werden, die einerseits nicht-reduktionistisch ist und trotzdem eine gehaltvolle Charakterisierung der Beziehung zwischen mentalen und neuronalen Eigenschaften abliefert. Auch gerade viele Psychologen äußern sich dahingehend, dass man, wenn es um die Klärung des Verhältnisses von Gehirn und Geist, Neurobiologie und Psychologie geht, sich auf das Prinzip der Emergenz stützen kann. Das Psychische sei demnach zwar auf eine bestimmte Art und Weise abhängig von seinen physischen Grundlagen, reduzierbar sei es aufgrund seines emergenten Charakters auf diese aber nicht. Eine besondere Attraktivität scheint insbesondere die zumeist vertretene Annahme zu haben, dass emergente Eigenschaften eigene Kausalkräfte besitzen, die eine abwärts gerichtete Kausalität ermöglichen (sodass mentale Eigenschaften physische Eigenschaften beeinflussen können).

6.4.1. Konsistente und inkonsistente Emergenzkonzepte

Zunächst einmal muss genauer geklärt werden, was man sinnvoll unter Emergenz verstehen kann und was nicht, denn es handelt sich hierbei um einen häufig verwendeten und dennoch recht unscharf gebrauchten Begriff. So äußert sich insbesondere Kim (2006) diesbezüglich sehr kritisch:

The term ‚emergence‘ seems to have a special appeal for many people; it has an uplifting, expansive ring to it, unlike ‚reduction‘ which sounds constrictive and overbearing. We now see the term being freely bandied about, especially by some scientists and science writers, with little visible regard for whether its use is underpinned by a consistent, tolerably unified, and shared meaning (and if so what it is). This has created situations in which those discussing emergence, even face to face, more often than not talk past each other. (S. 547f)

Wenn man sich nun die Verwendung des Emergenzkonzeptes genauer anschaut, kann man im Kern²²⁶ zwischen einer moderaten und starken bzw. radikalen Form der Emergenz unterscheiden (vgl. Van Gulick, 2001; Kim, 2006; Stephan, 2006). Gemeinsam ist allen Varianten, dass sie grundsätzlich einen (physischen) Monismus akzeptieren bzw. Substanzdualismus oder -pluralismus ablehnen.

6.4.1.1. Emergente Systemeigenschaften oder „moderate“ Emergenz

Wie bereits thematisiert, sind die Emergenztheoretiker bereits frühzeitig explizit von Ebenen- oder Schichtenmodellen der Wirklichkeit ausgegangen. Wie Metzinger (2007d) sagt, kann man den „ontologischen Kerngedanken der Emergenztheorie auch als ‚Schichtenrealismus‘ bezeichnen“ (S. 275). Die Emergenztheoretiker haben auf der Grundlage der Annahme natürlicher Organisationsebenen mereologischer Prägung (vgl. Kap. 5.1.2.1) einen „systemtheoretisch inspirierten“ Ansatz formuliert:

It is clear that, on the constructive philosophy of emergent evolution which I seek to develop, there are levels or orders of reality in respect both of intrinsic and of extrinsic relatedness. This does not, of course, imply a scale of more or less reality, as such, for relatedness as a mark of reality obtains at all levels. It does, however, imply (1) that there is increasing complexity in integral systems as new kinds of relatedness are successively supervenient; (2) that reality is, in this sense, in process of development; (3) that there is an ascending scale of what we may speak of as richness in reality; and (4) that the richest reality that we know lies at the apex of the pyramid of emergent evolution up to date. (Morgan, 1923, S. 203)

²²⁶ Nach Van Gulick (2001) gibt es auch noch eine weitere Form, *die Emergenz für spezielle Ausprägungen* eines Merkmals („specific value emergence“). Gemeint ist damit, dass wenn der ganze Gegenstand z.B. ein Gewicht von einem Kilogramm hat, dann ist diese Ausprägung des Gewichtes etwas, was keinem seiner Teile zukommt, obwohl jeder Teil durchaus ein Gewicht hat. Da es sich hierbei offenkundig um einen recht trivialen Fall von Emergenz handelt, werde ich ihn nicht weiter berücksichtigen.

Wie im Zusammenhang der Besprechung von Ebenenmodellen bereits erwähnt wurde, postuliert Morgan fünf Ebenen, die die gesamte natürliche Ordnung aller Dinge umfassen sollen (vgl. Abb. 26).

Menschen
Tiere
Pflanzen
Moleküle
Atome

Abb. 26: Hierarchie natürlicher Ordnungsebenen nach Morgan (1923)

Des Weiteren gehen die klassischen Emergenztheoretiker davon aus, dass alle Gegenstände oder Systeme, aus denen unsere Welt besteht, eine hierarchische Gliederung mit einer ansteigenden Komplexität aufweisen. Die so entstehenden hierarchisch strukturierten Systeme weisen Emergenz als kennzeichnende Eigenschaft auf.

Das Besondere hieran ist nun, dass auf den Makroebenen solcher Systeme Eigenschaften hervortreten, die auf den Mikroebenen desselben Systems nicht vorhanden sind. Genauer gesagt weist das Ganze Merkmale auf, die verschieden sind von denjenigen, die die *einzelnen* Teile des Ganzen aufweisen. So weist Wasser die Eigenschaften auf, flüssig und durchsichtig zu sein, obwohl diese Eigenschaften nicht den Molekülen, Atomen oder Elementarteilchen zukommen, aus denen Wasser besteht. Ebenso sind Farbempfindungen, räumliche Orientierungsleistungen Beispiele für höherstufige psychische Eigenschaften, die nur ein System als Ganzes (z.B. ein Mensch) besitzt und die keinem seiner Bestandteile zukommen. Stephan (2006) spricht in diesem Zusammenhang von der

These der *systemischen Eigenschaften*, die für alle Spielarten von Emergenz kennzeichnend ist. Sie besagt, dass emergente Eigenschaften systemische bzw. kollektive Eigenschaften in dem Sinne sind, dass eine „Eigenschaft [...] *genau dann systemisch [ist], wenn ein System als Ganzes sie hat, aber wenn kein Bestandteil des Systems eine Eigenschaft dieses Typs hat*“ (S. 148; Hervorh. im Orig.).

Eine weitere zentrale Bedeutungsfacette ist die „*synchrone Determiniertheit*“ (ebd., S. 149), die das Verhältnis zwischen der Mikrostruktur eines Systems und dessen emergenten Eigenschaften näher bestimmt. Sie besagt: „*Die Eigenschaften und Verhaltensdispositionen eines Systems hängen nomologisch von dessen Mikrostruktur, das heißt den Eigenschaften seiner Bestandteile und deren Anordnung ab. Es kann keinen Unterschied in den systemischen Eigenschaften geben, ohne dass es zugleich einen Unterschied in den Eigenschaften der Systembestandteile gibt*“ (ebd.; Hervorh. im Orig.).

Eine besondere Spielart der schwachen Emergenz ist der *schwache diachrone Emergentismus*²²⁷, der die vorherigen Begriffskonstituenten um die *These der Neuartigkeit* erweitert: „*Im Universum kommt es immer wieder zur Entstehung von genuin Neuartigem. Bereits bestehende Entitäten fügen sich zu neuen Konstellationen; sie bilden neue Strukturen aus, die neue Systeme mit neuen Eigenschaften und Verhaltensdispositionen konstituieren*“ (ebd. S. 150; Hervorh. im Orig.). Wichtig ist, dass durch das bloße Hinzufügen der Neuartigkeitsthese keine Irreduzibilität höherstufiger

²²⁷ Der schwache diachrone Emergentismus lässt sich nach Stephan (2006) dadurch verstärken, dass man die Annahme einer *Struktur-Unvorhersehbarkeit* hinzufügt. Diese bedeutet, dass das Aufkommen von Strukturen und der mit ihnen verbundenen Eigenschaften unter chaostheoretischen Voraussetzungen prinzipiell unvorhersehbar und damit emergent ist. Da diese Form für die weitere Arbeit nicht relevant ist, soll auf sie nicht weiter eingegangen werden.

Eigenschaften angenommen wird. Für diese wäre eine *prinzipielle* Unvorhersehbarkeit vonnöten.

In der heutigen Terminologie wird die – bereits von den Vertretern des klassischen Emergentismus propagierte – Emergenz systemischer Eigenschaften auch als *moderate* oder *schwache* Form der Emergenz bezeichnet. „Moderat“ ist diese Deutung des Begriffes insofern, als dass sie keine weitergehenden Unbestimmtheitsannahmen zwischen dem Ganzen und seinen Teilen hinzufügt. Wenn man allerdings die Frage aufwirft, ob es möglich ist, auf der Grundlage dieses Ansatzes eine gehaltvolle Gehirn-Geist-Theorie zu entwickeln, ist der alleinige Bezug auf die moderate Form der Emergenz nicht sehr hilfreich (vgl. Stephan, 2006). Wenn es z.B. heißt, dass bestimmte psychische Eigenschaften im schwachen Sinne emergent sind, dann charakterisiert das die Beziehung dieser Eigenschaften zu ihren physischen Grundlagen nicht wirklich nennenswert, da schwache Emergenz gewissermaßen ein „Allerweltsphänomen“ darstellt:

Die Welt ist voller Eigenschaften, die nur auf der Systemebene und nicht auf der Ebene der Systemkomponenten auftreten. Auch die Härte eines Diamanten ist in diesem Sinne eine emergente Eigenschaft, da sie nur dem Gesamtsystem und nicht den Kohlenstoffatomen zukommt, die in etwas anderer Anordnung den sehr viel weicheren Graphit instantiiieren. Der schwache Emergentismus ‚schneidet‘ die Natur zwar an ihren ‚Scharnieren‘, doch sind diese zu zahlreich. (ebd., S. 150)

Die Frage ist daher, inwieweit stärkere Formen der Emergenz in dieser Hinsicht Abhilfe schaffen können.

6.4.1.2. Radikale Emergenz

Die *radikale* Form der Emergenz behauptet zunächst auch, dass eine emergente Eigenschaft dadurch gekennzeichnet ist, dass sie – als Makro-

eigenschaft des Ganzen – verschieden ist von den Mikroeigenschaften, die die Teile des Ganzen aufweisen (These der systemischen Eigenschaften). Darüber hinaus wird aber auch noch angenommen, dass eine emergente Eigenschaft, „whose nature and existence is not necessitated by the features of its parts, their mode of combination and the law-like regularities governing the features of its parts” (Van Gulick, 2001, S. 17). Eine solche Form der Emergenz nimmt also gegenüber der traditionellen Bestimmung zusätzlich an, dass emergente Eigenschaften des Ganzen nicht durch die Eigenschaften und Beziehungen der Teile determiniert werden. Wie Kim (2006) zu Recht herausstellt, wird hiermit die Supervenienzbeziehung ausgeschlossen, sodass nicht mehr gesagt werden kann, dass die emergente Eigenschaft über der Mikrostruktur/den Mikro-eigenschaften eines Objektes superveniert. Die Folge ist: „two wholes may have identical microstructure (i.e., composed of identical basic physical constituents configured in an identical structure) and yet differ in respect of their emergent properties” (ebd., S. 549). Unter diesen Voraussetzungen wäre es also möglich, dass es zwei auf der molekularen Ebene identische Individuen gibt, von denen eines die Makroeigenschaft hat, ein Lebewesen wie eine Maus zu sein, während das andere diese Eigenschaft nicht aufweist. Dies scheint nach allem, was die heutige Wissenschaft dazu sagen kann, eine völlig unannehmbare Konsequenz zu sein. Wenn in Bezug auf die Körper-Geist-Beziehung die Verbindung zwischen z.B. Schmerzen und seinen neuronalen Substraten tatsächlich so willkürlich und zufällig ist, dann lässt sich natürlich die Frage stellen, warum Schmerzen überhaupt aus bestimmten neuronalen Bedingungen „emergieren“ sollen und nicht aus anderen (vgl. ebd., S. 550). Eine solche radikale Emergenz kommt also nicht als eine respektable Körper-Geist-Theorie infrage, da sie letztendlich gar keinen echten Zusammenhang, z.B. in Form eines Abhängigkeits-, Determinations- oder

Identitätsverhältnisses, zulässt und daher keine inhaltlich konsistente psychophysische Theorie darstellt. Selbst eine Position wie der Epiphänomenalismus wäre streng genommen auf der Grundlage einer solchen radikalen Emergenz nicht einsichtig zu machen.

6.4.1.3. Starke (synchrone) Emergenz

Wenn ein moderater Emergentismus allein wenig aussagekräftig und ein radikaler Emergentismus empirisch unplausibel ist, inwieweit kann dann dem Emergenzbegriff überhaupt Sinn abgerungen werden? Ein Mittelweg zwischen diesen beiden Polen stellt der (starke) *synchrone* Emergentismus²²⁸ dar, der zugleich die heutige Kernbedeutung des Emergenzbegriffs ausmacht. Der entscheidende Unterschied zur radikalen Form des Emergentismus ist, dass wir im Gegensatz zu dieser eine *epistemische* (statt ontologische) Ausformulierung der Emergenzrelation vorliegen haben. Der synchrone Emergentismus zielt also nicht – wie bei der radikalen Variante – darauf ab, die synchrone Determiniertheit höherstufiger Eigenschaften abzulehnen, sondern darauf, dass eine solche Determiniertheit nicht *einsichtig* gemacht werden kann (obgleich sie *de facto* durchaus vorhanden sein mag). Die hier vertretene These ist also, dass höherstufige Eigenschaften (wie das Haben von psychischen Zuständen) nicht durch den Rekurs auf deren physische Grundlagen *reduktiv erklärt* werden können. Der starke Emergentismus akzeptiert die ontologischen Thesen des schwachen Emergenzbegriffs, sagt allerdings, dass emergente Eigenschaften solche sind, die (epistemisch) irreduzibel und damit *synchron emergent* sind (vgl. Stephan, 2006, S. 151). Um zu verstehen, was

²²⁸ Wenn im Folgenden allgemein von Emergentismus die Rede ist, bezieht sich dies ausschließlich auf den synchronen Emergentismus.

mit Irreduzibilität an dieser Stelle gemeint ist, dürfte es sinnvoll sein, sich die historischen Hintergründe für das Aufkommen des hier thematisierten Emergenzbegriffs stärker vor Augen zu führen; und zwar sind diese vor allem mit Charlie Dunbar Broad's (1925) Werk *The Mind and Its Place in Nature* verbunden.

Broad ging es in der bereits angesprochenen Vitalismusdebatte darum, zwischen mechanistischen Ansätzen und vitalistischen Ansätzen zu vermitteln. Die Emergenztheorie lehnt – genau wie der Mechanismus – eine solche „Lebens-Substanz“ ab. Ebenso teilt sie mit diesem die Annahme einer synchronen Determiniertheit. Den entscheidenden Unterschied zwischen Emergentismus und mechanistisch-reduktionistischer Theorie erläutert Broad (1925) folgendermaßen:

Abstrakt gesprochen behauptet die Emergenztheorie, daß es bestimmte komplexe Gegenstände gibt, die, sagen wir, aus den Komponenten A, B und C bestehen, die in der Relation R zueinander stehen; daß alle komplexen Gegenstände, die aus Komponenten der gleichen Art A, B und C bestehen, die zueinander in der gleichen Art von Relation R stehen, bestimmte charakteristische Eigenschaften besitzen; daß A, B und C in anderen Arten von komplexen Gegenständen vorkommen können, in denen die Relation nicht von der gleichen Art wie R ist; und daß die charakteristischen Eigenschaften des Ganzen R (A, B, C) nicht einmal im Prinzip aus der vollständigen Kenntnis der Eigenschaften abgeleitet werden können, die A, B und C isoliert oder in anderen komplexen Gegenständen²²⁹ haben, die nicht die Form R (A, B, C) besitzen. Der Mechanismus bestreitet den letzten Teil dieser Behauptung (S. 61; zit. n. Beckermann, 1999, S. 218f; Hervorh.: M.L.).

²²⁹ Die Formulierung „Eigenschaften, isoliert oder in anderen Anordnungen“ wird von Broad verwendet, damit man bei der geforderten Ableitung nicht auf so genannte *Ad hoc-Eigenschaften* zurückgreifen kann (z.B. dass bestimmte Komponenten in einer bestimmten Anordnung ein Ganzes bilden, das die fragliche Makroeigenschaft aufweist). Nach Broad dürfe man bei der Ableitung nur auf die *allgemeinen* Naturgesetze zurückgreifen, die für die Teile des untersuchten Systems gelten, und das unabhängig davon, in welcher speziellen Anordnung sich die Teile befinden (vgl. Beckermann, 1999, S. 221f).

Emergente Eigenschaften sind also dadurch charakterisiert, dass sie einerseits – wie auch mechanisch erklärbare Eigenschaften – eine synchrone Determiniertheit durch die Mikrostruktur des Systems aufweisen, andererseits aber *prinzipiell* nicht aus der vollständigen Kenntnis der Eigenschaften der Mikrostruktur des betroffenen Systems *abgeleitet* werden können. Übertragen auf die gegenwärtige Debatte um den Status des Psychischen bedeutet dies, dass das Auftauchen bestimmter mentaler Eigenschaften selbst auf der Basis einer perfekten Kenntnis aller Eigenschaften ihrer neuronalen Grundlagen nicht *vorhersehbar* oder *erklärlich* wäre. „Emergentists will say that the fact that pain emerges from a certain neural state (say, C-fiber excitation) is a brute fact that cannot be explained, and that the full and ideally complete knowledge of the neurophysiology of the brain does not suffice for prediction of conscious states. Thus, determination is one thing; explainability and predictability are quite another” (Kim, 2006, S. 551).

Das, was Broad als Möglichkeit zur „Ableitung“ bestimmt hat, macht in der heutigen Diskussion die Möglichkeit einer *reduktiven Erklärbarkeit* aus. Wie Kim (2006) aufzeigt, handelt es sich hierbei nicht um die Möglichkeit zur Formulierung von Brückenprinzipien, wie sie die Theorie-reduktion von Nagel (1961) zulässt (vgl. Kap. 8.2.1 dieser Arbeit). Die Nagelsche Reduktion ist zwar eine Reduktion, die – wie Broad – eine *Ableitung* höherstufiger Eigenschaften erfordert, indem über Brückenprinzipien als Hilfsprämissen eine Verbindung der höherstufigen (zu reduzierenden) Eigenschaften mit solchen der niederstufigen Theorie, auf die reduziert werden soll, erfolgt. Genau dies ist aber nun für reduktive Ableitungen, wie sie Broad versteht, unzulässig. Für die Emergentisten ist es gerade der zu erklärende Aspekt, *warum* solche besonderen Brückenprinzipien (z.B. die Verbindung von Schmerzen mit C-Faser-Reizungen) bestehen. Zusatzprämissen, die eine solche Verbindung *ad*

hoc herstellen, sind deshalb nicht akzeptabel. „Why does pain arise from, or correlate with, C-fiber excitation, but not another kind of neural state? Why does pain, not itch or tickle, correlate with C-fiber excitation? Nagelian reduction allows the use of psychoneural correlations as unexplained additional premises of reductive derivations, and this, by the emergentist's light, begs the question on hand” (Kim, 2006, S. 552).

Daher wird in diesem Zusammenhang in der gegenwärtigen Debatte vor allem die so genannte *funktionale Reduzierbarkeit* diskutiert. Wie im Zusammenhang mit der Realisierungstheorie (vgl. Kap. 6.5) aufgezeigt werden wird, ist eine Eigenschaft nach diesem Reduktionsmodell reduktiv erklärbar, wenn erstens über funktionale Beschreibungen ihre kausale Rolle erfasst werden kann und zweitens die neuronalen Realisierer hierzu ermittelt werden können (vgl. Levine, 1993; Kim, 1998, 2005, 2006). Was durch diese Prozedur gewährleistet werden soll, ist „den Übergang von der Komponenten-Ebene zur Systemebene zu ermöglichen. Schlägt dies fehl, so scheitert die angestrebte reduktive Erklärung“ (Stephan, 2006, S. 152). Damit ist zugleich der zentrale Sinngehalt der Irreduzibilitätsannahme der starken Emergenz freigelegt: Wenn gezeigt werden kann, dass solche reduktiven Erklärungen aus prinzipiellen Gründen scheitern, dann ist die systemische Eigenschaft *synchron emergent*.

6.4.2. Zu den Grenzen der Reichweite des Emergenzkonzeptes

Auch wenn die starke Emergenz ein inhaltlich kohärentes Konzept sein mag, muss kritisch angemerkt werden, dass die von Broad und den anderen britischen Emergentisten für emergent gehaltenen chemischen und biologischen Eigenschaften im Zuge der voranschreitenden naturwissenschaftlichen Forschung mittlerweile nicht mehr als synchron emergent bezeichnet werden können (vgl. McLaughlin, 1992). Mittlerweile scheint

der einzig ernsthaft diskutierte Fall einer möglichen synchronen Emergenz der des phänomenalen Bewusstseins, also phänomenaler Eigenschaften, zu sein (für kognitive Eigenschaften wird die reduktive Erklärbarkeit von einem überwiegenden Teil der Autoren angenommen).

Ebenso kann kritisiert werden, dass eine synchrone Emergenz, die durch die Annahme der *Irreduzibilität* gekennzeichnet ist, nichts Genuines oder Informatives im Hinblick auf die Beziehung zwischen Eigenschaften verschiedener Ebenen aussagt (vgl. Kim, 2006). Die Annahme einer Reduzierbarkeit ist eine – in epistemischer bzw. explanatorischer Hinsicht – *positive* Charakterisierung der Beziehung zwischen Eigenschaften, die einen klaren Informationsgehalt hat. „If we know that X is reducible to Y, we know something interesting and important about the relationship between X and Y. [...] I believe we can take reducibility as a genuine relation characterizing two domains of properties, or theories” (ebd., S. 556). Die Annahme einer Nicht-Reduzierbarkeit ist dagegen nur eine negative Charakterisierung. Sie sagt nur, wie die Beziehung zwischen Eigenschaften verschiedener Ebenen *nicht* ist, scheint aber keine positive Bestimmung dafür abzugeben, *wie* sie ist. „What we have in supervenience [d.h. synchrone Determiniertheit; M.L.] and irreducibility, therefore, are two essentially negative conditions, and they do not amount to a positive account of what emergence really is. They tell us what emergence is not; they do not tell us anything – at least, not much – about what it is” (ebd., S. 557).

Ein weiterer wichtiger Aspekt ist, dass der Emergentismus annimmt, dass emergente Eigenschaften eigene Kausalkräfte besitzen, die nicht auf die Kausalkräfte der Eigenschaften der Entitäten auf der Mikroebene zurückzuführen sind. Wenn er das nicht annehmen würde, wären emergente Eigenschaften vermeintlich nur Epiphänomene von einem fragwürdigen Realitätsstatus. Die Folge ist, dass der Emergentismus auch die Möglich-

keit zu einer so genannten *abwärts gerichteten Kausalität* (downward causation) einräumen muss, was ebenfalls Kim (2006) analytisch klar dargelegt hat:

Suppose then that an instance of an emergent property, *M*, causes another emergent property *M** to instantiate. This, we might say, is an instance of ‘same-level’ causation. Now, *M**, as an emergent, must have a basal (physical) property *P** from which it emerges; *M** cannot be instantiated unless some appropriate basal condition, say *P**, is present; moreover, the presence of *P** by itself guarantees that *M** will be instantiated at that time, *no matter what has preceded this occurrence of M**. That is, as long *P** is there at the time, *M** will be there at the same time *whether or not M*'s purported cause, M, had been there at all* – unless, that is, *M* had something to do with *P**'s presence at that time. In fact, the only way to save the claim that *M* causes *M** appears to be say that *M* caused *M** *by causing P**. [...] This is downward causation because *P** is property at the base level; if *M* and *M** are mental properties; *P** would be a physical/neural property. (S. 557f; Hervorh. im Orig.)

Es ist daher nicht verwunderlich, dass die Möglichkeit einer abwärts gerichteten Kausalität bei nahezu allen Vertretern des Emergentismus eingeräumt wird (vgl. Alexander, 1920; Morgan, 1923; Broad, 1925; siehe auch McLaughlin, 1992; Fazekas & Kertész, 2011). Das Problem ist dabei allerdings, dass man, wenn man im Zuge der obigen Analyse Kims eine abwärts gerichtete Kausalität einräumt, damit zwangsläufig eine *kausale Überdeterminiertheit* heraufbeschwört:

In our schematic example above, we concluded that *M* causes *M** by causing *P**. So *M* causes *P**. Now, *M*, as an emergent, must itself have an emergence base property, say *P*. [...] If causation is understood as nomologically (law-based) sufficiency, *P*, as *M*'s emergence base, is nomologically sufficient for it, and *M*, as *P**'s cause, is nomologically sufficient for *P**. It follows that *P* is nomologically sufficient for *P** and hence qualifies as its cause. (Kim, 2006, S. 558)

Wir landen also erneut bei einer kausalen Überdeterminiertheit, die in Bezug auf naturwissenschaftliche Grundprinzipien (kausale Geschlos-

senheit; Energieerhaltungssatz) nicht hinnehmbar ist. Damit scheint dann aber auch die gesamte Position des Emergentismus ins Wanken zu geraten: „If downward causation goes, so goes emergentism“ (ebd.). Wenn nämlich unter den Voraussetzungen des Emergentismus in allen Fällen von abwärts gerichteter Kausalität eine kausale Überdeterminiertheit vorliegt, können die Emergentisten ihre entscheidende Prämisse bzgl. einer eigenen Kausalität emergenter Eigenschaften nicht aufrecht halten.

Allerdings haben Campbell und Bickhard (2011) aufgezeigt, dass der von Kim betriebene Diskreditierung des Emergenzbegriffs keine Notwendigkeit zukommt. Im Gegenteil sei es selbst unter den Voraussetzungen Kims der Fall, dass weiterhin eine relevante Form von Emergenz impliziert würde. Zentral sei hierfür allerdings, dass man die Supervenienz- bzw. Determinationsbeziehung nicht als notwendigen Bestandteil von Emergenz betrachtet. Hiermit soll allerdings nicht die radikale Emergenz vertreten werden, die eine völlige Beliebigkeit in der Beziehung von Mikro- und Makroeigenschaften eines Systems annimmt. Was sie vorschlagen, ist, dass es keine vollständige, sondern nur eine *teilweise* Determination durch die Mikroebenen (bzw. ihre Entitäten oder Eigenschaften) gibt: „All that an advocate of emergence need accept is that *some* physical base is necessary, but not sufficient, for the emergent property“ (S. 44; Hervorh. im Orig.).

Campbell und Bickhard (2011) geben Kim recht, dass die Mikrostruktur eines Objekts dem Objekt als Ganzem zukommt und nicht den Konstituenten, ziehen daraus aber andere Schlussfolgerungen:

If causal powers of a stable system emerge at the macro level from the holistic organization of its micro-constituents, it is precisely that macro *organization* which gives the system its distinctive properties, properties that its micro-constituents (and their properties) do not have. In that case, those macro properties would be

genuinely novel – properties not possessed by its simpler constituents. Therefore, according to his starting point, they are emergent. (S. 37; Hervorh. im Orig.)

Dieser Rettungsversuch wirkt im Hinblick auf die Annahme einer nur teilweisen Mikro-Makro-Determination recht konstruiert, denn weitere Begründungen für eine solche Annahme bleiben die Autoren schuldig. Des Weiteren reicht die Form der Emergenz, die Campbell und Bickhard durch ihre Argumentation abzusichern anstreben, letztendlich nicht wesentlich über die allgemein hin unproblematische Annahme emergenter Systemeigenschaften hinaus, die auch Kim keineswegs infrage stellt. Zusammenfassend kann man sagen, dass die Emergenztheorie zunächst einen sehr wichtigen Gedanken liefert. In Form der „moderaten“ Emergenz verdeutlicht sie den offensichtlichen Fakt, dass Systeme, die durch natürliche Organisationsebenen im Sinne kompositorischer Ebenen gekennzeichnet sind, auf ihren höheren Ebenen Eigenschaften aufweisen, über die die (agierenden) Entitäten der niedrigeren Ebenen nicht verfügen. So wie Neurone Eigenschaften aufweisen, die die einzelnen Zellbestandteile oder Moleküle, aus denen sie bestehen, nicht haben, ließe sich daher auch sagen, dass dem Psychischen Eigenschaften zukommen, die die einzelnen Neuronen (Molekülen, Atomen...), aus denen es besteht, nicht aufweisen.

Auf die Frage nach der Ergänzung der Supervenienzthese gibt sie (im Sinne der starken Emergenz) allerdings eine eher ernüchternde Auskunft: die Beziehung zwischen der neuronalen Ebene und der mentalen Ebene bleibt unerklärlich. Das Auftauchen mentaler Eigenschaften auf einer höheren Ebene ist aufgrund der bloßen Kenntnis der Entitäten und ihre Konfigurationen auf den zugrunde liegenden physischen Ebenen *prinzipiell* nicht vorhersehbar. Die Emergenztheorie ist damit aber noch keine eigenständige Leib-Seele-Theorie. Sie besagt nur, dass das Auftau-

chen der phänomenalen Ebene nicht völlig aus den Entitäten und Beziehungen zwischen diesen auf den niedrigeren Ebenen *abgeleitet* werden kann. Die einzig inhaltlich konsistente Bestimmung von Emergenz, die über die bloße Feststellung der Existenz von systemischen Eigenschaften hinausgeht, ist von *epistemischer* Qualität. Obgleich es eine synchrone Determiniertheit der Makroeigenschaften eines Systems durch dessen Mikroeigenschaften (z.B. aufgrund einer psychophysischen Identitätsbeziehung oder eines epiphänomenalistischen Verhältnisses) geben mag, ist diese Relation nicht einsichtig zu machen, da reduktive Erklärungsversuche scheitern.

Ontologisch bleibt die Emergenzrelation indifferent. Sie ist durchaus mit Theorien, die psychophysische Identitätsannahmen beinhalten, kompatibel. Auch wenn solche Annahmen gerade in der Psychologie kursieren, kann der Emergentismus somit prinzipiell nicht als Alternative zu einer (kompositorisch begründeten) Identitätstheorie fungieren. Die Emergenzrelation beinhaltet eine negative Bestimmung, die nur aussagt, wie die Beziehung zwischen den Eigenschaften verschiedener Ebenen *nicht* gestaltet ist. Obgleich „Unerklärbarkeit“ ein elementarer Bestandteil des emergenztheoretischen Konzeptes sein mag, muss trotzdem konstatiert werden, dass es ihr ähnlich wie der Supervenienzrelation an einer positiven Darstellung, einer Erklärung des von ihr betrachteten Zusammenhangs, mangelt.

6.5. Funktionalismus und physische Realisierung

Die *Realisierungstheorie* kann als die derzeit dominierende Theorie des Geistes angesehen werden. Sie hat sich (in ihrer „klassischen“ Variante) insbesondere aufgrund der Kritik an der Identitätstheorie in Form des

Argumentes der multiplen Realisierbarkeit herausgebildet. Das Zutreffen der Realisierungsrelation wird von vielen Autoren als ein Garant eines *nicht-reduktiven* Physikalismus angesehen, da eine solche Beziehung ohne Identitäten (auf Typen- oder Eigenschaftsniveau) auskommt, die als Brückenprinzipien im klassischen Nagelschen Modell der Reduktion gebraucht würden. Kim (1998, 2005) entwickelt dagegen ein Realisierungsmodell, das über eine Funktionalisierung mentaler Eigenschaften eine *reduktive Erklärbarkeit* mentaler Eigenschaften absichern soll und damit als ein reduktiver Ansatz der Realisierung fungiert (vgl. Kap. 8.4). Auch wenn ich im Folgenden primär auf Kim eingehe, hat dies lediglich damit zu tun, dass sein Ansatz in meinen Augen der am besten ausformulierte Ansatz zur Beschreibung der Realisierungsrelation ist. Es ist an dieser Stelle zunächst nicht entscheidend, ob man die reduktionistische Interpretation seiner Ausführungen mitträgt oder nicht. Wie sich zeigen wird, bezieht sich dieser „klassische“ Ansatz der Realisierung auf eine eindimensionale (*intra-level*) Relation zwischen einer funktional charakterisierten mentalen Eigenschaft und ihrem physischen Realisierer. Die Teil-Ganzes-Beziehung bleibt dabei weitestgehend unberücksichtigt. Einen alternativen Ansatz der Realisierung vertritt Gillett (2002, 2003), der genau diese Lücke schließen möchte, indem er ein multidimensionales (*inter-level*) Modell der Realisierung entwirft, das die Beziehung zwischen den Entitäten und Eigenschaften verschiedener (kompositorisch bestimmter) natürlicher Organisationsebenen betrifft. Es wird sich allerdings zeigen, dass Gilletts Ansatz in vielen Punkten unverständlich bleibt und inhaltliche Inkonsistenzen aufweist.

Im Allgemeinen nehmen physische Realisierungstheorien an, dass Instantiierungen einer mentalen (funktional analysierbaren) Eigenschaft bei Menschen für gewöhnlich durch Instantiierungen bestimmter physischer Eigenschaften realisiert werden. Das Aufkommen und die gegen-

wärtig sehr starke Dominanz der Vorstellung, dass mentale Eigenschaften durch physische Eigenschaften realisiert werden, speist sich dabei aus drei grundsätzlichen Vorbedingungen.

Physische Realisierungstheorien können *erstens* als unmittelbare Konsequenz des Argumentes der multiplen Realisierbarkeit an der „klassischen“ Identitätstheorie (vgl. Kap. 6.1.4.2) gesehen werden. Es ist nicht eine bestimmte physische Eigenschaft, mit der eine ganz bestimmte mentale Eigenschaft identisch ist, sondern es können verschiedene physische Eigenschaften sein, die diese realisieren. Mit anderen Worten: Die physischen Realisierer einer mentalen Eigenschaft können variieren.

Zweitens kann die Realisierungstheorie auch als Versuch angesehen werden, die abstrakten funktionalen Bestimmungen des Geistes, wie wir sie im computationalen Ansatz zur Analyse kognitiver Zustände vorfinden (vgl. Kap. 4.2.1.1), mit den neuronalen Grundlagen der menschlichen Informationsverarbeitung in Verbindung zu bringen. Computationale Modelle des Geistes sind relativ abstrakte theoretische Modelle psychischer Abläufe, die aufgrund bestimmter bekannter Eingangsvariablen zuverlässige Prognosen zukünftigen Verhaltens eines informationsverarbeitenden Systems wie des Menschen ermöglichen. Was aber fraglich bleibt, ist ihr Realitätsstatus. Sind diese kognitionswissenschaftlichen Modelle, die die Blackbox der Behavioristen gefüllt haben, letztendlich wirklich nur *Modelle* oder gibt es tatsächlich etwas in unseren Köpfen, was nach den postulierten Prinzipien abläuft? Genau hier kommt die Realisierungsrelation ins Spiel. Sie soll eine Brücke zwischen den theoretisch-abstrakten Auffassungen des Geistes und seiner konkreten physischen Seinsweise schlagen, wie sie in Marrs (1982) *Drei-Ebenen-Modell* herangezogen werden.

Drittens soll über die Realisierungsrelation die benötigte explanatorische Erweiterung der Supervenienzkonzeption hergestellt werden. Die rein

deskriptiven Zusammenhänge einer Gehirn-Geist-Korrelation (oder Mikro-Makro-Determination) sollen durch das Prinzip der physischen Realisierung erklärt werden. Die Stoßrichtung ist hierbei, dass das Mentale über dem Physischen superveniert, weil mentale Eigenschaften *funktionale Eigenschaften zweiter Ordnung* sind, die physisch realisiert werden (vgl. Kim, 1998).

Gemeinsamer Hintergrund für diese Entwicklung ist die grundsätzliche Akzeptanz eines *funktionalistischen* Verständnisses des Psychischen, das – vor allem aufgrund der Annahme der multiplen Realisierbarkeit – eine weitestgehende Unabhängigkeit der abstrakten funktionalen Eigenschaften von ihren konkreten physischen Grundlagen gewährleisten soll. Es ist der Funktionalismus, der sowohl in der Philosophie nach dem Scheitern der Identitätstheorie die dominierende „Theorie des Geistes“ war und in weiten Teilen noch immer ist als auch in der Psychologie als dominierendes Forschungsparadigma im Rahmen der so genannten „kognitiven Wende“ den Behaviorismus (zumindest *pro forma*) abgelöst hat. So wie Marr (1982) in seinem Ebenenmodell zur visuellen Wahrnehmung davon ausgeht, dass die computationalen Abläufe auf der höchsten Ebene durch einen bestimmten Algorithmus, also durch eine formale Prozedur zur Aufgabendurchführung, auf der mittleren Ebene realisiert werden, der wiederum durch physische Basisprozesse auf der untersten Ebene realisiert wird, nimmt auch Kim (1998, 2005) an, dass man mentale Eigenschaften über ihre jeweilige kausale Rolle (als funktionale Eigenschaften) individuieren kann, um sie dann mit ihren physischen bzw. neuronalen Realisierern in Verbindung zu bringen.

6.5.1. Realisierung als Identität von Eigenschaftsinstantiierungen

Die beiden hauptsächlichsten Verfechter einer psychophysischen Realisierungsrelation sind Jaegwon Kim (1997, 1998, 2005) und Sydney Shoemaker (2001, 2003, 2007). Die Kernidee hierbei ist, dass mentale Eigenschaften funktionale Eigenschaften zweiter Ordnung sind, die über physische Realisierer verfügen. Mentale Eigenschaften werden gemäß der Doktrin des Funktionalismus über ihre kausalen Rollen bestimmt:

Functionalism takes mental properties and kinds as functional properties, properties specified in terms of their causal roles as causal intermediaries between sensory inputs and behavioral outputs, and the physicalist form of functionalism takes physical properties as the only potential occupants, or ‘realizers,’ of these causal roles. To use a stock example, for an organism to be in pain is for it to be in some internal state that is typically caused by tissue damage and that typically causes groans, winces, and other characteristic pain behavior. (Kim, 1998, S. 19)

Für eine „Funktionalisierung“ (von Kim auch *funktionale Reduktion*²³¹ oder reduktive Erklärung genannt) mentaler Zustände muss man also zunächst konzeptuell arbeiten und die physischen bzw. verhaltensbezogenen Eigenschaften von z.B. Schmerzen aufdecken, die für die kausale Rolle von Schmerzen charakteristisch sind. In einem zweiten Schritt muss dann die empirische Forschung zeigen, welche neuronalen Konfigurationen genau diese kausale Rolle (beim Menschen oder bei vielleicht allen Wirbeltieren) besetzen oder erfüllen²³². Kim (2005) fügt später noch

²³¹ Ob auf der Grundlage einer Realisierungstheorie wirklich eine Reduktion mentaler Eigenschaften erfolgen kann, ist in der Philosophie des Geistes und Wissenschaftstheorie ein strittiger Punkt. Nach Kim (1998, 2005) sei dies gewährleistet. Die meisten anderen Autoren und Verteidiger eines nicht-reduktiven Physikalismus (z.B. Shoemaker, 2001, 2003, 2007) lehnen dies ab, da keine Reduktion auf Typen-Niveau möglich ist, sondern höchstens auf der Ebene von Token bzw. Instantiierungen von Eigenschaften, aber selbst das wird von anderen Autoren (z.B. Pereboom & Kornblith, 1991; Pereboom, 2002) noch abgelehnt.

²³² Für gewöhnlich wird die Formulierung dieses Prinzips Levine (1993) zugeschrieben,

einen dritten Schritt hinzu, der allerdings bereits durch den vorherigen impliziert wird. Dieser besagt, dass ein explanatorischer Ansatz dafür geschaffen werden muss, wie es die Realisierer des mentalen Zustandes (des Schmerzes) vermögen, diese kausale Rolle auszufüllen.

Potentiell ist die üblicherweise mit dem Funktionalismus²³³ in Zusammenhang gebrachte Realisierungstheorie damit allerdings auch mit dualistischen Theorien verträglich, was bereits Putnam (1967) festgestellt hat. So könnte es theoretisch auch eine geistige Substanz sein, die mentale Eigenschaften realisiert. Eine *physikalistische* Position wird die Realisierungstheorie erst, wenn die einschränkende Zusatzannahme hinzugefügt wird, dass es nur *physische* Eigenschaften sind, die mentale Eigenschaften realisieren (vgl. Kim, 1997, S. 280). Hierbei scheint es sich allerdings um eine ad-hoc-Annahme oder induktive Verallgemeinerung der bisherigen wissenschaftlichen Beobachtungen bzw. ihrer Interpretation zu handeln.

Ob eine bestimmte physische Eigenschaft *P* ein Realisierer einer mentalen Eigenschaft *M* ist, hängt von der Art des Systems ab, in das *P*

obwohl sich in der Literatur schon viel früher ganz ähnliche Gedanken finden. So heißt es z.B. bei Deutscher (1967): „Amongst the knowledge we think we have about our mental states is some knowledge of the sort of the causes they have, and the sort of behavioural effects. Thus to try to identify the mental with a brain state is to try to find a brain state with those causes and effects“ (S. 71). Auch David Armstrong (1966/1970) definierte einen mentalen Zustand bereits als „a state of the person apt for producing certain ranges of behaviour“ (S. 75), was später wiederum vor allem David Lewis (1966/2007, 1980) aufgegriffen und präzisiert hat.

²³³ M.E. handelt es sich hierbei um keine notwendige Beziehung, auch wenn häufig dieser Anschein erweckt wird. Die Akzeptanz einer (physikalischen) Realisierungstheorie impliziert nicht notwendigerweise die Akzeptanz des Funktionalismus. Grundsätzlich könnten es auch nicht-funktionale Bestimmungen des Mentalen sein, die in einem System realisiert werden (vgl. Shoemaker, 2007, S. 5). Ebenso ist ein Funktionalismus – wie die Fachgeschichte der Psychologie gezeigt hat – auch ohne physikalische Realisierung existenzfähig, obgleich ein solcher „frei schwebender“ Ansatz zu recht vielfach kritisiert wurde (siehe Kap. 4.2).

eingebettet ist, „since in psychology the input-output behavior of the total system is what is of concern, and the causal role that *P* plays will depend on the makeup (,causal wiring‘) of the system as a whole“ (Kim, 1998, S. 22). Ob also z.B. eine Gewebeverletzung nozizeptive Neuronen zum Feuern bringt, hängt von der neuronalen Organisation des betroffenen Organismus ab. Ob das Feuern dieser Nervenfasern ein entsprechendes Fluchtverhalten auslöst, hängt wiederum von den neuronalen und motorischen Systemen des Organismus ab. Wenn dieselbe Eigenschaft *P* in anderen Systemen implementiert wird, könnte es daher sein, dass sie *M* nicht realisiert (vgl. ebd.). Fakt ist damit auch, dass es eine kontingente oder empirische Frage und keine von metaphysischer bzw. begrifflich-notwendiger Qualität ist, ob eine bestimmte (physische) Eigenschaft eine funktionale Eigenschaft realisiert oder nicht.

Da mentale Eigenschaften über ihre kausalen Rollen und damit über ihre typischen Ursachen und Wirkungen definiert werden, müssen sie „extrinsische“ bzw. *relationale* Eigenschaften sein: „Whether or not a given property qualifies as an occupant of a specific role – that is, whether or not it is a realizer of a functional property – depends essentially on its causal/nomological relations to other properties, not on its intrinsic character“ (S. 21). In einem mentalen Zustand zu sein, bedeutet also primär, in einem Zustand mit bestimmten typischen Ursachen und typischen Wirkungen zu sein. Der „intrinsische“ Charakter, womit Kim meint, wie sich ein mentaler Zustand für uns anfühlt, hat nur noch die Bedeutung, dass er hilft, zu entscheiden, ob jemand einen bestimmten funktionalen Zustand realisiert oder auch nicht.

Wie Walter (2009) sagt, hat sich das „Realisierungsideom“ (S. 691) in dem Sinne eingebürgert, dass es gegenwärtig üblich ist zu sagen, „mentale Eigenschaften seien durch physikalische Eigenschaften realisiert, aber nicht mit ihnen identisch“ (ebd.). Kritisiert werden kann, dass es genau

genommen unklar bleibt, was mit einer Realisierungsbeziehung überhaupt gemeint ist, vor allem, wenn physische Realisierung derart strikt von psychophysischer Identität abgegrenzt wird. Inwiefern werden mentale Eigenschaften durch neuronale Eigenschaften *realisiert*, wenn dies nicht auf eine psychophysische Identität zurückgeht? Meinem Eindruck nach besteht eine untrennbare Brücke zwischen der Identitätsbeziehung und der Realisierungsrelation. Löst man diese auf, bleibt die Realisierungsrelation unverstandlich und verliert ihre Ressourcen zur Integration der Moglichkeit mentaler Verursachung.

Naturlich grenzt sich die Realisierungstheorie dadurch von der Identitatstheorie ab, als dass sie eine *asymmetrische* Relation ist (die Identitatstheoretiker mussten hierzu Zusatzannahmen einfuhren, da Identitaten per se symmetrischer Natur sind). Die Realisierungsrelation impliziert also nicht, dass eine mentale Eigenschaft immer durch ein und dieselbe neuronale Eigenschaft realisiert wird, was Beckermann (1999) sehr anschaulich am Beispiel der Temperatur erlautert:

Wir wissen heute, da Temperatur *nicht identisch* ist mit der mittleren kinetischen Energie der Molekule eines Stoffes; vielmehr gilt nur, da Temperatur *in Gasen* durch die mittlere kinetische Energie ihrer Molekule *realisiert* ist. In *festen Korpern* dagegen ist sie durch die Schwingungen von Molekulen realisiert; im *Plasma* auf noch andere Weise; und Vakuum schlielich durch die Verteilung elektromagnetischer Strahlung. Dennoch kame niemand mehr auf die Idee, anzunehmen, da es neben der mittleren kinetischen Energie von Molekulen oder neben der Verteilung elektromagnetischer Strahlung noch eine weitere eigenstandige Eigenschaft, die Eigenschaft der Temperatur, gibt – eine Eigenschaft mit eigenen Ursachen und Wirkungen. (S. 225f; Hervorh. im Orig.)

Die Eigenschaft „Temperatur“ kann also auf verschiedene Art und Weise realisiert werden. Es kann daher nicht einfach gesagt werden, dass Temperatur mit der kinetischen Energie der Molekule eines Stoffes identisch ist, da dies nur fur die Realisierung von Temperatur in Gasen gilt, wohin-

gegen es unter anderen Bedingungen auch andere Realisierungsformen von Temperatur gibt. Die Realisierungsrelation verneint demnach eine Identität auf Typen-Niveau. Was heißt es nun aber genau, dass die Temperatur in festen Körpern durch die Schwingung von Molekülen *realisiert* wird? Offenbar heißt es, dass die Temperatur mit der Schwingung von Molekülen identisch ist, zumindest wenn es sich um eine in festen Körpern realisierte Eigenschaft handelt (genauso wie die Temperatur im Vakuum mit der Verteilung elektromagnetischer Strahlung identisch ist etc.). Auch wenn eine allgemeine Identitätsaussage zur Temperatur nicht möglich zu sein scheint, liegt zumindest für jeden Fall einer *Instantiierung*²³⁴ von Temperatur offenbar ein Identitätsverhältnis mit seinen physischen Realisierern vor. Wenn dies nicht der Fall ist, wie soll das Prinzip der Realisierung sonst zu verstehen sein?

Zumindest Kim räumt die Implikation der Identität von Eigenschaftsinstantiierungen aufgrund der von ihm propagierten Realisierungstheorie explizit ein. Zentral hierfür dürfte insbesondere der Grund sein, die Möglichkeit mentaler Verursachung theoretisch absichern zu können. Allgemein kann man sagen, dass sich Ausbuchstabierungen der Realisierungsrelation, die die kausale Wirksamkeit mentaler Eigenschaft zu gewährleisten versuchen, in der Regel auf die *Theorie der kausalen Eigenschaften* von Shoemaker (1980, 1998) stützen. Diese Theorie besagt im Kern, dass Eigenschaften über die Kausalkräfte ihrer Träger individuiert werden. Demnach sind zwei Eigenschaften *F* und *G* dieselbe Eigenschaft, wenn „ihre Instantiierungen unter allen möglichen Umständen [...] zu

²³⁴ Ebenso könnte es natürlich sein, dass sich die globale Eigenschaft *Temperatur* in verschiedene Subtypen, wie sie von Beckermann vorgestellt werden, aufspalten lässt, die jeweils durch andere physische Größen realisiert werden. Dies ist zugleich eine Antwort auf das Argument der multiplen Realisierbarkeit und wird an späterer Stelle noch genauer erörtert werden (vgl. Kap. 7.5.2.2).

denselben Kausalkräften ihrer Träger beitragen” (Walter, 2009, S. 697). Hieraus leitet Kim (1993) bereits in einer frühen Arbeit das so genannte *Prinzip der kausalen Vererbung* ab: „If mental property *M* is realized in a system at *t* in virtue of physical realization base *P*, the causal powers of this instance of *M* are *identical* with the causal powers of *P*“ (S. 326; Hervorh.: M.L.). Wenn also z.B. ein bestimmter Wunsch mich dazu veranlasst, eine Handlung auszuführen, so verdankt sich die Wirkmächtigkeit meines Wunsches der Wirkmächtigkeit der neuronalen Konfiguration bzw. Eigenschaften, die diesen Wunsch physisch realisieren. Die Kausalkraft meines Wunsches ist identisch mit der Kausalkraft der diesen realisierenden neuronalen Konfigurationen²³⁵.

Wenn es also – um die Möglichkeit mentaler Verursachung integrieren zu können – notwendig erscheint, die Kausalkräfte von mentalen Eigenschaften auf die Kausalkräfte ihrer physischen Realisierer zurückzuführen und man gleichzeitig annimmt, dass Eigenschaften sich über die Kausalkräfte ihrer Träger individuieren, dann ist folgende Erklärung des Vorliegens psychophysischer Korrelationen Kims (1998) durchaus folgerichtig:

[M]ental properties are second-order functional properties with physical realizers (and no non-physical realizers). And we have an explanation of mental-physical correlations. Why is it that whenever *P* is realized in a system *s*, it initiates mental property *M*? The answer is that by definition, having *M* is having a property with causal specification *D*, and in systems like *s*, *P* is the property (or one of the properties) meeting specification *D*. For systems like *s*, then, having *M* consists in having *P*. [...] [H]aving *M*, for these systems, is ‘nothing over and above’ having *P*. (Kim, 1998, S. 24; Hervorh. im Orig.)

²³⁵ Zu einer solchen Annahme der Identität der physischen und psychischen Kausalkräfte sah sich bereits Wundt (1894) zur Absicherung der Möglichkeit mentaler Verursachung gezwungen, obgleich er sich ansonsten programmatisch ontologischer Bestimmungen (über eine bloße heuristische psychophysische Identitätsannahme hinaus) enthalten wollte.

Wenn das Haben einer bestimmten mentalen Eigenschaft darin besteht bzw. nichts anderes ist als das Haben eines physischen Realisierers, dann wird ersichtlich, dass die von Kim vertretene Variante der Realisierungstheorie eine psychophysische Identitätsannahme beinhaltet. Im Zuge dessen resümiert Kim, dass eine Identifikation der Instantiierungen mentaler Eigenschaften (hier: *M*) mit Instantiierungen neuronaler Eigenschaften als ihren Realisierern (hier: *P*) unaufhaltsam sein dürfte:

First, [...] the very conception of second-order properties and their realizers strongly incline us to deny that there can be anything to the instantiation of a second-order property over and above the instantiation of one of its realizers. For *M* to be instantiated on a given occasion *is* for one of its realizers to be instantiated on that occasion, and on this occasion it is *P* that realizes *M*. Second, if the causal powers of the *M*-instance are identical with those of its realizing *P*-instance, what reasons can we have for thinking there are here two events, not one? On a plausible view of properties, only causally relevant or efficacious properties should count as individuating properties, and it is in any case highly implausible to say that events that are indiscernible in respect of causal properties can yet be distinct events. (ebd., S. 56; Hervorh. im Orig.)

Der Witz an Kims Position ist, dass er sich mit dieser einerseits von einer Typen-Identitätskonzeption à la Place, Smart oder Feigl distanziert, die an der multiplen Realisierbarkeit mentaler Typen zu scheitern droht, andererseits aber die Notwendigkeit der Klärung der Beziehung mentaler Eigenschaften (die auch als gewisse mentale Typen verstanden werden können) betont. Seine Strategie besteht darin, ein *lokales* Reduktionsmodell zu entwerfen, das nur *Instantiierungen* von Eigenschaften und nicht die Beziehungen zwischen zwei *Eigenschaften überhaupt* thematisiert. Die Art von Reduktion, die Kim verteidigt, „can focus on the reduction of a mental property, or a group of them, for a specific population – that is, neural research on pain will aim at *local* reductions, not a one-shot *global* reduction [...]. Neural bases may differ for different instances of

pain, but individual pains must nonetheless reduce to their respective neural/physical realizers” (ebd., 2005, S. 25, Hervorh. im Orig.). Kim möchte sich also klar von einem nicht-reduktiven Physikalismus abgrenzen, wie ihn Davidson vertritt, in dem er eine reduktive physikalistische Position propagiert. Gleichzeitig gilt seine Reduktion aber auch nur für einzelne Instantiierungen von Eigenschaften oder Ereignissen, womit die von Kim suggerierte *Ontologie* genau genommen kaum von der verschieden ist, die der *Anomale Monismus* impliziert. Es sind nur in *epistemischer* Hinsicht grundverschiedene Schlussfolgerungen, die beide Autoren ziehen: Davidson leitet aufgrund der Anomalitätsthese trotz der akzeptierten Identität mentaler und physischer Ereignisse eine *Irreduzibilität* des Mentalen ab, während Kim auf der Basis ebendieser Identität im Zusammenhang mit der von ihm verteidigten reduktiven Erklärbarkeit mentaler Eigenschaften über funktionale Analysen die Möglichkeit einer psychophysischen *Reduktion* (zumindest in weiten Teilen²³⁶) als gegeben ansieht.

Was hier deutlich geworden sein sollte, ist, dass die Aussage, dass mentale Eigenschaften durch physische Eigenschaften realisiert, aber nicht mit diesen identisch seien, nur eingeschränkte Gültigkeit hat. Die Ausbuchstabierung der Realisierungsrelation nach Kim operiert durchaus mit Identitätsannahmen, nur eben nicht auf der Ebene von Eigenschaften oder Typen, sondern auf der Ebene von Eigenschaftsinstantiierungen.

²³⁶ Bei phänomenalen Eigenschaften könnte es sein, dass sie sich einer funktionalen Analyse widersetzen, wie sie für das Prinzip der reduktiven Erklärbarkeit mentaler Eigenschaften erforderlich wäre (vgl. Kim, 1998, S. 118f).

6.5.2. Realisierung als kausale Teilmengenbeziehung

Vielen Philosophen ist die Realisierungsrelation damit (immer noch) zu nah an die ungeliebte Identitätstheorie herangerückt. Eine prominente Alternative zu Kims Ansatz, die trotz einer grundsätzlichen Anerkennung des Prinzips der kausalen Vererbung (und der meisten von Kim getätigten Grundannahmen zur Charakterisierung der Realisierungsrelation) keine Identität der Kausalkräfte mentaler und neuronaler Eigenschaften annimmt, sondern nur eine *Teilmengenbeziehung* (*subset relation*) zwischen diesen postuliert, hat daher Sydney Shoemaker (2001) entwickelt: „In general, then, property *X* realizes property *Y* just in the case the conditional powers bestowed by *Y* are a subset of the conditional powers bestowed by *X*“ (S. 78; Hervorh.: M.L.). Mein Wunsch, eine bestimmte Handlung auszuführen, verdankt also seine kausale Wirkmächtigkeit einer *Teilmenge* der Kausalkräfte der ihn realisierenden neuronalen Konfiguration²³⁷. Eine ganz ähnliche Bestimmung findet sich auch bei Clapp (2001).

Was hierdurch beabsichtigt wird, ist eine Version physischer Realisierung zu etablieren, die auf keine psychophysische Identitätsannahme hinausläuft. Für Shoemaker, Clapp und viele andere Philosophen ist es inakzeptabel, dass es nicht die realisierte (mentale) Eigenschaft ist, die eine kausale Rolle spielt, sondern vielmehr seine diversen Realisierer. Insofern sagt Shoemaker selbst aus, dass seine Motivation darin besteht, eine Realisierungstheorie zu installieren, die solche Konsequenzen vermeidet: „We need an account of property-realization that assigns the relevant causal

²³⁷ Auch Kim hat in seinen späteren Arbeiten (1998, 2005) im Zuge seiner Ausformulierung einer Realisierungstheorie die Möglichkeit mit eingeschlossen, dass es sich um eine Teilmengenrelation bzgl. der Kausalkräfte handelt. Seine Vorstellungen bzgl. der Identitätsbeziehung bzw. Reduzierbarkeit von Instantiierungen mentaler Eigenschaften sieht er hierdurch offenbar aber nicht infrage gestellt.

role to the realized property itself, while acknowledging that it is in virtue of causal roles played by its realizers that it is able to play this causal role” (Shoemaker, 2007, S. 5). Er modifiziert also das Prinzip der kausalen Vererbung, auf das sich Kim stützt, zu einem Teilmengenkonzept der Realisierung. Die Kausalkräfte der realisierten Eigenschaft machen demnach nur noch eine Teilmenge der Kausalkräfte der diese realisierenden Eigenschaftsinstantiierung aus. Hieraus wird nun vollmundig die Nicht-Identität mentaler und physischer Eigenschaften abgeleitet:

But of course, if the causal powers of one property instance are a proper subset of those of another, the instances are not identical we can dismiss the argument of their identity to the causal impotence of the realized property. So I favor the way of viewing second-order properties according to which the instances of second-order properties are not identical with instances of their first-order realizers, and, what goes with this, second-order properties have causal profiles of their own, distinct from, although of course intimately related to, the causal profiles of their realizers. (ebd., 2007, S. 17)

Man kann zu Recht fragen, inwiefern physisch realisierte Eigenschaften ontologisch eigenständig sein können, wenn sie „in einem mehr oder minder wörtlichen Sinn ein Teil ihrer physikalischen Realisierer sind. Vielleicht sind die sie individuierenden Kausalkräfte für keinen ihrer Realisierer individuierend, aber es bleibt die Frage, in welchem Sinn ein Objekt, das eine realisierende Eigenschaft *P* hat, *darüber hinaus* auch noch eine von *P* realisierte Eigenschaft *F* hat“ (Walter, 2009, S. 700; Hervorh. im Orig.). Es spricht deshalb viel dafür, dass auch hier weiterhin eine versteckte Identitätsannahme vorliegt, die sprachlich-konzeptuell einfach nur – mehr oder weniger geschickt – kaschiert wird. Wie Shoemaker es selbst auch gar nicht leugnet, scheint es sich um eine *ad-hoc*-Konstruktion zu handeln, die über definitorische Bestimmungen (vor allem Shoemakers Theorie kausaler Eigenschaften, die über ihre Anforderungen an

eine Identität von Eigenschaften ein Tor hierfür öffnet) eine psychophysische Nicht-Identität nachzuweisen sucht. Denn auch wenn es zwar nur eine Teilmenge der kausalen Wirkmächtigkeit der neuronalen Eigenschaften ist, die mit der Kausalkraft meines Wunsches in Verbindung gebracht wird, scheint sich die Wirkmächtigkeit meines Wunsches weiterhin daraus zu ergeben, dass sie *identisch* mit einer nunmehr Teilmenge der Wirkmächtigkeit dieser neuronalen Eigenschaften ist.

Auf ein weiteres Problem weist Walter (2009) hin: Aus der Formulierung der Realisierung als Teilmengenrelation dürfte folgen, dass jede kausal wirksame Eigenschaft mehr als nur eine Eigenschaft realisieren kann. Wenn die Realisierung durch eine kausal wirksame Eigenschaft auf diese Weise bereits *hinreichend* für kausale Wirksamkeit ist, ergibt sich die problematische Konsequenz, dass es ausgeschlossen ist, dass ein und dieselbe realisierende Eigenschaft einerseits eine kausal wirksame und andererseits eine kausal unwirksame Eigenschaft realisiert. Im Fundus der Wissenschaften finden sich aber Fälle, bei denen sich klar das Gegenteil zeigt. So haben z.B. Jackson und Petit (1990) darauf hingewiesen, dass es eine ganz bestimmte Eigenschaft freier Elektronen in Metallen ist, die sowohl dessen thermische Leitfähigkeit als auch dessen elektrische Leitfähigkeit realisiert, aber nach dem Stand der Dinge gibt es Situationen, in denen die elektrische Leitfähigkeit von Metallen kausal wirksam ist, während es die thermische nicht ist, und umgekehrt (vgl. Walter, 2009, S. 702).

6.5.3. Eindimensionalität versus Multidimensionalität

Die Entfaltung der Gedanken Kims zu einer eindimensionalen versus multidimensionalen Betrachtungsweise findet vor allem vor dem Hintergrund der Beantwortung der Frage statt, ob seine Realisierungstheorie die

im Zusammenhang mit der Supervenienzthese aufgetretenen Probleme des kausalen Ausschlusses des Mentalen (*causal exclusion*) oder des Abdriftens der Kausalität (*causal drainage*) tatsächlich lösen kann.

Nochmals in Kürze: Nach Maßgabe der Realisierungstheorie soll die ontologische Eigenständigkeit von mentalen Eigenschaften über ihre multiple Realisierbarkeit durch verschiedene Varianten physischer Eigenschaften gewährleistet sein, während ihre kausale Wirksamkeit darüber abgesichert werden soll, dass sie funktionale Eigenschaften sind, die über ihre kausale Rolle individuiert werden. Allerdings lässt sich der letzte Punkt kritisch hinterfragen, denn was die kausale Rolle ausfüllt, ist gerade nicht die mentale Eigenschaft, sondern ihr physischer Realisierer, so dass die kausale Wirksamkeit genau genommen nicht der mentalen Eigenschaft, sondern ihrem physischen Realisierer zukommt (vgl. Block, 1990). Somit scheint es auch unter den Voraussetzungen der Realisierungsrelation offenbar zu einem Ausschluss von Kausalität auf den höheren Ebenen und damit auch auf der mentalen Ebene zu kommen. Nach Ansicht Blocks ist es sogar nicht unwahrscheinlich, dass sich der Funktionalismus aufgrund der Fraglichkeit der kausalen Wirkmächtigkeit von Eigenschaften zweiter Ordnung als eine bestimmte Variante des Epiphänomenalismus entpuppt.

Auch das Argument des *Abdriftens der Kausalität* (vgl. Block, 1990, 2003) lässt sich offenbar leicht auf die Realisierungsrelation anwenden. Wie Kim (1998) es selbst zum Ausdruck bringt:

[N]eural realizers of pain are themselves second-order with respect to certain lower-level properties (which preempt their causal powers, the latter in turn are second-order to their lower-level properties, and so on ad infinitum, until you reach the rock bottom level of microphysics (if such a level exists), which turns out to be the only level where genuine powers reside. [...] What's more: What if there is no bottom level (as Block challenges us to consider)? It looks as though if there is no bottom

level in this picture, causal powers would drain away into a bottomless pit and there wouldn't be any causation *anywhere*. (S. 81; Hervorh. im Orig.)

Folgt daraus nun wirklich, dass alle Kausalität auf höheren Wirklichkeits-ebenen eine bloße Illusion ist? Kim ist jedenfalls überzeugt, dass er mit seiner Ausformulierung der Realisierungstheorie die eigenständige Wirkmächtigkeit mentaler Eigenschaften absichern und ein Abdriften von Kausalität vermeiden kann. Wichtig ist hierbei, dass Kim zwischen „Ordnungen“ („orders“) und „Ebenen“ („levels“) differenziert (vgl. ebd., 1998, S. 83). Wenn Kim von Eigenschaften erster oder zweiter Ordnung spricht, sei hiermit noch keine Aussage über die Einordnung dieser Eigenschaften in eine Hierarchie von Ebenen getroffen. Eigenschaften unterschiedlicher Ordnung sind durchaus auf *ein und dieselbe* Ebene in einer Hierarchie natürlicher Organisationsebenen bezogen.

Alle funktionalen Eigenschaften (wie mentale Eigenschaften) sind demnach Eigenschaften zweiter Ordnung, die in Form von kausalen oder nomologischen Beziehungen zwischen Eigenschaften erster Ordnung bestimmt werden: „for a system x to have this property is for x to have some first-order property P that satisfies a certain condition D , where in the present case D specifies that P has pain's typical causes and typical effects“ (ebd., 1998, S. 19). Als ein Beispiel für eine solche Eigenschaft zweiter Ordnung, führt Kim (1997) die „dormitIVITY“, also die einschläfernde Wirkung von Medikamenten an (vgl. S. 281). Eine Substanz hat diese Eigenschaft, wenn sie chemische Eigenschaften aufweist, die Menschen zum Einschlafen bringen. Verschiedene Medikamente (wie Valium oder Seconal) weisen eine solche sedierende Wirkung auf, aber sie haben diese aufgrund unterschiedlicher chemischer Realisierer erster Ordnung (Diazepam vs. Secobarbital).

Eine entscheidende Bestimmung Kims ist nun, dass er der Mikro-Makro-Unterscheidung im Hinblick auf mentale Eigenschaften als Eigenschaften zweiter Ordnung und ihrem Verhältnis zu ihren physischen Realisierern widerspricht. Im Angesicht der von ihm propagierten Realisierungstheorie sei es nicht korrekt, neuronale und psychische Eigenschaften als Elemente zwei verschiedener Ebenen in einer Mikro-Makro-Hierarchie zu betrachten:

[T]his hierarchy does not parallel the micro-macro hierarchy – to put it another way, the realization relation does not track the micro-macro relation. The reason is simple: both second-order properties and their first-order realizers are properties of the same entities and systems. [...] It is evident that a second-order property and its realizers are at the same level in micro-macro hierarchy; they are properties of the very same objects. This is a simple and direct consequence of the very concepts of ‘second-order property’ and ‘realizer’: for something to have a second-order property is for it to have one or another of its realizers, that is, a first-order property satisfying the specification that defines the second-order property. Consequently, when we talk of second-order properties and their realizers, there is no movement downward, or upward, in the hierarchy of entities and their properties ordered by the micro-macro relation. (Kim, 1998, S. 82; Hervorh. im Orig.)

Die über eine funktionale Analyse erfassbaren mentalen Eigenschaften sind also auf derselben Ebene angesiedelt wie ihre physischen Realisierer, denn wie Kim immer wieder betont, besteht das Aufweisen eines bestimmten Schmerzzustandes *in nichts anderem als* dem Aufweisen eines bestimmten physischen (oder neuronalen) Zustandes. Die mentale Eigenschaft ist „nothing over and above“ gegenüber der sie realisierenden physischen Eigenschaft. Man kann also sagen, dass es sich bei der Realisierungsrelation um eine *eindimensionale* Konzeption des Verhältnisses von mentalen und neuronalen Eigenschaften handelt. Damit schlägt Kim also offenbar den Weg der „subjektivistischen“ Auflösung des Spinozistischen Dilemmas ein, der besagt, dass der Unterschied zwischen psychi-

schen und physischen (neuronalen) nur in der Art unseres sprachlich-konzeptuellen oder perzeptuellen Zugriffs ausschließlich auf ein und denselben Wirklichkeitsbereich zurückgeht (vgl. Kapitel 3).

Wie schon im Zusammenhang mit der Supervenienzthese besprochen, geht Kim somit in seiner Antwort auf das Problem des kausalen Ausschlusses davon aus, dass eine Identität zwischen den mentalen Eigenschaftsinstantiierungen und ihren neuronalen Realisierern besteht. Um dem Epiphänomenalismus zu entgehen, sollten wir uns seiner Ansicht nach für die reduktionistische Variante beim „Kausalitätsdilemma“ der Supervenienz entscheiden (vgl. Kap. 6.3.2). Die kausale Wirksamkeit (von Instantiierungen) mentaler Eigenschaften werde auf diese Art und Weise über ihre Reduktion auf (Instantiierungen) physische(r) Eigenschaften abgesichert. Allerdings räumt Kim selbst ein, dass bei supervenienten – physisch realisierten – psychologischen Eigenschaften, da sie auf der gleichen ontologischen Ebene wie die physischen bzw. neurobiologischen Eigenschaften angesiedelt sind, durchaus das Problem besteht, dass eine kausale Wirkung vollständig durch neuronale bzw. physische Eigenschaften erklärt werden kann (vgl. ebd., 1998, ebd., S. 82). Die kausale Wirksamkeit mentaler Eigenschaften scheint somit zwar nicht ausgeschlossen zu sein, verdankt sich jedoch der Wirkmächtigkeit ihrer physischen Träger.

Gemeinsam ist den vorgestellten Varianten der Realisierungsrelation nach Kim, Shoemaker, Clapp und anderen, dass sie das Prinzip der physischen Realisierung als eine *eindimensionale* (intra-level) Relation auffassen. Höherstufige (theoretische) Eigenschaften zweiter Ordnung werden durch niederstufige (natürliche) Eigenschaften erster Ordnung in die physische Welt gebracht oder eben durch diese *realisiert*. Beide Eigenschaftsklassen befinden sich auf *ein und derselben* natürlichen Organisationsebene. Bei der Bestimmung der Identität (oder kausalen Teilmen-

genbeziehung) zwischen Eigenschaften zweiter und erster Ordnung handelt es sich daher um eine ontologisch eindimensionale Bestimmung der Beziehung zweier Eigenschaften. Dennoch liefert insbesondere Shoemaker Bestimmungen dazu, wie sich die Eigenschaften auf den höheren Ebenen eines Systems zu der mikrostrukturellen Verfasstheit ebendieses Systems verhalten, vor allem um Antworten auf Probleme zu geben, die mit der Möglichkeit einer mentalen oder überhaupt höherstufigen Verursachung zusammenhängen.

Einige neuere Ansätze sprechen sich dezidiert für eine *multidimensionale* (inter-level) Sichtweise der Realisierung aus, die kompositorischen Prinzipien stärker Rechnung tragen soll. Sie erheben die Beziehung zwischen den Eigenschaften oder Entitäten auf den *verschiedenen* Organisationsebenen bzw. zwischen einer Eigenschaft einer höheren Ebene und deren mikrostrukturellen Verfasstheit zu den eigentlichen Fällen von Realisierung.

Einen solchen alternativen Realisierungsbegriff hat vor allem Carl Gillett (2002, 2003, 2003a, 2010, 2011) entwickelt und zusammen mit Kenneth Aizawa (Aizawa & Gillett, 2009) verteidigt. Nach Gillett kennzeichnet die von Kim oder (ursprünglich von) Shoemaker vertretene Realisierungstheorie der so genannte *Flat view* (also flache oder *eindimensionale* Ansicht), weil die hier herangezogenen Fälle von Realisierungen „vary in only one ontological dimension, involving different properties, but not distinct powers or individuals (Gillett, 2002, S. 317). Gillett kritisiert diese „flache“ Ansicht der Realisierung dahingehend, als dass unter ihren Voraussetzungen die eigentliche Kausalität bei den physischen Realisierern läge:

If we assume that the causal powers of individuals are not overdetermined, then appealing to Ockham's Razor the critic argues that we should accept the existence

of no more causally efficacious properties than we need to account for the causal powers of individuals. The proponent of this simple argument thus concludes that we should *only* accept that realizer/reasons are causally efficacious [...]. (ebd., 2003a, S. 36; Hervorh. im Orig.)

Das dem so ist, ergibt sich aus Gilletts Definition des Physikalismus, den er *de facto* als *Mikrophysikalismus* (vgl. Kap. 8.6.1) deutet: „All individuals are constituted by, or identical to, microphysical individuals, and all properties are realized by, or identical to, microphysical properties“ (ebd., 2003a, S. 36). Während er selbst die Konstitutions- und Realisierungsweiseart präferiert, hätte sich die „flache“ Realisierungstheorie hingegen für die *Identitätsvariante* des Physikalismus entschieden (denn ihr Begriff von Realisierung ist letztendlich eine Identitätsannahme auf dem Niveau von Eigenschaftsinstantiierungen), womit – nach Ansicht Gilletts – alle Eigenschaften, die es überhaupt gibt, letztendlich *mikrophysikalische* Eigenschaften, und alle Individuen, die es überhaupt gibt, letztendlich *mikrophysikalische* Individuen wären²³⁸. Damit wären mikrophysikalische Entitäten zugleich die einzigen Größen, die die Aussagen von Spezialwissenschaften wie der Psychologie wahr machen könnten, womit biologischen oder psychologischen Eigenschaften lediglich der Status von Prädikaten oder Begriffen zukäme (vgl. ebd., S. 37). Da all diese (vermeintlichen) Implikationen für Gillett inakzeptabel sind, möchte er einen alternativen *multidimensionalen* Ansatz der Realisierung entwickeln, „that allows for all dimensions that may be involved in cases of realization“ (ebd., 2002, S. 317). Gemäß der obigen Definition des Physikalismus entscheidet er sich hierbei für die Variante, dass alle Individuen durch

²³⁸ Dass diese Vermengung von Physikalismus und Mikrophysikalismus falsch ist, wird an späterer Stelle – im Zusammenhang der Besprechung der Reduktionismusdebatte – deziidiert nachgewiesen (vgl. Kap. 8.6.1).

mikrophysikalische Individuen *konstituiert* und alle Eigenschaften durch mikrophysikalische Eigenschaften *realisiert* werden.

Gillett stützt sich auf ein Bild der Realisierung, das von Lycan (2007) im Zusammenhang mit seiner Explikation des Ebenenbegriffs entwickelt wurde (vgl. Kap. 5.1.2.1) und eine Kette von Realisierungsrelationen zwischen Eigenschaften annimmt, die in *verschiedenen* Individuen instantiiert werden, die sich auf den *verschiedenen* natürlichen Organisations-ebenen (eines Objektes) befinden. Gilletts Intention ist dabei offensichtlich, ein Modell der Realisierung zu entwickeln, das als ontologische Fundierung des mechanistischen Ansatzes (vgl. Kap. 5.1.2.3) fungieren soll, indem es die Beziehung zwischen Entitäten und Eigenschaften erklärt, die auf den verschiedenen mechanistischen Ebenen vorliegen, die ein Phänomen kennzeichnen.

Er geht in seinen Arbeiten immer wieder auf ein Beispiel ein, das er als ein typisches Muster für viele von den Wissenschaften betrachtete Fälle ansieht. Es geht um einen Diamanten (oder Brillanten) s^* , der die Eigenschaft aufweist, extrem hart zu sein (H). Ein Diamant s^* hat als Konstituenten Kohlenstoffatome, wobei davon auszugehen ist, dass die einzelnen Kohlenstoffatome spezifische Eigenschaften haben bzw. in Relationen stehen, wie z.B. dass sie aneinander gebunden (*bonding*) und auf eine bestimmte Art und Weise angeordnet (*alignment*) sind. Die kausale Wirkmächtigkeit der Härte eines Diamanten (C) besteht darin, dass sie beispielsweise Risse in Glas verursachen kann. Gilletts Argument ist nun, dass die Wissenschaften eine sehr präzise und detaillierte Erklärung dafür abgeben, wie die Härte eines Diamanten aus den Eigenschaften und Beziehungen der einzelnen Kohlenstoffatome hervorgeht: „The very particular relations of alignment and bonding of the carbon atoms is such that they form a lattice-like structure in which the individual atoms have the characteristic power of causing other atoms to remain in a relatively

constrained location even under high temperatures and forces“ (ebd., 2002, S. 319). Hieraus resultiere die Härte eines Diamanten (H), da es aufgrund dieser Eigenschaften sehr schwierig ist, die relativen Positionen der Konstituenten eines Diamanten zu verändern. Für Gillett ist es damit klar, dass H nicht mit irgendeiner Eigenschaft bzw. irgendeiner Beziehung einzelner Kohlenstoffatome identisch sein kann, da H im Diamanten (als einem Ganzen) instantiiert wird, während die spezifischen Relationen der Bindung und Anordnung auf der Ebene seiner Teile, der Kohlenstoffatome, instantiiert werden. Daher würde die eindimensionale Realisierungstheorie à la Kim ausscheiden. Auch der Ansatz der eindimensionalen Realisierung als kausale Teilmengenbeziehung à la Shoemaker greife im vorliegenden Fall nicht, da der kausale Beitrag von H für s^* etwas ganz anderes als die kausale Wirkmächtigkeit ist, die durch die Relationen Bindung und Anordnung den Kohlenstoffatomen zukommt. „For example, the former contributes the power to cut glass, whilst the latter contribute the power to cause a contiguous carbon atom to remain in tight relative spatial range“ (ebd.). Wenn es also keine Identität (oder Teilmengenbeziehung Shoemakerscher Couleur) der Kausalkräfte ist, stellt sich die Frage, wie dann aus den Kausalkräften der Eigenschaften und Relationen der Kohlenstoffatome dennoch die Wirkmächtigkeit der Härte eines Diamanten hervorgehen soll (wobei außerdem die gegenteilige Ausrichtung der Beziehung ausgeschlossen werden soll). An dieser Stelle kommt seines Erachtens die Realisierungsrelation ins Spiel. Denn die Eigenschaften und Beziehungen der Kohlenstoffatome nehmen nach Ansicht Gilletts offensichtlich die kausale Rolle von H ein (und nicht andersherum), womit es folglich plausibel sei anzunehmen, dass H durch die Beziehungen und Eigenschaften der Kohlenstoffatome *realisiert* wird (vgl. ebd.). Allerdings sei es eben kein Fall einer Realisierung, wie diese vom *Flat view* oder der Standardansicht der Realisierung aufgefasst wird,

„since the properties/relations of the carbon atoms are instantiated in different individuals, and contribute *distinct* causal powers, from the properties of the diamond. The Flat view thus fails to cover a case of realization and it also erroneously classes the vast swathe of cases drawn from the sciences that are similar to this one” (ebd., S. 319f; Hervorh.: M.L.). Makroeigenschaften, wie die Härte eines Diamanten, könnten somit nicht in Beziehung zu Mikroeigenschaften, wie Bindung oder Anordnung funktionalisiert werden (vgl. ebd., 2003a, S. 39).

Aus diesen Gründen verwirft Gillett die Annahme, dass der Realisierer und die realisierte Eigenschaft gemeinsame (oder identische) Kausalkräfte haben und definiert seinen alternativen, multidimensionalen Ansatz der Realisierung wie folgt: „Property/relation instance(s) F_1-F_n realize an instance of a property G , in an individual s , if and only if s has powers that are individuating of an instance of G in virtue of the powers contributed by F_1-F_n to s or s 's constituent(s), but not vice versa” (Gillett, 2002, S. 322, Hervorh. im Orig.). Mit dieser etwas ungelassenen Definition möchte Gillett alle Fälle von Realisierungen berücksichtigen, sowohl also jene, die eine eindimensionale Relation betreffen, als auch jene, die sich auf ein multidimensionales Verhältnis beziehen (wobei der multidimensionalen Perspektive nach Ansicht Gilletts faktisch, also mit Bezug auf die von den Wissenschaften untersuchten Fälle, die größere Bedeutung zukommt). Während die „flache“ Realisierungsrelation, wie sie von Kim oder Shoemaker propagiert wird, die Beziehung zwischen einer Eigenschaft zweiter Ordnung und ihrem physischen Realisierer (als Eigenschaft erster Ordnung) thematisiert, fasst der multidimensionale Ansatz die Realisierungsrelation als *kompositorische Determinationsrelation* auf. Nach Ansicht von Aizawa und Gillett (2009) ist Realisierung offenkundig...

...a species of determination relation, but are rather different from causal relations. The 'horizontal' determination involved with causation is temporally extended, relates wholly distinct entities and often involves the transfer of energy and/or the mediation of force. In contrast, compositional relations are not temporal in nature, since their 'vertical' determination is instantaneous, does not relate wholly distinct entities, and does not involve the transfer of energy and/or the mediation of force. Composition is thus a variety of what has been termed 'non-causal' determination. (ebd., S. 187)

Die multidimensionale Realisierungstheorie sei von einer *nicht-reduktiven* Qualität (vgl. Gillett, 2003a, S. 42ff). Ähnlich wie Shoemaker möchte Gillett dies darüber gewährleisten, dass die mikrostrukturellen Eigenschaften und Relationen $P_1, P_2, P_3 \dots P_n$, die in den mikrophysikalischen Individuen $a_1, a_2, a_3 \dots a_n$ instantiiert sind, nur einen Teil ihrer konditionalen Kausalkräfte C^* dafür aufwenden, eine Instantiierung einer höherstufigen Eigenschaft H zu realisieren (die Kausalkräfte der niederstufigen Eigenschaften und Relationen sind konditional in dem Sinne, dass sie nur in diesem spezifischen Fall genau diese Kausalkräfte beitragen; in anderen Zusammenhängen könnte es sich ganz anders verhalten). Dass realisierte höherstufige Eigenschaften wie H eine eigene kausale Wirkmächtigkeit haben, begründet Gillett nun folgendermaßen:

The property H partially determines the contribution of a causal power to an individual, since P_1 only contributes C^* to individuals when realizing H . Assuming that P_1 is instantiated in a_1 , then a_1 's having C^* is accounted for only by ascribing the realized property H to an individual s that a_1 constitutes. The power C^* is contributed by the microphysical property P_1 , but the crucial point is that, in the particular circumstances, H is a necessary member of the properties which are only jointly sufficient for determining the contribution of C^* to a_1 . (ebd., S. 44)

Es würde daher *prima facie* einen Grund für die Annahme geben, dass realisierte Eigenschaften kausal wirksame Eigenschaften sind, weil sie den Beitrag an fundamentalen Kausalkräften, die einem Individuum zu-

kommen, teilweise determinieren. Denn dass P_1 seine konditionalen Kausalkräfte zum Individuum a_1 beiträgt, geht dieser Ansicht nach teilweise darauf zurück, dass P_1 H realisiert (was nach Gillett gleichbedeutend damit ist, dass H teilweise den kausalen Beitrag von P_1 determiniere). Die Wirkmächtigkeit realisierter höherstufiger Eigenschaften müsse man sich dabei ähnlich vorstellen wie die Wirkmächtigkeit räumlicher Relationen. Diese besitzen zwar selbst keine eigenen Kausalkräfte, die sie zu etwas beisteuern, aber sie könnten teilweise die Einbringung von Kräften in Individuen determinieren, die durch andere Eigenschaften oder Relationen erfolgt²³⁹ (vgl. ebd., S. 34f). Dennoch seien diese Annahmen grundsätzlich mit dem Physikalismus (bzw. der kausalen Geschlossenheit der physischen Welt) vereinbar, weil „the same fundamental microphysical forces or properties still exist when realized properties partially, non-causally determine their contributions of powers, and there would be no new non-physical forces, powers or properties” (ebd.).

Versuchen wir diesen immer noch sehr komplexen Zusammenhang an einem psychologischen Beispiel zu verdeutlichen: Meine Rot-Empfindung H als höherstufige Eigenschaft determiniert dieser Logik zufolge die sie realisierenden niederstufigen Eigenschaften und Relationen auf der neuronalen respektive mikrophysikalischen Ebene (P_1) teilweise dadurch, dass durch ihre Realisierung die Neuronen respektive Elementarteilchen (a_1), die durch P_1 gekennzeichnet sind, über die konditionale Wirkmächtigkeit C^* verfügen (die sie sonst nicht genauso hätten), die z.B.

²³⁹ Gillett beruft sich hierbei auf das so genannte *Eleatische Prinzip* Armstrongs (1997), das besagt, dass alles, was existiert, einen Unterschied in den Kausalkräften von irgendetwas ausmacht. Armstrong möchte damit sicherstellen, dass auch Phänomenen wie räumlichen Relationen eine Existenz zugesprochen werden kann.

dazu führt, dass die Neuronen feuern bzw. sich die Konfiguration der Elementarteilchen ändert²⁴⁰.

Auch Shoemaker (2007) fasst den multidimensionalen Ansatz mittlerweile als eine eigene (zusätzliche) Form der Realisierungsrelation auf, die er *Mikrorealisierung* nennt. Allerdings betont er, dass diese zwei Arten von Realisierung (in seiner Terminologie: Eigenschaftsrealisierung vs. Mikrorealisierung) kaum voneinander zu trennen sind: „But obviously these two sorts of realization are intimately related. If the instantiation of one property P realizes the instantiation of another property Q, it must be the case that the existence of the state of affairs that realized Q’s instantiation constitutively determines the existence of the state of affairs that realized Q’s instantiation – either because it is identical with it or because it contains it as a part” (ebd., S. 35).

Des Weiteren wird ein ähnlicher (kompositorische Prinzipien berücksichtigender) Vorschlag von Pereboom (2002) unterbreitet. Auch er weist Kims Prinzip der kausalen Vererbung zurück, da er befürchtet, dass die ontologische Eigenständigkeit mentaler Eigenschaften durch dieses gefährdet wird. Seinem Ansatz nach besteht anstatt eines Identitätsverhältnisses eine *Konstitutionsbeziehung* zwischen mentalen Eigenschaften und ihren physischen Realisierern: „If mental property *M* is realized in a system at *t* in virtue of physical realization base *P*, the causal powers of this instance of *M* are *wholly constituted* by the causal powers of *P*“ (ebd., S. 504; Hervorh.: M.L.).

²⁴⁰ Auch wenn dieses Kausalitätsverständnis einen gewissen Gehalt hat, stellt es die Kausalitätsbeziehung unnötig kompliziert dar, wie sich im folgenden Kapitel zeigen wird.

6.5.4. Stellenwert und Grenzen

Zunächst muss man festhalten, dass es ein großer Vorzug der Realisierungsrelation ist, dass sie sowohl als Antwortversuch auf die *ontologische* als auch *epistemische* Facette des Leib-Seele-Problems betrachtet werden kann. Die klassische „flache“ Version der Realisierung erweitert das von der Supervenienzrelation herausgestellte bzw. in kognitionswissenschaftlichen Modellen postulierte lediglich deskriptive psychophysische Abhängigkeitsverhältnis um die ontologische Annahme, dass mentale Eigenschaften durch physische Eigenschaften hervorgebracht werden, und zwar in dem Sinne, dass jede mentale Eigenschaftsinstantiierung mit einer physischen Eigenschaftsinstantiierung identisch ist²⁴¹. Eine Antwort auf die epistemische Frage, *warum* es gerade jene physische Eigenschaftsinstantiierung ist, mit der die mentale Eigenschaftsinstantiierung identisch ist, liefert die Realisierungstheorie über die funktionale Analyse mentaler Eigenschaften. Es ist jene physische Eigenschaftsinstantiierung und keine andere, weil genau sie es ist, die die kausale Rolle des funktionalen Profils der mentalen Eigenschaft einnimmt.

6.5.4.1. Problem der „Funktionalisierbarkeit“ phänomenaler Eigenschaften

Für kognitive Eigenschaften dürfte es nun in der Tat sehr gut möglich sein, ein solches Kausalprofil ausfindig zu machen. Schwierigkeiten liegen dagegen bei der Frage nach der generellen Verfügbarkeit von funktionalen Analysen *phänomenaler* Eigenschaften vor. Hier könnte

²⁴¹ Dies gilt für Kims Ansatz und im Grunde auch für den Ansatz Shoemakers. Obgleich Letzterer in Bezug auf die Kausalkräfte mentaler Eigenschaften nur von einer Teilmengenbeziehung ausgehen mag, besteht das „Wesen“ der Realisierung selbst dennoch in der Identität von mentalen und physikalischen Eigenschaftsinstantiierungen.

weiterhin die im Zusammenhang der Identitätstheorie erläuterte *Erklärungslücke* bestehen (vgl. Levine, 1983 bzw. Kap. 6.1.4.4 dieser Arbeit). Es ist das subjektive Erleben oder *Bewusstsein*, das in diesem Zusammenhang Schwierigkeiten bereitet.

Es muss jedoch die These hinterfragt werden, dass alltagssprachlich beschreibbares Verhalten dazu geeignet ist, den qualitativen Charakter phänomenaler Zustände zu erfassen (vgl. Pauen, 1999, S. 158). Obgleich es berechtigten Anlass zum Zweifel an vielen philosophischen Gedankenexperimenten gibt, werden diese in der Philosophie des Geistes gerade im Zusammenhang der Erklärungslückenproblematik häufig herangezogen. So wird insbesondere im Hinblick auf Farbempfindungen, bei denen man sich Personen mit fehlenden oder invertierten Qualia vorstellen könne, abgeleitet, dass viele Handlungsmuster, die das Ziel der funktionalen Analyse sind, nur sehr lose mit bestimmten phänomenalen Zuständen verbunden sind. Bezug nimmt man hierbei auf den Farbraum, dessen Abfolge der verschiedenen Farben (blau, grün, gelb, rot) sich aus der Reihenfolge der Farbe im Spektrum des sichtbaren Lichts ergibt (vgl. Abb. 27).



Abb. 27: Farbkreis (in Anlehnung an Goldstein, 2008, S. 158)

Bei invertierten Farbqualia wird dabei unterstellt, dass die Abbildung der Farbqualität sich unter Einhaltung der üblichen Beziehungen zwischen diesen Qualitäten gestaltet. Das Farbspektrum soll hierbei also nicht irgendwie, sondern um 180 Grad verschoben sein, sodass es zu einer Vertauschung von Komplementärfarben kommt (also z.B. Rot durch Grün vertauscht wird, wenn man sich auf das Gegenfarbenmodell nach Hering (1878) bzw. den hierauf aufbauenden Munsellschen Farbraum bezieht). Das Argument besagt nun jedenfalls, dass jemand mit solch einem invertierten Farbspektrum überall dort Grün sieht, wo etwas normalerweise als Rot wahrgenommen wird und umgekehrt, wobei derjenige mit einer roten Farbempfindung dasselbe kausale (Handlungs-) Profil zeigen wird, wie jemand (unter *ceteris paribus*-Bedingungen) mit einer grünen Farbempfindung. Nach Nida-Rümelin (2002) kann im Hinblick auf Farbempfindungen prinzipiell nicht ausgeschlossen werden, dass es verhaltensneutrale Inversionen des wahrgenommenen Farbspektrums gibt, so dass es auch einer womöglich „perfekten“ Wissenschaft der Zukunft wahrscheinlich verwehrt bleiben muss, zumindest einen Teil der phänomenalen Eigenschaften über Verhaltensmerkmale zu erfassen. Nicht einmal die betroffene Person selbst würde sich darüber bewusst werden können, dass bei ihr eine solche Inversion vorliegt, denn auf sprachlicher Ebene gebraucht sie dieselben Farbwörter wie eine „normale“ Person (nur ist eben das, was sie als „Rot“ bezeichnet, eine Farbe, die von „Normalsichtigen“ als „Grün“ bezeichnet würde und umgekehrt). Es stellt sich allerdings die Frage, ob man nicht doch erwarten könnte, dass sich die Personen mit einem normalen versus einem invertierten Farbspektrum bei z.B. einer Sortieraufgabe (nach Helligkeit, Sättigung) von entsprechenden Farbplättchen oder dergleichen anders verhalten würden, denn beispielsweise sind die Komplementärfarben Dunkelblau und Hellorange (siehe Abb. 27) in dieser Hinsicht so verschieden, dass

ihre Vertauschung bei solch einer Aufgabenstellung auffallen müsste. Damit würde dann aber auch ein Hinweis darauf vorliegen, dass sich diese Personen auch im Allgemeinen in ihrem kausalen (Handlungs-) Profil unterscheiden werden. Vielleicht mag dies im Alltag zwar nicht gleichermaßen zu Tage treten wie in dieser speziellen Experimentalsituation. Man könnte aber überlegen, ob nicht ganz andere emotionale Bewusstseinsfärbungen bei solchen Probanden mit einem invertierten Farbspektrum auftreten, denn z.B. sind die Farbeindrücke von Rot und Grün mit ganz anderen Kognitionen und Verhaltensintentionen verbunden (Erregung vs. Beruhigung), die nicht nur sozialisationsspezifisch erworben, sondern fest phylogenetisch verankert sein dürften. Insofern müsste sich aus dem invertierten Farbspektrum auch ein verändertes Verhalten ergeben. Das Argument des invertierten Farbspektrums weist aber bereits auf früherer Stufe unhaltbare Inkonsistenzen auf. So kritisiert Hardin (1987, 1997), dass eine Inversion des Farbspektrums dazu führen würde, dass durch die hiermit verbundene Verschiebung der Farben um 180 Grad gemischte Farben wie Pink oder Hellblau zu Primärfarben würden, während Rot und Blau nunmehr „gemischte“ Farben wären, was offensichtlich absurd ist. Möchte man dies vermeiden und führt nur eine Verschiebung um 90 Grad durch, sodass eine Primärfarbe die andere ersetzt, tun sich neue Probleme auf. Und zwar gibt es eine Besonderheit des Farbspektrums in Form einer gewissen Asymmetrie, die hierbei verloren ginge:

Es gibt zwei benachbarte Farben, Gelb und Orange, die einen anderen Farbton annehmen, wenn sie - relativ zum Kontext - abgedunkelt werden: Sie werden zu Braun. Da es hierzu bei anderen Farben keine Entsprechung gibt, ist keine Verschiebung des Farbspektrums möglich, die diese Besonderheit beibehält und gleichzeitig nur Primärfarben aufeinander abbildet. Die Spektrum inversion lässt sich also in der behaupteten Form nicht realisieren. (Pauen, 1999, S. 181)

Allerdings ist hiermit die Erklärungslücke keineswegs geschlossen. Denn auch wenn es einigermaßen zuverlässige Verbindungen zwischen den phänomenalen Qualitäten und bestimmten Verhaltensmustern gäbe, wäre es immer noch so, dass diese Verbindungen „offenbar nicht essenziell für die Erfahrung des Zustandes selbst [sind]“ (ebd., S. 158.) sind. Es mag so gesehen zwar typische Verhaltensweisen geben, die wir in einem Zustand von Freude ausführen, trotzdem sind es nicht diese Verhaltensweisen, die „Freude“ ausmachen, sondern ein charakteristisches *Erleben*, das als *Ursache* des gezeigten Verhaltens fungiert. In Bezug auf Schmerzen bringt dies Walter (2009) folgendermaßen auf den Punkt:

In genau dem Maß jedoch, wie sich Schmerz in verschiedenen Individuen unterschiedlich manifestiert, scheint es keine für die allgemeine Eigenschaft *hat Schmerzen* charakteristische Menge von Kausal Kräften zu geben. Menschen, die Schmerzen haben, stöhnen und rufen einen Arzt, Hunde winseln und kratzen sich, Tintenfische, sofern sie Schmerzen haben können, tun vermutlich wieder etwas anderes, und auch innerhalb einer Spezies scheint es kaum ein einheitliches kausales Profil einer mentalen Eigenschaft wie etwa *hat Schmerzen* zu geben. Letztlich ist das der Grund, warum funktionalistische Theorien intentionaler Eigenschaften plausibel sein mögen, funktionalistische Theorien phänomenaler Eigenschaften hingegen nicht. Phänomenale mentale Eigenschaften lassen sich ganz einfach nicht über ihre funktionale beziehungsweise kausale Rolle individuieren, und das bedeutet in letzter Konsequenz, dass sie sich nicht über Kausalkräfte individuieren lassen. (S. 701; Hervorh. im Orig.)

Auch wenn z.B. Pauen (1999, 2002) aufzeigt, dass hieraus nicht geschlossen werden kann, dass phänomenale Zustände *generell* nicht über funktionale Beschreibungen zu erfassen sind und für emotionale Zustände wie Furcht, Schmerzen und sogar Farbempfindungen nachweist, dass die gegenwärtigen wissenschaftlichen Modelle und Theorien der Psychologie und Neurowissenschaften durchaus einige Ansatzpunkte

für die Möglichkeit ihrer zumindest teilweisen funktionalen Analyse bieten, die darauf hoffen lassen, dass die zukünftige Forschung das nötige Potential zu ihrer Vervollständigung erlangen wird, besteht an einer vollständigen Auflösung dieser Problematik jedoch Grund zu zweifeln. So bleibt es nach z.B. Stephan (2002) fraglich, ob es überhaupt möglich ist, die Beziehung zwischen phänomenaler und neuronaler Ebene derart einsichtig zu machen, dass man aus dem wissenschaftlichen Zugriff in Form von funktionalen Analysen tatsächlich eine angemessene Erklärung phänomenaler Eigenschaften ableiten kann. Dass – wie Pauen es suggeriert – die funktionalen Implikationen phänomenaler Begriffe durch den Fortgang der empirischen Forschung zunehmend angereichert werden, bezweifelt Stephan, da es durch solche empirischen Erkenntnisse letztendlich nur zu einer Erweiterung unseres gegenwärtigen Wissensbestandes kommt, ohne dass damit ein Brückenschlag zu den erklärungsbedürftigen phänomenalen Qualitäten selbst vollzogen wäre. Was einer reduktiven Erklärbarkeit offenbar prinzipiell entgegensteht, ist also die fehlende Möglichkeit zu einer funktionalen Individuierung „des spezifischen *qualitativen Gehalts*“ (ebd., S. 360, Hervorh. im Orig.) phänomenaler Zustände. Wenn dem so ist, wären phänomenale Eigenschaften (epistemisch) synchron emergent (vgl. Kap. 6.4.1.3).

Kim selbst räumt ein, dass mit seiner physischen Realisierungstheorie die Konsequenz verbunden ist, dass man sich bezüglich der mentalen Eigenschaften, die sich einer Funktionalisierung entziehen, sprich: phänomenale Eigenschaften, zwischen der Annahme eines Epiphänomenalismus oder Eliminativen Materialismus entscheiden müsse, sodass man wieder die besagte Wahl zwischen „Pest und Cholera“ hätte (vgl. ebd., 1998, S. 119). Ohnehin würden sich die beiden Positionen aber kaum unterscheiden, da mit Verweis auf das Diktum Alexanders (1920)

sowieso alles, was keine kausale Wirkmächtigkeit hat, auf Dauer abgeschafft werden wird oder von vornherein gar nicht existiert hat.

Dass wir ein solches mentales (epiphänomenales) „Residuum“ in Form von Erlebnisqualitäten haben, sei aber für wissenschaftliche Belange und auch unser Selbstverständnis als Menschen nicht erheblich: „At least we can say this: If we can save intentional/cognitive properties, we can save our status as cognizers and agents. Saving itching isn't required for saving cognition or agency“ (Kim, 2005, S. 171). Phänomenale Eigenschaften (wozu neben dem „Jucken“, um das es in der Tat nicht sehr schade wäre, aber auch Farben, Töne oder Gefühle gehören) seien letztendlich nur so etwas wie Platzhalter: „without them there can be no qualia difference or similarities. Without content, there can be no form, no structure. You may now ask: Why are there just these qualia and not other possible ones? That remains a mystery“ (ebd., S. 173). Da Ähnlichkeits- und Unterschiedswahrnehmungen zwischen Empfindungsqualitäten „kausal gerettet“ werden könnten, würde zumindest Hoffnung dafür bestehen, dass das Haben von Bewusstsein nicht insgesamt kausal bedeutungslos ist. Was hingegen nicht (kausal) gerettet werden könne, ist dessen intrinsische Qualität (vgl. ebd., S. 174).

6.5.4.2. Die problematische Konzeption der eindimensionalen

Realisierung

Aber dies soll nicht der alleinige Grund sein, warum die Realisierungsrelation in der von Kim entworfenen Variante nicht für das zentrale Vorhaben dieser Arbeit, der Bestimmung des Verhältnisses von mereologisch-mechanistischen Ebenen unter Berücksichtigung des Psychischen, infrage kommt. Die Realisierungsrelation weist vielmehr eine Reihe grundsätzlicher begrifflich-konzeptueller Bestimmungen auf, die sich

primär aus den eingangs erläuterten Bedingungen ihres Aufkommens speisen und sie für ein solches Vorhaben disqualifizieren.

Zum einen muss man sich nochmals vor Augen führen, dass das Aufkommen der Realisierungsrelation primär eine Konsequenz des Argumentes der multiplen *Realisierbarkeit* an der Identitätstheorie war. Putnam, Fodor und andere haben an der Identitätstheorie bemängelt, dass sie dem Umstand nicht gerecht wird, dass ein und derselbe Typ von Schmerzen durch verschiedene physische Zustände *realisiert* werden kann. Ein Identitätsverhältnis zwischen mentalen und neuronalen Typen sei daher nicht haltbar. Durch diese entwicklungsgeschichtliche Verzahnung besteht nun aber weiterhin eine enge begriffliche Bindung des Realisierungsbegriffs an die „klassische“ Identitätstheorie, die in der heutigen Diskussion zumeist übersehen wird. Denn wenn kritisiert wird, dass Schmerzen entgegen der Identitätstheorie auf verschiedene Weise realisiert werden können, impliziert dies, dass die Identitätstheorie lediglich ein Spezialfall eines allgemeineren Konzeptes physischer Realisierung²⁴⁴ ist. Die Aussage der Identitätstheorie, dass eine Identität zwischen mentalen und physischen Eigenschaften besteht, wird auf diese Weise (implizit) gleichgesetzt mit der These, dass auch hier mentale Eigenschaften durch physische Eigenschaften *realisiert* werden, mit der einzigen Einschränkung, dass die klassische Identitätstheorie annimmt, dass es immer ein bestimmter Typ physischer Eigenschaften ist, der einen bestimmten Typ mentaler Eigenschaften *realisiert* (was diesen Spezialfall *empirisch* unplausibel macht).

²⁴⁴ Dass Shoemaker und einige andere eine kausale Teilmengenbeziehung gegenüber einer Identitätsbeziehung präferieren, ist an dieser Stelle irrelevant. Wie ich bereits zuvor dargelegt habe, sind alle Ausbuchstabierungen der Realisierungsrelation, die auf jedwede Annahme psychophysischer Identitäten verzichten möchten, äußerst fragwürdig.

Aufgrund dessen ist aber nun auch vieles von dem, was sich als problematischer begrifflicher Ballast der Identitätskonzeption der „klassischen“ Identitätstheoretiker herauskristallisiert hat, stillschweigend mit in den Realisierungsbegriff hinübergeflossen. So gilt weiterhin, dass mentale Eigenschaften dadurch „in die Welt gehoben“ werden, als dass sie eine *physische* Existenz haben. Das, was mentale Eigenschaften zu wissenschaftlich respektablen, ontologisch beständigen und kausal wirksamen Größen macht, ist in diesem Sinne ihre Hervorbringung durch (oder lokale Identität mit) physische(n) Eigenschaften. Dabei konstituiert das Psychische – entgegen den Annahmen dieser Arbeit – keine eigene natürliche Organisationsebene, sondern ist auf derselben Ebene angesiedelt wie seine neuronalen Realisierer. In gewisser Hinsicht gilt somit immer noch für die Realisierungstheorie, dass mentale Eigenschaften *nichts anderes als* physische Eigenschaften sind (wobei wir im Unterschied zur Identitätstheorie lediglich eine variierende physische Basis haben).

Die Identitätstheorie hatte den Vorzug, dass sie eine Erklärung für die Möglichkeit mentaler Verursachung bot. Wenn mentale Eigenschaften mit neuronalen Eigenschaften identisch sind, kann Psychisches qua seiner neuronalen Existenz kausal wirksam sein und stellt keine Bedrohung für die kausale Geschlossenheitsannahme der physischen Welt dar. In eben dieser Weise versucht nun auch die Realisierungstheorie weiterhin die Möglichkeit einer mentalen Verursachung zu retten. Der Grund dafür, dass mentale Eigenschaften trotz der kausalen Geschlossenheit der physischen Welt wirksam sein können, liegt gemäß Realisierungstheorie darin, dass sie selbst durch physische Eigenschaften in die Welt gehoben werden (egal ob nun durch eine vollständige oder teilweise Identität der Kausalkräfte). Dass ein Gedanke eine Wirkung in der Welt entfalten kann, geht einzig und allein auf die kausale Wirkmächtigkeit des (neuronalen) Realisierers ebendieses Gedankens zurück.

Ähnliches lässt sich sagen, wenn man die Bedeutung der Realisierungstheorie vor allem darin sieht, einen explanatorischen Ansatz dafür zu liefern, wie abstrakte, funktional charakterisierte Eigenschaften, wie sie von der Kognitionswissenschaft postuliert werden, ein Teil unserer physischen Welt sein können. Auch vor diesem Hintergrund betrachtet, zeichnet sich ab, dass es die neuronale Realität mentaler Zustände ist, durch die das (funktionalistisch interpretierte) Psychische einen „natürlichen“ Platz in der Welt erhält und kausal wirksam sein kann. Ein abstraktes, theoretisches²⁴⁵ Phänomen wird in diesem Sinne durch konkrete physische Gegebenheiten *realisiert*, wobei es unterschiedliche physische Systeme sein können, die es bewerkstelligen, den computationalen oder konnektionistischen Formalismus in die Welt zu heben. Mentale Eigenschaften sind demnach abstrakte Eigenschaften zweiter Ordnung, die man – zumindest programmatisch – mittels funktionaler Beschreibungen erfassen und deren eigentliche Existenz man ausfindig machen kann, indem man sich auf die Suche nach denjenigen physischen Realisierern als Eigenschaften erster Ordnung macht, die genau diese funktionale Rolle ausfüllen. Natürlich wird der (klassische) Funktionalist dem entgegen, dass es seiner Überzeugung nach gerade nicht die physische Seinsweise des Mentalen sein soll, die dessen eigentliches Wesen²⁴⁶ ausmacht, sondern dass dies vielmehr der abstrakten funktionalen Charakterisierung selbst zukommt, die sich in vielfältigsten physischen Systemen manifestieren kann. Dieses Bekenntnis ändert aber nichts an der Tat-

²⁴⁵ Keinesfalls wird von der Kognitionsforschung geleugnet, dass mentale Zustände eine bestimmte *subjektive* Erlebnisqualität aufweisen. Es bleibt nur fraglich, ob die abstrakt-theoretischen Beschreibungen mentaler Zustände (insbesondere phänomenaler Zustände oder Zustände kognitiven *Bewusstseins*) wirklich deren Realität einfangen können.

²⁴⁶ Dass der Versuch einer solchen Bestimmung der *Essenz* des Mentalen überhaupt fragwürdig ist, sei an dieser Stelle dahingestellt.

sache, dass es die physischen Realisierer sind, die Teil unserer Welt sind, während der Realitätsstatus einer abstrakten funktionalen Beschreibung äußerst fraglich bleibt (und etwas an den Platonischen Ideenhimmel erinnert). So bleibt es irgendwie mysteriös, was es überhaupt heißen soll, eine abstrakte Eigenschaft zweiter Ordnung zu sein und es drängt sich unweigerlich der Eindruck auf, dass mittels der Neueinführung einer solchen ontologisch recht nebulösen Größe eher eine Verschiebung des Problems als eine echte Lösung etabliert werden kann. Darüber hinaus nimmt aus guten Gründen mittlerweile kaum noch ein Funktionalist respektive Kognitionsforscher eine völlige Beliebigkeit der Realisierungsgrundlage an, wie sie für die frühe Kognitionswissenschaft üblich war (vgl. Kap. 4.2.1.1).

Die Realisierungsrelation ist problematisch, weil sie als eine Art „Ersetzungsrelation“ konzipiert ist. Auch wenn sogar gerade viele Philosophen (entgegen Kim) die Realisierungsrelation als ein Garant für eine Nicht-Reduzierbarkeit mentaler Eigenschaften ansehen möchten, da die physische Basis variieren kann bzw. sich die kausale Wirksamkeit mentaler Eigenschaften nur einer *Teilmenge* der kausalen Eigenschaften der physischen Realisierer verdankt, bleibt es dabei, dass Realisierung bei genauerer Betrachtung nur bedeuten kann, dass sich die Wirklichkeit und Wirkmächtigkeit einer Klasse von theoretischen Eigenschaften in der physischen Seinsweise erschöpft.

Was die Realisierungsrelation nach Maßgabe ihrer klassischen Formulierung also primär für das Vorhaben dieser Arbeit disqualifiziert, ist ihre Implikation, dass es sich beim Psychischen nicht um eine eigene natürliche Organisationsebene handelt, sondern nur um eine sekundäre Größe, die sich über ihre funktional beschriebene und physisch realisierte Rolle individuiert.

6.5.4.3. Die problematische Konzeption der multidimensionalen Realisierung

Die Varianten der Realisierungsrelation nach Gillett und Pereboom (bzw. die der Mikrorealisierung Shoemakers) sind im Kontext dieser Arbeit interessant, weil sie mit der angenommenen Kompositions- bzw. Konstitutionsbeziehung eine Verbindung zwischen der klassischen ontologischen Analyse des Leib-Seele-Problems (in Form der Realisierungsrelation) und dem mechanistischen Ansatz Bechtels und Cravers (vgl. Kap. 5.1.2.3) herstellen, der sich als beste Grundlage zur Individuation einer natürlichen Ebene des Psychischen herauskristallisiert hat. Die Frage ist, ob der multidimensionale Ansatz der Realisierung Gilletts tatsächlich als die gesuchte psychologische Basistheorie zur Erklärung der Beziehung zwischen den verschiedenen natürlichen Ordnungsebenen der Architektur des Geistes infrage kommt. Zunächst muss festgestellt werden, dass er prinzipiell Raum für die Annahme einer eigenen natürlichen Ebene des Mentalen lässt.

Anzumerken ist allerdings, dass die von Gillett in Bezug auf das Verhältnis von höherstufigen und niederstufigen Kausalkräften ins Feld geführte *in-virtue-of*-Relation relativ nebulös bleibt. Trotz einer Vielzahl an Publikationen, die er der näheren Erläuterung der multidimensionalen Realisierungsrelation gewidmet hat, bleibt letztendlich unklar, was genau damit gemeint ist, wenn er sagt, dass „powers that are individuating of an instance of G *in virtue of* the powers contributed by F₁-F_n...“ (Gillett, 2002, S. 322, Hervorh.: M.L.). Wenn die zentrale Bestimmung hier lautet, dass das Objekt, in dem eine Instantiierung einer Eigenschaft G realisiert ist, Kausalkräfte hat, die *kraft* (in virtue of) G selbst oder seiner Konstituenten individuierend für G sind, muss aufgezeigt werden, welche Art von Abhängigkeit zwischen Teilen und Ganzen damit gemeint ist. Dies bleibt Gillett weitestgehend schuldig. An dieser Stelle einfach den Realisie-

rungsbegriff einzuführen, wiederholt letztendlich nur das, was eigentlich genauer erklärt werden müsste, nämlich die *in-virtue-of*-Relation. Es wird etwas begrifflich eingeführt und gesetzt, was der eigentliche Gegenstand der Erklärung sein müsste. Hier erweist es sich insbesondere als Problem, dass Gillett die Bedeutung von Identitäten unterschätzt bzw. Identitätsbeziehungen zumeist (aufgrund einer ablehnenden Haltung gegenüber Reduktionen) vorschnell ablehnt.

Für mich bleibt zweifelhaft, ob Gillett mehr als nur ein bloßes Wortspiel betreibt, wenn er aus dem Umstand, dass eine Instantiierung einer mikrophysikalischen Eigenschaft/Relation P_1 eine höherstufige Eigenschaftsinstantiierung H realisiert, ableitet, dass H damit gleichzeitig teilweise auch determiniere, dass P_1 überhaupt einen bestimmten kausalen Beitrag C abliefern, und zwar weil P_1 eben nur dann C abliefern, wenn sie H realisiert (vgl. ebd., 2003a, S. 44f). Es wirkt sehr konstruiert, wenn man nun der Eigenschaft H die *Wirkmächtigkeit* zuordnet, kraft derer es dazu kommt, dass P_1 C entfaltet, nur weil man H ausschließlich einem Individuum s zuschreiben kann, dass eine mikrophysikalische Entität a_1 konstituiert, die über P_1 verfügt. Lässt sich hieraus tatsächlich eine eigene Kausalität höherstufiger Eigenschaften ableiten? Wohl kaum. Was genau ist der kausale Beitrag von H , wenn es letztendlich doch nur darum geht, dass ein System über H verfügen muss, damit es zu der geschilderten Beziehung kommt. Daraus, dass ein System über H verfügt, kann wohl kaum abgeleitet werden, dass H in diesem Fall einen *kausalen* Beitrag geleistet hat. Zumal es ohnehin sehr merkwürdig ist, dass Gillett einerseits vorgibt, dass realisierte Eigenschaften gemäß seinen Ausführungen eine *kausale* Wirkmächtigkeit haben, die sich dann allerdings als ein *nicht-kausaler determinierender* Einfluss herausstellt (vgl. ebd., S. 45). Auch wenn nicht alle Annahmen Gilletts grundsätzlich falsch sind, gelingt es

ihm nicht, ein klares und verständliches Kausalitätsmodell zu konzipieren.

Es stellt sich außerdem die Frage, ob das, was er beschreibt, wirklich als Realisierung bezeichnet werden kann und sollte. In diesem Sinne werfen Polger und Shapiro (2008) Gillett vor, dass er mit seiner Umwidmung der Realisierungsrelation eine Art Etikettenschwindel betreiben würde. Wie es bereits zuvor thematisiert wurde, insistieren sie darauf, dass die ursprüngliche Motivation für die Formulierung der Realisierungsrelation gewesen ist, einen Erklärungsansatz dafür zu schaffen, wie abstrakte, augenscheinlich nicht-physische, *funktional individuierte* Dinge wie Turing-Maschinen, Währungsumtäusche und eben auch mentale Zustände zu physisch akzeptablen Größen gemacht werden können (d.h. zu Größen, die sich trotz ihrer abstrakten Realität irgendwie in unser physikalisches Weltbild einpassen):

The idea was to show that there is a legitimate way of carving nature at joints other than or in addition to those carved by physics and chemistry. Realization is the special relation by which physical objects like brains can have “nonphysical” states, processes, or properties – perhaps those that are computational, economic, or mental. Physical objects can be instances of “nonphysical” kinds if being so is a matter of what physical objects can do rather than what they are made of. Such “functional” kinds are therefore not “nonphysical” in a way that requires the introduction of any new psychical substance. (ebd., S. 216)

Obgleich dies *de facto* nicht der einzige Hintergrund für das Aufkommen der Realisierungsrelation war, haben Polger und Shapiro insofern recht, als dass der Realisierungsansatz Gilletts tatsächlich eine entscheidende Ausgangsfrage des Realisierungsansatzes verfehlt, nämlich, wie eine abstrakte funktionale Beschreibung von etwas (z.B. Schmerzen) in der physischen Welt implementiert wird. Wenn man sich auf die funktionalistische Grundüberzeugung einlässt, dass mentale Eigenschaften funktio-

nale Eigenschaften sind, dann ist die Beziehung zwischen einer solchen funktional charakterisierten mentalen Eigenschaft und ihrem physischen Realisierer etwas, was offenbar nur mittels einer eindimensionalen (*intra-level*) Relation beschrieben werden kann. Ob man den funktionalistischen Grundgedanken nun gutheißt oder nicht, die Realisierungsrelation ist in diesem Sinne konzeptuell darauf festgelegt, die Beziehung zwischen einer funktionalen Rolle und der (physischen) Größe, die diese Rolle ausfüllt, zu beschreiben. Wenn Gillett dennoch seinen multidimensionalen (*inter-level*) Ansatz als Realisierung bezeichnet, scheint er damit einen Begriff mit einer bereits festgelegten Bedeutung für die Beschreibung eines ganz anderen Zusammenhangs zu missbrauchen (vgl. ebd., S. 219). Auch wenn ich denke, dass es einen bestimmten gesellschaftlichen Entwicklungen oder dem wissenschaftlichen Fortschritt geschuldeten Bedeutungswandel von (philosophischen) Fachtermini geben kann, was natürlich auch auf den Realisierungsbegriff zutreffen könnte, muss jedoch festgestellt werden, dass Gillett sich nicht auf einen solchen Vorgang beruft, sondern fälschlicherweise unterstellt, dass sich der Realisierungsbegriff schon immer auf „geschichtete kompositorische Hierarchien“, die den Wissenschaften entlehnt würden, bezogen hätte, während nur Kim auf die kühne Idee gekommen sei, Realisierung als Motor einer ontologischen Reduktion anzusehen (vgl. Gillett, 2007, S. 194).

Allgemein kritisieren Polger und Shapiro, dass es keinen nachvollziehbaren Grund für die von Gillett vorgenommene Bedeutungsverschiebung von der Rollen-Besetzungs-Relation zur Teil-Ganzes-Beziehung gäbe. Die ursprüngliche Absicht bei der Formulierung von der Realisierungsrelation sei es eben *nicht* gewesen, mereologische Beziehungen zu beschreiben. Des Weiteren würden Wissenschaftler und Philosophen bereits über einen ausdifferenzierten Begriffskatalog verfügen, um solche spezifischen mereologischen Abhängigkeitsbeziehungen zu beschreiben, so-

dass auch vor diesem Hintergrund der Ansatz Gilletts überflüssig sei: „Thus, it is no surprise that philosophers and scientists have availed themselves of concepts like composition, constitution, noncausal determination, and mereological supervenience in order to capture the relationships between wholes and parts. What does Gillett’s dimensioned realization add to this collection of relations?” (ebd., S. 219).

Was sich bei Gillett (2003, 2003a, 2011) bzw. Aizawa und Gillett (2009) in der Tat zeigt, ist, dass die multidimensionale Realisierung als eine (nicht-kausale) *kompositorische Determinationsbeziehung* (vgl. Aizawa & Gillett, 2009, S. 187) aufgefasst wird, bei der die Makroeigenschaften eines Systems durch dessen Mikroeigenschaften *determiniert* werden. Makroeigenschaften eines Systems werden in dem Sinne durch Mikroeigenschaften eines Systems realisiert, als dass die Eigenschaften der Entitäten auf der Mikroebene auf nicht-kausale Art und Weise die Kausalkräfte *beinhalten*²⁴⁷ (comprise), die für die Kausalkräfte der Eigenschaften der makroskopischen Entität individuierend sind (vgl. ebd., S. 186). Wenn der Begriff der Realisierung sich also auf kompositorische Verhältnisse verlagert und gleichzeitig keine Identitätsbeziehungen zwischen den betrachteten Größen zugelassen werden sollen, kann es tatsächlich auch wohl nur noch eine Art Determinationsbeziehung sein, die durch diesen Begriff gekennzeichnet wird. Wie ich in späteren Kapiteln mit Verweis auf Hüttemann (2004) nachweisen werde, ist ein solcher *Mikrophysikalismus* bzw. (einseitiger) *Mikrodeterminismus* aber inkonsistent. Stattdessen wird es darum gehen nachzuweisen, dass eine Identitätsannahme, die auf einer mereologisch-mechanistisch begründeten Koextensivität fußt, in diesem Kontext der zielführendere Ansatz ist.

²⁴⁷Wiederum bleibt unklar, was genau mit „beinhalten“ gemeint ist.

Es stellt sich ganz allgemein die Frage, ob der Ansatz, mentale Eigenschaften als funktionale Eigenschaften zweiter Ordnung aufzufassen, überhaupt korrekt ist. Wie in Kapitel 5 aufgezeigt, kann durchaus davon ausgegangen werden, dass das Psychische eine eigene natürliche Organisationsebene konstituiert und damit auch die Existenz mentaler Eigenschaften mit der Entfaltung einer mereologisch-mechanistischen Hierarchie bis hin zur psychischen Ebene zusammenhängt. Mentale Eigenschaften wären demnach nicht Eigenschaften zweiter Ordnung, die sich auf derselben Organisationsebene befinden wie neuronale Eigenschaften, sondern stellen genauso wie neuronale Eigenschaften Eigenschaften erster Ordnung dar. Genau genommen erscheint die gesamte Einführung von Eigenschaften unterschiedlicher Ordnungsgrade fragwürdig.

6.5.4.4. Die Analyse von Eigenschaftsrelationen ist ontologisch verfehlt

Polger und Shapiro (2008) tragen eine weitere interessante Kritik vor, die sie zunächst an Gillett richten, die aber genau genommen für den Realisierungsansatz überhaupt gilt, nämlich, dass die herangezogenen Bezugsgrößen der Realisierungsbeziehung inkohärent sind. Nach Maßgabe der Realisierungstheorie werden Eigenschaftsinstantiierungen durch andere Eigenschaftsinstantiierungen realisiert, was so nicht korrekt sein kann:

Of course to explain the realization relation we still need to talk about properties of brains and the properties of, say, pains. After all, the properties of brains help explain what brains do. But that does not entail that realization is a relation between properties themselves. A realizer of a property, state, or process P is a physical thing that meets the criteria that serve to identify instances of P, that does the P functions. Properties do not “do” anything, in this sense. [...] The simple point is that properties do not have, instantiate, or exemplify the relations characteristic of functional kinds, for example, causal relations. Properties do not stand in the relations that the

entities that instantiate them stand in. (ebd., S. 217, inkl. der Fußnote; Hervorh. im Orig.)

Insofern ist es genau genommen falsch, wenn man eine Realisierungsbeziehung so formuliert, dass sie auf beiden Seiten Eigenschaften als Relata aufweist. Die Realisierer sollten stattdessen in der Hinsicht *real* sein, als dass sie konkrete physische Vorkommnisse sind (vgl. ebd.).

Es ist daher auch nicht korrekt, wenn man annimmt, dass Eigenschaften selbst funktionale oder kausale Rollen einnehmen. Die Realisierung (mentaler) Eigenschaften setzt zwar die Analyse ihrer funktionalen Charakteristik voraus. Wenn eine solche funktional charakterisierte Eigenschaft realisiert wird, bedeutet das, dass es physische Größen gibt, die genau dieses funktionale Profil aufweisen. Es ist daher auch bei Kim und Shoemaker ein falscher Ansatz, wenn von ihnen die Beziehung zwischen *Eigenschaftsinstantiierungen* als Besprechungsgrundlage herangezogen wird (vgl. ebd., S. 218).

Des Weiteren könne man nach Polger und Shapiro (2008) auch nicht sagen, dass – wie Pereboom behauptet – Eigenschaften und ihre Kausalkräfte „aus etwas gemacht“ oder konstituiert sind (vgl. ebd.). Es sind immer irgendwelche Beschaffenheiten eines Systems, die charakteristisch sind für bestimmte Eigenschaften (wie es für Systeme charakteristisch ist, die Schmerzen haben, dass sie gereizte C-Fasern aufweisen).

Nimmt man an, was nach Polger und Shapiro ein trivialer Fakt ist, dass Eigenschaftsinstantiierungen nicht multipel realisiert werden können, lässt sich sogar eine innere Selbstwidersprüchlichkeit von Gilletts Ansatz nachweisen, da dieser zwangsläufig zur Annahme multipel realisierbarer *Instantiierungen* von Eigenschaften führt²⁴⁸. Ihr Argument ist, dass

²⁴⁸ Das genaue Argument lautet:

Gilletts Ansatz – insofern er Realisierung als Beziehung zwischen Eigenschaftsinstantiierungen konzipiert – mit dem Fakt inkompatibel ist, dass Dinge multipel realisiert werden können. Der Grund hierfür ist, dass *Instantiierungen* von Eigenschaften einmalig und zeitlich-räumlich lokalisiert sind. Damit sind sie Dinge, die nur zu einer Zeit an einem Ort auftreten können, und sind nicht wiederholbar. Wenn sich die Realisierungsrelation also streng nur auf *Instantiierungen* von Eigenschaften bezieht, wie Gillett es vorschlägt, dann besteht kein Raum mehr für die Möglichkeit multipler Realisierungen. Gillett kontert in späteren Arbeiten allerdings, dass es seiner Ansicht nach tatsächlich multiple Realisierungen von Eigenschafts*instantiierungen* geben kann (vgl. ebd., 2011). So sei es möglich, dass eine bestimmte Eigenschaftsinstantiierung von z.B. Schmerz multipel realisiert werde, wenn im zeitlichen Verlauf ihres Auftretens eine Veränderung in der zellulären Realisierungsbasis stattfindet. So könnte z.B. eine bestimmte Instantiierung von Schmerzen zu einem Zeitpunkt t_1 durch die C-Fasern 1, 2 und 3 realisiert sein, während es zu einem Zeitpunkt t_2 die C-Fasern 2, 3 und 4 sind, die ein und dieselbe Instantiierung von Schmerz realisieren. Obgleich auch hieran wiederum Kritik geübt wurde (z.B. Bradley, nicht-publiziertes Manuskript), soll dieser Aspekt an dieser Stelle nicht weiter verfolgt werden, da er für die vorliegende Arbeit nicht weiter relevant ist.

(P1) Everything that is realized is a property instance, and at least one property instance is realized. (Gillett's account)

(P2) Some things that are realized are multiply realized. (assumption)

(P3) No property instances are multiply realized. (trivial)

(C1) Some property instances are multiply realized. (from P1 and P2)

(C2) Not (P1). (reductio from P3 and C1) (Shapiro & Polger, 2008, S. 214)

6.6. Fazit

Es wurde gezeigt, dass das Konzept einer psychophysischen Identität *potentiell* als Grundlage einer gehaltvollen Theorie des Geistes fungieren kann, dass sich aber vor allem aufgrund ungünstiger Interpretationen der Identitätsrelation bzw. ihre Anbindung an bestimmte metaphysische oder epistemische Rahmenkonzepte ein sehr negatives Bild von psychophysischen Identitätsannahmen entwickelt hat, für das *der Sache nach* kein Anlass besteht.

Eine zentrale Ausgangsfrage dieses Kapitels war, ob die (metaphysischen) Relationen Supervenienz, Emergenz oder Realisierung als eine geeignete theoretische Grundlage für eine nähere Bestimmung der zuvor herausgestellten psychoneuralen Koextensivität fungieren können, die eine geringere „ontologische Nähe“ des Psychischen zu seinen neuronalen Grundlagen annimmt als eine Identitätsbeziehung. Wie sich nun gezeigt hat, ist es offensichtlich nicht möglich, auf Grundlage dieser Begrifflichkeiten eine konzeptuell konsistente und gleichzeitig inhaltlich gehaltvolle Theorie zur Beschreibung *und* Erklärung der Beziehung zwischen natürlichen (kompositorischen) Organisationsebenen zu etablieren. Vielmehr hat die genauere Betrachtung dieser Konzepte der psychophysischen Relation gezeigt, dass ohne die Hinzuziehung psychophysischer Identitätsannahmen keine gehaltvolle und konsistente Theorie der Gehirn-Geist-Beziehung formuliert werden kann.

Zunächst ist festzustellen, dass die untersuchten Gegenwartskonzeptionen der psychophysischen Relation explizit zumeist keine psychophysische Koextensivität annehmen. Eine solche ergibt sich in der hier vertretenen Form nur, wenn man eine Individuation des Psychischen über kompositorische Prinzipien vornimmt. Eine solche Herangehensweise

ist für die meisten Ansätze der Philosophie des Geistes allerdings untypisch, bei denen für gewöhnlich die Beziehung zwischen mentalen und neuronalen *Eigenschaften* als zwei verschiedenen Eigenschaftsklassen eines Systems analysiert wird, ohne dass Teil-Ganzes-Beziehungen hierbei explizit beachtet werden (implizit scheinen zumeist *aggregative* Organisationsprinzipien angenommen zu werden, die dann in Verbindung mit Identitätsannahmen in der Tat zu problematischen reduktionistischen Implikationen führen).

Die *Supervenienzrelation* stellt für sich genommen keine eigenständige Gehirn-Geist-Theorie dar. Ihre Grundaussage ist lediglich, dass es keinen Unterschied (oder keine Veränderung) auf der mentalen Ebene geben kann, ohne dass ebenso ein Unterschied (oder eine Veränderung) auf der physischen (neuronalen) Ebene vorliegt. Erweitert man dies um den Gedanken eines aus verschiedenen natürlichen (kompositorischen) Organisationsebenen konstituierten Bezugsobjektes, erhält man zwar das an sich interessante Prinzip einer *mereologischen* Supervenienz, das besagt, dass wenn zwei Menschen auf ihren verschiedenen Mikroebenen ununterscheidbar sind, sie auch die gleichen makroskopischen Eigenschaften (den gleichen mentalen Zustand) aufweisen müssen. Dennoch liefert die Supervenienztheorie einen bloß *deskriptiven* Zusammenhang zweier Variablen, ohne eine Erklärung der von ihr herausgestellten Korrelationen bzw. Abhängigkeitsverhältnisse anzubieten. Sie ist daher sowohl mit der gesamten Bandbreite physikalistischer als auch dualistischer Positionen (wie dem Epiphänomenalismus) vereinbar.

Ganz ähnlich verhält es sich im Hinblick auf das Prinzip der *Emergenz*. Zunächst gilt, dass die Annahme emergenter Systemeigenschaften einen wichtigen Aspekt im Hinblick auf Systeme aufzeigt, die durch mereologische Ebenen charakterisiert sind. So weisen die Entitäten höherstufiger Ebenen Eigenschaften auf, die *einzelnen* Entitäten der niedrigeren Stufen

ein und desselben Systems, fehlen. Dies ist in der Tat eine wichtige Feststellung. Mit der Möglichkeit zur Etablierung einer Gehirn-Geist-Theorie auf der Basis der Emergenztheorie hat sie hingegen nichts zu tun. Denn wie sich gezeigt hat, ist es lediglich die starke (synchrone) Emergenz, die den Status einer aussagekräftigen und inhaltlich konsistenten Position haben könnte. Hierbei handelt es sich allerdings um eine *epistemische* (und keine ontologische) Bestimmung der Emergenz. Diese zielt also nicht darauf ab, die synchrone Determiniertheit höherstufiger Eigenschaften abzulehnen, sondern nur darauf, dass eine solche Determiniertheit nicht einsichtig gemacht werden kann (obgleich sie *de facto* durchaus vorhanden sein mag). Das Aufweisen von psychischen Zuständen kann in diesem Sinne nicht durch den Rekurs auf deren (durch verschiedene kompositorische Ebenen gekennzeichnete) physische Grundlagen *erklärt* werden. Ontologisch bleibt die Emergenzrelation indifferent. Sie ist durchaus mit Theorien, die psychophysische Identitätsannahmen beinhalten, kompatibel. Insofern kann der Emergentismus nicht als Alternative zu einer (kompositorisch begründeten) Identitätstheorie fungieren, wie sie im nächsten Kapitel entwickelt wird. Wie auch der Supervenienzrelation mangelt es der Emergenzrelation an einer *positiven* Darstellung, einer Erklärung des von ihr betrachteten Zusammenhangs und kann daher auch nicht zu einer eigenständigen Theorie des Geistes reichen. Obgleich die Realisierungstheorie augenscheinlich eine vielversprechende, explanatorisch gehaltvolle Position ausmacht, gibt es dennoch bestimmte Pole, zwischen denen sie sich zu zerreißen scheint: Beschreibt sie gemäß ihres ursprünglichen Zwecks die Beziehung zwischen einer funktional charakterisierten Eigenschaft zweiter Ordnung und ihrem physischen Realisierer, wird sie zu einer Art „Ersetzungsrelation“, die eine eindimensionale (intra-level) Beziehung zwischen lediglich theoretisch-abstrakten Größen und der physischen Realität herstellt und bei der

unklar ist, ob sie überhaupt bei phänomenalen Eigenschaften (inkl. Zuständen kognitiven Bewusstseins) greift. Die Konstitution einer eigenen natürlichen Organisationsebene des Psychischen ist unter ihren Voraussetzungen offenbar ausgeschlossen. Wird der Realisierungsbegriff zu einem Begriff umetikettiert, der sich auf kompositorische Verhältnisse bezieht und gleichzeitig Identitätsannahmen ausschließt, mutiert er zu einer mikrodeterministischen Position, bei der einerseits hinterfragt werden kann, ob sie begrifflich überhaupt adäquat als Realisierung betrachtet werden kann und die andererseits konzeptuell inkonsistent zu sein scheint. Würde man Identitäten in letzterem Fall wiederum zulassen, stellt sich die Frage, ob man den Realisierungsbegriff überhaupt benötigt. Denn auf diese Weise sind es lokale Identitäten, die zwischen den ontologischen Bezugsgrößen (Entitäten, Eigenschaften oder Kausalkräften) der verschiedenen Organisationsebenen bestehen, die über das Anwenden des Realisierungsbegriffes auch nicht näher charakterisiert werden als über den Identitätsbegriff selbst (natürlich kann man den Fehler begehen und Identitäten als spezielle Form von Realisierungen betrachten und damit natürlich auch im Falle von solchen lokalen Identitäten von *realisierten* Größen sprechen; nur ist damit der Sache nach nichts gewonnen, außer dass man das Ausmaß an begrifflichen Konfusionen auf dieser Welt unnötig gesteigert hat).

Insbesondere ist die spätestens mit dem Aufkommen des Argumentes der multiplen Realisierbarkeit verbreitete Deutung der Identitätstheorie als Variante physischer Realisierung – so schlüssig und harmlos sie auch erscheint – keineswegs unproblematisch. Die Verbindung von Identität und Realisierung mag zwar begrifflich nicht falsch bzw. inkonsistent sein, aber sie ist keineswegs *notwendig*. Dass sie dennoch unreflektiert als Tatsache aufgefasst wurde (und wird), hat zu einer einseitigen Leseart sowohl der Identitäts- als auch der Realisierungsrelation geführt. Denn es

lässt sich eine psychophysische Identität auch auf ganz andere Weise (mit Rückgriff auf kompositorische Prinzipien) begründen, die ohne die Annahme einer physischen Realisierung auskommt.

Die Realisierungstheorie ist insofern der falsche Ansatzpunkt, um die Beziehung zwischen der mentalen und den verschiedenen physischen Ebenen verständlich zu machen. Sie leitet sich aus einem einseitigen Verständnis, einem auf Smart (1959) zurückgehenden Relikt der Identitätstheorie ab, das besagt, dass es sich bei der „strikten“ Identität respektive Realisierung um eine Art „Ersetzungsrelation“ handelt. Eine mentale Eigenschaft(sinstantiierung) *ist nichts anderes als* eine physische Eigenschaft(sinstantiierung). Mentale Eigenschaften als Eigenschaften zweiter Ordnung werden durch physische Eigenschaften in die Welt gebracht.

Dabei scheint die Relation einer kompositorisch begründeten Identität eine viel geeignetere Position zu sein. Es ist eine Identität, die nicht als Spezialfall der Realisierungsrelation angesehen werden sollte, sondern eine, die sich durch die notwendige Koextensivität von kompositorisch verfassten Objekten ergibt, also solchen, die in einer Teil-Ganzes-Beziehung stehen. Es besteht hier ein ganz anderer Grund, eine Identität zu etablieren. Die Intention ist nicht mehr zu zeigen, worin das *Wesen* des Psychischen besteht, indem sich mentale Eigenschaften als neuronale Eigenschaften entpuppen, sondern dass das Psychische in einer natürlichen Gliederung der Welt, die sich durch Teil-Ganzes-Beziehungen ergibt, eine eigene Komplexitätsstufe konstituiert, und über diese Rolle als höherstufige Ganzheit gegenüber ihren neuronalen (chemischen, atomaren, subatomaren...) Grundlagen in einem koextensiven Verhältnis zu ebendiesen stehen muss.

Wenn im Folgenden an der einen oder anderen Stelle nichtsdestotrotz von *realisierten* Eigenschaften oder *realisierten* mentalen Leistungen und dergleichen die Rede ist, ist hiermit keineswegs ein Bekenntnis zur Reali-

sierungstheorie und ihren Voraussetzungen intendiert. Dies ist vielmehr lediglich der sprachlichen Schwierigkeit (oder vielleicht meiner diesbezüglichen Einfallslosigkeit) geschuldet, ein solches Phänomen anders als „realisiert“ oder „hervorgebracht“ zu benennen.

7. Die Architektur des Psychischen: Eine mereologisch-mechanistisch fundierte psychophysische Identitätskonzeption

Spreche ich von meinen Empfindungen, so sind dieselben nicht räumlich in meinem Kopfe, sondern mein ‚Kopf‘ theilt vielmehr mit ihnen dasselbe räumliche Feld.
(Ernst Mach, 1886/1903, S. 22)

Ziel der folgenden – das Herzstück der vorliegenden Arbeit ausmachenden – Darstellung soll es sein, die Befunde der vorangegangenen Kapitel in eine begrifflich kohärente und empirisch plausible Theorie im Sinne eines Modells einer allgemeinen *Architektur des Geistes* zu integrieren, die es vermag, als ontologische bzw. begrifflich-konzeptuelle Basis der Psychologie zu fungieren. Es wird hierbei untersucht, ob es unter der Berücksichtigung kompositorischer Verhältnisse möglich ist, eine Theorie psychophysischer Identität zu postulieren, die zwar annimmt, dass zwischen mentalen Prozessen oder Eigenschaften und ihren physischen Grundlagen ein Identitätsverhältnis besteht, ohne jedoch ein krudes *Nichts-anderes-als-Verhältnis* unterstellen zu müssen, wie es die „klassische“ Identitätstheorie oder auch die Realisierungstheorie (in ihrer begrifflich konsistentesten Ausformulierung) suggerieren.

Um nicht den Überblick zu verlieren, erscheint es zunächst allerdings angezeigt, die bisherigen teilweise weit verzweigten Argumentationspfade und Befunde – in vereinfachter Form – zusammenzutragen und hieraus die Eckpfeiler der zu erfolgenden Präzisierung des hier entwickelten Ansatzes abzuleiten. Diese Vorbedingungen oder Constraints sind also bei der Ausformulierung einer ontologischen Basistheorie der Psychologie zu berücksichtigen. Sie sind gewissermaßen die zentralen und man-

chmal noch zu entwirrenden *Fäden des Weltknotens* (in der Schopenhauer-schen Bezeichnung für das Leib-Seele-Problem):

1. Das Psychische, auf das sich psychologische Konzepte und Theorien beziehen, konstituiert *keinen substantiellen* Wirklichkeitsbereich, der der „physischen Wirklichkeit“ in irgendeiner Form gegenübersteht (Ablehnung eines Substanzdualismus).
2. Die phänomenale Differenz zwischen dem Psychischen und seinen physischen Grundlagen geht auf die Repräsentation unterschiedlicher (und dennoch koextensiver) natürlicher Organisationsebenen des menschlichen Organismus zurück. Gehirn und Geist sind „Trigger“ divergierender konzeptueller Formen unseres kognitiv-perzeptuellen Systems.
3. Es besteht eine *Teil-Ganzes-Beziehung* mereologisch-mechanistischer Art zwischen psychischen und neuronalen Prozessen.
4. Das Psychische ist aufgrund dessen mit seinen physischen Grundlagen räumlich-zeitlich *koextensiv*.
5. Das Psychische ist dennoch bzw. gerade deshalb ein *reales* bzw. natürliches Phänomen.

Auch wenn in Kapitel 5 bereits aufgezeigt wurde, dass es grundsätzlich *möglich* ist, psychische Prozesse als etwas zu betrachten, das in eine natürliche Strukturgliederung des menschlichen Organismus eingepasst ist und hier eine eigene höherstufige natürliche Organisationsebene konstituiert, wurde noch nicht geklärt, *wie genau* man sich dies vorzustellen hat. Inwiefern ist dies also nicht nur *möglich*, sondern auch in begrifflich kohärenter und empirisch plausibler Weise verständlich zu machen, dass eine solche „Einpassung“ auf der Basis der getätigten Feststellungen *tatsächlich gegeben* ist. Es wird nachgewiesen, dass mentale Eigenschaften als

mikrobasierte Struktureigenschaften des menschlichen Organismus anzusehen sind. Neben ihrer psychischen (kognitiven und/der phänomenalen) Struktur ist eine mentale Eigenschaft durch eine (potentiell unendliche) Anzahl von koextensiven Mikrokonfigurationen gekennzeichnet. Aufgrund der ungenügenden ontologischen Eindeutigkeit einer Koextensivitätsbeziehung, die eine weitergehende Bestimmung notwendig macht, sowie im Angesicht der Tatsache, dass die als Alternativen infrage kommenden Konzepte der Supervenienz, Emergenz oder auch der physischen Realisierung – in ihren „schwächeren“ Formulierungen – prinzipiell keine ontologisch aussagekräftige Bestimmung der Gehirn-Geist-Relation vornehmen können bzw. – in ihren „stärkeren“ Formulierungen – grundsätzliche begriffliche Inkonsistenzen oder eine Inkompatibilität zu kompositorischen Ansätzen aufweisen, wird argumentiert, dass man diese Koextensivität als eine bestimmte Form psychophysischer *Identität* betrachten sollte. Es wird zugleich verdeutlicht, warum diese – mereologisch-mechanistisch basierte – Art einer Identitätsrelation nicht mit den negativen Konsequenzen verbunden ist, die für die „klassische“ Identitätstheorie (oder auch die identitätsbasierte Realisierung à la Kim) kennzeichnend sind. Es handelt sich hierbei um keine *Identifizierung* des Psychischen *als* etwas Neuronales, vielmehr stellt die neuronale Ebene nur eine unter vielen *Mikrostrukturen* dar, *mit* der das Psychische *identifiziert* werden kann. Auch wenn mentale Eigenschaften mikrobasierte Eigenschaften sind, bleibt ihre psychische Struktur für sie maßgeblich. Diese *strukturelle* Bestimmung einer mentalen Eigenschaft führt dabei nicht zu einer Diskrepanz zu funktionalistischen Theorien des Geistes, wie sie für die Kognitionswissenschaft charakteristisch sind. Wie bereits Aristoteles wusste, stellen Form (Funktion) und Materie (Struktur) eine unzertrennliche Einheit dar, was nun im Hinblick auf den hier verfolgten mereologisch-mechanistischen Ansatz dargelegt werden kann (im Kern:

die Struktur einer bestimmten Organisationsebene eines Mechanismus ist immer schon eine „geformte Materie“ und zwar so geformt, dass sie eine bestimmte Funktion erfüllen kann).

Es muss des Weiteren untersucht werden, inwiefern die Möglichkeit zu einer psychischen Kausalität bzw. mentaler Verursachung unter den hier getroffenen Voraussetzungen tatsächlich gewährleistet werden kann. Obgleich aufgezeigt wird, dass es genau genommen keine Top-down-Kausalität (aber genauso wenig Bottom-up-Kausalität) in kompositorisch verfassten Systemen gibt, sondern etwas, das man als „mechanistisch vermittelte Effekte“ bezeichnen kann, wird sich zeigen, dass das Psychische dennoch ein *kausal wirksames* Phänomen ist. Kausalität vollzieht sich in kompositorisch verfassten Systemen ebenenübergreifend, wobei das Psychische als „ebenendurchdringendes“ Phänomen in vielen Fällen für diesen Kausalitätsvollzug maßgeblich ist. Auch wenn dies nicht zu einem zentralen Themenschwerpunkt erhoben wird, soll in diesem Kontext zumindest in Form eines kurzen Abrisses verdeutlicht werden, warum die hier entwickelte Identitätstheorie eine *Handlungs- und Entscheidungsfreiheit* des Menschen keineswegs untergräbt, sondern untermauert.

Wenn aus der psychophysischen Koextensivität eine Identitätsannahme abgeleitet wird, muss sich diese, egal, ob sie sich nun in vielerlei Hinsicht von der „klassischen“ Identitätstheorie absetzen mag, mit den zentralen Einwänden gegen die Identitätstheorie, den Argumenten der multiplen Realisierbarkeit, der starren Designatoren und auch der explanatorischen Lücke, auseinandersetzen oder zumindest aufzeigen, warum das hier vorgestellte Identitätskonzept nicht von diesen Einwänden betroffen ist. Im Zuge dessen muss vor allem geklärt werden, ob die hier vorgestellte Konzeption überhaupt den Anspruch erheben kann, über eine Identität auf dem Niveau von Vorkommnissen bzw. Eigenschaftsinstantiierungen hi-

naus auch eine solche auf dem Niveau von Typen bzw. Eigenschaften (als Gattungsbegriff) etablieren zu können.

7.1. Geist, Eigenschaft und Struktur

Wenn zuvor die Koextensivität des *prozessualen* Geschehens auf den verschiedenen natürlichen Strukturebenen behauptet wurde und gesagt wurde, dass diese als eine bestimmte Form psychophysischer Identität zu interpretieren ist (vgl. Kapitel 5), muss doch konstatiert werden, dass hierfür noch keine weitergehende Erklärung abgeliefert wurde. Auch sind bislang mentale Eigenschaften nur wenig in die Diskussion einbezogen worden, obgleich sie – wie im letzten Kapitel aufgezeigt wurde – *die* „ontologische Basiskategorie“ in der philosophischen Gegenwartsdebatte sind²⁵¹. Es wird sich zeigen, dass beide Aspekte zusammenhängen und in ihrer gemeinsamen Betrachtung die bestehende Kluft überbrückt werden kann.

7.1.1. Der Strukturbegriff von Eigenschaften

In punkto der Analyse von Eigenschaften ist man zunächst wieder mit einer weiteren Spielart von „Spinozas Dilemma“ konfrontiert. Die Frage ist, inwieweit Eigenschaften überhaupt mehr als bloße *Attribute* im Sinne von Kennzeichnungen eines Objektes sind, die diesem lediglich „von außen“ beigelegt werden. Wenn es nicht eine bloße „subjektive“ Beile-

²⁵¹ Das hiermit zusammenhängende Problem, dass eine Identität von Eigenschaften bloße *aggregative* Verhältnisse impliziert, greift dabei nicht, da in dem hier verfolgten Ansatz explizit auf mereologisch-mechanistische Individuationskriterien natürlicher Organisationsebenen Bezug genommen wird.

gung ist, was genau an einem Objekt ist es dann, was zu der berechtigten Zuschreibung von bestimmten Eigenschaften führt?

Natürlich gilt auch hier, dass Eigenschaften in der Tat zunächst als „Zuschreibungen“ im Sinne psychischer Akte interpretiert werden sollten und damit unweigerlich etwas mit unserer eigenen kognitiven (emotionalen...) Verfasstheit zu tun haben. Insofern gibt es zweifelsohne Eigenschaften, die wir bestimmten Systemen zuschreiben, die diesen offensichtlich nicht selbst inhärent sind (man denke z.B. an die Versuche von Fritz Heider mit den Videoaufnahmen von sich bewegenden geometrischen Figuren, denen von den Probanden psychische Eigenschaften zugeschrieben wurden wie: „möchte die Nähe des anderen aufsuchen“, „weist ihn von sich“ und dergleichen; vgl. Heider & Simmel, 1944).

Die Frage ist also auch hier, wann man – statt von einer bloßen „Attribution“ von etwas – von Eigenschaften sprechen kann, die etwas herausgreifen, was auf eine tatsächliche Beschaffenheit des betreffenden Objektes zurückgeht. In Bezug auf den bisher erarbeiteten Kenntnisstand wird im Folgenden in einem ersten Schritt die These verfolgt, dass die faktisch vorhandenen Eigenschaften eines Systems Kennzeichnungen der strukturellen Verfasstheit oder *Konfiguration* der natürlichen (kompositorischen) Organisationsebenen darstellen. Die einem System selbst zukommenden Eigenschaften sind nichts anderes als dessen (mikro-)strukturelle Verfasstheit. So wie die Eigenschaft, Wasser zu sein, letztendlich darin besteht, die Mikrostruktur „H₂O“ zu besitzen, gilt für (beobachterunabhängige) Eigenschaften oder Eigenschaftsinstantiierungen eines Systems überhaupt, dass sie letztendlich mit der strukturellen Verfasstheit eines Systems (zu einem bestimmten Zeitpunkt) zusammenfallen.

Ein Beispiel wäre die Eigenschaft einer Person „sportlich zu sein“, die auf körperliche Verfasstheit (und das hiermit verbundene Vermögen) einer Person zurückgeht. In gleicher Weise gilt dies aber nun auch für mentale

Eigenschaften. So beziehen wir uns gemäß des hier verfolgten Ansatzes ebenfalls auf die *strukturelle*²⁵² Verfasstheit einer Person, wenn wir sagen, dass eine Person „nachdenklich“ oder „euphorisch“ ist und hiermit einen tatsächlich vorhandenen Fakt herausgreifen. Natürlich sind Eigenschaften zumeist relational, insofern als sich solche Eigenschaften nur durch den Vergleich mit anderen Personen in ihrer berechtigten Anwendung rechtfertigen lassen (es ist schwer zu sagen, dass es für eine Person *absolut* kennzeichnend ist, nachdenklich oder euphorisch zu sein). So gilt dann aber auch, dass sich zwei Personen, die sich faktisch in ihren mentalen Eigenschaften unterscheiden, auch in ihrer mikrostrukturellen Konfiguration (auf den verschiedenen natürlichen Organisationsebenen) unterscheiden müssen (mereologische Supervenienz).

Auch wenn die funktionalistische Auffassung von Eigenschaften zunächst vernachlässigt wird, heißt dies nicht, dass sie in diesem Zusammenhang unerheblich ist oder funktionalen Eigenschaften die Realität abgesprochen wird. Im Gegenteil: in einem zweiten Schritt soll aufgezeigt werden, dass der hier verfolgte Ansatz durchaus kompatibel mit der funktionalistischen „Weltsicht“ ist, obgleich die herkömmliche Realisierungsrelation modifiziert werden muss bzw. zugunsten einer lokalen (mereologisch-mechanistisch begründeten) Identität zu weichen hat. Struktur und Form (Funktion) lassen sich nicht gegeneinander ausspielen, sondern sind untrennbar miteinander verwoben, sodass es unsinnig wäre, eine Entscheidung *zwischen* einer identitätsbasierten oder funktionalistischen Theorie des Geistes treffen zu wollen.

²⁵² Mit einer solchen *Struktur* ist keineswegs gemeint, dass es die neuronale Struktur ist, auf die wir uns beziehen. Gemeint ist hier nur, dass eine Person, auf die diese Eigenschaft faktisch zutrifft, eine bestimmte Verfasstheit aufweisen muss, die diese Aussage wahr macht.

7.1.2. Kims Ansatz mikrobasierter Eigenschaften

Ein vielversprechender Ausgangspunkt sind die Ausführungen Kims (1998, 2003) zu *mikrobasierten* Eigenschaften, die als Makroeigenschaften eines Systems in ihrer Beschaffenheit auf dessen Mikrostruktur zurückgehen sollen. Nach Kims Darlegung zur Abwendung eines Abdriftens der Kausalität zu den niedrigeren Ebenen bzw. zur Basisebene sind die verschiedenen Mikrokonfigurationen im Sinne einer (potentiell) unendlichen Abfolge von kompositorischen Ebenen, die eine solche mikrobasierte Eigenschaft kennzeichnen, reduktiv identisch miteinander. Kim erläutert dies zwar im Zusammenhang seiner Realisierungstheorie, dennoch macht die Beziehung zwischen mikrobasierten Eigenschaften und ihren verschiedenen mikrostrukturellen Konfigurationen für ihn gerade *keinen* Fall einer physischen *Realisierung* aus.

Der Hintergrund für die Einführung mikrobasierter Eigenschaften ist bei Kim folgender: Wie erläutert, widerspricht Kim der Mikro-Makro-Unterscheidung im Hinblick auf mentale Eigenschaften als Eigenschaften zweiter Ordnung und ihrem Verhältnis zu ihren physischen Realisierern. Es ist immer eine (physische) Eigenschaft erster Ordnung, die das kausale Profil, über das die mentale Eigenschaft zweiter Ordnung gekennzeichnet ist, erfüllt. Woraus Kim folgert: „Consequently, when we talk of second-order properties and their realizers, there is no movement downward, or upward, in the hierarchy of entities and their properties ordered by the micro-macro relation“ (vgl. ebd., 1998, S. 82). Die über eine funktionale Analyse erfassbaren mentalen Eigenschaften sind auf der gleichen Ebene angesiedelt wie ihre physischen Realisierer, denn wie Kim immer wieder betont, besteht das Aufweisen eines bestimmten Schmerzzustandes *in nichts anderem als* dem Aufweisen eines bestimmten physischen (bzw. neuronalen) Zustandes. Es besteht eine psychophysische Identität auf dem Niveau von Eigenschaftsinstantiierungen.

Wenn die von Kim propagierte Realisierungstheorie also eine eindimensionale Beziehung zwischen einer funktionalen Eigenschaft zweiter Ordnung und ihrem physischen Realisierer beschreibt, stellt sich die Frage, wie dies mit der offenkundigen Existenz einer mereologischen und damit *multidimensionalen* Mikro-Makro-Relation von Entitäten und ihren Eigenschaften zusammengeht, wie sie Kim selbst als gegeben annimmt (vgl. ebd., 1998, S. 15). Für Kim handelt es sich beim Eindruck, dass zwischen einer mentalen Eigenschaft und ihrem physischen Realisierer eine Mikro-Makro-Bewegung in Bezug auf natürliche Organisationsebenen stattfindet, um eine Illusion, die dadurch zustande kommt, dass die Realisierer erster Ordnung etwas sind, was man als *mikrobasierte Eigenschaft* (microbased property/microstructural property) bezeichnen kann (vgl. ebd.). Dies sind Eigenschaften eines Ganzen, die durch seine Mikrostruktur charakterisiert sind. Während zwischen der Instantiierung einer mentalen Eigenschaft und ihrem neuronalen Realisierer ein eindimensionales (Identitäts-)Verhältnis besteht, ist der neuronale Realisierer selbst eine mikrobasierte Eigenschaft, die in ihrer Beschaffenheit auf die Mikrostruktur des betreffenden Systems zurückgeht. Erst in der Beziehung zwischen dem Realisierer und der ihn bedingenden Mikrostruktur kommt nach Kim die multidimensionale (mehrere natürliche Organisationsebenen betreffende) Betrachtungsweise ins Spiel und führt zu dem Eindruck, dass auch die Beziehung zwischen mentaler Eigenschaft und ihrem Realisierer durch eine solche Differenz der natürlichen Organisationsebene gekennzeichnet ist.

Nach Kim dürfe man nun solche *mikrobasierten* Eigenschaften aber nicht als Mikroeigenschaften missverstehen, denn sie können nicht einzelnen Konstituenten der Mikrostruktur zugeschrieben werden, sondern nur dem Ganzen (wie dem Realisierer einer mentalen Eigenschaft). Mikrobasierte Eigenschaften sind *Makroeigenschaften* eines Systems. Als ein Bei-

spiel gibt Kim die physische Eigenschaft „1 kg“ an. Dies sei eine mikro-basierte Eigenschaft, da ihre Konstituenten, physische Eigenschaften und Beziehungen sind. Außerdem hat die Masse „1 kg“ eine kausale Kraft, die geringe Massen nicht haben, sodass es gerechtfertigt ist, sie als eine eigenständige (mikrobasierte) Makroeigenschaft gegenüber den Mikro-eigenschaften zu verstehen (vgl. ebd., 1998, S. 113). Das Beispiel ist allerdings etwas unglücklich gewählt, da hier eine rein aggregative Beziehung zwischen den Makro- und Mikro-eigenschaften besteht, die keineswegs notwendig ist und auch faktisch eher nur eine Ausnahmerecheinung ist. Für mikrobasierte Makroeigenschaften eines Systems, wie auch ihre kausale Wirkmächtigkeit, gelte des Weiteren, dass sie über grundlegendere physischen Eigenschaften auf einer niedrigeren Ebene des betreffenden Gegenstandes supervenieren. Es ist die *mereologische Supervenienz* angesprochen, was bedeutet, dass es jede Veränderung einer mikrobasierten Makroeigenschaft eines Systems eine Variation der Mikrostruktur dieses Systems voraussetzt.

Während bei mentalen Eigenschaften als funktionalen Eigenschaften zweiter Ordnung das im Zusammenhang der Supervenienz erläuterte Problem des kausalen Ausschlusses besteht, indem ihre kausale Wirkung vollständig durch neuronale Eigenschaften als ihren Realisierern erklärt werden kann (Kim spricht hier auch vom „supervenience argument“), ergäben sich bei mikrobasierten Eigenschaften solche Schwierigkeiten nicht, weil diese nicht wie Eigenschaften zweiter Ordnung über den sie konstituierenden Eigenschaften erster Ordnung supervenieren, sondern...

...they supervene on specific mereological configurations involving these microproperties – for a rather obvious and uninteresting reason: *they are identical with these micro-configurations*. [...] It follows then that we must grant novel causal powers to micro-based properties at higher levels – novel in the sense that these causal powers

are not had by any lower-level properties that constitute them. And, as we saw, the supervenience argument does not apply to them, and their causal roles are not threatened by the supervenience argument. (S. 117f; Hervorh.: M.L.)

Aus diesem Grund blieben mikrobasierte Eigenschaften vom Problem der *causal drainage* (des Abdriftens der Kausalität zu immer niedrigeren Ebenen, bis sie sich auf der Basisebene der Mikrophysik gänzlich zu verflüchtigen droht) unberührt, was Kim (2003) am Beispiel der Beziehung von Wasser und H₂O erläutert. Die Identität der Eigenschaft „Wasser zu sein“ und der Eigenschaft „H₂O zu sein“ wird für gewöhnlich akzeptiert. Die Eigenschaft „H₂O zu sein“ ist nun eine mikrobasierte Eigenschaft von Wasser auf der atomaren bzw. molekularen Ebene: „Being water is to have this microstructure. Having this microstructure is the microstructural essence of water, and being water just is having this structure“ (ebd.). Formalisiert kann man daher sagen, dass H₂O durch eine bestimmte Mikrostruktur auf der atomaren Ebene L gekennzeichnet ist, sodass gilt:

(1) Wasser zu sein = das Haben von M_L

Wenn man nun die kompositorischen Ebenen weiter „nach unten“ verfolgt, ist man auf der nächstniedrigeren Ebene L-1 wahrscheinlich auf der Ebene der vom derzeitigen „Standardmodell“ postulierten Elementarteilchen angelangt. Es kann angenommen werden, dass die Wasserstoff- und Sauerstoffatome der Ebene L auf dieser Ebene eine ganz bestimmte Zusammensetzung (M_{L-1}) haben, sodass wiederum gilt:

(2) Wasser zu sein = das Haben von M_{L-1}

Wenn es nun tatsächlich der Fall sein sollte, dass die Elementarteilchen entgegen den Implikationen des physikalischen Standardmodells weiter

zerlegbar sind, dann könnte es sein, dass Wasser eine Mikrostruktur auf der Ebene L-2 hat.

(1) Wasser zu sein = das Haben von M_{L-2}

Diese Reihe ließe sich nun tatsächlich bis ins Unendliche fortsetzen, so dass sich eine (potentiell unendliche) Abfolge von Identitätsbeziehungen ergibt:

(2) $M_L = M_{L-1} = M_{L-2} = M_{L-3} \dots$

Die Eigenschaft „Wasser zu sein“ ist damit etwas, was durch verschiedene mikrostrukturelle Konfigurationen verschiedener natürlicher, kompositorisch individuierter Organisationsebenen gekennzeichnet ist. „Wasser zu sein“ ist also nicht nur identisch damit, die Mikrostruktur M_L zu haben (also „ H_2O zu sein“), sondern genauso mit allen anderen mikrostrukturellen Ebenen (M_{L-1} , M_{L-2} , $M_{L-3} \dots$).

Kim führt dies aus, um zu zeigen, dass die Identität aller mikrostrukturellen (kompositorischen) Ebenen ein *causal drainage*, ein Abdriften der Kausalität zu den immer niedrigeren Ebenen, stoppen kann, denn Kausalität vollzieht sich notwendigerweise auf allen in einer Identitätsbeziehung stehenden Ebenen zugleich. Wie im Falle des kausalen Ausschlusses ist hier eine identitätsbasierte Reduktion die „Rettung“ und das trotz der Möglichkeit, dass es keine Basisebene geben könnte und Materie prinzipiell unendlich teilbar sein könnte: „Reduction is the stopper that will plug the cosmic hole through which causal powers might drain away“ (Kim, 2003, S. 175).

Fast im „Nebensatz“ formuliert er hiermit eine Position, die ein weit größeres Potential haben dürfte, als ihm selbst bewusst geworden zu sein scheint. Dass Kim diesen Ansatz nicht weiterverfolgt, liegt offensichtlich

daran, dass sein Ziel ist, eine physikalistische Theorie *mentaler* Eigenschaften zu entwickeln, die nach seinen Ausführungen keine mikrostrukturellen Eigenschaften sind. Es geht ihm primär also um die Charakterisierung der Beziehung zwischen mentalen Eigenschaften als Eigenschaften zweiter Ordnung und ihren physischen Realisierern. Hierbei wird – vereinfacht gesagt – das Verhältnis zweier Eigenschaften eines Systems diskutiert und nicht die Teil-Ganzes-Relation (vgl. Hüttemann, 2004, S. 2). Nur indirekt, also über ihren physischen (neuronalen) Realisierer erster Ordnung, lassen sie sich in einen Bezug zu natürlichen Organisationsebenen mit verschiedenen strukturellen Beschaffenheiten setzen²⁵³. Auch wenn ich grundsätzlich davon überzeugt bin, dass Kims Ansatz zu mikrobasierten Eigenschaften ein guter Ansatzpunkt zur Erklärung der Beziehung zwischen verschiedenen kompositorischen Ebenen darstellt, wirkt er insgesamt gesehen aber noch unausgereift. Er führt hier – mehr notgedrungen als systematisch hergeleitet – Identitäten zur Rettung mikrobasierter Eigenschaften vor einem *causal drainage* ein, ohne diese genauer zu begründen. Warum gibt es eine Identität zwischen den verschiedenen mikrostrukturellen Ebenen? Was genau bedeutet eine solche Identität? Ist hiermit überhaupt eine strikte Reduktivität oder gar ein Reduktionismus verbunden, wie Kim sie als zwangsläufig folgend ansieht? Auch scheint mir der Ansatz mikrobasierter Eigenschaften begrifflich missverständlich zu sein, da trotz der von Kim später angenommenen Identität der mikrostrukturellen Konfigurationen eine gewisse kausale Höhergewichtung der Mikroebenen angedeutet scheint. Denn Kim

²⁵³ Allerdings lässt sich dies auch in Zweifel ziehen. Denn wenn unter seinen Voraussetzungen gilt, dass zwischen einer mentalen Eigenschaft und ihrem physischen Realisierer ein Identitätsverhältnis (auf dem Niveau von Eigenschaftsinstantiierungen) besteht und der Realisierer eine mikrobasierte Eigenschaft ist, dann muss auch die mentale Eigenschaft genau genommen eine mikrobasierte Eigenschaft sein.

(1998, S. 116f) geht davon aus, dass die Kausalkräfte der Entitäten der höheren Ebenen durch die Kausalkräfte der Entitäten der niedrigeren Ebenen und ihrer Beziehungen einseitig *determiniert* werden (was höchst problematisch ist, wie sich zeigen wird; vgl. Kap. 8.6). Zumindest ist Kims Ansatz in der Hinsicht nicht wirklich klar ausformuliert, als dass er nicht aufzeigt, was genau das bedeutet.

7.1.3. Mentale Eigenschaften als Struktureigenschaften

Nach der hier vertretenen Vorstellung sind mentale Eigenschaften – in der Terminologie Kims ausgedrückt – keine abstrakten Eigenschaften zweiter Ordnung, sondern selbst (mikrobasierte) Kennzeichnungen der strukturellen Verfasstheit oder Konfiguration eines Systems. Mentale Prozesse sind auf einer eigenen höherstufigen Komplexitätsebene angesiedelt, die durch neuronale Netzwerke konstituiert wird. Ein mentaler Prozess bildet hierbei die Spitze einer aufgespannten lokalen Teil-Ganzes-Hierarchie und ist koextensiv mit den agierenden Entitäten aller ihm zugrunde liegenden natürlichen Organisationsebenen. Eine mentale Eigenschaft ist in ebendiesem Sinne eine mikrobasierte Struktureigenschaft eines informationsverarbeitenden Systems, die durch eine (potentiell unendliche) Abfolge von koextensiven Mikrokonfigurationen gekennzeichnet ist.

Durch den Einbezug des Ansatzes mikrobasierter Eigenschaften kann dieser Zusammenhang auf eine noch profundere Art und Weise erklärt werden. Dies soll an den bereits zuvor erläuterten Beispielen der räumlichen Orientierung und des Farbensehens erläutert werden.

7.1.3.1. Kognitive Eigenschaften

Es ist an dieser Stelle sinnvoll, sich nochmals die verschiedenen natürlichen Organisationsebenen im Falle des räumlichen Gedächtnisses vor Augen zu führen, wie sie Craver (2007) herausgestellt hat (vgl. Abb. 28). Wir haben hier das psychische Zielphänomen des *räumlichen Gedächtnisses*, das eine Maus oder einen Menschen dazu befähigt, durch ein Labyrinth zu navigieren. Dieses Zielphänomen wird nun durch einen Mechanismus bewerkstelligt, der sich auf verschiedenen Ebenen einer natürlichen Organisation von in einer Aktivität befindlichen Entitäten genauer untersuchen lässt.

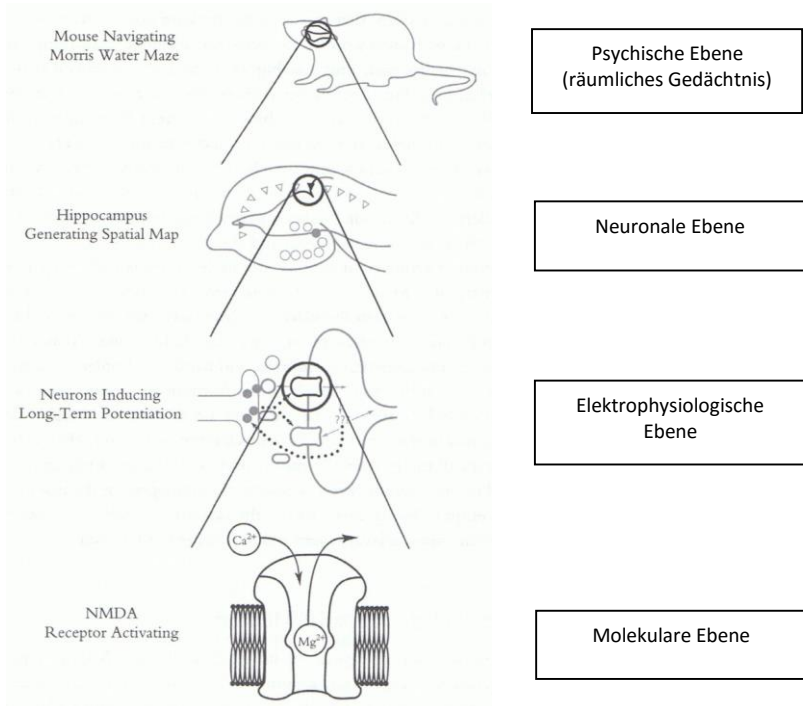


Abb. 28: Mechanistische Ebenen des räumlichen Gedächtnisses (Aus: Craver, 2007, S. 166)

Die verschiedenen Ebenen sind durch eine Teil-Ganzes-Beziehung charakterisiert, bei der sich die agierenden Entitäten einer höheren Ebene des Mechanismus auf solche niedrigerer Ebenen zerlegen lassen. Dies sei zugleich die Aufgabe der Wissenschaftler, die das Verhalten eines Mechanismus in das Verhalten seiner Teile zerlegen, um darauf wiederum das Verhalten dieser Teile in das Verhalten ihrer eigenen Teile zu zerlegen usw. usf. (vgl. Craver & Bechtel, 2007, S. 549).

Im Hinblick auf die Eigenschaft „räumliche Orientierung“ (bzw. räumliches Gedächtnis) muss man sich zunächst vergegenwärtigen, dass die Instantiierung von räumlicher Orientierung in der beschriebenen Situation eine bestimmte kognitive (und teilweise auch phänomenale) Struktur hat. Sie wird in einer ganz konkreten Instantiierung z.B. darin bestehen, sich die räumliche Umgebung (bildhaft²⁵⁴ oder eher abstrakt) vorzustellen, seine eigene gegenwärtige Position und die zuvor selbst durchlaufenen Pfade mit dieser Repräsentation abzugleichen, um eine Entscheidung darüber zu treffen, welcher Weg nun eingeschlagen werden sollte, um aus dem Labyrinth zu gelangen. Es sind also Prozesse, die vor allem durch kognitionspsychologische Modelle (wie u.a. das Arbeitsgedächtnismodell Baddeleys) beschrieben werden können.

Man könnte also sagen, das Haben von räumlicher Orientierung einer Person x zum Zeitpunkt t_1 besteht in der Existenz einer ganz bestimmten (primär) kognitiven Struktur, nennen wir sie R , sodass gilt:

- (1) Das Haben räumlicher Orientierung (von x zu t_1) = das Haben der kognitiven Struktur R (von x zu t_1)

²⁵⁴ Es wird sich im Falle bildhafter Vorstellungen durchaus um einen Prozess handeln, der einen bestimmten Erlebnischarakter hat.

Dass Craver das Phänomen der räumlichen Orientierung auf der neuronalen Ebene mit der durch den Hippocampus bewerkstelligten räumlichen Kartierungsfunktion in Verbindung bringt, muss an dieser Stelle genauer betrachtet werden. Der von Bechtel und Craver propagierte mechanistische Erklärungsansatz funktioniert so, dass er sich die (neurowissenschaftlichen) Erklärungsmodelle psychischer Leistungen anschaut und diese in ein Modell natürlicher, kompositorisch individuierter, mechanistischer Ebenen einpasst. Dabei wird die derzeitige wissenschaftliche Befundlage sehr verallgemeinert. Zu sagen, dass es lediglich die (in einer Aktivität befindliche) Hippocampus-Formation ist, die auf der neuronalen Ebene mit der Hervorbringung der räumlichen Orientierungs- oder Gedächtnisleistung betraut ist, ist also nur insofern korrekt, als dass man nach dem gegenwärtigen Kenntnisstand der kognitiven Neurowissenschaft sagen kann, dass die Aktivität der Hippocampus-Formation für die Hervorbringung dieser mentalen Leistung ein *relevanter Faktor* ist. Des Weiteren richtet sich die Auswahl der Ebenen bei Craver danach, wo die Wissenschaft bestimmte Erklärungen zu bieten hat. Wenn hier von der *nächstniedrigeren* Ebene gesprochen wird, ist dies also nicht in einem strengen Sinne zu verstehen, sondern nur eine – sich nach den bestehenden wissenschaftlichen Erklärungsmodellen richtende – *mögliche* Auswahl.

Ogleich es also vielleicht sogar *der* entscheidende Faktor sein mag, muss für den hiesigen Kontext die Annahme Cravers dahingehend präzisiert werden, dass wir auf der neuronalen Ebene eine komplexe Mikrostruktur der Eigenschaftsinstantiierung „des Habens räumlicher Orientierung“ vorfinden werden, die als wahrscheinlich zentralen Bestandteil die (in einer ganz bestimmten Aktivität befindliche) Hippocampus-Formation zum Gegenstand hat. Formalisiert man dies in der von Kim vorgeschlagenen Weise, kann man sagen, dass die räumliche Orientierung durch

eine bestimmte Mikrostruktur MR auf der neuronalen Ebene L gekennzeichnet ist, sodass gilt:

- (2) Das Haben räumlicher Orientierung (von x zu t_1) = das Haben von MR_L (von x zu t_1)

Es gilt wiederum, dass die nächste Ebene, die elektrophysiologische Ebene, nicht streng genommen, die faktisch mögliche nächstniedrigere Ebene darstellt, sondern die nächste Ebene, auf der ein weiterer kausal relevanter Befund angesiedelt ist. Es gilt hier also, dass die agierenden Entitäten der neuronalen Ebene (die primär im Hippocampus angesiedelt sein könnten und zusammen genommen MR_L konstituieren), in eine Struktur zerlegt werden können, bei der die im synaptischen Spalt zwischen Neuronen stattfindende Langzeitpotenzierung eine zentrale Rolle innehat. Es gilt daher, dass die räumliche Orientierungsleistung durch eine bestimmte Mikrostruktur MR auf der elektrophysiologischen Ebene L-1 gekennzeichnet ist, sodass gilt:

- (3) Das Haben räumlicher Orientierung (von x zu t_1) = das Haben von MR_{L-1} (von x zu t_1)

Und schließlich haben wir auf der molekularen Ebene das kausal relevante Phänomen der NMDA-Rezeptor-Aktivierung, in das die an der Langzeitpotenzierung beteiligten agierenden Entitäten zerlegt werden könnte, obgleich auch hier gilt, dass dies nur eine (kausal wirksame) Teilstrukturkomponente in der tatsächlichen kompositorischen Hierarchie sein wird. Es kann also angenommen werden, dass die (in einer Aktivität befindliche) Struktur der Ebene L-1 auf dieser molekularen Ebene eine ganz bestimmte Mikrokonfiguration (MR_{L-2}) aufweist, bei der die NMDA-Rezeptor-Aktivierung von primärer Bedeutung ist, sodass wiederum gilt:

- (4) Das Haben räumlicher Orientierung (von x zu t_1) = das Haben von M_{RL-2} (von x zu t_1)

Wie in Kims Beispiel lässt sich auch hier diese Reihe potentiell bis ins Unendliche fortsetzen oder auch durch verschiedene, hier nicht erwähnte kompositorische Zwischenebenen (z.B. der Zellorganellen, Atome...) erweitern. Ungeachtet der konkreten (beliebig veränderbaren) Nummerierungsweise ergibt sich also auch hier eine (potentiell unendliche) Abfolge von Identitätsbeziehungen:

- (5) $M_{RL} = M_{RL-1} = M_{RL-2} = M_{RL-3} \dots$ (von x zu t_1)

Wie Wasser ist somit auch die Eigenschaft der räumlichen Orientierung etwas, was durch verschiedene mikrostrukturelle Konfigurationen natürlicher – mechanistisch individuierter – Organisationsebenen gekennzeichnet ist. Durch die Annahme, dass (beobachterunabhängige) Eigenschaften eines Systems (mikro-)strukturelle Eigenschaften sind und den Fakt, dass eine Eigenschaft auf jeder natürlichen (kompositorisch individuierten) Organisationsebene eines Systems auf eine ganz bestimmte Mikrostruktur rekurriert, ist es nun in der Tat plausibel – wenn nicht sogar zwingend – anzunehmen, dass zwischen den koextensiven Mikrostrukturen ein *Identitätsverhältnis* besteht.

Räumliche Orientierung ist also nicht nur identisch damit, die Mikrostruktur M_{RL} zu haben, sondern genauso mit allen anderen mikrostrukturellen Konfigurationen der niedrigeren Organisationsebenen (M_{RL-1} , M_{RL-2} , M_{RL-3} ...). Die realisierte Zieleigenschaft selbst ist dabei gewissermaßen der „Kit“, der aus den verschiedenen koextensiven strukturellen Gliederungsebenen eine Einheit bildet.

Ein zentraler Unterschied zu Kim ist, dass es genau genommen nicht nur die verschiedenen Mikrostrukturen sind, die miteinander identisch sind,

sondern auch die kognitive Zieleigenschaft selbst. Sie ist selbst eine höherstufige Struktureigenschaft (R) und steht in einer kompositorisch verfassten Hierarchie, eben als Zieleigenschaft, an deren Spitze. Es gilt also vielmehr:

- (6) Das Haben räumlicher Orientierung (von x zu t_1) = R = MR₁ = MR₁₋₁ = MR₁₋₂
= MR₁₋₃... (von x zu t_1)

Das hier vorgeschlagene Modell weicht außerdem insofern von der Theorie Kims ab, da dessen Ansatz zu mikrobasierten Eigenschaften in eine funktionalistisch konzipierte Realisierungstheorie eingebettet ist, während dies hier nicht der Fall ist. So nimmt Kim an, dass eine mentale Eigenschaft durch ein bestimmtes kausales Profil (Inputs und Outputs auf der Verhaltensebene des Organismus) gekennzeichnet ist und dass dieses kausale Profil von einem physischen Realisierer erfüllt wird, so dass das Haben der mentalen Eigenschaft *nichts anderes als* das Haben des physischen Realisierers ist. Hier wird hingegen angenommen, dass eine mentale Eigenschaft wie „das Haben räumlicher Orientierung“ eine bestimmte kognitive Struktur aufweist, die selbst auf der höherstufigen natürlichen Organisationsebene des Psychischen angesiedelt ist. Sie ist daher keine Eigenschaft zweiter Ordnung (im Sinne Kims), sondern eine strukturgebundene, mikrobasierte Eigenschaft, die auf einer potentiell unendlichen Abfolge natürlicher Organisationsebenen eine spezifische Konfiguration aufweist.

Es gilt damit nicht, dass sich die mentale Eigenschaft darin erschöpft, eine bestimmte Mikrostruktur auf der *neuronalen* Ebene zu haben, wie es die Identitätstheorie annimmt und wie es letztlich in das Konzept der Realisierung mit hinübergeflossen ist. In erster Linie besteht die mentale Eigenschaft der räumlichen Orientierung nach dem hiesigen Ansatz da-

rin, eine ganz bestimmte *kognitive* Struktur zu haben. Erst in zweiter Linie ist festzustellen, dass diese kognitive Struktur mit einer (potentiell unendlichen) Abfolge mikrostruktureller Konfigurationen, die durch ein spezifisches Aktivitätsmuster gekennzeichnet sind, identisch ist.

7.1.3.2. Phänomenale Eigenschaften

Es bleibt zu klären, ob das vorgeschlagene formale Modell mikrostruktureller „Identifizierungen“ auch im Hinblick auf phänomenale Eigenschaften durchführbar ist. Ziehen wir hierzu das bereits zuvor schon einmal aufgegriffene Beispiel einer Farbempfindung heran. Zunächst muss festgehalten werden, dass auch jedes reale phänomenale Geschehen, jedes bewusste (Alltags-)Erlebnis von einer ganz bestimmten *strukturellen* Qualität ist. So ist jeder mentale Vorgang mit Erlebnisqualität eine aus vielen Einzelkomponenten (wie den verschiedenen Sinnesqualitäten, Emotionen, Stimmungen, Aspekten kognitiven Bewusstseins wie Erwartungen, Überzeugungen...) zusammengesetzte Ganzheit. Deshalb ist in einem gewöhnlichen mentalen Erlebniseindruck auch eine Vielzahl nicht-farbbezogener Elemente enthalten. Man denke allein an die verschiedenen Kanten, Formen, Schattierungen usw., die einen Wahrnehmungseindruck kennzeichnen oder auch an die phänomenalen Eindrücke, die wir praktisch in jedem Moment von *allen* Sinnesorganen erhalten. Mit der Hervorbringung eines solchen (in der Alltagsrealität ausschließlich vorkommenden) komplexen mentalen Zustands zu einem bestimmten Zeitpunkt ist nun ebenso eine weitverzweigte, in einer bestimmten Aktivität befindliche, neuronale Struktur in Form eines komplexen neuronalen Netzwerkes betraut. Nennen wir sie V_{cor}. Eine funktionale Teilstruktur dieser neuronalen Gesamtstruktur ist das informationsverarbeitende System, das als neuronales „Korrelat“ der Farbwahr-

nehmung fungiert. Zur Erinnerung: Laut aktuellem Forschungsstand sind dies vor allem die so genannten Blob-Regionen des Parvo-Blob-Systems, welches Fasern von der okzipitalen Region V1 in das Areal V2 und von dort in den ventralen Okzipitalkortex V4 entsendet. Insbesondere die Region V4 zeigt hierbei eine farbsensitive Aktivierung und wird für gewöhnlich als primärer „Ort“ des Farbensehens betrachtet (vgl. Zeki, 1990; Engel, Zhang & Wandell, 1997). Es bleibt allerdings dabei, dass für die hier verfolgte Zielstellung die Exaktheit der empirischen Daten nicht entscheidend ist. Entscheidend ist allein, dass es ein informationsverarbeitendes neuronales Netzwerk in Form eines *Mechanismus* gibt, der das Farbensehen hervorbringt. Da es ein Teilnetzwerk des gesamten neuronalen Netzes V_{COR} zu t_2 ist, nennen wir es wie zuvor V_{CORF} .

Für unseren Farbeindruck zu t_2 gilt nun, dass dieser eine ganz bestimmte Zusammensetzung von verschiedenen Farben beinhaltet, die ein phänomenales Ereignis kennzeichnen, z.B. den in Abbildung 29 dargestellten Blick auf einen Blumengarten, der eine Vielzahl unterschiedlicher Farben mit zugleich verschiedener Intensität oder auch einnehmender „Fläche“ im visuellen Gesamteindruck zeigt.



Abb. 29: Blumengarten

Somit gilt zunächst, dass die Eigenschaftsinstantiierung „die Farbempfindung (der Person y zum Zeitpunkt t_2) zu sein“ darin besteht, eine ganz bestimmte „farbbezogene Erlebnisstruktur F “ aufzuweisen. Womit gilt:

- (1) Farbempfindung (von y zu t_2) = das Haben der farbbezogenen Erlebnisstruktur F (von y zu t_2)

Nunmehr gilt, dass die Farbempfindung (von y zu t_2) als Mikrostruktur das neuronale Netzwerk V_{CORF} aufweist²⁵⁵. Die Eigenschaftsinstantiierung „ V_{CORF} zu sein“ ist nun eine mikrobasierte Eigenschaft des Farbensehens auf der neuronalen Ebene. Formalisiert man dies in der von Kim vorgeschlagenen Weise, kann man sagen, dass das Farbensehen durch eine bestimmte Mikrostruktur (MF) auf der neuronalen Ebene L gekennzeichnet ist, sodass gilt:

- (2) Farbempfindung (von y zu t_2) = das Haben von MF_L (von y zu t_2)

Das Farbensehen (bzw. die neuronale funktionale Einheit V_{CORF}) hat nun wiederum auf den verschiedensten niedrigstufigeren Ebenen eine ganz bestimmte Zusammensetzung. So muss angenommen werden, dass die (in einer Aktivität befindlichen) Neurone der Ebene L auf der z.B. molekularen Ebene eine ganz bestimmte Zusammensetzung (MF_{L-1}) haben, so dass wiederum gilt:

- (3) Farbempfindung (von y zu t_2) = das Haben von MF_{L-1} (von y zu t_2)

²⁵⁵ Das neuronale Geschehen im Zusammenhang visueller Informationsverarbeitung ist sehr komplex. Um sich hier nicht in neurowissenschaftlichen Details zu verlieren, wird auf ihre Darstellung zu Gunsten der Klarheit der Argumentation verzichtet. Eine umfassende Darstellung der Mechanismen des visuellen Prozessierens findet sich bei z.B. Bechtel (2008, S. 89ff).

Verfolgt man die kompositorischen Ebenen bis „ganz nach unten“, ist man auf der Ebene der vom physikalischen „Standardmodell“ postulierten Elementarteilchen angelangt, die nach dem jetzigen Kenntnisstand der Naturwissenschaften die Basisebene darstellt. Wiederum gilt, dass das Farbsehen (bzw. die neuronale funktionale Einheit V_{CORF}) auch auf dieser mikrostrukturellen Ebene eine ganz bestimmte Zusammensetzung (MF_{L-2}) hat, sodass wiederum gilt:

$$(4) \quad \text{Farbempfindung (von } y \text{ zu } t_2) = \text{das Haben von } MF_{L-2} \text{ (von } y \text{ zu } t_2)$$

Auch hier lässt sich diese Reihe potentiell bis ins Unendliche fortsetzen oder durch verschiedene hier nicht erwähnte kompositorische Zwischenebenen (z.B. der Zellorganellen, Atome...) erweitern. Ungeachtet der konkreten (austauschbaren) Nummerierungsweise ergibt sich also auch hier wie in Kims Beispiel eine (potentiell unendliche) Abfolge von Identitätsbeziehungen:

$$(5) \quad MF_L = MF_{L-1} = MF_{L-2} = MF_{L-3} \dots \text{ (von } y \text{ zu } t_2)$$

Wie die Eigenschaften „Wasser zu sein“ oder „räumliche Orientierung zu haben“ ist somit auch die Eigenschaft des Farbsehens etwas, was durch verschiedene mikrostrukturelle Konfigurationen der bestehenden natürlichen – kompositorisch individuierten – Organisationsebenen gekennzeichnet ist. „Die Farbempfindung von y zu t_2 zu sein“ ist also nicht nur identisch damit, die Mikrostruktur V_{CORF} bzw. M_L zu haben, sondern genauso mit allen anderen mikrostrukturellen Ebenen (MF_{L-1} , MF_{L-2} , $MF_{L-3} \dots$), wobei auch hier gilt, dass die Eigenschaft „eine Farbempfindung zu sein“ wesentlich darin besteht, eine *phänomenale* Struktur zu besitzen. Die Identität mentaler und neuronaler Eigenschaften, wie sie die Identitätstheorie oder auch die Realisierungstheorie in ihrer begrifflich konsis-

testesten Ausformulierung annimmt, ist demnach also zwar grundsätzlich nicht falsch, aber sie greift erstens viel zu kurz und nimmt zweitens eine der Sache nach unberechtigte (vereinsseitige) Interpretation des psychophysischen Identitätsverhältnisses vor.

7.2. Von psychophysischer Koextensivität zu lokaler Identität

Zunächst könnte sich natürlich die Frage stellen, warum man überhaupt den Übertritt von einer Koextensivitätsannahme zu einer Theorie psychophysischer Identität wagen soll. Der Grund hierfür liegt vor allem darin, dass eine reine Koextensivitätsannahme zu unbestimmt ist, um das Verhältnis zwischen bestimmten Wirklichkeitsausschnitten zu beschreiben. Z.B. könnte es ein Magnetfeld geben, das genau den Bereich meines Gehirns abdeckt, sodass man auch hier sagen könnte, Gehirn und Magnetfeld sind koextensiv. Ein solches Verhältnis ist es aber gerade nicht, das hier vertreten wird, nämlich, dass es *zwei verschiedene* Dinge gibt, die „zufällig“ dasselbe räumliche Feld zur gleichen Zeit einnehmen. Im Gegenteil: Wenn hier von der Koextensivität eines Moleküls und der Atome, aus denen es besteht, gesprochen wird, dann handelt es sich hierbei um *einen* Gegenstand, der verschiedene Ebenen einer natürlichen Organisation aufweist. Das Molekül ist kein zusätzliches Phänomen, das zu einem bestimmten Zeitpunkt „zufällig“ den gleichen räumlichen Bereich einnimmt, den auch seine Atome einnehmen. In gewisser Hinsicht *ist* das Molekül dasselbe wie die Atome, aus denen es besteht. Das Molekül und die es konstituierenden Atome sind koextensiv-identisch. Sie sind koextensiv, da sie denselben Raum zur selben Zeit einnehmen und identisch, da sie mereologische Strukturebenen *ein und desselben* Gegenstan-

des widerspiegeln. Und ebenso ist ein mentaler Prozess kein hinzukommender Vorgang, der in eine (kausale) Beziehung zu seinen neuronalen (molekularen, atomaren, subatomaren) Grundlagen bzw. Basisprozessen gesetzt werden muss. In gewisser Hinsicht *ist* der mentale Prozess zugleich der mit ihm koextensive neuronale (molekulare, atomare, subatomare) Prozess. Eine mentale Eigenschaft ist *identisch* mit einer (potentiell unendlichen) Abfolge (mikro-)struktureller Konfigurationen.

Der Begriff der Identität mag problematisch sein, zumal er – wie im letzten Kapitel aufgezeigt – in der Philosophie des Geistes eine unzuträgliche Auslegung erfahren hat. Dennoch scheint er der einzige Begriff zu sein, der ontologisch über eine bloße Koextensivität hinausgeht und dabei die Beziehung zwischen den natürlichen Organisationsebenen adäquat einfängt. Der Begriff der Konstitution, der hier manchmal als Alternative zu einer Identitätsannahme gehandelt wird, wird der Aufgabe jedenfalls nicht gerecht, da er – wie sich im Folgenden zeigen wird – einen anderen Zusammenhang zum Gegenstand hat bzw. über Merkmale gekennzeichnet ist, die für mereologisch-mechanistische Ebenen gerade nicht zutreffen. Natürlich ließe sich weiterhin auch von Realisierung sprechen. Ist diese um die Notwendigkeit funktionaler Analysen bereinigt, ist sie allerdings nicht mehr als ein Synonym für eine bestimmte Identitätsaussage. Als Schwierigkeit für psychophysische Identitätsaussagen wird z.B. häufig der Leibniz (1704/1996) zugeschriebene Grundsatz genannt, dass identische Dinge *ununterscheidbar* sein müssen. Die identifizierten Phänomene müssen demnach alle Eigenschaften gemeinsam haben. Da sich aber Gehirn und Geist offensichtlich stark (phänomenal) unterscheiden und ihnen andere Eigenschaften zugeschrieben werden, so ließe sich ableiten, können sie nicht identisch sein. Dass diese Argumentation hinfällig ist, sollte in Kapitel 3 allerdings bereits hinlänglich aufgezeigt worden sein. Im Kern heißt Identität in der Tat zwar $A = A$. Natürlich kann auch

gelten $A = B$. Dann haben wir z.B. den Fall vorliegen, dass wir herausgefunden haben, dass etwas, was wir als B bezeichnet haben, in Wirklichkeit A ist. Es ist die von Frege genannte und von den Identitätstheoretikern aufgegriffene Identität von Morgenstern und Abendstern, die in Wirklichkeit beide den Planeten Venus bezeichnen.

Einen ähnlichen und für die hier zu entwickelnde Identitätskonzeption ausschlaggebenderen Fall betrifft die vor allem von Fechner propagierte Tatsache, dass wir unterschiedliche Perspektiven auf einen Gegenstand haben, das heißt, dass wir einen Gegenstand immer nur nach Maßgabe einer bestimmten Perspektive wahrnehmen und ein unterschiedliches Erscheinungsbild dieses Gegenstandes haben. Aus der divergierenden – unserem eigenen Wahrnehmungs- und Kognitionssystem geschuldeten – Variation der Repräsentationsweise zu schließen, dass es sich nicht um einen, sondern verschiedene Gegenstände handelt, wäre offensichtlich absurd. Wenn es sich bei bestimmten Perspektiven auf einen Gegenstand aber nun um dessen natürliche Gliederung (in bestimmte) Strukturebenen handelt, geht die Divergenz seiner Erscheinungsweise zwar nicht einfach nur auf uns selbst zurück. Nichtsdestotrotz handelt es sich weiterhin um *ein und denselben* Gegenstand, auch wenn dieser zugleich (berechtigterweise) als ein Gefüge von Zellen oder auch Atomen betrachtet werden kann. In gewisser Hinsicht sind es weiterhin Perspektiven auf *ein und denselben* Gegenstand, die in diesem Fall aber nicht nur auf uns selbst, sondern auf dessen eigene kompositorische Verfasstheit zurückgehen. Es ist allerdings keine problematische *substantielle* Verschiedenheit von neuronalen und psychischen Prozessen, sondern nur eine solche der Divergenz einer nach mereologischen Prinzipien individuierten Organisationsebene.

Für eine jede Teil-Ganzes-Hierarchie gilt notwendigerweise eine räumlich-zeitliche *Koextensivität*. Denn in einer solchen Hierarchie, wie sie der

mechanistische Ansatz annimmt, gibt es eine Abfolge von strukturellen/funktionalen Ganzheiten, die *prinzipiell* in Teile (in Form von agierenden Entitäten) zerlegt werden können. Grundsätzlich gilt hierbei, dass ein Ganzes und seine Teile dasselbe räumlich-zeitliche Feld einnehmen und damit koextensiv sind. Wenn ein Ganzes in Teile zerlegt werden kann, dann muss für die Teile also gelten, dass sie zur selben Zeit das gleiche räumliche Feld einnehmen wie das Ganze. Die Annahme eines Teils, das außerhalb des räumlichen Feldes des Ganzen vorkommt, ist kein kohärenter Gedanke. Ebenso wäre es unsinnig anzunehmen, dass etwas Teil eines Ganzen sein kann, obwohl es nicht zur selben Zeit wie das Ganze existiert.

Von zentraler Bedeutung ist nun, dass das Koextensivitätsprinzip auch für das Psychische gilt. Es nimmt dasselbe räumliche Feld zur selben Zeit ein wie seine physischen Grundlagen auf den verschiedenen natürlichen Organisationsebenen. Denn wie in Kapitel 5 aufgezeigt wurde, lässt sich das Psychische selbst als etwas auffassen, das eine eigene höherstufige natürliche Organisationsebene bildet, die nach denselben Kriterien als eine solche Einheit individuiert werden kann wie „herkömmliche“ physische Ganzheiten. Ein mentaler Prozess, der eine räumliche Orientierungsleistung beinhaltet, ist selbst eine höherstufige Ganzheit gegenüber seinen physischen Grundlagen. So gibt es ein (in einer ganz bestimmten Aktivität befindliches) neuronales Netzwerk, das sich unter anderem auf den Hippocampus bezieht, dessen Aktivitätszusammenhang und *funktionale Einheit* überhaupt erst dadurch zustande kommt, dass es genau diese mentale Eigenschaft des räumlichen Gedächtnisses bzw. der räumlichen Orientierung (mit einer ganz bestimmten psychischen Struktur) hervorbringt. Erst durch die Hervorbringung der räumlichen Orientierungsleistung werden aus dem relativ undifferenzierten Geflecht zellulärer Gegebenheiten oder „blinder“ Aktivitätsmuster, die unser Gehirn ausmachen,

zusammenhängende Netzwerke oder „Bereiche“ im Sinne einer mechanistischen Teil-Ganzes-Hierarchie geformt. Es sind neuronale Netzwerke, die bestimmte neue (systemisch emergente) Eigenschaften hervorbringen, die auf allen niedrigeren Organisationsebenen gewissermaßen nicht vorhanden sind.

Mentale Eigenschaften rekurren in einer entscheidenden Hinsicht auf die strukturelle Verfasstheit eines Systems, das neben der genuin psychischen Struktur der instantiierten mentalen Eigenschaft auch eine hiermit verbundene (potentiell unendliche Abfolge) mikrostruktureller Konfigurationen aufweist. Da die Instantiierung einer mentalen Eigenschaft durch ein System damit identisch ist, dass dieses System eine bestimmte ebenenübergreifende (mikro-)strukturelle Konfiguration hat, ist gleichzeitig eine lokale Identität der mikrostrukturellen Konfigurationen der verschiedenen mereologischen Ebenen (des betreffenden Mechanismus) begründet.

7.2.1. Die Identifikation „mit“ statt der Identifikation „als“

Diese Art der Identität ist keine, wie sie von den „klassischen“ Identitätstheoretikern propagiert wurde und besagt, dass ein mentaler Vorgang *nichts anderes als* ein Vorgang in unserem Gehirn ist oder mentale Eigenschaften *nichts anderes als* neuronale Eigenschaften sind. Diese Art der Identität besagt nur, dass ein ganz bestimmter mentaler Vorgang in ein Gefüge natürlicher Organisationsebenen eingebettet ist, dessen (kausal agierende) Entitäten in einem mereologischen und damit koextensiven Verhältnis stehen. Betrachtet man einen ganz bestimmten mentalen Prozess, so gilt für diesen, dass er dasselbe räumlich-zeitliche Feld einnimmt wie seine physischen Grundlagen und deshalb auch koextensiv-identisch mit ebendiesen ist. Ein mentaler Prozess ist nicht ablösbar von seinen

physischen Grundlagen und ist über die Teil-Ganzes-Beziehung fest in dieser Welt verortet. Er ist eben genau dort lokalisiert, wo seine neuronalen, mikrozellulären, molekularen, atomaren und subatomaren Komponenten existieren und eine Teil-Ganzes-Hierarchie bilden. Eine solche Teil-Ganzes-Hierarchie ist *ein* Gebilde, auf das wir mittels verschiedener wissenschaftlicher Techniken und Analysemethoden zugreifen können und dabei auf eine natürliche strukturelle Gliederung dieses *einen* Objektes stoßen. Wie Craver (2001) treffend sagt, sind mechanistische Beschreibungen „distinct perspectives on an activity in a hierarchically organized mechanism“ (S. 67).

Dennoch bleibt es dabei, dass ein mentaler Prozess auf einer eigenen höherstufigen Komplexitätsebene angesiedelt ist, von der nicht einfach gesagt werden kann, dass sich ihre Entitäten auf diejenigen der zugrunde liegenden Organisationsebenen im Sinne des „Nichts-anderes-als-Prinzips“ reduzieren lassen. Auch wenn ein Molekül koextensiv-identisch mit den Atomen ist, aus denen es besteht, bedeutet dies nicht, dass ein Molekül *nichts anderes als* zwei Atome ist. Es handelt sich um eine Entität von höherer Komplexität, die Eigenschaften aufweist, die einzelnen Atomen nicht zukommen. Und so ist auch die Existenz mentaler Eigenschaften an das Vorliegen bestimmter physischer Größen gebunden. Dennoch erschöpft sich ihre Existenz nicht einfach in ihren neuronalen Grundlagen, die dessen kausale Rolle einnehmen, wie es die Realisierungstheorie à la Kim postuliert. Ein mentaler Prozess konstituiert vielmehr in einer Teil-Ganzes-Hierarchie eine eigene höherstufige Komplexitätsebene, ist aber gleichzeitig durch eben diese Teil-Ganzes-Relation mit den (kausal agierenden) Entitäten der neuronalen Ebene, aber auch mit denjenigen aller niedrigeren Ebenen koextensiv-identisch. Der mentale Vorgang beim Anblick eines Blumengartens hat phänomenale Eigenschaften (farbliche Erlebnisqualitäten), der mentale Prozess des „Aus-dem-Irrgarten-Findens“

hat kognitive Eigenschaften (räumliche Orientierung), die einzelne Neuronen, Moleküle oder Atome nicht aufweisen. Nur in einer bestimmten Zusammensetzung/Konfiguration von Neuronen (respektive Molekülen respektive Atomen oder anderen herausgreifbaren agierenden Entitäten der verschiedenen Ebenen eines Mechanismus) kommt es zur Ausbildung von phänomenalen oder kognitiven Eigenschaften, die ein solches System (als Ganzes) kennzeichnen. Mentale Eigenschaften sind damit etwas, das erst entstehen kann, wenn eine mereologische Hierarchie bis hin zur mentalen Ebene aufgespannt wird, was von der Komposition und Konfiguration der Entitäten der physischen Organisationsebenen abhängt.

Um auf die Unterscheidung von Deutscher (1967) zurückzukommen: Während die „klassische“ Identitätstheorie und auch die Realisierungstheorie (zumindest in der von Kim vorgeschlagenen Form) den Charakter einer *Identifikation als* hat, hat die hier vorgestellte Konzeption vielmehr den Charakter einer *Identifikation mit*. Während erstgenannte Ansätze im Kern die These vertreten, dass mentale Prozesse/Eigenschaften in Wirklichkeit neuronale Prozesse/Eigenschaften sind, wird hier eine andere Feststellung gemacht. Über das mit mereologischen Verhältnissen verbundene Koextensivitätsprinzip werden mentale Prozesse *mit* neuronalen (molekularen, atomaren, subatomaren) Prozessen „identifiziert“ und keineswegs mentale Prozesse *als* neuronale Prozesse herausgestellt. Ebenso werden mentale Eigenschaften als makroskopische Struktureigenschaften eines Systems verstanden, die mit einer (potentiell unendlichen) Abfolge von mikrostrukturellen Konfigurationen identifiziert werden können. Die Annahme, dass mentale Eigenschaften *nichts anderes als* neuronale Eigenschaften sind, ist schon daher falsch, da man genauso gut – wenn man sich denn unbedingt einer solchen ontologischen Vereinseitigung schuldig machen möchte – sagen kann, dass mentale Eigenschaf-

ten „nichts anderes als“ molekulare oder atomare oder subatomare... Eigenschaften sind, da eine mentale Eigenschaft auf jeder dieser Mikroebenen eine bestimmte strukturelle Verfasstheit hat. Es ist nicht mehr die Identität von Schmerzen und C-Faser-Reizungen, die hier behauptet wird, sondern die Identität der agierenden Entitäten einer Teil-Ganzes-Hierarchie, an deren „Spitze“ mentale Prozesse angesiedelt sind (womit das System als Ganzes über mentale Eigenschaften verfügt). Insofern ist es genau genommen keine Theorie einer psychoneuronalen Identität, die hier verteidigt wird, sondern *de facto* eine Theorie einer *psychophysischen* Identität.

7.2.2. Psychophysische Identität und systemische Emergenz

Des Weiteren ist von zentraler Bedeutung, dass über eine „klassische“ Identitätsaussage festgestellt werden soll, dass das „Wesen“ des Psychischen darin besteht, eine bestimmte *neuronale* Struktur zu haben. Entgegen diesem „neuronalen Chauvinismus“ (vgl. Shoemaker, 1981, S. 280) bezieht sich die hier entwickelte Identitätskonzeption genauso auf die psychische Ebene selbst, in dem zwar trivial anmutenden Sinne, dass das Haben einer mentalen Eigenschaft vor allem darin besteht, eine ganz bestimmte *psychische* Struktur zu besitzen (die wiederum identisch mit den verschiedenen mikrostrukturellen Konfigurationen ist). Trotz der psychophysischen Identitätsannahme besteht das Wesen des Psychischen auch und in gewisser Hinsicht sogar primär darin, *psychisch* zu sein (und hiermit bestimmte durch psychologische Theorien beschreibbare Charakteristika zu haben). Mentale Prozesse/Eigenschaften sind das „Zielphänomen“ mereologisch-mechanistisch verfasster Systeme. Es ist die zu „realisierende“ Funktion (z.B. räumliches Gedächtnis), die durch die in einer charakteristischen Aktivität befindlichen Entitäten der ver-

schiedenen natürlichen Ebenen des Mechanismus konstituiert wird und überhaupt erst eine funktionale Einheit aus dem undifferenzierten physischen Geflecht erschafft. Die hier entwickelte psychophysische Identitätskonzeption ist zwangsläufig ontologisch neutral. Es ist nicht möglich, von irgendeiner Ebene zu behaupten, dass sie die eigentliche Ebene der Existenz (eines Phänomens) oder die Herberge der kausalen Wirkmächtigkeit ist (zum letztgenannten Punkt an späterer Stelle mehr; vgl. Kap. 7.4).

Wenn also die Herausstellung mentaler Eigenschaften als neuronaler Eigenschaften zu kurz greift, stellt sich aber die Frage, ob nicht zumindest gilt, dass mentale Eigenschaften im Sinne des Materialismus bzw. Physikalismus *per se* physische Eigenschaften sind. In gewisser Hinsicht mag dies berechtigt sein. Nach Boyd (1980) behauptet der Materialist, dass „all natural phenomena, all events, processes, objects and so forth, are in fact physical: all objects are composed solely of matter and all events and processes consist merely in interactions between material things“ (S. 85). Es mag zwar sein, dass sich mentale Eigenschaften oder Prozesse von „herkömmlichen“ physischen Eigenschaften oder Prozessen unterscheiden, aber der Unterschied besteht lediglich darin, dass Systeme mit mentalen Eigenschaften/Prozessen über eine *besondere Konfiguration* von Materie oder physikalischer Kräfte verfügen, die andere (nicht-mentale) Systeme eben nicht aufweisen. „Pains are quite different from, for instance, earthquakes; but the difference is configurational, not constitutional. They are made of the same sorts of stuff“ (ebd.).

Ebenso ist die „klassische“ Identitätstheorie mit dem Problem konfrontiert, dass ihre These der Identität mentaler und neuronaler Eigenschaften suggeriert, dass es sich hierbei um eine aggregative Beziehung handelt. In diesem Sinne wäre eine mentale Eigenschaft nichts anderes als eine *Ansammlung* von neuronalen Eigenschaften, wie die Eigenschaft eines Objektes „1 kg zu sein“ nichts anderes als ein Agglomerat der Massen

der Teile ebendieses Objektes ist. Dabei ist die Organisation oder Konfiguration der Teile des Systems unerheblich, sodass es die Makroeigenschaft nicht beeinflusst, ob man Teile aus dem System durch andere (mit der gleichen Masse) ersetzt oder die Mikrostruktur eines Objektes verändert. In der Möglichkeit einer solchen Interpretationsweise liegt nach Wimsatt (2007) ein Hauptproblem der herkömmlichen identitätstheoretischen oder reduktiven Ansätze, die unnötigerweise in die Nähe eines Eliminativismus oder kruden Reduktionismus rücken (vgl. ebd., S. 174). Faktisch gibt es nur einige wenige Fälle, in denen aggregative Beziehungen überhaupt gelten. Nach Wimsatt sind es genau genommen nur die Eigenschaften, die die Erhaltungsgesetze der Physik aufgreifen: Masse, Energie, Ladung usw. (vgl. ebd., S. 175). Für eine Identitätsannahme, die auf *mechanistisch* konzipierten Teil-Ganzes-Hierarchien fußt, ist nun gerade die *Mikrostruktur* entscheidend: „Mechanistic explanations of phenomena commonly involve highly differentiated parts and behavior that depend on their mode of organization. Some people see such explanations as very threatening. But they misidentify the enemy: only aggregativity, not mechanism, would justify the claim they fear. And aggregative claims are false” (vgl. ebd., S. 177).

Mentale Eigenschaften sind insofern *emergente Systemeigenschaften*, da sie erst auftreten, wenn eine mereologisch-mechanistische Hierarchie bis hin zur Ebene des Psychischen bzw. neuronaler Netzwerke aufgespannt wird, während keine einzelne der (agierenden) Entitäten, die dieses Netzwerk konstituieren, sprich: Neuronen, Zellbestandteile, Moleküle, Atome etc., diese Eigenschaft aufweist. Es ist die Konfiguration dieser agierenden Entitäten der verschiedenen Organisationsebenen, die für das Hervorbringen der mentalen Eigenschaft verantwortlich ist. Dennoch ist hiermit kein problematischer Eigenschaftsdualismus verbunden, der „andersartige“ psychische Eigenschaften den gewöhnlichen physikalischen

Eigenschaften gegenüberstellen würde. Denn wie gezeigt wurde, liegt eine Identität der agierenden Entitäten der verschiedenen (mikro-)strukturellen Organisationsebenen vor, wobei das Psychische selbst mit der (potentiell unendlichen) Abfolge von Konfigurationen agierender Entitäten in einer solchen Strukturhierarchie identisch ist.

An dieser Stelle lohnt es sich, die – unter Anwendung des bei Davidson selbst unbeliebten Eigenschaftsbegriffs ableitbare – These des Anomalen Monismus aufzugreifen, nämlich, dass ein Ereignis mehrere Eigenschaften instantiieren kann. Wenn diese These also besagt, dass Ereignisse, die als Instantiierung *einer* Makroeigenschaft aufgefasst werden können, „damit gleichzeitig ein ganzes Bündel von Mikroeigenschaften [instantiieren]“ (Pauen, 1999, S. 36), wird nun erst richtig verständlich, wie dies möglich bzw. inwieweit dies überhaupt eine plausible Annahme ist. Denn instantiiert ein Ereignis qua seiner strukturellen Verfasstheit eine bestimmte mentale Makroeigenschaft, dann sind hiermit notwendigerweise zugleich eine Vielzahl an Mikroeigenschaften instantiiert, weil eben das Haben dieser Makroeigenschaft – neben dem Aufweisen einer bestimmten psychischen Struktur – darin besteht, eine (potentiell unendliche) Abfolge von materiellen Mikrokonfigurationen und damit Mikroeigenschaften dieses Systems zu sein.

7.2.3. Die „Aufstapelungssillusion“ mereologischer Hierarchien

Obwohl es trivial erscheinen mag, muss an dieser Stelle thematisiert werden, dass der Begriff einer mereologisch-mechanistischen *Hierarchie* zu ungünstigen bildhaften Vorstellungen einlädt, die eine gewisse Distinktheit oder gar „Übereinander-Reihung“ der verschiedenen Ebenen eines Mechanismus suggeriert. Auch hier könnte eine Wurzel für die verbreitete Überzeugung vorliegen, dass die agierenden Entitäten einer solchen

Hierarchie nicht identisch miteinander sein können. Das sich in der bildhaften Vorstellung aufdrängende Abheben der einzelnen Ebenen voneinander entspricht aber keineswegs der Wirklichkeit, sondern ist vor allem unserem begrenzten (stark analog operierenden) Vorstellungsvermögen geschuldet. Durchläuft man gedanklich (bildhaft) die verschiedenen Gliederungsebenen, bleibt die räumliche Struktur bzw. Ausdehnung des herangezogenen Objektes natürlich hiervon unberührt. Dies lässt sich anhand des in Abbildung 30 auf der linken Seite dargestellten Neuronenverbandes erläutern.

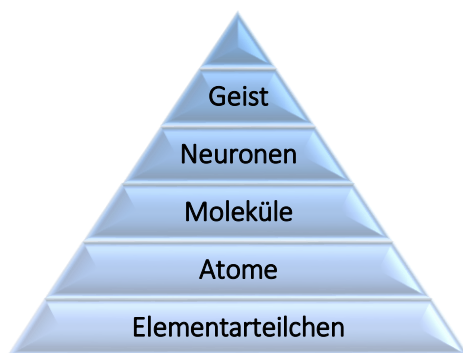
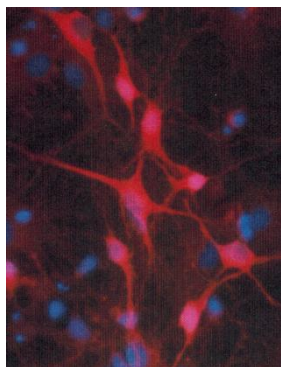


Abb. 30: Neuronales Geflecht (links) versus hierarchische Anordnung in Form einer Pyramide (rechts)

Jedes Neuron kann zunächst natürlich als eine eigenständige, gegenüber seiner Umwelt abgrenzbare Entität betrachtet werden. Man kann aber auch die Vernetzung mit anderen auf diese Weise abgrenzbaren Einheiten, also anderen Neuronen oder vielleicht Gliazellen fokussieren, und dabei ein bestimmtes (neuronales) Netzwerk erfassen. Es ist dabei aber eigentlich nichts an dem betrachteten Wirklichkeitsausschnitt verändert

worden. Wir haben lediglich erfolgreich unser kognitives Konzept einer Teil-Ganzes-Beziehung auf diesen Wirklichkeitsausschnitt angewandt. Dass es sich um ein echtes (beobachterunabhängiges) „höherstufiges“ Ganzes handelt, ist aber nur dann gerechtfertigt, wenn ein solcher Verbund bestimmte Leistungen erbringen kann bzw. bestimmte Eigenschaften aufweist, die einzelne dieser Entitäten nicht erbringen können bzw. innehaben (systemische Emergenz).

Verlagern wir unsere Perspektive nun so, dass wir den Auflösungsgrad verkleinern, könnten wir z.B. feststellen, dass dieses Netzwerk wiederum ein Teil des Hippocampus ist. Der Hippocampus wäre (wie in den mechanistischen Modellen der räumlichen Orientierung) also wiederum eine „höherstufige“ Struktureinheit, die mit neuen Leistungen/Eigenschaften assoziiert ist, obwohl auch hier genau genommen, keinerlei „Aufstapelung“, „Auffaltung“ oder überhaupt Veränderung des betrachteten Wirklichkeitsausschnittes stattfindet. Genauso ließe sich der Auflösungsgrad im Hinblick auf die abgebildeten Neuronen auch erhöhen und wir würden dann auf bestimmte Zellorganellen stoßen, die wiederum in ihrer Gesamtheit eine Nervenzelle bilden, oder – bei noch stärkerem „Zoom“ – auf die Moleküle, aus denen wiederum diese Zellorganellen bestehen usw. Wenn wir nun Moleküle betrachten, haben wir die höherstufige Ganzheit eines Neurons oder gar des Hippocampus zwangsläufig gänzlich aus dem Blick verloren. Wir neigen daher dazu, eine solche Veränderung wahrzunehmen, weil wir mit neuen strukturellen Einheiten auf den höheren Ebenen konfrontiert sind und die Entitäten der niedrigeren Ebenen gewissermaßen verschwunden sind, wenn wir den Auflösungsgrad bei der Betrachtung eines Objektes stark genug verändern. Es findet in Wirklichkeit aber keinerlei „Auffächerung“ statt, wie man sie bei der visuellen Darstellung hierarchischer Gebilde verwendet (vgl. die Pyramide auf der rechten Seite in Abb. 30). Es gibt faktisch keine *übereinander liegenden*

natürlichen Organisationsebenen, obwohl wir mereologische Verhältnisse für gewöhnlich auf diese Weise zu Anschauungszwecken abbilden. Damit dürfte auch das Unternehmen, mentale Prozesse in eine solche mereologisch-mechanistische Hierarchie einpassen zu wollen, weitaus weniger mysteriös wirken. Wenn nämlich davon die Rede war, dass eine Teil-Ganzes-Hierarchie bis hin zur mentalen Ebene „aufgespannt“ wird (oder nicht), war dies natürlich nicht wortwörtlich zu nehmen. Mentale Prozesse heben sich in keinerlei Hinsicht von den ihnen zugrunde liegenden physischen Ebenen ab. Wir mögen sie in der Tat nicht auf die Weise in unser visuelles Abbild einer mereologisch-mechanistischen Hierarchie einpassen können, wie dies mit „gewöhnlichen“ materiellen Entitäten der Fall ist. Gewissermaßen verlieren wir die agierenden Entitäten der niedrigeren Ebenen aus dem Blick, wenn wir uns einen phänomenalen Prozess wie z.B. das Sehen von Farben vorstellen. Aber wir verlieren eben auch jedwede kortikale Struktur aus dem Blick, wenn wir bei der Visualisierung ebendieses Wirklichkeitsausschnittes bei der Ebene der Moleküle oder Atome angekommen sind. Man mag sich das Psychische also zwar nicht als „gewöhnliche“ materielle Entität vorstellen können (die es auch gar nicht ist, denn es hat gegenüber klassischen materiellen Entitäten in der Tat eine ganze Reihe *ungewöhnlicher* Eigenschaften), dennoch liegt hier kein prinzipieller Hinderungsgrund vor, von einer eigenständigen, „höherstufigen“ mentalen Ebene zu sprechen, die sich in eine solche „materielle Hierarchie“ einpasst. Sie durchdringt gewissermaßen alle „niedrigeren“ Ebenen und ist selbst nur gegeben, wenn eine Konfiguration eines materiellen Wirklichkeitsausschnittes so beschaffen ist, dass eine strukturelle Verfasstheit vorhanden ist, die für die Existenz mentaler Eigenschaften und Prozesse notwendig ist (wobei eine solche mereologisch-mechanistische Hierarchie hierzu offenbar eine neuronale Ebene bzw. eine Ebene neuronaler Netze aufweisen muss).

7.3. Die Einheit von Struktur und Form/Funktion

Bisher wurde die These verfolgt, dass mentale Eigenschaften Struktureigenschaften sind. Wenn wir einer Person berechtigterweise eine mentale Eigenschaft zuschreiben, berechtigt in dem Sinne, dass es wirklich etwas an der Person gibt, dass diese Eigenschaftszuschreibung wahr macht, wurde gesagt, dass es die (mikro-)strukturelle Verfasstheit der Person ist, auf die wir hierbei rekurren. Des Weiteren wurde herausgestellt, dass – da die Instantiierung einer mentalen Eigenschaft durch ein System damit identisch ist, dass dieses System eine bestimmte ebenenübergreifende (mikro-)strukturelle Konfigurationen aufweist – gleichzeitig eine lokale Identität der (mikro-)strukturellen Konfigurationen der agierenden Entitäten der verschiedenen mereologischen Ebenen (des betreffenden Mechanismus) angenommen werden muss.

7.3.1. Funktionalismus versus Identitätstheorie

Was bis hierhin nicht explizit angesprochen wurde, ist das funktionalistische Verständnis mentaler Eigenschaften. Dass lediglich auf Strukturen rekuriert wird, könnte daher so verstanden werden, dass – wie bei der „klassischen“ Identitätstheorie – lediglich das materielle Substrat mentaler Eigenschaft ihr „eigentliches Wesen“ charakterisieren würde. Es wären dann vielleicht nicht mehr das rein neuronale „Material“, was hier fokussiert wird, sondern auch das „Material“, aus dem sich Neuronen (auf den verschiedenen natürlichen Organisationsebenen des Gehirns) formen.

Der Funktionalismus verfolgt bekanntermaßen die These, dass mentale Eigenschaften keine neuronalen Eigenschaften oder überhaupt irgendwelche genuin physische Eigenschaften sind (wie es die „klassische“

Identitätstheorie behauptet), sondern dass das Wesen des Psychischen darin besteht, funktionaler Natur zu sein. Das, was eine mentale Eigenschaft zu einer mentalen Eigenschaft macht, ist nicht seine stofflich-strukturelle Grundlage, sondern es ist vielmehr die durch diese (oder eben eine beliebig andere) stofflich-strukturelle Grundlage realisierte *Funktion*. Diese Annahme stützt sich vor allem auf die von Shoemaker (1980, 1998) entwickelte *Theorie der kausalen Eigenschaften*, die letztendlich besagt, dass alle kausalen Eigenschaften (und damit auch mentale Eigenschaften) funktionale Eigenschaften sind²⁵⁶ und das *Argument der multiplen Realisierbarkeit*, wie es von Putnam (1967) oder Fodor (1974) vertreten wird. Nach Polger (2004) kann man dies als einen *metaphysischen Funktionalismus* bezeichnen (vgl. ebd., S. 183).

Von den Funktionalisten wird nun proklamiert, dass der Funktionalismus einer „Gehirnzustandstheorie“ überlegen ist, da das Zentrale an mentalen Eigenschaften nicht der Stoff, sondern sein kausales (funktionales) Profil bzw. die kausale Rolle sei, die sie in einem Organismus spielen, eine Überzeugung, die vor allem Lycan (1987/2007) sehr pointiert dargestellt hat:

Und tatsächlich stellen wir fest, dass unsere Mentalität intuitiv unverändert bliebe, wenn eine Neurochemie, die nicht auf Kohlenstoff-, sondern auf Siliziumverbindungen basiert, dieselbe Rolle spielen bzw. dieselben Funktionen realisieren würde, oder wenn unsere einzelnen Neuronen Stück für Stück durch elektronische Prothesen ersetzt würden, die dieselben Aufgaben erfüllen. Entscheidend ist die Funktion, nicht der Funktionär; das Programm, nicht das realisierende Material; die Software, nicht die Hardware; die Rolle, nicht ihr Träger. So entwickelte sich der Funk-

²⁵⁶ Nach Van Gulick (1982) lässt sich die Theorie der kausalen Eigenschaften aber bereits dahingehend kritisieren, dass – wenn sie vorgibt, dass alle kausalen Eigenschaften funktionale Eigenschaften sind – die These, dass mentale Eigenschaften funktionale Eigenschaften sind, an Aussagekraft verliert, denn über die funktionale Charakterisierung würde nichts eingefangen werden, was spezifisch für mentale Eigenschaften wäre.

tionalismus mit der ihm zugrunde liegenden Unterscheidung von „funktionalen“ und „strukturellen“ Zuständen oder Eigenschaften eines Organismus. (S. 432)

Allerdings meint Lycan, dass diese Gegenüberstellung den Tatsachen nicht gerecht würde und liefert einen ersten Ansatz zur Integration von Struktur und Funktion, der allerdings doch die Struktur der Funktion in gewisser Weise einverleibt. Funktionen lägen demnach in bestimmten graduellen Abstufungen für alle Objekte vor, sodass alles zu einem bestimmten *Grad* als funktional anzusehen sei. Die strikte Unterscheidung zwischen Funktion und Struktur geht nach ihm auf das *Zwei-Ebenen-Dogma* („Two-Levelism“; S. 449) zurück, das besagt, dass es im Grunde nur zwei Ebenen gibt: *die* strukturelle und *die* funktionale Ebene. In Wirklichkeit gäbe es aber eine Vielzahl von Organisationsebenen, auf die sich ein Funktionalist beziehen kann. Die verschiedenen natürlichen Organisationsebenen (Organe, Zellen, Moleküle, Atome...) eines Organismus sind für Lycan also primär funktional spezifizierte Ebenen, womit aber auch die Unterscheidung zwischen Funktionalismus und der („klassischen“) Identitätstheorie aufgehoben wird:

Doch wenn wir auch meine Behauptung akzeptieren, dass homunktionale²⁵⁷ und physiologische Beschreibungen der Zustände von Personen lediglich unterschiedliche Abstraktionsebenen innerhalb einer sie umgebenden funktionalen Hierarchie oder Kontinuums reflektieren, dann können wir den Funktionalisten nicht mehr in irgendeiner absoluten Weise vom Identitätstheoretiker unterscheiden. „Neuron“ bspw. kann entweder als physiologischer Begriff (als Bezeichnung für eine Art von menschlichen Zellen) verstanden werden, oder als (teleo-)funktionaler Begriff (als Bezeichnung für einen Übermittler elektrischer Ladungen). In *beiden* Fällen steht es für eine instantiierbare Entität – bzw. für eine Rolle, die durch die eine Gruppe

²⁵⁷ Mit dem *homunktionalen* Ansatz des Funktionalismus meint Lycan subpersonale funktionale Beschreibungen, wie sie in der (kognitiven) Psychologie üblich sind (Sprach- oder Kontrollzentren; Problemlösekomponenten und dergleichen; vgl. ebd., 1987/2007, S. 435ff). In der Theoriebildung oder funktionalen Beschreibung eines Systems wird hier gewissermaßen ein interner *Homunkulus* eingeführt, der bestimmte Ziel verfolgt.

fundamentalerer Gegenstände gespielt wird. Damit *wird auch der Identitätstheoretiker zum Funktionalisten* – der mentale Entitäten auf einer sehr niedrigen Abstraktionsebene ansiedelt. (S. 450f; Hervorh. im Orig.)

Auch wenn es vom Grundsatz her ein richtiger Schritt ist, den Lycan unternimmt, wenn er Identitätstheorie und Funktionalismus zu vereinen sucht, denkt er ihn aber nicht konsequent zu Ende. So mag es sein, dass man auf verschiedenen Abstraktionsebenen einer funktionalen Hierarchie ansetzen kann und ein Organismus so nicht nur eine funktionale Beschreibungsebene hat, sondern potentiell unendlich viele. Es ist aber ebenso eine Tatsache, dass auch die strukturelle Seite nicht nur aus einer Ebene besteht. Genauso wie es also nicht *die* funktionale Ebene gibt, gibt es also auch nicht *die* strukturelle Ebene. Natürlich hat Lycan hingegen recht, dass die „klassische“ Identitätstheorie eine Ebene herausgreift und behauptet, dass nur auf dieser – wie er sagt – neuroanatomischen Ebene, das Mentale angesiedelt sei.

Wie sich gezeigt hat, ist eine solche Vereinseitigung und ontologische Fixierung auf eine bestimmte Organisationsebene aber nun keineswegs nötig. Vielmehr gilt die hier entwickelte psychophysische Identitätskonzeption für alle Ebenen (bzw. ihre Prozesse und Eigenschaften) in der mereologischen Hierarchie eines Mechanismus. Das Psychische ist nicht auf irgendeiner bestimmten Ebene allein angesiedelt, sondern durchdringt quasi alle Ebenen zugleich, weil die Ebenen selbst eine Einheit darstellen. Vielleicht lässt es sich mittels funktionaler Modelle besser (greifbarer, anschaulicher...) analysieren, daraus sollten aber keine *ontologischen* Schlüsse im Sinne eines metaphysischen Funktionalismus gezogen werden (die bessere Analysierbarkeit könnte sich viel eher als ein Artefakt unserer eigenen kognitiven Verfasstheit herausstellen). Während Lycan also behauptet, dass die hierarchische Struktur funktionaler Ebenen etwas über die Ontologie der Natur aussagt, nämlich, dass sie

essentiell funktionalistisch verfasst ist, bleibt der mechanistische Ansatz ontologisch neutral. Er ist sowohl mit funktionalistischen als auch identitätstheoretischen Annahmen kompatibel (vgl. Polger, 2004, S. 209).

Es ist der Sache allerdings nicht zuträglich, wenn man entgegen Lycan versucht, die („klassische“) Identitätstheorie *gegen* den Funktionalismus zu verteidigen, wie Polger (2004) dies unternimmt und im Kern weiterhin die – letztendlich auch nach Maßgabe seiner eigenen Ausführungen, bei denen er die Implikationen des mechanistischen Ansatzes berücksichtigt – zweifelhafte These *Place'* vertritt, dass mentale Zustände Gehirnzustände seien (vgl. Polger, 2004, S. 241, S. 181ff). Diese These ist nicht unbedingt falsch, aber wenn man sich wie Polger als Verteidiger der „klassischen“ Identitätstheorie gebärdet, führt dies unweigerlich auch zur „klassischen“ (und abwegigen!) Interpretation der psychophysischen Identität, nämlich, dass mentale Zustände *nichts anderes als* Gehirnzustände sind.

7.3.2. Aristoteles' Vermächtnis

Wenn sich die Struktur also nicht einfach in der Funktion auflösen lässt, wie lässt sich ihre Beziehung dann charakterisieren? Es ist der mechanistische Erklärungsansatz, der wohl am besten vor Augen führt, dass Struktur und Funktion keine Gegensätze sind, sondern in einem solchen Mechanismus zugleich verwirklicht sind. Es sind *agierende Entitäten* verschiedener mereologischer Komplexitätsstufen, die hierbei qua ihrer ebenenübergreifenden mikrostrukturellen *Konfiguration* eine bestimmte (mentale) Leistung oder eben *Funktion* erzeugen. Struktur und Funktion sind hier untrennbar miteinander verbunden.

Ein schönes illustrierendes Beispiel findet sich bei Dörner (1998/2008), bei dem er den Mechanismus zur Hervorbringung des Geistes mit dem

Mechanismus einer Windmühle zum Mahlen von Korn vergleicht. So gilt hier, dass die Aktivität der Windmühle durch ein Zusammenspiel verschiedener Variablen reguliert wird. Über ein schräg geneigtes, mit den Schäften der Windmühlenflügel gekoppeltes Rüttelbrett wird dem Mahlwerk Korn zugeführt. Wenn sich die Flügel schnell drehen, kommt es zu einem starken Rütteln und hoher Zufuhr an Korn. „Viel Korn im Mahlwerk aber erzeugte einen großen Widerstand für das Antriebsaggregat, nämlich die Flügel“ (S. 27), was mit einer Verringerung von deren Umdrehungsgeschwindigkeit einhergeht. Dies führt dann aber wiederum zu einer Abnahme der Rüttelung und Verringerung der Kornzufuhr, was auch den Widerstand für die Flügel mindert.

Wenn man den ganzen Mechanismus geschickt einstellt, ergibt sich aus ihm ein ruhiger Gleichlauf des Mahlwerks. Es stellt sich also ein Gleichgewicht ein; die Mühle beherrscht sich selbst. Aber es sitzt kein *andersartiger* Geist in ihr. Der ‚Geist‘ der Windmühle ist die Gestaltung der Materie, das spezifische Muster der Kausalbeziehungen zwischen den Flügeln, der Rüttelaktivität und dem Mahlwerk. Er ist also nicht ein bestimmtes Organ, keine bestimmte Instanz, sondern liegt in der spezifischen Form der Zusammenfügung der Materie. (ebd., S. 27f; Hervorh. im Orig.)

Und eine solche Beziehung von Struktur und Form kennzeichne nun auch das Körper-Geist-Verhältnis, wie es der mechanistische Erklärungsansatz aufzeigt. Die hier beschriebene Einheit von Struktur und Form kann man als ‚Vermächtnis Aristoteles‘ (1995a) ansehen, der sehr weitsichtig ausgesagt hat, dass der Geist als das *formende Prinzip eines natürlichen Körpers* fungiert. Hierüber kann Aristoteles den Platonischen Dualismus zurückweisen, da Form (Funktion) und Materie (Struktur) eines Einzeldings nicht zwei verschiedene Objekte sind, sondern Aspekte *ein und desselben* Gegenstandes (vgl. S. 29/ 412b). So wie man über ein Haus aussagen kann, aus welchem Material es gebaut ist und welche konkrete

Form es (zur Erfüllung eines bestimmten Zwecks: Schutz vor Witterungseinflüssen...) besitzt, besteht auch die Seele aus einem bestimmten Material in einer ganz bestimmten Form, die ihr Wesen, ihre Zweckmäßigkeit ermöglicht (vgl. ebd., S. 4f/403b). Was allerdings nicht heißt, dass Struktur und Form dasselbe sind. Das wäre begrifflich in der Tat unsinnig, aber sie sind zugleich in einem jeden Objekt verwirklicht, bilden eine Einheit und charakterisieren hierdurch insbesondere die Körper-Geist-Beziehung. So ist die Seele nicht dasselbe wie der Körper, dennoch ist sie körperlich, weil Zustände der Seele immer auch Zustände des Körpers sind.

Allerdings sollte man die These, dass die Seele das Formprinzip des Körpers ist, nicht in der Hinsicht missverstehen, dass es nun doch die Form selbst ist, die das Geistige ausmacht, ganz im Sinne der funktionalistischen Interpretation, dass die Materie nur ein beliebig austauschbarer und damit irrelevanter Träger („Hardware“) des Psychischen sei und der Geist vielmehr die implementierte und ihn in seinem Wesen bestimmende „Software“. Aristoteles sollte also nicht so verstanden werden, dass er die Form selbst als Seele bezeichnet hat, sondern es ist die *geformte Materie* und zugleich die – wenn man möchte – *materialisierte Form*. Die Seele ist also in der Tat nicht einfach der Körper, sondern sie ist der Körper, der eine ganz bestimmte Form qua seiner (mikro-)strukturellen Verfasstheit aufweist. Denn wie die allgemeingültige Bestimmung eines Mechanismus lautet: „A mechanism is a *structure performing a function* in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena“ (Bechtel, 2008, S. 13; Hervorh. M.L.). Genau genommen sagt der Begriff der Struktur eigentlich selbst schon aus, dass wir es hier nicht einfach mit einem beliebigen „Klumpen Materie“ zu tun haben,

sondern dass hier eine bestimmte Form (zur Verwirklichung einer Funktion) vorliegt.

Wenn also zuvor die Rede davon war, dass wir bei dem Herausgreifen von Eigenschaften eines Systems, die diesem faktisch zukommen, auf (mikro-)strukturelle Konfigurationen dieses Systems rekurren, dann ist hiermit nicht gemeint, sich auf eine relativ statische (z.B. neuroanatomische) Größe hierbei zu beziehen. Vielmehr ist die Struktur einer bestimmten Organisationsebene (eines Mechanismus) immer schon eine „geformte (und aktive) Materie“ und zwar so geformt, dass sie eine bestimmte Funktion erfüllen kann.

7.3.3. Identität versus Realisierung

Natürlich ist es möglich, weiterhin davon zu sprechen, dass hier eine bestimmte Funktion *realisiert* wird. Man kann also sagen, dass der mechanistische Ansatz die Annahme beinhaltet, dass eine bestimmte kausale Rolle durch einen oder mehrere Mechanismen instantiiert oder eben auch *realisiert* werden kann. Mit der funktionalistisch konzipierten (eindimensionalen) Realisierungstheorie, wie sie im letzten Kapitel behandelt wurde, hat dies aber nicht mehr viel zu tun. Wie Polger (2004) sagt, gibt sich der funktionalistische Ansatz der Realisierung damit zufrieden, dass er mechanistische Rollen identifiziert. Eine funktionalistische Analyse ergibt daher eher eine „mechanistische Skizze“ als eine ausgereifte mechanistische Erklärung (vgl. ebd., S. 200). Eine rein funktional spezifizierte Entität sei wie eine Blackbox, über die nicht viel gesagt werden könne. Wenn man eine solche Rolle isoliert betrachtet, könnte sie in der Tat von vielen Realisierern besetzt werden. „But integrating the isolated role description into a multilevel mechanism may well provide – for example,

spatial and temporal constraints – that provide information about the lower-level mechanisms that fill the role in a particular system“ (ebd.).

Nach der Terminologie des Kausale-Rollen-Funktionalismus postuliert der mechanistische Ansatz eine mechanistische Hierarchie, bei der zwischen einer kontextbezogenen Rolle auf der höherstufigen Ebene, einer isolierten Aktivität auf einer mittleren Ebene und wiederum den diese Aktivität konstituierenden Vorgängen auf der niederstufigen Ebene unterschieden wird, wobei die Bezeichnung höher-, mittel- und niederstufig natürlich nur relative Gültigkeit in Abhängigkeit von der Perspektive hat, die man gegenüber der Aktivität auf einer bestimmten Ebene der mechanistischen Hierarchie einnimmt (vgl. Craver, 2001, S. 67). Der Kausale-Rollen-Funktionalist hat deshalb insofern recht, als dass die Idee einer Ebene einen relativen und keinen absoluten Gehalt hat. Falsch liegt er hingegen bei der Annahme, dass man eine komplette Erklärung eines Phänomens auf einer einzigen Ebene erhalten kann. Der mechanistische Ansatz umfasst grundsätzlich eine *ebenenübergreifende* Erklärungsstruktur (vgl. Polger, 2004, S. 201f).

Dies mag der Anlass für Gillett gewesen sein, eine multidimensionale Form der Realisierung in die Welt zu rufen (vgl. Kap. 6.5.3). Gemäß seiner Theorie werden alle Individuen durch mikrophysikalische Individuen *konstituiert* und alle Eigenschaften durch mikrophysikalische Eigenschaften *realisiert*. Er nimmt also eine Kette von Realisierungsrelationen zwischen Eigenschaften an, die in *verschiedenen* Individuen instantiiert werden, die sich auf den *verschiedenen* natürlichen Organisationsebenen (eines Objektes) befinden. Gillett formuliert seinen Ansatz dabei aber weiterhin nach dem Muster des Kausale-Rollen-Funktionalismus, wenn er bei seinem Beispiel des Diamanten sagt, dass die Eigenschaften und Beziehungen der Kohlenstoffatome die kausale Rolle der Härte des Diamanten einnehmen, womit die Härte des Diamanten folglich durch die

Beziehungen und Eigenschaften der Kohlenstoffatome *realisiert* wird (vgl. ebd.). Damit untergräbt er offenbar wiederum die Tatsache, dass die Ebene der Kohlenstoffatome nur eine bestimmte Ebene in einer mechanistischen Hierarchie ist, die die Eigenschaft der Härte des Diamanten hervorbringt. Natürlich könnte man hier davon sprechen, dass die Eigenschaft der Härte des Diamanten, die – gemäß der hiesigen Argumentation – als Beschaffenheit einer makroskopischen Strukturebene aufgefasst werden müsste, insofern durch die atomare Ebene, auf der eine bestimmte mikrostrukturelle Konfiguration von Kohlenstoffatomen vorliegt, realisiert wird, indem zwischen den verschiedenen strukturellen Ebenen eine *Identität* vorliegt. Aber es ist genau diese Form einer Identitätsaussage, die Gillett vehement bestreitet, weil er glaubt, dass eine Identitätsannahme bzgl. der agierenden Entitäten der verschiedenen Ebenen eines Mechanismus implizieren würde, dass letztendlich alle Entitäten *nichts anderes als* mikrophysikalische Entitäten seien und alle Eigenschaften eines Systems letztendlich *nichts anderes als* dessen Mikroeigenschaften. Dass diese Sorge unberechtigt ist, wurde bereits thematisiert. Da er nun zwangsläufig vor der Aufgabe steht abzusichern, dass höherstufige Eigenschaften wie das Psychische über Kausalkräfte verfügen, die sich nicht auf ihre mikrophysikalischen „Realisierer“ zurückführen lassen (denn nach dem weitgehend akzeptierten Diktum Alexanders (1920) muss alles Existierende über eigene Kausalkräfte verfügen), entwickelt er hierfür sein begriffliches „Monstrum“, das äußerst undurchsichtig und inhaltlich fragwürdig ist bzw. den „Geruch“ einer Ad-hoc-Konstruktion zur Absicherung seiner (auf falschen Prämissen fußenden) Weltansicht hat: „Property/relation instance(s) F_1 - F_n realize an instance of a property G, in an individual s , *if and only if* s has powers that are individuating of an instance of G in virtue of the powers contributed by F_1 - F_n to s or s 's constituent(s), but not vice versa” (Gillett, 2002, S. 322, Hervorh. im Orig.).

Auch wenn diese konkrete Form der Ausbuchstabierung einer (multiplen) Realisierungstheorie missglückt ist, sollte die grundsätzliche Idee einer multiplen Realisierung allerdings nicht schon dadurch scheitern, dass wir mit dem Begriff der Realisierung für gewöhnlich eine eindimensionale Rollen-Ersetzungsbeziehung bezeichnen und daher der Begriff hierauf festgelegt sei, wie Polger und Shapiro (2008) es Gillett vorwerfen. Natürlich könnte man den Realisierungsbegriff in der oben vorgeschlagenen Form bezüglich der Härte des Diamanten vertreten und den Grundsatz formulieren, dass höherstufige Eigenschaften eines Systems Struktureigenschaften sind, die dadurch realisiert werden, dass dieses System bestimmte mikrostrukturelle Konfigurationen auf den verschiedenen kompositorischen Organisationsebenen aufweist. Wenn wir aber zu Recht davon ausgehen, dass es unter Maßgabe des mechanistischen Ansatzes nicht mehr zwangsläufig nötig ist, eine Analyse des kausalen Profils einer mentalen Eigenschaft (in Form von Verhaltensinputs und -outputs) zu betreiben, um ihren Realisierer zu bestimmen, würde ein solcher multi-dimensionaler Ansatz der Realisierung letztendlich nur die Aussage beinhalten, dass eine psychophysische Identität auf dem Niveau von Eigenschaftsinstantiierungen (bzw. lokaler Prozesse oder Zustände) vorliegt und scheint bis auf eine Erweiterung unserer Sprachvielfalt durch die Möglichkeit eines synonymen Gebrauchs von Realisierung und (lokaler) Identität aber keine weitere Bedeutung zu haben.

7.3.4. Identität versus Konstitution

Die modernen Mechanisten sagen zu Recht, dass mechanistische Erklärungen ontologisch neutral sind und nichts über die ontologische Natur der Entitäten aussagen, deren Aktivitäten mittels eines Mechanismus beschrieben werden (vgl. z.B. Machamer, Darden & Craver, 2000). Wie

schon erwähnt, ermöglicht dies, dass man sie zugleich funktional als auch identitätstheoretisch interpretieren kann. Manchmal wird aber auch die These vertreten, dass man die Beziehung der agierenden Entitäten oder strukturellen Konfigurationen der verschiedenen Ebenen eines Mechanismus nicht als Identitäten kennzeichnen dürfe, da Identitätsausagen im Hinblick auf Teil-Ganzes-Hierarchien, um die es sich hier handelt, generell scheitern würden (vgl. z.B. Bechtel, 2012, S. 46).

Since the same thing is not at two different levels, identity claims do not cross levels. For example, they do not relate something and its parts. The brain activity in question is not at a lower level of organization than the cognitive activity. Rather, the very same thing is being picked out in two different ways by the two descriptions. The two descriptions may pick out the process from different perspectives [...]. But it is still the operation that is being identified with itself. (ebd.)

Bechtel verfällt hier offensichtlich der Ansicht, dass die – in der Terminologie Deutschers – *Identifikation als* die einzige Möglichkeit einer Identitätsbeziehung darstellt bzw. dass die „Ungleichheit“ der miteinander zu identifizierenden Größen nur „subjektivistisch“ (im Sinne divergierender perspektivengebundener Repräsentationen) aufgelöst werden kann bzw. muss. Gegen einen solchen kategorischen Ausschluss einer „mereologischen Identität“ spricht zunächst, wie im Vorfeld hinlänglich aufgezeigt worden sein sollte, dass die Instantiierung einer höherstufigen (mentalen) Eigenschaft verschiedene mikrostrukturelle Konfigurationen aufweist, die durchaus *miteinander* identifiziert werden können bzw. sogar identifiziert werden müssen.

Stattdessen heißt es bei den Mechanisten häufig, dass der korrekte(re) Begriff der der Konstitution wäre und dass dieser eine im Hinblick auf Teil-Ganzes-Beziehungen adäquatere Alternative zum Identitätsbegriff darstellen würde (vgl. Craver, 2007a). Im Hintergrund dürfte auch hier

der „schlechte Ruf“ der „klassischen“ Identitätstheorie stehen, die – in der sich durchsetzenden Formulierung durch Smart – eine einseitige Interpretation in Form eines kruden Neuroreduktionismus offeriert, obwohl dieser Interpretation keine Notwendigkeit, geschweige denn (empirische) Plausibilität zukommt. Nach dieser Denkweise sei Konstitution (im Gegensatz zu Identität) eine asymmetrische Relation, das heißt, ein Klumpen Lehm konstituiert die Statue, aber nicht andersherum (vgl. Baker, 1997).

Ein erstes Hindernis für eine solche Ansicht besteht allerdings bereits darin, dass Konstitution für gewöhnlich als eindimensionale (*intra-level*) Beziehung verstanden wird (vgl. Paul, 2007). Denn was Konstitution für gewöhnlich aussagt, ist, dass ein Gegenstand aus einer bestimmten *Art von Material* geformt ist, so wie eine Statue z.B. aus dem Material Ton geformt ist. Unzweifelhaft beinhaltet eine solche Konstitutionsaussage etwas völlig anderes als eine Identitätsaussage (zumindest in der hier verfolgten Form).

Dennoch findet sich bei Craver und auch Bechtel manchmal die Ansicht, dass auch wenn man – entgegen dem normalen Begriffsverständnis – eine (*inter-level*) Konstitutionsbeziehung zwischen den agierenden Entitäten der verschiedenen Ebenen annehmen könnte, die zugleich symmetrisch wäre: „The relation is *symmetrical* precisely because the mechanism as a whole is fully *constituted* by the organized activities of its parts: a change in the parts is manifest as a change in the mechanism as a whole²⁵⁸, and a change in the mechanism is also a change in at least some of its component parts“ (Craver & Bechtel, 2007, S. 554; Hervorh.: M.L.).

²⁵⁸ Hier liegt eine gewisse Ungenauigkeit vor, denn nicht jede Veränderung auf Seiten der Komponenten muss zu einer Veränderung des Ganzen führen. Eine solche („aufwärts gerichtete“) mereologische Supervenienz funktioniert nur mit der Zusatzannahme, dass es sich um eine *kausal relevante* Komponente der Mikrostruktur handelt (vgl. Kap. 6.3.1).

Dieses Konzept einer Konstitution hat nun aber nichts mehr mit seiner herkömmlichen Bedeutung zu tun, sondern ist offensichtlich eine kaschierte (lokale) Identitätsbehauptung. Denn statt einer Auskunft darüber, dass ein Mechanismus durch eine bestimmte Materialart realisiert wird oder dass das Material den durch den Mechanismus gekennzeichneten Gegenstand in irgendeinem Sinne konstituiert, nehmen Craver und Bechtel vielmehr an, dass es *eine* bestimmte mechanistische Struktur gibt, die aus mehreren organisierten kompositorischen Ebenen agierender Entitäten besteht und eine bestimmte (mentale) Leistung hervorbringt. Wie Craver (2001) selbst an anderer Stelle sagt, sind mechanistische Beschreibungen „distinct perspectives on an activity in a hierarchically organized mechanism“ (S. 67). Es ist *ein* Geschehen, das auf verschiedenen natürlichen kompositorisch individuierten Organisationsebenen herausgreifbar ist. Dass eine solche Kompositionsbeziehung etwas anderes ist als eine Konstitutionsbeziehung, hat Envine (2011) systematisch nachgewiesen.

In diesem Sinne kritisieren auch Fazekas und Kertész (2011) die Interpretation, dass es sich bei der *inter-level*-Beziehung eines Mechanismus um eine Konstitutionsbeziehung handle, da die organisierte Ganzheit durch den mechanistischen Ansatz so erklärt wird, dass die Teile zusammengenommen genau das tun, was das Ganze selbst tut: „the overall behaviour of the parts is *identical* with the behaviour of the whole“ (S. 372; Hervorh. im Orig.). Der zentrale Gehalt des mechanistischen Ansatzes sei ihres Erachtens vielmehr, dass er erklärt, wie ein System eine bestimmte Leistung erbringt, indem er aufzeigt, auf welche Art und Weise seine Teile organisiert sind, dass sie *genau diese* Aufgabe erfüllen können. Wenn die organisierte Struktur der Entitäten der niedrigeren Ebenen eine andere Aktivität vollzöge als die der Entitäten der höheren Ebenen, dann wäre der Ansatz nicht dazu in der Lage, das sich auf den höheren Ebenen

zeigende Phänomen (mittels einer Beschreibung der Verfasstheit der niedrigeren Ebenen) zu erklären, und die organisierte Aktivität der Entitäten der niedrigeren Ebenen würde nicht die Ganzheiten der höheren Ebenen konstituieren. „In this sense, mechanistic explanations *identify* the activity of the spatial and temporal organization of the parts with the activity of the higher level whole” (ebd., S. 372; Hervorh. im Orig.). Es sei daher absurd anzunehmen, dass die Teile der niedrigeren Ebenen das Ganze der höheren Ebenen konstituierten, wenn ihre Aktivitäten nicht identisch sind. Damit gelte dann aber auch: „the constitution claim the mechanistic approach makes amounts to an identity claim: the causal role played by the organisation of the parts is the very same causal role what is played by the whole” (ebd., S. 373).

Auch wenn Fazekas und Kertész mit der Schlagrichtung ihres Arguments recht haben, dass es sich bei der Konstitutionsannahme um ein bloßes Lippenbekenntnis handelt, das den eigenen Voraussetzungen der Mechanisten widerspricht, zeigt sich sogleich, warum diese (zumeist²⁵⁹) davor zurückschrecken, eine Identitätsbeziehung anzunehmen. Denn eine solche Annahme führt schnell zu den Vorwürfen und Ressentiments, die gegenüber der „klassischen“ Identitätstheorie bestehen, und diese Position mit einem kruden Reduktionsmus gleichsetzen: „...by claiming that the whole is constituted by the organised activity of its parts the mechanistic approach inherently commits itself to the claim that whatever the

259 Entgegen dem obigen Zitat finden sich an anderen Stellen bei Bechtel oder auch McCauley durchaus (heuristische) Identitätsannahmen (vgl. Bechtel, 2001, 2007; McCauley & Bechtel, 2001). Diese beziehen sich allerdings nicht auf eine Identität der Ebenen einer mereologisch-mechanistischen Hierarchie, sondern auf eine Identifizierung *psychischer* Eigenschaften im Gehirn (ähnlich es die physische Realisierungstheorie annimmt, scheinen auch für Bechtel et al. mentale Eigenschaften so etwas wie Eigenschaften zweiter Ordnung zu sein, die sich nicht direkt in die Hierarchie natürlicher Organisationsebenen eines Organismus einordnen lassen, sondern nur indirekt über ihren neuronalen „Realisierer“).

whole does is something what is done by the organization of its parts” (Fazekas & Kertész; 2011, S. 373f). Aus der Kombination einer mereologisch-mechanistischen Hierarchie und einer Identitätsannahme wird so ein *Mikrophysikalismus* oder Mikroreduktionismus abgeleitet, der alles Geschehen und alle kausale Wirkmächtigkeit auf den höheren Ebenen auf die Mikrostruktur des betreffenden Systems zurückführt. Schließlich würde doch die Identitätsthese behaupten, dass die Struktur der höheren Ebenen (bzw. die Konfiguration der agierenden Entitäten) *nichts anderes als* die Struktur der niedrigeren Ebenen (bzw. deren Konfiguration agierender Entitäten des Mechanismus) sei. Ganz im Gegenteil wird ein solcher kausaler Drift, den Fazekas und Kertész hier unterstellen, gerade durch eine Identitätsannahme verhindert. Denn aus einer Identitätsannahme zwischen den verschiedenen Ebenen eines Mechanismus kann eben – entgegen dem verbreiteten Vorurteil, dem die „klassische“ Identitätstheorie Raum verschafft hat – *nicht* abgeleitet werden, dass es eine bestimmte (Mikro-)Ebene ist, die die „eigentliche Wirklichkeit“ ausmacht oder auf der die „eigentliche Kausalität“ stattfinden würde. Eine sozusagen *mit sich selbst identische* mereologisch-mechanistische Struktur kann keine Ebene aufweisen, die die „eigentliche Wirklichkeit“ darstellt oder auf der die „eigentliche Kausalität“ stattfinden würde. Das kausale Geschehen durchdringt alle Ebenen des betreffenden Mechanismus gleichermaßen. Das Scheitern der Vorstellung eines Primats der Mikroebenen unter Maßgabe einer Identitätsbeziehung wird im nächsten Kapitel noch eingehender behandelt.

Was jedenfalls klar gezeigt worden sein sollte, ist, dass sich der mechanistische Ansatz nicht hinter einer gekünstelten Konstitutionsbeziehung verstecken muss, um seinen Grundsatz der ontologischen Neutralität einlösen zu können. Es ist die zu Unrecht geschmähte Identitätsbeziehung,

die neben ihrer ohnehin mehr oder weniger zwingenden Gültigkeit unter den Voraussetzungen der Mechanisten auch den Grundsatz der ontologischen bzw. kausalitätsbezogenen Neutralität wahrt.

7.4. Zur Realität psychischer bzw. psychophysischer Kausalität

Reductionistically writers frequently claim to find this kind of talk mysterious, especially when it is claimed that lower levels are controlled from above, or even more, that there is a hierarchy of levels of control [...], with higher levels controlling lower levels. I suspect that they are bothered here by the feeling that some entity under one description is said to be the cause of its own behavior under another description. (Wimsatt, 1976, S. 210)

Bei der Besprechung der mereologisch-mechanistisch fundierten Theorie psychophysischer Identität ist bisher ein entscheidender Aspekt weitestgehend unbeachtet geblieben: die Frage nach der Möglichkeit psychischer bzw. psychophysischer Kausalität. Was bedeutet es unter den Voraussetzungen dieser Theorie, dass wir mittels unserer Gedanken Pläne schmieden und diese buchstäblich in die Tat umsetzen können, indem wir selbst (als geistige Wesen) als Steuerungsinstanz unseres Körpers auftreten? Inwiefern ist es also möglich, das Phänomen *mentaler Verursachung* hinreichend zu integrieren, das dem ontologischen Dualismus so schwerwiegende Probleme bereitet (vgl. Kap. 3.1)? Ist auf der anderen Seite der Preis für eine mit naturwissenschaftlichen Constraints kompatible Erklärung der Möglichkeit mentaler Verursachung dessen Degradierung zu einem rein neuronalen oder physischen Geschehen? Um eine hartnäckige philosophische Intuition aufzugreifen: Würde unter den Voraussetzungen einer psychophysischen Identitätsannahme in der „materiellen Welt“

nicht alles ganz genauso ablaufen, ohne dass es Geist und Bewusstsein gibt? Und wird aus einem psychologisch beschreibbaren Handlungsgrund so nicht vielmehr ein rein physisch determiniertes, naturgesetzlich-*ursächliches* Geschehen, das keinen Raum für die Freiheit von Entscheidungen und Handlungen lässt?

7.4.1. Inter-Level-Kausalität als mechanistisch vermittelter Effekt

Einen für den hiesigen Kontext interessanten Fall schildert Patricia S. Churchland (1993), den sie aus einem in den USA verbreiteten Kochbuch (*Betty Crocker Cookbook*) entnommen hat: Hier beschreibt die Autorin Betty Crocker zunächst annäherungsweise korrekt, dass Mikrowellen ihre Funktion der Beschleunigung der Molekülbewegungen in den verwendeten Lebensmitteln verdanken. Falsch liegt sie allerdings, wenn sie erklärt, dass die Stimulation der Molekülbewegung bzw. das Aneinanderreiben der Moleküle und die hierbei entstehende Spannung das Phänomen der Wärme *erzeugt*. Die Fehlannahme besteht darin, dass sie annimmt, dass Wärme *kausal* durch die Erhöhung der mittleren kinetischen Energie der Moleküle hervorgerufen wird. Tatsächlich *ist* Wärme die mittlere kinetische Energie (in Gasen) bzw. das Schwingungsmuster (in festen Stoffen) von Molekülen (vgl. Kap. 6.5.1; Becker-mann, 1999). Hier liegt eine *Identitätsbeziehung* vor, sodass Betty Crockers Annahme einer Kausalrelation eine abwegige Selbstverursachung bedeuten würde, wie sie auch Wimsatt im obigen Zitat beschrieben hat.

In Bezug auf Inter-Level-Kausalität in mechanistischen Systemen haben Craver und Bechtel (2007) daher einen Ansatz begründet, der aufzeigt, dass es sich hierbei nicht um Kausalität im gewöhnlichen Sinne handelt, sondern um ein Phänomen, dass sie als *mechanistisch vermittelten Effekt* bezeichnen (vgl. ebd., S. 547). Sie legen dar, dass es keine echte Verursa-

chung „von oben nach unten“ oder „von unten nach oben“ innerhalb eines Mechanismus geben kann, weil es unsinnig ist anzunehmen, dass sich eine agierende Struktur gewissermaßen selbst verursacht. Denn wenn die agierenden Entitäten der verschiedenen Ebenen eines Mechanismus in einer Identitätsbeziehung stehen, dann ist es absurd zu sagen, dass das Verhalten der Teile das Verhalten des Ganzen *verursacht* hat oder andersherum.

Causation is frequently described as a kind of cement, glue, spring, string, or some other physical transmission or exchange from one object, process or event to another through contact action or through a propagated signal [...]. Such a conception of causation as a physical connection between two things does not accommodate interlevel causes between mechanisms and their components because mechanisms and their components are not distinct events, objects, or processes. Given the compositional relations between mechanisms and their components, the space-time path of the mechanism includes the space-time path of its components. They coexist with one another, and so there is no possibility of their *coming to* spatiotemporally intersect with one another. (ebd., S. 551f; Hervorh. im Orig.)

Wenn hier festgestellt wird, dass die Anwendung des Kausalitätsbegriffes auf Top-Down-Verhältnisse (innerhalb eines Mechanismus) unangebracht ist, scheint hiermit eine Verursachungsrichtung ausgeschlossen zu werden, die aber doch einen Normalfall in der wissenschaftlichen Erklärungspraxis darzustellen scheint. Nämlich, dass es bestimmte Ereignisse auf den Mikroebenen eines Systems gibt, die dessen makroskopische Beschaffenheit oder dessen Verhalten als Ganzes verursachen. So verändert die Einnahme von bestimmten Psychopharmaka beispielsweise den Stimmungszustand der betroffenen Person. Warum soll hier nun keine Kausalbeziehung vorliegen?

Craver und Bechtel (2007) erläutern dies mit Bezug auf Kandel et al. (1996) am Beispiel der Informationsaufnahme und -verarbeitung im Auge. Auf dessen höchster Ebene findet eine Umwandlung von Licht in

neuronalen Aktivität des optischen Nerven statt, wobei dieser Prozess in Komponenten der niedrigeren Ebenen und ihre Aktivitäten zerlegt werden kann. Das Licht erreicht das Auge und wird durch die Linse invertiert und fokussiert. Es findet eine Projektion in die Retina statt, wo die im Licht enthaltene Information in ein Muster neuronaler Aktivität im optischen Nerv umgewandelt wird. Diese Umwandlung des Lichts in neuronale Aktivität kann wiederum selbst in verschiedene Komponenten zerlegt werden, insbesondere die Stäbchen und Zapfen, die ihren elektrischen Ladungszustand in Abhängigkeit von den Charakteristika des Lichtstimulus (vor allem Wellenlänge und Intensität) verändern. Auf einer noch niedrigeren Ebene wird auch die Aktivierung der Stäbchenzelle durch einen bestimmten Mechanismus aufrechterhalten. Licht wird absorbiert und aktiviert Rhodopsin, das darauf G-Proteine stimuliert. Diese G-Proteine aktivieren die zyklische GMP-Phosphodiesterase, welche die Umwandlung des second Messengers cGMP zu 5'GMP befördert. Die Absenkung der cGMP-Konzentration bewirkt eine Schließung der Natrium-Kanäle, reduziert den Natriumeinstrom und hyperpolarisiert somit die Zelle. Nach Craver und Bechtel (2007) enthüllt hierbei jede neue Zerlegung bzw. Dekomposition eines Mechanismus in seine Teile einen neuen Mechanismus einer niedrigeren Ebene, bis der Mechanismus mit einzelnen Items, für die keine weitere Zerlegung mehr möglich sei, seine Talsohle erreicht (vgl. ebd., S. 549).

Nun stellt die Veränderung der Konformation von Rhodopsin eine Phase in Bezug auf den Pfad der Signalweiterleitung dar. Obwohl aber die strukturelle Veränderung von Rhodopsin die Ursache des von der Zelle erzeugten elektrischen Signals ist, ist sie nicht als die *Ursache* der Signalweiterleitung anzusehen. Wie Craver und Bechtel (2007) sagen, könne man zwar niemanden daran hindern, von Ursache oder Verursachung zu sprechen, aber dann wird der Ursachenbegriff auf Kontexte ausgedehnt, die

bereits vollkommen durch kompositorische (mereologische) Verhältnisse charakterisiert werden können (vgl. ebd.). Man macht sich so also einer gewissen Begriffsverbiegung schuldig, die schnell zu Missverständnissen führt und mit Widersprüchlichkeiten beim grundsätzlichen Verständnis von Kausalität verbunden sein kann.

Wie Craver und Bechtel (2007) daher zu Recht betonen, ist Kausalität eine *asymmetrische* Relation (A bewirkt B, während B aber nicht zugleich A bewirkt), während alle wissenschaftlich interessanten Fälle einer Inter-Level-Verursachung *symmetrische* Relationen sind, denn aufgrund der mereologischen Beziehungen innerhalb mechanistischer Hierarchien gilt: „components act as they do because of factors acting on mechanisms, and mechanisms act as they do because of the activities of their lower-level components“ (ebd., S. 553).

Ein Problem, das den Diskurs über Inter-Level-Kausalität im Kontext von Mechanismen so schwierig macht, könnte sein, dass die Experimente, die zur Erkundung, ob eine bestimmte Komponente Teil eines Mechanismus ist, denselben Aufbau haben wie solche, die zur Untersuchung von Kausalrelationen auf einer Ebene herangezogen werden (vgl. Craver, 2002). Craver und Bechtel (2007) erläutern dies an Korrelationsbefunden als auch Läsionsstudien (vgl. ebd. S. 553f) und ziehen hieraus folgendes Fazit:

Such experimental techniques involve either intervening to alter a component of the mechanism and observing the behavior of the mechanism as a whole or intervening to alter the behavior of the mechanism as a whole and observing the behaviors of one or more of its parts. In using both sorts of strategies, or in simply seeking correlations between the presence of components and the behavior of whole mechanisms, scientists seem to be exploiting a symmetrical relation. Yet, typically experimental techniques for testing causal claims work from cause to effect and not (generally) vice versa. One cannot generally intervene to change the effect and there-

by alter the cause. Those who endorse the asymmetry of causation will thus have another reason to find interlevel causation mysterious. (ebd., S. 554)

Aus diesen Gründen sprechen Craver und Bechtel (2007) in diesem Zusammenhang statt von einer Inter-Level-Kausalität von *mechanistisch vermittelten Effekten*, die sie vom herkömmlichen Begriff der Kausalität abzugrenzen versuchen: „Mechanistically mediated effects are hybrids of constitutive and causal relations in a mechanism, where the constitutive relations are interlevel, and the causal relations are exclusively intralevel“ (S. 447). Hierbei handelt es sich um eine symmetrische Relation, weil der Mechanismus als ein Ganzes vollkommen durch die organisierte Aktivität seiner Teile konstituiert wird (bzw. ein Identitätsverhältnis vorliegt). Hier zeigt sich, dass für mechanistische Ebenen notwendigerweise das Prinzip der *mereologischen Supervenienz* gilt. Es kann also keine Veränderung einer mikrobasierten Makroeigenschaft eines Systems geben, ohne dass zugleich eine Variation der Mikrostruktur dieses Systems stattgefunden hat. Oder anders ausgedrückt: Jede Veränderung der Teile manifestiert sich als eine Veränderung des Mechanismus als Ganzem, während jede Veränderung des Gesamtmechanismus auch mit einer Veränderung bestimmter Komponenten des Mechanismus verbunden ist. Es besteht dabei kein Anlass, in solchen Fällen von Verursachung zu sprechen, weil diese hinreichend ohne den Rekurs auf Kausalität beschrieben werden können (vgl. ebd., S. 554).

Wenn man nun erneut den im menschlichen Auge installierten Mechanismus betrachtet, mag man im Hinblick auf *aufwärts gerichtete „Kausalität“* zwar in Versuchung sein zu sagen, dass die Aktivierung der cGMP-Phosphodiesterase, die die Umwandlung von GMP in 5'GMP katalysiert, die Hyperpolarisierung der Stäbchenzelle *verursacht*, welche wiederum *verursacht*, dass das Auge Licht in neuronale Aktivität umwandeln kann.

Dies ist dann aber nichts anderes als Betty Crockers Fehlschluss, denn die Aktivierung der c-GMP-Phosphodiesterase ist ein Teil der Hyperpolarisationsaktivität, welche Teil der vom Auge bewerkstelligten Umwandlung von Licht in neuronale Aktivität ist (vgl. ebd., S. 555). „However, by attending to the constitutive relationship and mechanistic mediation, we can both trace the intra-level causal processes within the eye and note that the transduction process is constituted by these sub-processes“ (ibd.).

In Bezug auf eine *abwärts gerichtete* „Kausalität“ scheint es naheliegend zu sein, dass die Aktivität der Teile durch das Agieren des gesamten Mechanismus verursacht wird: Wenn Licht auf das Auge fällt und verursacht, dass dieses einen visuellen Stimulus in neuronale Aktivität überführt, verursacht es die die Depolarisation der Stäbchenzellen. Dabei handelt es sich auch hier um verschiedene mechanistische Ebenen, die durch Teil-Ganzes-Beziehungen und die hiermit verbundene mereologische Supervenienz charakterisiert sind. Statt einer Kausalbeziehung handelt es sich um einen mechanistisch vermittelten Effekt: „Again, however, by tracking both the constitutive [resp. identity; M.L.] relation between the eye and its parts and tracing the intra-level causal processes within the eye, we can describe what is going on without needing to invoke interlevel causes“ (ibd.).

Gibt es aber nicht dennoch bestimmte Fälle, in denen unzweifelhaft eine echte Inter-Level-Kausalität vorliegt? Craver und Bechtel betrachten hierzu eine Vielzahl an Beispielen, von denen hier einige sehr aussagekräftige wiedergegeben werden sollen. Sie beschreiben den Fall, dass eine Generalin einen Herzinfarkt erleidet und hieran stirbt (vgl. ebd., S. 557). Auch hier scheint es das Aussetzen der Aktivität der Herzmuskulatur (als Komponente des Blutzirkulationssystems, das wiederum eine Komponente der Person der Generalin ist) zu sein, die ihren Tod *verursacht*. Statt nun zu schließen, dass es sich hier um aufwärts gerichtete Kausalität handelt,

kann man leicht aufzeigen, dass es sich auch hier um einen mechanistisch vermittelten Effekt handelt. Wenn das Herz aufhört zu schlagen, hört es auf, Sauerstoff und Nährstoffe in andere Gewebe und Organe des Körpers zu transportieren, sodass diese ihre Funktion einstellen. Folglich ist hiermit erklärt, wie eine Vielzahl an physiologischen Mechanismen aufhört zu funktionieren. „And insofar as that non-functioning constitutes the general's death, we explain her death. Notice that when we reach the state of the mechanism that constitutes the state of death, we do not say, with Betty Crocker, that it causes death. It just *is* death“ (ebd.; Hervorh.: M.L.) .

Ein weiterer Fall ist der Hotdog-Verkäufer Ignatius, der seinen mit Hotdogs befüllten Wagen mit großer Anstrengung zu einer bestimmten Häusercke bugsiiert. Wenn man nun die Frage stellt, wer oder was die Hotdogs (und damit die Moleküle, Atome etc., aus denen die Hotdogs bestehen) dazu gebracht hat, an der Ecke anzugelangen, scheint die Antwort klar zu sein: Ignatius. Womit offenbar zugleich ein Fall abwärts gerichteter Kausalität gegeben ist. Aber auch hier lässt sich aufzeigen, dass es sich um einen Fehlschluss bzw. zumindest um eine unpräzise Feststellung handelt:

In the sense that we normally think about mechanisms, this is a deviant case. The cart is not doing anything in virtue of which the hotdogs move. But if we are curious why the hot dogs moved with the cart (whereas neutrinos that happened to be in the cart would not necessarily move with it) we need to focus on the interaction of the hotdogs with the walls of the cart, and so on. Even though this is a deviant case in which the parts are not working parts, it nonetheless introduces a common way in which constituents are affected by higher-level changes: they are *carried along for the ride*. The hotdogs are merely along for the ride in the cart [...]. (ebd., S. 558; Hervorh. im Orig.)

Es ist daher unangebracht, wenn man annimmt, dass die Bewegung des Wagens die Bewegung der Hotdogs verursacht hat. Insofern die Hotdogs

ein Teil dessen sind, was Ignatius zur Ecke gebracht hat, bedarf es keinerlei Bezug auf eine mysteriöse Interaktion zwischen dem Ganzen und seinen Teilen. Vielmehr – und dies ist in diesem Kontext zentraler Grundsatz – *ist die Bewegung des Ganzen zugleich die Bewegung der Teile* (vgl. ebd.). Auch wenn es hier eine mechanistisch zu analysierende Beziehung zwischen dem Wagen und den Hotdogs gibt, sind hier auch nicht die Mechanismen der niedrigeren Ebenen maßgeblich. Das Geschehen ist also nicht durch die Atome, aus denen die Hotdogs bestehen, die Hotdogs selbst oder auch den Wagen verursacht. Letztendlich ist es (primär) Ignatius, der das Ereignis hervorruft. Auch hier ist es nach Craver und Bechtel (2007) wieder ein Hybrid aus Kausalität und Konstitution (bzw. Identität), der das Geschehen korrekt erklärt: „There is a perfectly intelligible *etiological* relationship (Ignatius pushing the cart) and a perfectly intelligible *constitutive* relationship (the cart’s containing the hotdogs, their molecules, etc.)“ (S. 559; Hervorh.: M.L.).

Zu guter Letzt wird eine weitere Fallgeschichte dargestellt, die aufgrund der hier thematisierten *psychologischen* Struktur für die vorliegende Arbeit sehr aufschlussreich ist. Die Autoren greifen hierzu auf einen Aspekt des Filmdramas *Citizen Kane* des US-amerikanischen Regisseurs Orson Welles zurück:

A boy’s mother buys him a sled. When later he is forcibly removed from both his mother and the sled, the event leaves a permanent memory trace. Through his adolescence and adulthood, this memory comes to stand for him as a singular emblem of motherly love and profound loss. At the moment of his death, the now old man recalls a winter afternoon on his sled, utters the word ‘Rosebud,’ and falls silent. Here again is an apparent case of top-down causation. The young boy’s tragic experience leaves a permanent trace in the neural and molecular circuitry of his brain. (Craver & Bechtel, 2007, S. 560)

Mit Verweis auf die Theorie des episodischen Gedächtnisses, wobei sie sich an der Darstellung von Eichenbaum (2002) orientieren, erläutern die Autoren, dass es sich auch hier um einen mechanistisch beschreibbaren Vorgang handelt: Der sensorische Input zur Zeit des Ereignisses führte – in Abhängigkeit von den Merkmalen der verschiedenen Charakteristika des Ereignisses – zu einem Aktivitätsmuster in verschiedenen Hirnregionen von Kane (des Protagonisten der Geschichte). Manche dieser Merkmale sind eher abstrakt (wie Gefühle gegenüber der Mutter, dem Vater oder bezüglich der Vorstellung von Besitz, Macht und Kontrolle), andere beziehen sich auf sehr spezifische und flüchtige Charakteristika von Kanes zu diesem Zeitpunkt vorhandener „Welt“ (z.B. die Kühle des Windes, der Klang einer Stimme usw.).

However this representation is achieved, it is preserved in the neural networks of the hippocampus (by virtue of the influence of cortical input on the strength of synaptic connections in that brain region). Over time, repeated activation of the hippocampus (perhaps during REM sleep and surely during future cases of remembering) re-excites these cortical patterns. The repeated re-excitation of this basic pattern (certainly shaped and altered with each occurrence) consolidates the representation in the cortex by changing the strengths of connections between cortical cells under this ‘tutoring’ activity of the hippocampus. So Kane leaves his childhood home for the city, his hippocampus starts to encode representations, the synapses in the hippocampus and the cortex begin to change their strengths, and myriad molecular mechanisms within the cells begin to churn in their own distinctive ways. And why did all of this happen? Because Kane lost his mother and his sled in the same crushing event. (ebd., S. 560f)

Auch hier gibt es also *ein* bestimmtes Kausalereignis, wobei zwischen den beteiligten Ebenen eine Konstitutions- bzw. Identitätsrelation besteht. Der Kortex von Kane muss dabei eine Organisation aufweisen, die eine Sensibilität für die Integration subtiler und abstrakter Merkmale der sozialen Umwelt aufweist wie die Beziehung zu Mutter und Vater,

Freundschaften, Hochzeiten oder sein Bedürfnis nach Kontrolle und Macht. Des Weiteren muss sein Gedächtnissystem dazu fähig sein, sich auf verschiedene komplexe Weisen zu verhalten. Es muss bestimmte Erinnerungen verstärken, verzerren oder sie mit anderen Erinnerungen in Verbindung setzen können. Außerdem muss es bestimmte Auswirkungen auf andere kognitive Systeme haben können wie auf diejenigen, die für Selbstbeobachtung und Planung verantwortlich sind. „So Kane’s episodic memory mechanism is put into a new environmental situation, one to which the mechanism as a whole is organized to be responsive, and the effects of this change in environment ramify through the other components of the memory system, including the hippocampus, its cells, and their constituent molecules. And [...] so the hippocampus, the cells and the molecules are enlisted in Kane’s ride down memory lane“ (ebd., S. 561).

7.4.2. Mentale Verursachung als „ebenendurchdringende“ Kausalität

Wenn Craver und Bechtel auf diese Weise in Bezug auf mechanistisch beschreibbare Systeme recht schlüssig die Möglichkeit einer echten Inter-Level-Kausalität zugunsten mechanistisch vermittelter Effekte aufgeben, schließen sie dann nicht auch die Möglichkeit einer echten psychischen bzw. psychophysischen Kausalität aus, die für gewöhnlich als *abwärts gerichtete* Kausalität aufgefasst wird? Denn wenn es mein Gedanke sein soll, der mich dazu bewegt, mich aufzurichten und zum Kühlschrank zu gehen, um mir eine Cola zu holen, dann wird doch auch hier eine Situation geschildert, bei der die vermeintliche Top-Down-Verursachung in Wirklichkeit „nur“ ein mechanistisch vermittelter Effekt ist und damit keine Verursachung im eigentlichen Sinne. Ist nicht auch die von den Autoren vorgenommene Interpretation des Falls von Citizen Kane in ebendieser

Weise zu deuten, dass es sich im Fall von psychischen Vorgängen letztendlich um „blinde“ neuronale, molekulare bzw. atomare Vorgänge handelt? Auch wenn hier von einer neuronalen Repräsentation mentaler (intentionaler) Zustände gesprochen wird, bleibt ihre kausale Wirklichkeit unklar oder wird sogar insgesamt infrage gestellt.

Auch wenn dies wahrscheinlich nicht von Craver und Bechtel so intendiert ist, laden sie dennoch zu einer solchen Interpretation ein, wenn sie sagen, dass ein mechanistisch vermittelter Effekt ein Hybrid ist aus einer Konstitutionsbeziehung (bzw. Identitätsbeziehung), die ebenenübergreifend (interlevel) ist, und einer *Kausalrelation*, die *ausschließlich ebenenspezifisch* (intralevel) sein soll (vgl. ebd., S. 447). Dabei steht eine rein *ebenen-spezifische* Kausalität zu ihren eigenen Aussagen in Widerspruch, denn an anderer Stelle sagen sie explizit, dass die Bewegung des Ganzen zugleich die Bewegung der Teile *ist* (vgl. ebd., S. 558). Es handelt sich also in allen geschilderten Fällen immer um ein Kausalgeschehen, das mehrere bzw. alle Ebenen umfasst, obgleich es tatsächlich keine Kausalität *zwischen* den beteiligten mechanistischen Ebenen gibt. Im Hintergrund dürfte hier wiederum die Vermeidung der Annahme eines Identitätsverhältnisses zwischen den agierenden Entitäten der verschiedenen Ebenen des Mechanismus stehen, obgleich eine solche Annahme erstens mehr oder weniger zwangsläufig aus den Prämissen des mechanistischen Ansatzes folgt und zweitens keineswegs mit den negativen Konsequenzen verbunden ist, die man in sträflicher Weise seit der „klassischen“ Identitätstheorie mit dem Konzept der psychophysischen Identität verbindet. Wenn sie deshalb nur von einer Konstitutionsbeziehung zwischen den Ebenen sprechen, die keine Identität implizieren soll, dann müssen sie sich zwangsläufig damit schwertun, den Kausalitätsbegriff adäquat auf mechanistische Hierarchien anzuwenden, was sich in der irreführenden An-

nahme einer *ebenenspezifischen* Kausalität als Teil ihres Hybridmodells zeigt.

Unter Heranziehung der hier entwickelten Identitätskonzeption wird diese Missverständlichkeit beseitigt. Jede mentale Eigenschaft hat eine bestimmte mentale Struktur und ist wiederum mit einer (potentiell unendlichen) Abfolge von Ebenen agierender Entitäten, wie sie der mechanistische Ansatz beschreibt, identisch. Damit ist aber auch ein mentales Geschehen, das z.B. eine körperliche Bewegung *verursacht*, zugleich ein neuronales, molekulares oder auch atomares Geschehen. Jede mentale Regung ist notwendigerweise zugleich eine „Regung“ der Zellen, Moleküle, Atome etc., aus denen meine Gedanken- und Gefühlswelt bestehen. Eine aufschlussreiche Aussage findet sich bei Stanislas Dehaene (2010), in der er sich gegen dualistische Gehirn-Geist-Vorstellungen bzw. die Skepsis gegenüber der Wirkmächtigkeit psychologischer Interventionen wendet:

Aber ich behaupte, dass es für einen Gegensatz zwischen wissenschaftlicher Psychologie und der Wissenschaft vom Gehirn keinerlei Grundlage gibt. Die Verschränkung der Organisationsebenen innerhalb unseres Gehirns ist so ausgeprägt, dass jede psychologische Intervention sich in unseren Hirnschaltkreisen bis hinunter zur Ebene der Zellen, Synapsen und Moleküle niederschlägt – das geht sogar bis zu einer Modifikation der Genexpression. Wenn also eine Erkrankung auf einer mikroskopischen neurobiologischen Ebene angesiedelt ist, heißt das noch lange nicht, dass sie nicht durch eine psychologische Intervention kompensiert werden könnte. (S. 291)

Psychologische Interventionen schlagen sich deshalb bis zur Ebene der Moleküle oder gar Atome nieder, weil die *Verschränkung* der Organisationsebenen, von der Dehaene spricht, dadurch begründet ist, dass die (kausal agierenden) Entitäten der verschiedenen Ebenen in einer Teil-Ganzes-Relation stehen, für die ein Verhältnis der Koextensivität bzw.

Identität besteht. Wenn etwas Teil eines Ganzen ist, dann muss es notwendigerweise in demselben räumlichen Feld und zum selben Zeitpunkt existieren wie das Ganze. Da ein Molekül aus Atomen besteht, muss eine Veränderung auf der molekularen Ebene immer auch ein Pendant auf der atomaren Ebene haben. Verändert sich irgendein Parameter des Moleküls, muss auch eine Veränderung bei den Atomen vorliegen, aus denen es besteht. Jede Veränderung eines mentalen Zustandes muss daher genauso mit einer Veränderung auf der neuronalen, mikrozellulären, molekularen, atomaren und subatomaren Ebene verbunden sein. Insofern hat Dehaene vollkommen recht, wenn er sagt, dass die Organisationsebenen so stark miteinander verschränkt sind, dass sich jede psychische Veränderung bis zur Ebene mikrophysikalischer Entitäten herab niederschlägt. Kommt es – z.B. im Rahmen einer Psychotherapie – zu einer kognitiven Umstrukturierung des semantischen Netzwerkes einer Person, dann muss auch eine Modifikation der Konfiguration aller niedrigeren Ebenen der entsprechenden Mechanismen stattgefunden haben. Denn eine mentale Eigenschaft ist etwas, das *alle* Ebenen des sie hervorbringenden Mechanismus „durchdringt“. Sie ist genauso ein mentales Erlebnis mit einer phänomenalen Qualität oder einem kognitiven Gehalt wie auch etwas, das gewissermaßen in den Atomen dieser mechanistischen Hierarchie „steckt“.

Hier zeigt sich erneut die Bedeutung des Prinzips der *mereologischen Supervenienz*. Auch wenn es selbst nicht dazu genügt, eine eigenständige Theorie des Geistes zu begründen (vgl. Kap. 6.3), liefert es dennoch eine sehr wichtige Beschreibung der Beziehung zwischen den Entitäten der verschiedenen Ebenen einer mereologischen bzw. mechanistischen Hierarchie. So kann es keine Veränderung einer mikrobasierten Makroeigenschaft eines Systems geben, ohne dass zugleich eine Variation der Mikrostruktur dieses Systems stattgefunden hat. Damit gilt aber auch,

dass es keine Veränderung der kognitiven und vor allem auch der phänomenalen Eigenschaften geben kann, ohne dass hiermit eine Veränderung der Mikroeigenschaften, d.h. der mikrostrukturellen Konfiguration des Systems verbunden ist.

Wenn nun der Eindruck bestehen mag, dass es mit der Ausräumung einer Inter-Level-Kausalität innerhalb einer mechanistischen Hierarchie zugleich keine mentale Verursachung geben kann, liegt das vor allem daran, dass eine gewisse Mehrdeutigkeit des Ursachenbegriffs und – hiermit zusammenhängend – die *zeitliche Dimension* vernachlässigt wird (vgl. Lüdmann, 2009, S. 122ff; 2010, S. 56ff). Der Kausalitätsbegriff, der sich in der Philosophie und auch den Naturwissenschaften durchgesetzt hat²⁶⁰, ist derjenige, den bereits Aristoteles (1995) als *causa efficiens* (Wirkursache) bezeichnet hat. Hierunter versteht Aristoteles primär den äußeren Anstoß einer Bewegung. Es sind hier also zwei bestimmte Größen, von denen eine auf die andere kausal einwirkt (natürlich kann es sich hierbei auch um Entitäten, Zustände oder Prozesse innerhalb eines Systems handeln), womit sich eine *in einem zeitlichen Verlauf* ereignende Verknüpfung von Ursache und Wirkung ergibt. So sei die Ursache „dasjenige, wovon her die Veränderung oder die Ruhe ihren ersten Anfang nimmt; so ist z.B. [...] der Vater die Ursache des Kindes, und überhaupt das Hervorbringende die Ursache des Hervorgebrachten, das Verändernde die Ursache des Veränderten“ (S. 90/1013a). Neben weiteren, für den hiesigen Zusammenhang weniger relevanten Begriffen (*causa formalis*, *causa finalis*) hat Aristoteles hiervon die *causa materialis* (Materialursache) unterschied-

²⁶⁰ Was nicht heißt, dass ihn die Naturwissenschaften auch immer als adäquate Beschreibung der Wirklichkeit akzeptieren. Bekanntermaßen negiert die Kopenhagener Deutung der Quantenmechanik eine streng deterministische Kausalitätsvorstellung. Nichtsdestotrotz ist auch hier der Gegenstand der Kausalitätskepsis der Begriff der *causa efficiens*.

den. Hiermit ist gemeint, dass die zu *einem ganz bestimmten Zeitpunkt*²⁶¹ vorliegende (makroskopische) Eigenschaft eines Gegenstandes ihren Ursprung in dessen materieller Struktur hat. „Ursache wird [...] der immanente Stoff genannt, woraus etwas wird, so ist das Erz der Bildsäule, das Silber der Schale Ursache“ (S. 89f/1013a). Demnach ist die Ursache der Zerbrechlichkeit von Glas darin begründet, dass Glas eine ganz bestimmte materielle Beschaffenheit hat. Auch wenn die Rede von „Verursachung“ an dieser Stelle nicht prinzipiell falsch erscheint, ist es eben keine solche im Sinne einer Wirkursache und damit keine im heutigen Verständnis von Kausalität, für das die Zweiteilung in ein Verursachendes und Verursachtes (als voneinander unterscheidbare Größen) in einem zeitlichen Verlauf charakteristisch ist, die im Fall der *causa materialis* nicht vorliegt.

Dennoch ist der Begriff der Materialursache für Mechanismen nicht ganz zutreffend, da er eher so etwas wie eine *asymmetrische Bottom-Up*-Determination beschreibt, während in Bezug auf Mechanismen oder überhaupt mereologische Beziehungen vielmehr ein *symmetrisches* Verhältnis im Sinne einer *wechselseitigen* „Determination“ zwischen dem Ganzen und seinen Teilen besteht (vgl. Hüttemann, 2004; hierzu an späterer Stelle mehr; vgl. Kap. 8.6.2). Durch eine Vermengung dieser beiden Konzepte wird nun für eine hochproblematische *kausale Überdeterminierung* eines Ereignisses Tür und Tor geöffnet, die bereits im Zusammenhang der Besprechung der Supervenienz-Relation thematisiert wurde und eine Gefahr für die Wirkmächtigkeit höherstufiger Eigenschaften darstellt (vgl. Kap. 6.3.2).

²⁶¹ Natürlich bleibt die Materialursache auch für gewöhnlich über einen längeren Zeitraum bestehen. Eine solche Stabilität in der Zeit ist aber etwas anderes wie die *Veränderung* in der Zeit, die die *causa efficiens* zum Gegenstand hat.

Dass unter Berücksichtigung der zeitlichen Dimension die Möglichkeit einer mentalen Verursachung bzw. die Realität psychophysischer Kausalität gewahrt bleibt, soll nun an dem bereits in Kapitel 3.1 herangezogenen Beispiel erläutert werden (vgl. Abb. 31).

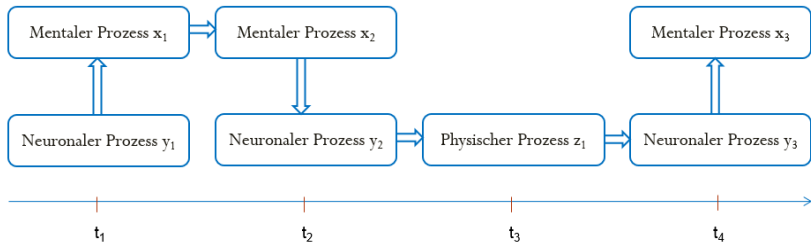


Abb. 31: Psychophysische Kausalität im Sinne des Interaktionismus

Zur Erinnerung hier nochmals die intuitive Beschreibung der psychophysischen Wechselwirkung beim „Cola-Beispiel“: Wenn zu t_1 osmosensible Neurone im vorderen Hypothalamus (y_1) „registrieren“, dass eine erhöhte Salzkonzentration im Blut vorhanden ist, spüre ich Durst und habe den Wunsch, etwas zu trinken (x_1), sodass ich beschließe, mir eine Cola aus dem Kühlschrank zu holen (x_2 zu t_2). Im Zuge meiner Bewegungsinitiation kommt es zu einem neuronalen Gesamtzustand y_2 , der u.a. die hierfür charakteristischen Aktivierungen im motorischen Kortex beinhalten wird. Zum Zeitpunkt t_3 bewege ich mich zum Kühlschrank (physischer Prozess z_1). Dann trinke ich die Cola, was im weiteren Verlauf (t_3 zu t_4) eine Veränderung des neuronalen Gesamtzustandes (y_3) bewirkt. So kommt es durch die Flüssigkeitszufuhr zu einer Absenkung der Salzkonzentration im Blut, wodurch sich die Aktivität der osmosensiblen Zellen im Hypothalamus vermindert. Infolgedessen nimmt mein Durst ab (mentaler Prozess x_3).

Wie nun klar geworden ist, handelt es sich hierbei nicht um distinkte Ereignisreihen, die in der beschriebenen Art und Weise miteinander wechselwirken, sondern um das Herausgreifen von zwei Ebenen einer mereologisch-mechanistischen Hierarchie, wobei die hier jeweils angesprochene mentale Eigenschaft bzw. Eigenschaftsinstantiierung eine mentale Struktur hat, die mit der Struktur der agierenden Entitäten aller niedrigeren Ebenen identisch ist und so also auch mit der hier angesprochenen neuronalen Ebene.

Betrachtet man nun einen ganz bestimmten Zeitpunkt wie z.B. t_2 , dann kann man daher tatsächlich nicht sagen, dass mein Gedanke die zeitgleich vorhandenen agierenden Entitäten der mit ihm assoziierten agierenden Entitäten des Mechanismus *verursacht* hat. Wenn ich eine Bewegung initiiere und hierbei eine Aktivität im motorischen Kortex zu verzeichnen ist, dann ist es einfach falsch, wenn man sagt, dass die Bewegungsinitiierung die neuronale Aktivität im motorischen Kortex hervorgerufen hat, genauso wie die gegenteilige Annahme falsch wäre, dass die neuronale Aktivität die mentale Bewegungsinitiierung bewirkt hat. Dies wäre nur unter den Voraussetzungen eben jenes intuitiv einsichtigen, faktisch aber hochproblematischen psychophysischen Dualismus möglich (vgl. Kap. 3.1.1). In Anbetracht einer (mereologisch-mechanistisch fundierten) psychophysischen Identitätsannahme *ist* die Bewegungsinitiation zum Zeitpunkt t_2 zugleich die Aktivität im motorischen Kortex und auch die Aktivität der agierenden Entitäten aller niederstufigen Ebenen des hier installierten Mechanismus. Wenn nun aber zu jedem beliebigen Zeitpunkt ein kausales Einwirken „des Geistes“ auf die „physische Welt“ *nicht* möglich ist, weil Geist und Physis *de facto* Ebenen einer mit sich selbst identischen mechanistischen Struktur sind und eine Selbstverursachung begrifflicher Nonsens wäre, mag prima facie tatsächlich der

Schluss naheliegen, dass das Psychische ganz im Sinne des Epiphänomenalismus kausal wirkungslos ist.

Betrachtet man aber einen *zeitlichen Verlauf*, dann zeichnet sich ein ganz anderes Bild ab. Natürlich besteht über die verschiedenen Zeitpunkte der beschriebenen Ereigniskette hinweg ein *kausales* Geschehen. Dies ist allerdings keineswegs ein Intra-Level-Phänomen, sondern – wie sich auch im Hinblick auf das von Dehaene erläuterte Beispiel der Psychotherapie gezeigt hat – ein „*ebenendurchdringendes*“ Phänomen, also ein Phänomen, bei dem das kausale Geschehen auf allen Ebenen des jeweiligen (Gesamt-)Mechanismus abläuft. Auch wenn also für einen ganz konkreten Zeitpunkt nicht angenommen werden kann, dass das Geschehen auf einer Ebene für das Geschehen auf der anderen Ebene kausal (im Sinne der *causa efficiens*) verantwortlich ist, ist es aber natürlich vollkommen korrekt, wenn man ableitet, dass z.B. das Geschehen in der mechanistischen Hierarchie zu t_1 das Geschehen in der herausgegriffenen mechanistischen Hierarchie zu t_2 verursacht hat. Allerdings lässt sich dieses *eine* kausale Geschehen auf unterschiedlichen Ebenen betrachten. Natürlich war es also *mein* Durst bzw. *mein* Wunsch nach einer Cola und *meine* Entscheidung, mir eine solche aus dem Kühlschrank zu holen, der mich dazu gebracht hat, ebendieses zu tun. Genauso war es das „Feuern“ der osmosensiblen Zellen im Hypothalamus, das die Aktivität im motorischen Kortex ausgelöst hat, was wiederum eine koordinierte Aktivität der Muskeln in den Armen und Beinen initiiert hat usw. Es bestehen hier also nicht zwei unabhängig voneinander ablaufende Kausalreihen oder -netze. Der psychische Kausalnexus *ist* der physische Kausalnexus in einer anderen Perspektive bzw. auf einer anderen natürlichen Organisationsebene einer mit sich selbst identischen mechanistischen Hierarchie (vgl. Lüdmann, 2009, S. 124). Insofern ist meine körperliche Bewegung zum Kühlschrank genauso durch meinen Wunsch nach etwas zu trinken und mei-

nen Beschluss, mir eine Cola zu holen, verursacht, wie sie auch durch die Aktivität im Hypothalamus und motorischen Kortex verursacht ist. Es zeigt sich also klar, dass die hier vorgestellte Identitätskonzeption gegenüber den Fallgruben höherstufiger Kausalität, den Problemen des kausalen Ausschlusses sowie des Abdriftens der Kausalität zu den Basisebenen, vollkommen immun ist.

Liegt aber nicht zumindest eine gewisse Willkür vor, welcher Organisationsebene man die Verursachung zuschreibt oder sind wirklich alle gleichermaßen maßgeblich bei der Hervorbringung eines Ereignisses? Ist also die Anerkennung einer mentalen Verursachung physischer Ereignisse gewissermaßen von einer „wohlwollenden“ Interpretation abhängig und keineswegs zwingend? Das ist nicht der Fall. Denn bei genauerer Betrachtung *muss* der psychischen Ebene bei einer mereologisch-mechanistischen Hierarchie, die bis hin zum Psychischen aufgespannt ist, eine gewisse Maßgeblichkeit für das Auftreten des nachfolgenden Ereignisses zugeordnet werden²⁶². Denn aufgrund der systemischen Emergenz, also aufgrund der Tatsache, dass das Ganze Eigenschaften aufweist, die seine Teile entbehren, folgt, dass ein solches ebenendurchdringendes Kausalgeschehen wie meine Bewegungsinitiation *kraft* der instantiierten mentalen Eigenschaft zustande gekommen ist. Auch wenn es zwar keine Ebene gibt, auf der *allein* sich der Kausalprozess ereignet, ist die Ebene des Mentalen bei einem Vorgang, bei dem ein bestimmtes Ziel verfolgt wird, insofern maßgeblich, als dass erst hier die eigentlichen Gründe für diese

²⁶² Dies bedeutet nicht, dass die psychische Ebene, sobald sie in einer solchen Hierarchie als Ebene aufgespannt wird, *immer* diejenige ist, die in der hier beschriebenen Form maßgeblich ist. Natürlich können in anderen Fällen auch die mikroskopischen Ebenen für die nachfolgenden Prozesse oder Systemzustände maßgeblich sein, so wie es in dem herangezogenen Beispiel die Flüssigkeitszufuhr zu t_3 ist, die darauf zu t_4 einen mentalen Zustand auslöst, bei dem das Durstgefühl bzw. der Wunsch nach etwas zu trinken nachlässt.

Handlung gegeben sind. Würden die mentalen Eigenschaften nicht vorliegen, würde die gesamte mechanistische Hierarchie aufgrund der mereologischen Supervenienzrelation zwischen den Ebenen nicht die Struktur aufweisen können, die sie bei Vorhandensein der mentalen Eigenschaft hat. Denn das Psychische *ist* – um es etwas plakativ auszudrücken – gemäß den Ausführungen des letzten Unterkapitels *geformte Materie* (die auf der psychischen Ebene aber auch eine ganz bestimmte kognitive resp. phänomenale Struktur hat). Die philosophische Intuition, dass sich unter den Voraussetzungen psychophysischer Identitätsannahmen alles physische Geschehen in der Welt genauso vollziehen würde, auch wenn es keinen Geist in ebendieser gäbe, ist daher schlichtweg falsch bzw. selbstwidersprüchlich. Wenn Materie auf eine solche Art und Weise „geformt“ ist, dass eine mechanistische Hierarchie agierender Entitäten bis hin zur Ebene neuronaler Netze innerhalb eines Gehirns aufgespannt wird, dann ist hiermit *zwangsläufig* das Vorhandensein geistiger Eigenschaften verknüpft, denn das Psychische *ist* letztendlich eben genau diese auf eine bestimmte Art und Weise konfigurierte Materie.

Wenn die durch die Identitätstheorie nahegelegte Sicht der Dinge dennoch der konventionellen Art und Weise der Interpretation solcher Sachverhalte entgegen zu stehen scheint, weil man für gewöhnlich das zu einem Zeitpunkt gegebene mentale Geschehen als durch das zu diesem Zeitpunkt vorhandene neuronale Geschehen verursacht ansieht, dann mag ein Grund hierfür im intuitiven Dualismus als psychologischer Alltagskonzeption liegen, der vor allem auf die psychischen Strukturmerkmale der Perspektivität und phänomenalen Transparenz zurückgeht (vgl. Kap. 3.4). Hiermit verbunden ist die kognitive Schwierigkeit, zwei solch stark phänomenal differente Organisations- bzw. Betrachtungsebenen in ihrer Gleichzeitigkeit als Einheit aufzufassen. Man kann zunächst nur eine (bestimmte physische) Ebene zentrieren und mitunter

eine Veränderung entdecken. Darauf wechselt man mit seiner Aufmerksamkeit zur anderen Ebene und kann das Auftreten eines bestimmten mentalen Ereignisses feststellen. Es liegt somit nahe, das eine als Ursache und das andere als Folge zu interpretieren.

Auch eine Illustration anhand evolutionstheoretischer Entwicklungsprozesse ist hier hilfreich. Wenn sich im Laufe der Evolution einzellige Organismen zu Mehrzellern zusammengeschlossen haben, haben sich diese „im Kampf ums Dasein“ bewährt, weil sie bestimmte Fähigkeiten hervorbrachten, die Einzeller nicht aufweisen. Mehr- und Vielzeller unterscheiden sich dabei von Einzellern oder auch von Kolonien einzelliger Lebewesen durch einen gemeinsamen Stoffwechsel des Zellverbands oder auch eine mehr oder weniger ausgeprägte Aufgaben- bzw. Arbeitsteilung zwischen den Zellen (vgl. Campbell & Reece, 2011, S. 692ff). Hierdurch haben sie u.a. Vorteile bei der Nahrungssuche oder beim Schutz gegenüber Feinden.

Betrachtet man einen solchen Mehrzeller nun unter ontologischen Gesichtspunkten, dann hat dieser als ganzer Organismus (höherstufige) Eigenschaften, die Einzeller bzw. seine einzelnen Zellen nicht aufweisen. Eigenschaften als strukturelle Kennzeichnungen eines Systems, die durch eine höherstufige Systemebene realisiert werden, sind daher nicht auf der Ebene der Teile dieses Systems vorhanden (z.B. funktionale Spezialisierung in Keim- und Körperzellen). Obgleich auch hier gilt, dass die Zellen, aus denen ein Mehrzeller besteht, und der Mehrzeller als Ganzes in einem zeitlich-räumlichen Koextensivitätsverhältnis stehen und daher eine mit sich selbst identische mereologisch-mechanistische Hierarchie begründen und ein jeder Kausalvorgang auch hier im Sinne einer ebenendurchdringenden Kausalität auf allen beteiligten Ebenen (Zellen und Organismus) zugleich abläuft, ist es die Ebene des Ganzen, *kraft* der der mehrzellige Organismus die Eigenschaften besitzt, die ihn

von einem Einzeller unterscheiden. Erst durch das Zusammenspiel der verschiedenen Zellen wird u.a. die erhöhte Wehrhaftigkeit gegenüber natürlichen Feinden hervorgebracht.

In ebendieser Weise spricht viel dafür, dass es in vielen Fällen die höchste in einer mechanistischen Hierarchie realisierte Ebene ist, die für den ebenendurchdringenden Kausalitätsvollzug ausschlaggebend ist. Wenn ein Mensch dazu fähig ist, zu sprechen, sich räumlich zu orientieren, sich in seine Artgenossen emotional hineinzusetzen, Pläne für die Zukunft und zur Lösung komplexer Probleme zu schmieden, eine feine Ironie in einer Aussage zu erkennen und sich hieran zu erfreuen, dann kann er dies nur *kraft* der psychischen Fähigkeiten bzw. der Eigenschaften, die hierfür notwendig sind, auch wenn sich die kausale Aktivität bei diesen mentalen Leistungen auf allen Ebenen der hierzu installierten Mechanismen vollzieht. Die mereologisch-mechanistischen Hierarchien dieser Vorgänge sind nach Maßgabe der realisierten mentalen Eigenschaften gestaltet. Ohne die mentalen Eigenschaften gäbe es auch die jeweiligen „genuin physischen“ Ereignisketten nicht in dieser Form.

Zur Erinnerung: Kernproblem der funktionalistisch konzipierten Realisierungstheorie ist, dass eine Funktionalisierung phänomenaler Eigenschaften in Form von alltagssprachlich beschreibbarem Verhalten kaum zu verwirklichen ist (vgl. Kap. 6.5.4.1). So scheint es insbesondere im Hinblick auf Farbempfindungen gegeben zu sein, dass Personen mit unterschiedlichen (oder auch fehlenden) Farbeindrücken dasselbe kausale (Handlungs-)Profil aufzeigen werden, sodass das Programm der funktionalen Reduktion keine Handhabe hat, phänomenale Differenzen solcher Art abzubilden und hieran insgesamt zu scheitern droht. Phänomenale mentale Eigenschaften lassen sich offenbar nicht über ihre funktionale (kausale) Rolle individuieren.

Dagegen ist nach Maßgabe der hier entwickelten psychophysischen Identitätstheorie jede Veränderung der Farbeindrücke einer Person zwangsläufig mit einer Veränderung ihrer mikrostrukturellen Verfasstheit verbunden. Wenn also in meinem in Abbildung 29 dargestellten Wahrnehmungseindruck die Farbe einer Blüte durch eine andere Farbe ausgetauscht wird, ist hiermit zwangsläufig eine Veränderung *aller* (mikro-)strukturellen Ebenen des Mechanismus verbunden, der das Farbsehen hervorbringt. Die agierenden Entitäten der verschiedenen Organisations-ebenen sind qua der zu realisierenden Funktion eine Einheit oder miteinander identisch, sodass es widersprüchlich wäre anzunehmen, dass eine Veränderung des phänomenalen Eindrucks ohne eine Veränderung seiner (auf diversen Ebenen potentiell herausgreifbaren) Mikrostruktur erfolgen kann. Ob hiermit zugleich eine beobachtbare Verhaltensänderung verbunden ist, bleibt irrelevant, da die hier vorgeschlagene Konzeption nicht an die Analyse des kausalen (Verhaltens-)Profils mentaler Eigenschaften gebunden ist.

Es sei angemerkt, dass dies aber natürlich in der gegenteiligen Richtung möglich ist: Es kann also durchaus mikrostrukturelle Veränderungen geben, die sich mental bzw. phänomenal nicht niederschlagen. Nicht jede agierende Entität einer Mikroebene ist ein kausal relevanter Teil einer (gerade hervorgebrachten) mentalen Eigenschaft. Nur mit dieser einschränkenden Zusatzannahme ließe sich also das Prinzip der mereologischen Supervenienz (trivialerweise) auch in gegenteiliger Richtung formulieren: Jede Veränderung einer kausal relevanten Komponente oder mikrostrukturellen Verfasstheit eines Mechanismus zur Hervorbringung einer mentalen Eigenschaft ist notwendigerweise mit einer Veränderung ebendieser Eigenschaft verbunden. Vorausgesetzt wir können die Gültigkeit des Mechanismus zur Hervorbringung von farblichen Erlebnisqualitäten induktiv verallgemeinern, gilt damit auch, dass eine Person, die über ein

verändertes Farbspektrum oder fehlende „Farb-Qualia“ verfügt, nicht über dieselbe mikrostrukturelle Konfiguration verfügen kann, wie eine Person, die dies tut, obgleich sich diese beiden Personen im Hinblick auf ihr kausales (Verhaltens-)Profil nicht unterscheiden mögen.

7.4.3. Ein Abriss zur Handlungs- und Entscheidungsfreiheit

Seit Benjamin Libets berühmten Versuchen wird die Frage nach der sogenannten *Willensfreiheit* in der Philosophie und Psychologie aber auch gesamtgesellschaftlich hitzig diskutiert. Dabei hat sich schnell gezeigt, dass die von Libet untersuchte *Spontaneitätswillensfreiheit* nur einen bestimmten Sonderfall von Freiheit aufgreift, der für die Beurteilung der Freiheit eines Menschen eher irrelevant ist (vgl. Laucken, 2005). Diese Art von Freiheit beinhaltet lediglich – was sich direkt aus der Originalversuchsanordnung von Libet (1985) ergibt – den Umstand, dass die Versuchspersonen „freely and capriciously“ (S. 530) ihr Handgelenk beugen sollten, wobei Libet ein Bereitschaftspotential (eine spezifische elektrische Veränderung im Gehirn) jeweils 550 Millisekunden vor solch einer „Handlung“ bzw. 350 – 400 Millisekunden vor der bewussten Entscheidung zur Handgelenkbeugung feststellen konnte. Dies wurde von Libet zunächst als Bestätigung der Nicht-Existenz menschlicher Willensfreiheit interpretiert (und eben auch von einer Vielzahl von Gelehrten, wie die Neurowissenschaftler Gerhard Roth (2004) oder Wolf Singer (2004), die hieraus u.a. weitreichende Konsequenzen für die Schuldfähigkeit von Personen zogen), bis er selbst – in fundamental-dualistischer Manier – die Möglichkeit eines geistigen „Vetos“ gegenüber der vom Gehirn getroffenen Entscheidung postuliert hat und bemüht war, ein solches nachzuweisen.

Wie vor allem Peter Bieri (2003) äußerst klar dargelegt hat, ist diese Art von „Freiheit“ aber irrelevant für unser Selbstverständnis, frei im Entscheiden und Handeln zu sein, und sogar in weiten Teilen selbstwidersprüchlich. Denn – sehr grob zusammengefasst – sollten wir uns als freiheitsliebende Menschen einen von unserer eigenen Physis und personalen Vorgeschichte losgelösten Willen im Sinne eines „Ersten Bewegers“, also eines Bewegers, der selbst durch gar nichts determiniert ist, keinesfalls wünschen. Ganz im Gegensatz zu der verbreiteten Intuition, dass Freiheit auf Zufälligkeit basiert und Unfreiheit auf Determiniertheit, kann gerade bei der *Zufälligkeit* des Zustandekommens meiner Entscheidungen und Handlungen (wie es bei einem solchen „Ersten Beweger“ der Fall wäre) nicht mehr die Rede von Freiheit sein. Denn wenn ich selbst durch nichts bei meinen Entscheidungen und Handlungen festgelegt bin, können sich diese auch nicht aus der Vorgeschichte meiner Person, meiner Persönlichkeit, meinen Erfahrungen oder meinen emotionalen und kognitiven Einstellungen ergeben. Statt einer *selbstbestimmten* Handlung wäre mein Agieren nur noch ein zufälliges, für mich selbst unerklärliches *Geschehen*. Meine Willensentscheidungen und Handlungen würden mir also lediglich „zustoßen“ oder „widerfahren“, statt sich als Konsequenz aus meinen Überlegungen und Prägungen zu ergeben. Es erschiene unter diesen Voraussetzungen genau genommen sogar unangebracht zu sagen, dass es sich hier um *meine* Entscheidungen und Handlungen handelt, denn der hierfür notwendige Bezug zu *meiner* Person ist offenkundig gar nicht gegeben. Letztendlich ist die einzig relevante und zugleich begrifflich konsistente Art von Freiheit daher die einer *Handlungs- und Entscheidungsfreiheit* im Sinne der Selbstbestimmung oder *Selbstdetermination* (vgl. auch Pauen, 2004).

Vor dem Hintergrund der vorliegenden Arbeit ließe sich nun allerdings problematisieren, dass die bewussten Abwägungen, die zur Ausbildung

eines Willens und wiederum einer Handlung führen, durch *unbewusste* Prozesse *vorbestimmt* sein könnten. Wenn Freiheit die Möglichkeit zur Selbstdetermination sein soll, dann scheint eine Handlung im Falle solcher unbewussten Prozesse gerade nicht *selbst*-determiniert zu sein. Diese Tatsache führte daher auch Gerhard Roth (2004) zu der häufig zitierten Schlussfolgerung: „Nicht mein bewusster Willensakt, sondern mein Gehirn hat entschieden!“ (S. 73), wonach freie Entscheidungen und Handlungen insgesamt hinfällig sind, da sie alle (unbewusste) neuronale Vorläuferprozesse (wie das von Libet ermittelte Bereitschaftspotential) aufweisen würden.

Bei eingehenderer Betrachtung wird das Verständnis von Freiheit im Sinne einer Entscheidungs- und Handlungsfreiheit aber auch durch (unbewusste) neuronale Vorläuferprozesse von freien Entscheidungen oder Handlungen nicht ernsthaft in Frage gestellt. Natürlich könnte zwar die Gewissheit, dass jede meiner Erwägungen durch zuvor ablaufende neuronale Prozesse festgelegt ist, auf eine Revision meiner Selbstzuschreibung von Freiheit hinauslaufen, denn ich mag mich dann zwar frei fühlen, aber die diesem „Gefühl“ zugrunde liegende Selbstzuschreibung basiert womöglich nur auf der (vielleicht auch nur vorübergehenden) Unkenntnis der Existenz solcher Vorläuferprozesse. Diese vordergründig schlüssige Argumentation erweist sich nur auf den ersten Blick als Gefahr für das hier verfochtene Freiheitsverständnis. Wie herausgestellt wurde, muss sich eine als frei verstehbare Handlung aus unserer Vorgeschichte, aus unseren Überlegungen, Wahrnehmungen etc. ergeben und zu uns, zu unserer Persönlichkeit, zu unseren Emotionen, Präferenzen und Gewohnheiten passen, so dass die von uns vollzogene Handlung mitsamt ihren Beweggründen von uns selbst potentiell nachvollziehbar und bejahbar ist. Und dies alles wird durch die Hinzuziehung der neuronalen Ebene und durch die hier angesiedelten Vorläuferprozesse gar nicht in

Frage gestellt. Denn was sind unsere Persönlichkeit, unsere emotionale Verfasstheit und unsere Präferenzen überhaupt? Mit Verweis auf die hier entwickelte identitätstheoretische Antwort auf das Leib-Seele-Problem kann geschlussfolgert werden, dass es sich hierbei um Mechanismen handelt, die (auch) eine bestimmte neuronale Struktur aufweisen, die allerdings nicht – zumindest nicht durchgehend – mit einer konkreten phänomenalen Repräsentation verknüpft ist (also „unbewusst“ wirkt). Daher wird es sich bei den unbewussten, einer freien Entscheidung oder Handlung eines Menschen vorausgehenden neuronalen Prozessen wie dem Bereitschaftspotential um Facetten *seiner* Persönlichkeit, *seiner* Fähig- und Fertigkeiten sowie *seiner* emotionalen Verfasstheit handeln. Dies alles sind Dinge, von denen man durchaus wollen sollte, dass sie entscheidungsrelevant wirken. Denn letzten Endes bin ich (auch) mein Gehirn (vgl. Cruse, 2004). Die Frage, ob ich *oder* mein Gehirn entschieden hat, beinhaltet einen absurden Dualismus. Denn auch für das psychologische Konstrukt eines „Ichs“ als einer Entscheidungs- und Handlungsinstanz gilt, dass es sich hierbei um etwas handelt, das eine bestimmte mentale Struktur aufweist, die mit einer (potentiell unendlichen) Abfolge von mechanistischen Ebenen in einer mereologisch verfassten Hierarchie identisch ist und somit auch mit einer bestimmten neuronalen Konfiguration. Es sollte aber nunmehr evident sein, dass hiermit der Realität einer geistigen Person und ihren Entscheidungs- und Handlungsprozess keinerlei Abbruch getan ist, sondern vielmehr die Realität ebendieser psychischen Größen untermauert wird.

Des Weiteren ist hinlänglich bekannt, dass im Rahmen der kognitiven Entlastung Automatisierungsprozesse unabkömmlich sind. Ein Beispiel hierfür liefert eine interessante Studie von Spelke, Hirst und Neisser (1976). Probanden, die zwei Aufgaben (Lesen und Schreiben) gleichzeitig ausführen sollten, konnten nach sechs Wochen Übung mit normaler

Geschwindigkeit einen Text lesen und nebenher trotzdem die vom Versuchsleiter diktierten Wörter notieren. Dabei war für die Teilnehmer das Lesen während des Niederschreibens der Wörter nicht schwieriger als wenn sie während des Gehens lesen mussten. Höchstinteressant ist nun, dass sie berichteten, dass sie sich dessen, was sie schrieben, nicht bewusst wären. Vergleichbar mit den automatisierten Prozessen beim Auto- oder Radfahren nahmen die Probanden die automatisierte Aktivität zwar nicht bewusst wahr, sie konnte aber im Zuge nachfolgender Experimente durchaus bewusst gemacht werden.

Dinge, die nicht in meinem gegenwärtigen Bewusstsein sind, aber im Einklang mit mir als Person stehen, können also durchaus freie Entscheidungen sein! Wenn man bestimmte Sportarten wie Squash spielt, bei denen man sehr schnell agieren muss, „hinkt“ das Bewusstsein mitunter auch den eigenen Handlungen hinterher – man ist deshalb aber keineswegs unfrei in seinem Handeln. Ganz im Gegenteil kann das Freiheitsgefühl in einem solchen „Flow-Erlebnis“ besonders groß sein. Es lässt sich also festhalten, dass es nicht entscheidend ist, dass man sich zu jeder seiner inneren Regungen oder äußeren Bewegungen bewusst entschieden hat. Entscheidend ist, dass man potentiell sein Denken und Tun bejahen würde, wenn man sich dieses bewusst vergegenwärtigen würde. So sind unbewusstes Schreiben, unbewusstes Autofahren oder auch „nachbewusstes“ Squash spielen ohne Probleme als freie Handlungen zu verstehen.

7.5. Zur Möglichkeit typentheoretischer psychophysischer Identitäten

Die bisher ausgearbeitete Theorie einer psychophysischen Identität basiert auf der Koextensivität der strukturellen Konfigurationen aus agierenden Entitäten der verschiedenen Ebenen eines konkreten Mechanismus. Statt einer bloßen psychoneuralen Identität ist es eine komplexe Architektur des Psychischen, die sich im Zuge dessen enthüllt hat. Es sind daher nur rein *lokale* psychophysische Identitätsbeziehungen impliziert, die die verschiedenen mereologisch-mechanistischen Ebenen agierender Entitäten in einer kompositorisch individuierten Hierarchie kennzeichnen. Der ebenenübergreifende Mechanismus zeigt dabei auf, wie genau es ein *bestimmter* Organismus *in einer konkreten Situation* schafft, eine (mentale) Leistung zur Lösung eines bestimmten Problems zu erbringen. Es ist die Maus, die gerade einen Weg aus dem Labyrinth sucht, in das sie von den Wissenschaftlern gesetzt wurde oder der Mensch, der einen Weg aus einem Irrgarten sucht.

Vom *ontologischen* Standpunkt betrachtet, ist eine Konzeption rein lokaler Identitäten durchaus befriedigend. Die hiermit verbundene Möglichkeit multipler Realisierungen führt keineswegs zu einer Schwächung der Aussagekraft einer solchen psychophysischen Identitätstheorie. Wenn wir die noch zu besprechenden Differenzen zwischen dem Identitäts- und Reduktionsbegriff einmal beiseitelassen, hat diesen Aspekt bereits Kim (2005) auf den Punkt gebracht:

For most antireductionist philosophers, multiple realizability has long been a mantra, an all-purpose antireductionist argument applied across the board to all special science properties. They see multiple realization everywhere, and this leads them to see irreducibility everywhere. I believe, however, that the notion of 'realization' as it is often invoked in this context is too loose and ill-formed, and that when realization

is properly understood, *multiple realization only leads to reducibility to multiple reduction bases, not to irreducibility.* (S. 55f; Hervorh.: M.L.)

Für eine inhaltlich konsistente und aussagekräftige Theorie des Geistes ist es also gar nicht erforderlich, eine Identität auf dem Niveau von Typen oder Eigenschaften (anstatt Eigenschaftsinstantiierungen) zu formulieren. Die Möglichkeit multipler Realisierungen mentaler Eigenschaften führt keineswegs dazu, dass mentale Eigenschaften Gefahr laufen, (substanz-)dualistisch oder epiphänomenalistisch aufgefasst werden zu können (was – wie dargelegt – mit vielen Problemen verbunden wäre). Obgleich es also in der einen Situation diesen und einer weiteren Situation einen anderen Mechanismus zur Realisierung von räumlicher Orientierung (oder auch Schmerz) geben mag, führt dies keineswegs dazu, dass diese mentalen Eigenschaften von ihren physischen Grundlagen „ablösbar“ oder als „nicht-physisch“ zu klassifizieren wären. Wenn es die Möglichkeit multipler Realisierungen mentaler Eigenschaften gibt, zeigt sie im Fall kognitiver Leistungen zunächst den Spielraum an Möglichkeiten auf, ein bestimmtes Problem zu bewältigen. So gibt es natürlich eine Vielzahl an verschiedenen Strategien im Zusammenhang mit räumlicher Orientierung, die einen Weg aus einem Labyrinth weisen mögen. Jede dieser verschiedenen Plan- und Verhaltensstrategien muss dabei zwangsläufig mit einem anderen oder veränderten Mechanismus verbunden sein, der daher auch auf der neuronalen Ebene ein anderes „Profil“ aufweisen muss. Dennoch weist jeder einzelne dieser verschiedenen Mechanismen zur Hervorbringung einer räumlichen Orientierung dabei eine Identität der agierenden Entitäten seiner verschiedenen (mikro-)strukturellen Ebenen auf.

Es könnte jedoch eingewendet werden, dass eine reine Einzelfallbetrachtung *wissenschaftlich* unbefriedigend ist, da der wissenschaftliche An-

spruch grundsätzlich mit der Erfassung oder Erklärung eines möglichst *allgemeingültigen* Zusammenhangs verbunden ist. Es kann also erwartet werden, dass zumindest eine Auskunft darüber abgegeben wird, welche Beziehung zwischen dieser lokalen zu einer *globalen* Theorie des Geistes besteht, auch wenn die Forderung nach einer solchen globalen Theorie in bestimmter, noch genauer zu spezifizierender Hinsicht, fehlgeleitet sein mag. Es muss also die Möglichkeit von Aussagen zu psychologischen Typen oder Eigenschaften (und nicht nur deren Instantiierungen) geklärt werden. Inwieweit besteht über eine Identität lokaler Ereignisse bzw. Eigenschaftsinstantiierungen hinaus auch eine Möglichkeit zur Erweiterung dieses Ansatzes auf dem Niveau mentaler Typen bzw. Eigenschaften, ohne dabei dem Argument der multiplen Realisierbarkeit anheim zu fallen?

Zunächst muss man eingestehen, dass das Argument der multiplen Realisierbarkeit, wie es von Putnam (1967) oder Block und Fodor (1972), Fodor (1974) oder auch Lycan (1987/2007) vorgetragen wird, eine hohe augenscheinliche Durchschlagskraft hat, die eine Formulierung einer solchen Theorie auszuschließen scheint. Denn wenn z.B. behauptet wird, dass die mentale Eigenschaft „Schmerzen zu haben“ mit der neuronalen Eigenschaft „C-Faser-Reizungen zu haben“ identisch sein soll, dann scheint man hiermit zwangsläufig zu behaupten, dass jeder nur denkbare Organismus, der über Schmerzen verfügt, die gleichen neuronalen Basisprozesse, also C-Faser-Reizungen, aufweisen muss, was empirisch höchst unplausibel wäre. Denn warum sollte jedes schmerzempfindende Wesen im gesamten Universum unbedingt C-Faser-Reizungen aufweisen müssen, um ebensolche Bewusstseinsindrücke zu haben? Nicht weniger schlagkräftig erscheint die Intuition, dass wir im Zuge des weiteren Fortschritts in den Neurowissenschaften in die Lage kommen könnten, eine Bewusstseinswelt zu erschaffen, die nicht mehr auf Kohlenstoff basiert,

sondern auf Silizium (oder anderen stofflich-elektronischen Grundlagen), sodass wir bei einem unverändert bleibenden mentalen Geschehen sukzessive unsere einzelnen Neuronen Stück für Stück durch künstliche (elektronische) Prothesen ersetzen könnten. Aber vielleicht muss man gar nicht so weit gehen. Denn ist nicht allein schon die neuronale Plastizität, die wir beim Menschen (z.B. nach Schlaganfällen) vorfinden, bei der bestimmte mentale Leistungen aufgrund einer Läsion von anderen Hirnbereichen als den für gewöhnlich hiermit betrauten übernommen werden, ein Beweis darauf, dass die Formulierung einer psychophysischen Identitätstheorie auf Typen-Niveau von vornherein zum Scheitern verurteilt ist?

Zugleich darf man nicht vergessen, dass – wenn eine psychophysische Identität auf Typen-Niveau postuliert wird, die beansprucht, eine wesentliche Kennzeichnung eines mentalen Prozesses bzw. einer mentalen Eigenschaft zu sein – diese gemäß den Ausführungen Kripkes (1971, 1972) einen *notwendigen* Charakter haben muss. Wenn also gelten soll, dass Schmerzen C-Faser-Reizungen sind, dann muss diese Beziehung *notwendigerweise* wahr sein. Die scheinbare Kontingenz dieser Aussage muss daher ausgeräumt werden, was sich im Angesicht des von Kripke vorgeschlagenen epistemischen Modells des Nachweises einer Erscheinungs-Realitäts-Diskrepanz als sehr schwierig herausgestellt hat (vgl. Kap. 6.1.4.3).

Es sei angemerkt, dass der Terminus „multiple Realisierbarkeit“ im Angesicht der hier entwickelten Identitätskonzeption nicht ganz treffend ist. Denn es ist genau genommen nicht die Vorstellung einer abstrakten geistigen Eigenschaft, die durch verschiedenste physische Realisierer in die Welt gehoben werden kann (und die in diesem Sinne identisch mit ebendiesen Realisierern ist), sondern eine Identität der verschiedenen (mikro-)strukturellen Konfigurationen der agierenden Entitäten eines *kompo-*

sitorisch individuierten Mechanismus. In diesem Sinne ist der zu klärende Aspekt vielmehr, wie die hier entwickelte Konzeption mit der Möglichkeit einer *kompositorischen Variabilität* umgeht, das heißt mit der Möglichkeit, dass eine bestimmte höherstufige (mentale) Eigenschaft durch verschiedene Mechanismen hervorgebracht werden könnte, also solche, die sich in ihrer konkreten Zusammensetzung aus agierenden Entitäten der verschiedenen Organisationsebenen unterscheiden²⁶³.

Im Folgenden werden zwei Ansätze besprochen: Einerseits besteht die Möglichkeit, per induktiver Verallgemeinerung von Einzelbeobachtungen bzw. idealtypischer Mechanismen eine heuristische Identität auf Typen-Niveau abzuleiten. Es kann gezeigt werden, dass sich ein solches Vorgehen auch als sinnvolle Forschungspraxis bewährt hat. Mit einer bloßen heuristischen Identitätsannahme ist kein Anspruch auf Notwendigkeit der unterstellten Identitätsaussagen verbunden. Man kann andererseits aber auch einen Schritt weiter gehen und ausloten, inwieweit das Argument der multiplen Realisierbarkeit überhaupt greift, wenn man nicht nur von einer Forschungsheuristik ausgeht, sondern zugesteht, dass es für z.B. (menschliche) Schmerzen *wesentlich* ist, auf der neuronalen Ebene aus C-Faser-Reizungen zu bestehen. Ein metaphysischer Abweg ist dabei allerdings, mittels einer Identitätsaussage die *exklusive Essenz* von Schmerzen bestimmen zu wollen. Vielmehr ist für die Möglichkeit einer typentheoretischen Identitätskonzeption entscheidend, wie genau man eine (natürliche) Art oder Eigenschaft bestimmt bzw. welchen Allgemeinheitsgrad und ontologischen Status man solchen Universalien beimisst.

²⁶³ Aufgrund seiner Gebräuchlichkeit wird der Begriff der multiplen Realisierbarkeit im Folgenden weiter verwendet und zwar synonym zur kompositorischen Variabilität.

7.5.1. Heuristische Identitäten als Forschungsstrategie

Natürlich kann eine Beobachtung oder ein zunächst für den Einzelfall konzipiertes Modell *induktiv verallgemeinert* werden, ein Vorgehen, wie es für die Naturwissenschaften charakteristisch ist. Wenn hier im Einzelfall oder in einer bestimmten Anzahl von Einzelfällen die gleichen mechanistischen Prinzipien herausgestellt werden können, spricht viel dafür, dass es sich um einen Mechanismus handelt, der *grundsätzlich* die Erfüllung dieser bestimmten Leistung wie der räumlichen Orientierung (oder Schmerzen, Farbempfindungen...) gewährleistet. Man könnte daher ableiten, dass der beobachtete Zusammenhang also nicht nur für diese eine Instantiierung einer Eigenschaft gilt, sondern der allgemeingültige Mechanismus zur Hervorbringung ebendieser Eigenschaft ist (beim Menschen oder vielleicht allen Säugetieren, was wiederum induktiv zu erforschen wäre). Eine bestimmte Instantiierung einer solchen Eigenschaft kann in diesem Sinne zu einer Art *Idealmodell* oder *Prototyp* erhoben werden. Wenn man die Identität der Strukturebenen eines Mechanismus akzeptiert, wäre dies ein induktiver Schluss auf die Identität von Eigenschaften selbst, was einer (idealisierten) Identitätsannahme auf Typen-Niveau gleichkommt. Aufgrund der lediglich induktiven – keine Notwendigkeit implizierenden – Verallgemeinerung von Einzelbeobachtungen eines Idealmodells bestünde erstens genügend Spielraum für eine kompositorische Variabilität, die dennoch die grundsätzliche Gültigkeit des postulierten Mechanismus unangetastet ließe. Zweitens greift das Kripkesche Argument der starren Designatoren an dieser Stelle nicht, denn bei lediglich heuristischen Identitäten muss es sich nicht um *notwendige* Identitäten *a posteriori* handeln.

In ebendieser Weise kann die Annahme von psychophysischen Identitäten auf Typen-Niveau zunächst als eine sinnvolle Forschungsstrategie angesehen werden. So werden für gewöhnlich ganz spezifisch geistige

Funktionen bestimmten Hirnbereichen zugeordnet, obgleich es (aufgrund der neuronalen Plastizität oder aufgrund einer „Funktionsverlagerung“ infolge bestimmter Hirnläsionen etc.) Fälle geben mag, bei denen die typischen Aktivierungsmuster bei der Ausführung einer bestimmten geistigen Leistung nicht vorliegen. Bekannte Fälle solcher Zuordnungen sind bereits die Entdeckung der „Sprachareale“ durch Broca oder Wernicke. So zeigten sich bei Patienten, die unter spezifischen Ausfällen sprachbezogener Kompetenzen litten (Defizite in der Sprachproduktion vs. Defizite beim Sprachverständnis), Läsionen in ganz bestimmten Hirnzentren, die heute nach ihren Entdeckern als Broca- bzw. Wernicke-Areal benannt sind. Durch die nunmehr vorhandene Möglichkeit, neuronale Aktivierungsmuster mittels bildgebender Verfahren darzustellen, kann gezeigt werden, dass es nach dem aktuellen Stand der Wissenschaft tatsächlich diese Areale sind, die bei gesunden Probanden bei der Sprachproduktion vs. der Sprachanalyse bzw. dem Sprachverständnis aktiv sind. Wenn man nun davon ausgeht, dass diese sprachbezogenen Fähigkeiten (beim Menschen) für gewöhnlich mit genau diesen neuronalen Regionen assoziiert sind, dann kann man dies als eine *heuristische* Identitätsannahme auf Typen-Niveau bezeichnen. Wie Bechtel und McCauley (vgl. Bechtel & McCauley, 1999; McCauley & Bechtel, 2001; McCauley, 2007, 2012; Bechtel, 2008) aufzeigen, sind solche Annahmen in der neurowissenschaftlichen Forschung sehr verbreitet und haben sich – wie im Fall der „Sprachzentren“ – als Forschungsstrategien bewährt: „Hypothesizing cross-scientific identities is a pivotal engine of scientific development. Hypothetical identities in interlevel contexts serve as valuable heuristics of discovery for inquiry at both of the explanatory levels involved“ (McCauley & Bechtel, 2001, S. 751).

Die Autoren verdeutlichen dies des Weiteren am Beispiel der Erforschung der (neuronalen) Mechanismen des Sehens (vgl. ebd., S. 745ff).

Bereits im 19. Jahrhundert wurde der Okzipitallappen als Ort des Sehens identifiziert, da die Kartierung neuroanatomischer Pfade und auch Läsionsstudien an Tieren und Menschen diesen Ort nahelegten. Ein Meilenstein in der Erforschung der neuronalen (okzipitalen) Mechanismen des Sehens war die Arbeit von Hubel und Wiesel (1962). Aufgrund der Anwendung verhaltensbezogener und neurophysiologischer Methoden entwickelten sie wegweisende Hypothesen bezüglich der beteiligten Hirnregionen als auch der implementierten Informationsverarbeitungsmechanismen. Sie konzentrierten sich hierbei zunächst auf die Region des – heute so bezeichneten – *primären* visuellen Kortex (V1), kamen aber allmählich zu dem Schluss, dass diese Region nur einen kleinen Teil des gesamten (kortikalen) Mechanismus des Sehens ausmachen kann. Infolgedessen kam es zu einer weiteren Forschungswelle, die zur *Identifizierung* der „topografischen Karten“ der Areale V2 und V3 führte (vgl. McCauley & Bechtel, 2001, S. 748). Schließlich konnte Zeki in einer Arbeit aus dem Jahr 1969 systematische Projektionen in zwei weitere Regionen feststellen, die heute als V4 und V5 (bzw. MT) bezeichnet werden (vgl. ebd.). Durch die Erforschung mittels Einzelzelleitungen bei der Darbietung verschiedener Stimuli ermittelte Zeki des Weiteren, dass es die Applikation unterschiedlicher Wellenlängen des Lichts („Farben“) ist, die zu einer Aktivität in V4 führt, während die Zellen in V5 selektiv auf Bewegungen reagieren. Diese Regionen wurden daher in der Forschungspraxis mit den genannten Funktionen *identifiziert*. Auf der Basis dieser funktionalen Kartierung kam es wiederum zu einer Ausdifferenzierung weiterer („kortikal detektierbarer“) Teilaufgaben des Sehens, wobei nunmehr mindestens 33 verschiedene Areale und Pfade identifiziert werden konnten, die an der Verarbeitung visueller Informationen beteiligt sind (van Essen & Gallant, 1994; vgl. McCauley & Bechtel, S. 748). Die wohl prominenteste dieser späteren Entdeckungen war die Annahme

zweier sich an V1 anschließender unterschiedlicher Pfade der Verarbeitung visueller Informationen durch Ungerleider und Mishkin (1982). Demnach gibt es einen *dorsalen* Pfad (Wo-Pfad), der mit der Lage und Bewegung von Objekten im Raum verbunden ist und von V2 über V5/MT in den posterioren parietalen Kortex verläuft; und einen *ventralen* Pfad (Was-Pfad), der über V4 in die inferotemporale Region verläuft und mit der Identifikation dessen, was ein Objekt ist, beschäftigt ist. Trotz einiger (berechtigter) Kritik an der Unabhängigkeit dieser beiden Pfade gilt diese Hypothese heute sowohl in der Neurowissenschaft als auch in der Psychologie als anerkannter Fakt (vgl. McCauley & Bechtel, S. 748).

Führt man sich diesen Fortgang der Forschung vor Augen, sieht man, dass die ursprüngliche Identifizierung von V1 als visuelles Sehzentrum keineswegs ein Hindernis für die spätere Forschung war oder dass – wie häufig behauptet wird – Identitätsannahmen nicht mehr als psychophysische Korrelationen behaupten (vgl. z.B. Chalmers, 1996, S. 115). Vielmehr ist die Forschung gerade durch solche Funktionszuordnungen im Sinne heuristischer Identitäten vorangetrieben worden: „However, rather than undermining the strategy of hypothesizing identities, Hubel and Wiesel’s determining the functions of cells in V1 led to more hypothetical identities that were even more precise about the visual and the neural processes identified and about the additional brain areas involved“ (McCauley & Bechtel, 2001, S. 479). Interessant ist des Weiteren, dass die Befunde zu den verschiedenen am Sehen beteiligten Hirnregionen eine Schlüsselrolle bei der Formulierung neuer Hypothesen geführt haben und das sowohl auf der neurophysiologischen als auch psychologischen Ebene der visuellen Informationsverarbeitung bzw. des menschlichen Kognitionssystems schlechthin (vgl. ebd.). Ein Beispiel hierfür wäre die Wahrnehmungs- und Kognitionspsychologie von Ulric Neisser (1994), die auf der Unterscheidung einer Was- und Wo-Verarbeitung aufbaut.

Ein weiteres Beispiel einer heuristischen psychophysischen Identität ist die Wahrnehmung von Gesichtern, von der angenommen wird, dass sie von einer hierauf spezialisierten Gehirnregion, dem im okzipitotemporalen Bereich gelegenen Gyrus fusiformis (*fusiform face area*; FFA), bewerkstelligt wird und auch hier eine hypothetische Identität zwischen der Gesichtswahrnehmung und der FFA-Aktivität postuliert wird (Kanwisher, McDermott & Chun; 1997; McCauley, 2012).

So zeigt sich nicht nur, dass die Annahme psychophysischer Identitäten auf Typen-Niveau eine sinnvolle Forschungsstrategie ausmacht, sondern auch, dass neurowissenschaftliche Entdeckungen keineswegs in Konkurrenz zu psychologischen Theorien stehen müssen oder psychophysische Identitätsannahmen keine Bedrohung für die Psychologie und ihre theoretischen Modelle und Konzepte sein müssen, sondern dass ganz im Gegenteil eine sehr fruchtbare Wechselwirkung zwischen der Psychologie und Biologie erfolgen kann.

Allerdings bleibt der alleinige Rekurs auf eine heuristische Identität auf Typen-Niveau in gewisser Hinsicht unbefriedigend bzw. es besteht wiederum die Gefahr einer einseitigen Einschätzung des Identitätskonzeptes. Denn wenn man wie Bechtel (2012) in solchen wissenschaftlichen *Identifizierungen* mentaler Prozesse im Gehirn bereits den vollen Gehalt einer psychophysischen Identitätstheorie ausfindig gemacht zu haben glaubt, muss es nicht verwundern, dass man (wie er) zu der Einschätzung gelangt, dass Identitätsaussagen keine ebenenübergreifenden Beziehungen herausstellen können (vgl. ebd., S. 46). Vielmehr bestünde die Existenz mentaler Eigenschaften darin, *nichts anderes als* eine bestimmte neuronale Konfiguration bzw. ein neurophysiologischer Prozess zu sein. Der alleinige oder zumindest primäre Rekurs auf die bestehende Forschungspraxis (in den kognitiven Neurowissenschaften) bei der Reflexion über die

Reichweite von Identitätsbeziehungen suggeriert ausschließlich eine *Identifizierung* des Mentalen in seinen physischen Grundlagen:

When an fMRI study relates a cognitive operation with a brain region (e.g., recognizing faces with activity in the fusiform face area [...]), when an evoked response potentials study *identifies* a cognitive operation with a change in electrical potentials [...], the intended claim is that the neural activity *is* the performance of the cognitive operation: increased activity in the fusiform gyrus *is* the recognition of faces, etc. (Bechtel, 2012, S. 46)

Mit dieser Feststellung läuft Bechtel daher offensichtlich Gefahr, die problematische Identitätskonzeption der „klassischen“ Identitätstheorie bzw. der physischen Realisierungstheorie zu übernehmen und letztendlich bei einer Aussage zu landen wie: „Schmerzen sind *nichts anderes als* C-Faser-Reizungen“. Er verspielt damit die weitergehenden Möglichkeiten, die der mechanistische Ansatz im Zusammenspiel mit adäquaten Identitätsaussagen zur Klärung des Leib-Seele-Problems – über den Verweis auf heuristische Identitäten als Forschungspraxis hinaus – beinhaltet.

7.5.2. Psychophysische Typen-Identitäten jenseits von Heuristiken?

Im Folgenden soll geklärt werden, inwieweit es möglich und sinnvoll ist, über bloße heuristische, forschungsstrategische Annahmen hinaus eine psychophysische Identität auf Typen- bzw. Eigenschaftsniveau postulieren zu wollen, obgleich eine kompositorische Variabilität der Mechanismen zur Hervorbringung mentaler Eigenschaften als Faktum betrachtet werden muss. Oder anders ausgedrückt: Kann es für Schmerzen *wesentlich* sein, (auf der neuronalen Ebene) mit C-Faser-Reizungen identisch zu sein, obwohl es empirisch gesehen nicht abwegig erscheint, dass

auch andere neuronale oder sogar künstliche Mechanismen Schmerzen hervorbringen könnten?

7.5.2.1. Das exklusive Wesen des Mentalen – ein metaphysischer Abweg

Bieri (1981/1997) sagt, das zentrale Problem einer jeden Identitätstheorie sei, dass „die Auskunft, daß ein bestimmtes mentales Phänomen, beispielsweise ein Schmerz, hier das ist und anderswo das ist, unbefriedigend [sei], denn wir wollen gerade wissen, was *Schmerz* ist: was *allen* Fällen von Schmerzen gemeinsam ist. Wir wollen etwas über das Universale, die Eigenschaft des Schmerz-Habens erfahren“ (S. 41; Hervorh. im Orig.). Das Dilemma einer jeden Identitätstheorie sei daher: „Entweder ist sie eine Theorie über mentale Universalien [wie die „klassische“ Identitätstheorie; M.L.]; dann scheint sie empirisch unplausibel. Oder sie ist empirisch plausibel [wie die heuristische Identitätstheorie; M.L.]; dann sagt sie uns nicht das, was wir wissen wollen“ (ebd.). Diese Einschätzung dürfte in der Tat eine verbreitete Überzeugung (in der Philosophie) wiedergeben. Einerseits sei es die multiple Realisierbarkeit mentaler Eigenschaften, die eine Identitätstheorie auf Typen-Niveau empirisch unhaltbar erscheinen lässt, andererseits wird der Anschein erweckt, dass eine Identitätsaussage auf Token-Niveau bzw. lokale Identitätsannahmen mehr oder weniger belanglos wären.

Es mag sein, dass sich Verteidiger der „klassischen“ Identitätstheorie von dieser Kritik angesprochen fühlen (müssen), wenn sie in der Annahme einer psychophysischen Identität tatsächlich so etwas wie eine *exklusive Wesensbestimmung* des Psychischen sehen. Wenn also gelten soll: „Schmerzen sind C-Faser-Reizungen“ und hiermit die *strikte* Identität à la Smart gemeint ist, sodass gilt: „Schmerzen sind *nichts anderes als* C-Faser-Reizungen“ und Schmerzen in diesem Sinne *als* C-Faser-Reizun-

gen *identifiziert* werden, dann muss die sicherlich nicht un plausible Möglichkeit, dass Schmerzen (in anderen Systemen oder aufgrund von Läsionen etc. in ein und demselben System) durch eine abweichende Mikrostruktur gekennzeichnet sein könnten (z.B. „D-Fasern“), unweigerlich zu der Einschätzung führen, dass *diese* Identitätsannahme falsch ist. Und da die „klassische“ Identitätstheorie in genau ebendieser Weise verstanden wurde, erstaunt ihr (zunächst) schnelles Ableben aus der philosophischen Debatte nicht.

Im Angesicht der hier entwickelten Identitätskonzeption, die explizit auf mereologisch-mechanistischen Prinzipien fußt, entpuppt sich der Versuch, das Wesen von Schmerzen oder anderer psychischer Prozesse und Eigenschaften *exklusiv* in der Mikrostruktur *einer ganz bestimmten* Ebene ausmachen zu wollen, als ein unsinniges Unterfangen. Ein bestimmtes (menschliches) Schmerzerlebnis mag auf der neuronalen Ebene typischerweise die Mikrostruktur „C-Faser-Reizungen“ aufweisen. Gleichzeitig weist ebendieses Geschehen eine potentiell unendliche Abfolge mikrostruktureller Konfigurationen auf den verschiedenen Ebenen des zur Hervorbringung dieses Schmerzerlebnisses vorhandenen Mechanismus auf. Die Identitätsaussage gilt daher für *alle* Ebenen dieser mereologisch-mechanistischen Hierarchie gleichermaßen. Das Haben von Schmerzen besteht demnach keineswegs einfach darin, C-Faser-Reizungen aufzuweisen, sondern genauso in dem Haben *aller* mikrostrukturellen Konfigurationen der mereologisch-mechanistischen Hierarchie und – das sollte keineswegs vergessen werden – vor allem natürlich auch darin, eine ganz bestimmte *mentale* Struktur (bzw. hier: phänomenale Erlebnisqualität) aufzuweisen. Die Möglichkeit, dass Schmerzen im Zusammenhang einer anderen mereologisch-mechanistischen Hierarchie auf der neuronalen Ebene die Mikrostruktur „D-Faser-Reizungen“ aufweisen mögen, führt weder dazu, dass die Identitätskonzeption empirisch unplausibel

wird, noch dazu, dass die hier entwickelte Konzeption ihr Ziel verfehlt (denn dieses besteht keineswegs darin, die Frage nach der Essenz des Mentalen zu beantworten).

Für denjenigen, der also von der Identitätstheorie tatsächlich eine Antwort auf die Frage nach dem Wesen des Mentalen erhalten möchte, mag der alleinige Rekurs auf die (mikro-)strukturelle Verfasstheit eines mentalen Ereignisses unbefriedigend sein. Aber es ist gerade dieses Verlangen nach einer Bestimmung der Essenz des Mentalen, das in die Irre führt. Denn hier liegt ein – im engeren bzw. „dunklen“ Sinne des Wortes – *metaphysisches* Interesse vor und nicht der Wunsch nach einer wissenschaftlichen Erklärung. Statt einer Annäherung an eine Lösung des Leib-Seele-Problems landet man hierdurch unweigerlich in einer begrifflich-konzeptuellen Sackgasse und unentscheidbare, wenig weiterführende Begriffsstreitigkeiten sind vorprogrammiert. Denn jeder metaphysische Ansatz in der Philosophie des Geistes behauptet auf diese Weise – gewissermaßen recht willkürlich – das Wesen des Geistes exklusiv bestimmen zu können. Nach der („klassischen“) Identitätstheorie bestünde es darin, ein Vorgang in unserem Gehirn zu sein, nach dem Funktionalismus darin, dass es ein auf verschiedene Weise zu realisierender funktionaler Zustand ist, oder nach dem Eliminativismus darin, dass es letztendlich etwas gar nicht Existierendes ist.

Die hier entwickelte psychophysische Identitätstheorie beansprucht dagegen keineswegs, ein für alle Mal bzw. in exklusiver Weise zu klären, was genau Schmerz ist. Die metaphysische (und wahrscheinlich unbeantwortbare) Frage nach dem *exklusiven Wesen* von Schmerz gehört nicht in eine Metatheorie der Wissenschaften bzw. nicht zum Gegenstandsgebiet einer Theoretischen Psychologie. Wie aufgezeigt wurde, besteht zwischen identitätstheoretischen (strukturbezogenen) und funktionalistischen

(formbezogenen) Ansätzen auch keineswegs eine strenge Trennung, wie sie die meisten philosophischen Lehrbuchdarstellungen²⁶⁴ nahelegen. Häufig wird in diesem Zusammenhang von Verteidigern identitätstheoretischer Konzepte die Annahme vertreten, dass sich die Mikrostruktur eines Systems verändern und dabei nichtsdestotrotz ein mentaler Vorgang *desselben Typs* hervorgebracht werden kann. So seien Schmerzen zwar identisch mit C-Faser-Reizungen (neuronale Ebene), ihre Mikrostruktur auf der zellulären, molekularen oder auch (sub-)atomaren Ebene kann aber variieren (vgl. z.B. Pauen, 1999; Shapiro & Polger, 2012). Der Hintergrund dieser Strategie ist, dass hier – gemäß dem Prinzip der exklusiven Wesensbestimmung des Mentalen als etwas Neuronales – grundsätzlich nur die Identität von Schmerzen mit den (agierenden) Entitäten oder Eigenschaften der neuronalen Ebene angenommen wird, denn laut der „klassischen“ Identitätstheorie sind Schmerzen *nichts anderes als* C-Faser-Reizungen. Die (agierende) materielle Entität oder Struktur der C-Faser-Reizungen kann aber nun – dieser Denkweise zufolge – durch verschiedene mikrostrukturelle Beschaffenheiten konstituiert sein, denn die Kriterien dafür, ein solcher „neuronaler Typ“ zu sein, lassen eine gewisse mikrostrukturelle Variabilität zu. Hiermit würde ermöglicht, eine Typen-Identität vertreten zu können und zugleich einen gewissen Spielraum für multiple Realisierungen zu lassen.

Im Angesicht der in dieser Arbeit getätigten Feststellungen muss diese Strategie kritisch beleuchtet werden. Wie aufgezeigt wurde, handelt es sich bei den Teil-Ganzes-Beziehungen, die der mechanistische Ansatz zur Erklärung psychischer Leistungen heranzieht, zunächst nur um Ein-

²⁶⁴ Diese orientieren sich für gewöhnlich daran, welche Antworten die verschiedenen Theorien des Geistes auf die Frage nach dem Wesen des Mentalen geben, wodurch tiefe Gräben zwischen diesen gezogen werden, die der Sache nach gar nicht vorliegen.

zelfallbetrachtungen, die (im Sinne des Ansatzes heuristischer Identitäten) induktiv verallgemeinert werden könnten. Dabei gilt für jede Instantiierung einer mentalen Eigenschaft wie Schmerzen, dass sie nicht nur mit den agierenden Entitäten der neuronalen Ebene (sprich: einer bestimmten Instantiierung von C-Faser-Reizungen) koextensiv-identisch ist, sondern notwendigerweise mit allen agierenden Entitäten der verschiedenen in einer Teil-Ganzes-Beziehung stehenden Ebenen, die dieses Ereignis hervorbringen. Insofern weist jede Instantiierung von Schmerzen eine potentiell unendliche Abfolge von miteinander identischen Mikrostrukturen auf.

Aufgrund der Ähnlichkeit, die verschiedene Instantiierungen von Schmerzen miteinander aufweisen, erscheint es allerdings in der Tat möglich und auch sinnvoll zu sein, sie zu einem bestimmten allgemeinen *Typ* (des Habens) von Schmerzen zu gruppieren. Da zwischen den Mitgliedern einer solchen Kategorie keine vollständige Gleichheit, sondern nur eine mehr oder weniger stark ausgeprägte *Ähnlichkeit* erwartbar ist, liegt es nahe, dass bei den einzelnen konkreten Fällen, die solch einem Typ zugeordnet werden können, die mikrostrukturelle Verfasstheit variiert. Es darf aber nicht vergessen werden, dass sich jede Veränderung von agierenden Entitäten in einer mechanistischen Hierarchie gemäß dem Prinzip der mereologischen Supervenienz *ebenenübergreifend* niederschlägt (vgl. Kap. 6.3.1). Jede Veränderung der phänomenalen Qualität eines Schmerzzustandes ist notwendigerweise mit einer Veränderung auf der neuronalen, mikrozellulären, molekularen, atomaren, etc. Ebene verbunden. Ebenso muss auch jede Veränderung einer kausal relevanten mikrostrukturellen Komponente eines Mechanismus zu einer Veränderung der Beschaffenheit der höherstufigen agierenden Entitäten bzw. der instantiierten Eigenschaft führen. Insofern wird z.B. eine Veränderung der molekularen Struktur einer Instantiierung von Schmerzen dazu füh-

ren, dass sich die „phänomenale Struktur“ des Schmerzens-Habens bzw. die strukturelle Verfasstheit der neuronalen Ebene verändert. Trotzdem mag ebendieses modifizierte Ereignis, wenn die Modifikation geringfügig genug ausfällt, immer noch eine Instantiierung einer Eigenschaft sein, die man als das Haben von Schmerzen bzw. als das Haben von C-Faser-Reizungen klassifizieren kann. Denn die einzelnen konkreten Exemplare, die eine Art konstituieren, sind keineswegs völlig gleichförmig, sondern weisen nur ein ausreichendes Maß an Ähnlichkeit auf²⁶⁵. Soweit scheint die erwähnte Argumentation also zunächst zu greifen. Dennoch läuft eine solche Identitätsannahme bzw. Verteidigungsstrategie, die sich ausschließlich auf das Verhältnis zwischen neuronaler und mentaler Ebene bzw. mentalen und neuronalen Typen/ Eigenschaften bezieht, Gefahr, begrifflich widersprüchlich zu sein oder eine in Bezug auf das Argument der multiplen Realisierbarkeit irrelevante Feststellung zu treffen. Denn wie gezeigt wurde, liegt bei einer Variation kausal relevanter mikrostruktureller Komponenten eines Mechanismus notwendigerweise zugleich eine Veränderung der höherstufigen Entitäten der mereologischen Hierarchie des Mechanismus und damit der realisierten Zieleigenschaft vor. Mit einer solchen Variation der Mikrostruktur des „Schmerz-Mechanismus“ unterliegen also auch die neuronale Struktur der C-Faser-Reizungen sowie die phänomenale Struktur des Erlebens von Schmerzen zwangsläufig einer Veränderung. Die suggerierte Konstanz höherstufiger Typen trotz der Variation ihrer Mikrostruktur hat also – wenn denn überhaupt – nur eine sehr begrenzte Gültigkeit. Es ist ausgeschlossen, dass sich etwas auf der neuronalen oder mentalen Ebene

²⁶⁵ Die Grenzen einer (natürlichen) Art und dabei vor allem auch mentaler Eigenschaften sind nur selten scharf umrissen, sondern weisen einen hohen Grad der Abhängigkeit von Vorannahmen, situationsbezogenen praktischen Erwägungen etc. auf.

ähnlich ist, sich aber auf der molekularen oder atomaren Ebene stark unterscheidet. Phänomene, die sich als C-Faser-Reizung (oder auch als das Haben von Schmerzen) klassifizieren lassen, müssen also auch auf den niedrigeren mereologisch-mechanistischen Ebenen ein hinreichend hohes Maß an Ähnlichkeit aufweisen. Man könnte das Argument zwar zu modifizieren versuchen, indem man annimmt, dass es sich lediglich um eine Variation von solchen mikrostrukturellen Teilen des Systems handelt, die für die Beschaffenheit von C-Faser-Reizungen respektive Schmerzen irrelevant sind. Hierdurch bliebe ihre neuronale bzw. mentale Beschaffenheit unangetastet. Einerseits muss dann aber hinterfragt werden, ob ein solcher Teil wirklich als ein *Bestandteil* des „Schmerz-Mechanismus“ angesehen werden kann. Er mag zwar ein mikrostruktureller Fakt bzw. eine mikrostrukturelle Entität eines *Organismus* sein, der Schmerzen aufweist, aber offensichtlich ist er kein Bestandteil eines Mechanismus zur Hervorbringung einer bestimmten mentalen Leistung im Sinne einer (kausal!) agierenden Entität. Damit wäre ein solcher Fall von kompositorischer Variabilität weitestgehend irrelevant. Keinesfalls werden die verschiedenen inhaltlichen Facetten des Argumentes der multiplen Realisierbarkeit hiermit ausgeräumt.

7.5.2.2. Artspezifität und Granularitätsparität

Ungeachtet dessen dürfte sich etwas ganz Entscheidendes gezeigt haben, nämlich, dass die Schlagkraft des Argumentes der multiplen Realisierbarkeit entscheidend davon abhängt, welchen *Allgemeinheitsgrad* oder *ontologischen Status* man Universalien wie Typen, Eigenschaften oder (natürlichen) Arten beimisst. Nimmt man eine Aussage wie „Schmerzen sind C-Faser-Reizungen“ wortwörtlich, ist sie empirisch in der Tat äußerst unplausibel, denn sie ist auf einem viel zu hohen Allgemeinheitsniveau

formuliert. Die wissenschaftliche Entdeckung, dass das Auftreten von Schmerzen beim Menschen (oder vielleicht auch bei bestimmten Tierarten) mit C-Faser-Reizungen korreliert ist, lässt eigentlich nur die induktive Verallgemeinerung zu, dass (zumindest) *menschliche* Schmerzen (auf der neurobiologischen Strukturebene) C-Faser-Reizungen sind²⁶⁶. Entscheidend ist also, dass in dem Argument der multiplen Realisierbarkeit eine Verallgemeinerung enthalten ist, die in dieser Form überhaupt nicht von einer psychophysischen Identitätstheorie impliziert sein muss. Nur wenn man von einer Identitätstheorie erwartet, dass sie ein für alle Mal die metaphysische Frage zu beantworten habe, was *Schmerzen überhaupt* sind, welches Charakteristikum der Eigenschaft, Schmerzen zu haben, insgesamt zu Eigen ist, kommt das Argument der multiplen Realisierbarkeit wirklich voll zum Tragen. Wie aufgezeigt wurde, ist der Versuch einer exklusiven Wesensbestimmung des Mentalen aber abwegig. Die Aussage, dass (menschliche) Schmerzen C-Faser-Reizungen sind, sollte nicht als exklusive Wesensbestimmung aufgefasst werden, sondern vielmehr als Angabe der bzw. – genauer gesagt – *einer* Mikrostruktur von Schmerzen, ohne dass hiermit beansprucht wird, das Wesen von Schmerzen „in seinem tiefsten Inneren“ (was auch immer das heißt) charakterisiert zu haben.

Der Anspruch, dass Schmerzen eines jeden möglicherweise existierenden Außerirdischen, wie Putnam (1967) sagt, nach Maßgabe identitätstheoretischer Vorstellungen dieselbe neurophysiologische Basis haben müssten wie beim Menschen, ist daher irreführend. Denn ist es wirklich *derselbe* Typ Schmerz, den ein möglicherweise existierender Außerirdischer fühlen würde wie ein Mensch? Gibt es die hier extrem verallgemei-

²⁶⁶ Bereits Paul Churchland (1979, 1982) hat auf die Möglichkeit der Formulierung empirisch plausibler speziesspezifischer Identitätsannahmen hingewiesen.

nerte Art oder Eigenschaft „Schmerz“ bzw. „das Haben von Schmerzen“ überhaupt als solche? Auch wenn sich Schmerzen für einen solchen „Marsmenschen“ sicherlich ähnlich anfühlen könnten wie menschliche Schmerzen, wäre es ein Zufall gigantischen Ausmaßes, wenn sie *ein und dieselbe* phänomenale Qualität aufweisen würden. Es besteht jedenfalls keinerlei Notwendigkeit hierfür und damit auch keine dafür, dass ein solches Wesen mit derselben Neurophysiologie ausgestattet sein müsste wie wir, um Schmerzen zu empfinden. „Marsmenschen-Schmerz“ und „Erdenmenschen-Schmerz“ mögen beide gemeinsam haben, unangenehme Empfindungen zu sein, dennoch müssen sie keineswegs *ein und dieselben* phänomenalen Qualitäten bzw. die hiermit verbundenen kausalen Rollen aufweisen (vgl. Lewis, 1980). Dasselbe gilt natürlich auch für Schmerzen bei Tieren. Natürlich mögen diese ähnliche Schmerzempfindungen haben wie wir. Trotzdem ist es völlig abwegig anzunehmen, dass *jedes* Tier, dem wir geneigt sind, die Möglichkeit zur Empfindung von Schmerzen zuzuschreiben, nach Maßgabe einer psychophysischen Identitätstheorie auf Typen-Niveau C-Faser-Reizungen aufweisen muss, nur weil dies für menschliche Schmerzen gilt. Weinbergschnecken-Schmerz muss keineswegs ein und dasselbe sein wie Menschen-Schmerz (wenn es erstgenannten denn gibt). In Abbildung 32 wird eine solche speziespezifische Identität (rechte Seite) im Vergleich zu der gewöhnlichen, zu stark verallgemeinernden Identitätsannahme (linke Seite) aufgezeigt. Gleichzeitig ist es allerdings interessant, sich vor Augen zu führen, dass das Vorhandensein eines Nervensystems bzw. die Ähnlichkeit eines solchen zu dem des Menschen ein zentrales Kriterium für die Zuschreibung mentaler Zustände bei Tieren in der Wissenschaft ist²⁶⁷ (siehe z.B.

²⁶⁷ Natürlich kann ein solches Kriterium trotzdem falsch sein. Man mag außerdem einwenden, dass das Verhalten eines Organismus die entscheidende Rolle bei der Zuschreibung

Searle, 1992/1996). So wird für gewöhnlich die phänomenale Qualität tierischer Schmerzempfindungen der unseren umso ähnlicher beurteilt, je ähnlicher ihre (neuro-)physiologische bzw. (neuro-)anatomische Beschaffenheit der unseren ist.

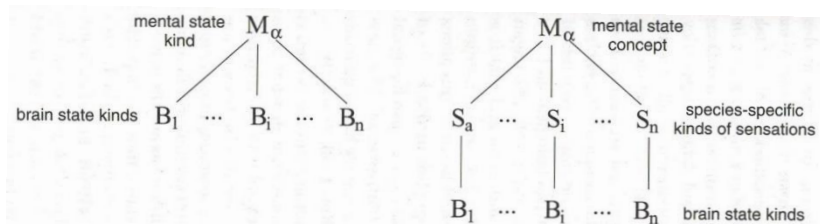


Abb. 32: Verallgemeinertes (links) versus speziesspezifisches Modell multipler Realisierungen (rechts) (Aus: Polger, 2004, S. 27)

Aber was ist mit der multiplen Realisierbarkeit bzw. kompositorischen Variabilität, die wiederum im Falle *menschlicher* mentaler Prozesse oder Eigenschaften festzustellen ist? Solche Fälle multipler Realisierbarkeit scheinen sehr verbreitet zu sein, sodass „even within a species brains differ. Even within an individual over time there are differences (neurons die, connections are lost, etc.). Thus multiple realizability seems to arise within species (including our own) and even within individuals“ (Bechtel & Mundale, 1999, S. 177). Obwohl menschliche Gehirne in ihrer ganz konkreten Beschaffenheit recht stark variieren können, bringen sie offenbar *dieselben* mentalen Eigenschaften hervor. Selbst Menschen mit star-

mentaler Zustände spielt. Das Verhalten (im Sinne vermeintlicher zielgerichteter Bewegungen) ist aber als alleiniger Indikator äußerst fraglich. Natürlich *kann* man das Verhalten eines Regenwurms als zielgerichtet interpretieren (wie auch das „Verhalten“ der geometrischen Formen in Heiders Experimenten), aber kann man deshalb schon annehmen, dass ein Regenwurm über mentale Zustände bzw. Bewusstsein verfügt? Ein Vergleich der Nervensysteme von Mensch und Regenwurm dürfte eine verlässlichere Entscheidungsbasis bereitstellen.

ken Läsionen des Gehirns (infolge von Verletzungen oder Schlaganfällen) zeigen häufig ein mehr oder weniger normales Verhalten und können infolge der neuronalen Plastizität des Gehirns einmal verloren gegangene mentale Leistungen häufig wieder erlernen (vgl. z.B. Knecht & Ringelstein, 1999), sodass hier offenbar andere Hirnregionen *dieselben* mentalen Leistungen oder phänomenalen Eigenschaften erzeugen wie im „Normalfall“. Ein anderes Phänomen wären Phantom-Schmerzen, wie sie z.B. häufig bei Patienten mit einer abgetrennten Hand oder anderen abgetrennten Gliedmaßen auftreten. Nach dem Stand der aktuellen Forschung kommt es hierbei zu einer neuronalen Neukartierung, bei der sich die Nervenverbindungen, die für gewöhnlich die Felder im Umkreis der kortikalen Handrepräsentation besetzen (vor allem repräsentative Felder für Gesicht und Oberkörper), in das zuvor von der Hand besetzte Feld ausdehnen. Die Patienten haben daher bei Berührung dieser Regionen im Gesicht oder z.B. am Oberarm präzise lokalisierbare Empfindungen in Bezug auf ihre „Phantomhand“ bzw. ihre „Phantomfinger“ (vgl. Kandel & Kupfermann, 1996, S. 340f). Bezogen auf das diskutierte Beispiel der menschlichen Schmerzen im Allgemeinen besteht also das Problem, dass auch bei menschlichen Schmerzen starke interindividuelle oder sogar intraindividuelle Variationen bezüglich ihrer physischen Grundlagen auftreten können, sodass eine typentheoretische Identitätsannahme zwischen menschlichen Schmerzen und C-Faser-Reizungen verfehlt erscheint. Bei dieser Art von Variation dürfte es sich zugleich um den eigentlich ernst zu nehmenden Kern des Argumentes der multiplen Realisierbarkeit handeln.

Auch hier muss aber zunächst festgestellt werden, dass es *den* Schmerz (oder *den* Hunger) genau genommen gar nicht gibt (vgl. Bechtel & Mundale, 1999; Bechtel, 2007). Zahnschmerzen fühlen sich anders an als Kopfschmerzen oder Rückenschmerzen. Auch wiederum meine Rücken-

schmerzen fühlen sich heute vielleicht ganz anders an als noch gestern. Die Schlagkraft des Argumentes der multiplen Realisierbarkeit hängt daher stark von der Bestimmung dessen ab, was ein mentaler Typ überhaupt sein soll. Gleichzeitig muss hinterfragt werden, ob es einen klar abgegrenzten mentalen Typ für z.B. Schmerzen bei eingehender Betrachtung überhaupt gibt, wie er im alltäglichen oder auch wissenschaftlichen Sprachgebrauch in dieser Form verwendet wird. Auch wenn die Identitätstheorie also behauptet, dass mentale Zustände mit physischen Zuständen identisch sind, bedeutet dies nicht, dass alle mentalen Zustände einer Art selbst miteinander identisch sein müssen. Was sie zu einer Art macht, ist, dass sie in der Eigenschaft oder in den Eigenschaften übereinstimmen, die charakteristisch für diese Art sind, so wie alle Schmerzzustände die Eigenschaft aufweisen, sich „schmerzartig“ anzufühlen (vgl. Polger, 2004, S. 11). So hat bereits Kim (1972/1980) festgestellt, dass „the mere fact that the physical bases of two nervous systems are different in material composition or physical organization with respect to a certain scheme of classification does not entail that they cannot be in the same physical state with respect to a different scheme“ (S. 234f). Insofern ist hiermit zunächst ein weiteres Argument dafür gegeben, dass man selbst solchen Tieren oder Außerirdischen die Eigenschaft, Schmerzen haben zu können, zuschreiben kann, die keine C-Faser-Reizungen aufweisen. C-Faser-Reizungen und die vielleicht existierenden alternativen physischen Grundlagen von Schmerzen aller zu Schmerzempfindungen fähigen Wesen müssen nach Maßgabe einer typentheoretischen Identitätstheorie hierzu lediglich *bestimmte* physische Eigenschaften gemeinsam haben. Keineswegs müssen sie in allen Eigenschaften übereinstimmen (genauso wie eben einzelne Schmerzzustände nicht in *allen* phänomenalen Eigenschaften übereinstimmen müssen, um als Schmerzen klassifiziert werden zu können).

Die in Identitätsaussagen zumeist verwendeten Arten (wie Schmerzen und wahrscheinlich auch C-Faser-Reizungen) sind also als *Idealisierungen* tatsächlicher Arten aufzufassen (vgl. Klein, 2008). So sind wissenschaftliche Artbegriffe häufig Abstraktionen von tatsächlichen Einzelereignissen oder von konkreten Entitäten (bzw. von viel eng umrisseneren natürlichen Arten). In ebendieser Weise erläutern auch Shapiro und Polger (2012), dass...

...scientific kinds are not simple and unstructured. The members of scientific kinds are not indiscernible from one another: they have many individual differences in both their kind-irrelevant and kind-relevant properties. Eyes vary in their mass, but also in their size, ability to focus, light-collection capacity, and so forth. The mere fact of these variations is not reason to think that the kinds of one science cannot be identified with the kinds of another any more than it is evidence against the existence of scientific kinds at all. (S. 276f)

So mag ein Element dieselbe Anzahl von Atomen aufweisen und damit ein Kriterium für eine Typen-Identität abliefern. Was daraus aber nicht folgt, ist, dass alle konkreten Einzelexemplare, die unter diesen Typ fallen bzw. die wir aus bestimmten Gründen unter diesen Typ ordnen, qualitativ völlig identisch miteinander sind (vgl. ebd., S. 271). Insofern kann das Argument der multiplen Realisierbarkeit hier gar nicht greifen, da eine kompositorische Variation innerhalb der Grenzen einer wissenschaftlich bestimmten Art ein ganz natürliches Phänomen ist, das allerdings bei der Bestimmung einer solchen Art zumeist vernachlässigt wird: „In forming scientific kinds, one often ignores or ‚substracts‘ variations among the members, even kind-relevant variations, and even potentially kind-busting anomalies. In these cases, the presence of variability in the composition of these kinds [...] does not demonstrate multiple realization because the kind might not be ‚ontologically committing‘“ (ebd., S. 279). So mögen wir zwar insofern von *den* (menschlichen) Schmerzen im All-

gemeinen sprechen und hierbei eine gewisse Homogenität unterstellen, genau genommen fallen aber unter diesen Begriff eine Vielzahl qualitativ unterschiedlicher einzelner Schmerzereignisse oder bestimmter Subtypen (wie Zahnschmerzen, Kopfschmerzen oder Rückenschmerzen). „Particular pain events are highly structured in space, time, and quality; and different pain events are associated with different stimuli, different time courses, different receptors, and so on“ (ebd., S. 278; vgl. Polger & Sufka, 2006).

Wenn der Begriff einer (mentalen) Art nicht nur eine sprachlich-begriffliche Konvention darstellen soll, sondern etwas faktisch Vorhandenes herausgreifen soll, das (individuenübergreifend) im Gehirn identifiziert werden kann, darf seine Extension, sein Begriffsumfang, offensichtlich nicht zu breit sein. Wie Bechtel und Mundale (1999) aufzeigen, besteht ein zentrales Problem für psychophysische Identitätsaussagen auf Typen-Niveau daher darin, dass unterschiedliche *Granularitätsmaßstäbe* in den Taxonomien von psychologischen und neuronalen Zuständen bestehen (vgl. ebd., S. 202). Während psychologische Zustände – wie sich schon in Bezug auf Schmerzen gezeigt hat – in relativ groben Klassen zusammengefasst werden, werden Gehirnzustände relativ feinkörnig bestimmt bzw. voneinander abgegrenzt. Wenn man nun das Granularitätsmaß auf beiden Seiten einer Identitätsaussage konstant hält, verliert das Argument der multiplen Realisierbarkeit endgültig seine Durchschlagskraft. Entschieden man sich für eine grobkörnige Analyse, erhält man psychologische Zustände, die sich über verschiedene Individuen und Arten hinweg zeigen. Bei gleichem Granularitätsmaß erhält man eine entsprechende (relativ grob gefasste) Aktivität in verschiedenen Hirnarealen, was die Möglichkeit einer Eins-zu-eins-Korrespondenz bewahrt (vgl. ebd.). Wenn man eine feinkörnige Granularität heranzieht und zwischen verschiedenen psychologischen Zuständen über Individuen hinweg differenziert

oder sogar bei einem Individuum über verschiedene Zeitpunkte hinweg, dann muss man auch bei der Analyse des Gehirns ein feines Granularitätsmaß anwenden und auch hier werden sich so die psychologischen Differenzen in neuronalen Differenzen und die neuronalen Differenzen in psychologischen Differenzen widerspiegeln (vgl. ebd.). Auch die herausgestellte Artspezifität mentaler Zustände lässt sich nach Maßgabe dieses Ansatzes analysieren: Schmerzen beim Menschen versus Schmerzen bei einer bestimmten Tierart (oder einem Außerirdischen) wären demnach feinkörnige Unterscheidungen innerhalb eines grobkörnig konzipierten mentalen Zustandes (vgl. Polger, 2004, S. 28). Jedenfalls verdeutlicht diese Granularitätsparität zugleich die These Donald Davidsons, dass es keine strengen psychophysischen Gesetze geben kann.

Aufgrund dieser Differenz der Körnung zwischen funktionalen mentalen Typen und neurobiologischen Typen schlagen Esfeld und Sachse (2007) bzw. Soom, Sachse und Esfeld (2010) vor, dass man im Vokabular der Psychologie präzisere funktionale mentale Subtypen mit einem "Körnungsgrad" definieren sollte, der letztendlich dem der neurobiologischen Typen entspricht. Ihre weitergehende Annahme dabei ist, dass die verschiedenen Mechanismen, die die Effekte produzieren, die einen bestimmten mentalen Typ kennzeichnen, systematisch unterschiedliche Nebeneffekte hervorbringen: „For any of these differences, there are physically possible conditions under which those side-effects manifest themselves in a functional salient way, that is, in a way that is salient for psychology. Hence, it is possible to grasp those differences in a functional manner by making the functional, mental types of psychology more precise in their own vocabulary” (Soom, Sachse & Esfeld, 2010, S. 11). Der mentale Typ M ist insofern eine Abstraktion – oder wie Klein (2008) sagt – eine Idealisierung, die aus verschiedenen funktional charakterisierbaren mentalen Subtypen M_1, M_2, \dots, M_n besteht, die jeweils mit anderen

neurophysiologischen (Sub-)Typen N_1, N_2, \dots, N_n koextensiv-identisch sind. Dabei gilt, dass die vorgeschlagene Prozedur wiederum auf dem Niveau der Subtypen selbst wiederholbar ist, das heißt, dass auch diese potentiell wieder in noch feinkörnigere Subtypen zerlegt werden könnten und diese wiederum etc., bis an irgendeinem Punkt keine weitere multiple Realisierbarkeit bzw. kompositorische Variabilität mehr auftritt (vgl. Soom, Sachse & Esfeld, 2010, S. 12).

Folgt man dem Subtypisierungsmodell in der hier vorgeschlagenen Form, zeigt sich, dass es nicht nur die Eingrenzung auf eine bestimmte Spezies, sondern auch die Bestimmung von möglichst feinkörnigen Typen ist, die als Konsequenz einer Verträglichkeit von kompositorischer Variabilität und Identitätsaussagen auf Typenniveau zu akzeptieren sind. Allerdings meinen Bechtel und Mundale (1999): „There are, after all, cases for which there are no alternative possible realizations for a set of functional properties (e.g., the functional properties of water can only be realized, given physical law as we understand it, by H_2O), and it is possible that mentality is one of these cases” (S. 204). Insofern wäre es natürlich in der Tat möglich, dass geistige Zustände immer und überall durch neuronale Zustände realisiert würden bzw. Schmerzen im ganzen Universum an die Existenz von C-Faser-Reizungen gebunden wären. Evidenz dafür oder dagegen gibt es (bisher) bekanntlich nicht.

Der Versuch einer grafischen Zusammenfassung realistischer Fälle kompositorischer Variationen bzw. multipler Realisierungen findet sich bei Polger (2004) und ist in Abbildung 33 dargestellt. In diesem Modell sind bestimmte Arten von mentalen Zuständen S_1, \dots, S_r identisch mit bestimmten Arten von Gehirnzuständen B_1, \dots, B_r . Diese Arten der mentalen Zustände lassen sich (z.B. im Rahmen der psychologischen Forschung oder auch im allgemeinen Sprachgebrauch) zu grobkörnigeren

und allgemeineren speziesspezifischen Arten mentaler Zustände gruppieren (z.B. M_α oder M_β).

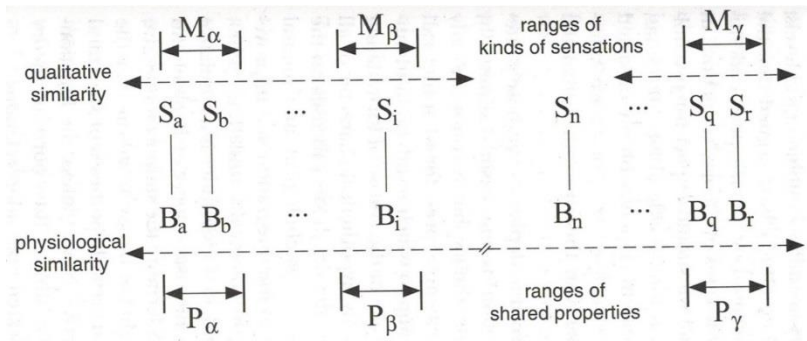


Abb. 33: Modell der multiplen Realisierbarkeit (Aus: Polger, 2004, S. 31)

Dabei kann erwartet werden, dass die Exemplare dieser Arten eine Übereinstimmung in bestimmten physischen Eigenschaften aufweisen (z.B. P_α). Damit gilt, dass Geschöpfe, die sich physisch ähnlich sind (wie die, die P_α oder P_β aufweisen) ähnliche mentale Eigenschaften aufweisen werden wie eben M_α und M_β (vgl. Polger, 2004, S. 30). Insofern werden Menschen mit Läsionen und Funktionsverlagerungen im Gehirn in der Tat ähnliche mentale Zustände aufweisen wie alle anderen Menschen, obgleich sie einen besonderen Subtyp darstellen und wahrscheinlich gewisse qualitative Unterschiede zum „Normalfall“ aufweisen werden²⁶⁸. Außerdem sollten wir in der Tat erwarten, dass „human beings and higher primates to have similar conscious mental states because their

²⁶⁸ So gilt für den Regelfall des Wiedererlernens bestimmter mentaler Leistungen nach z.B. einem mit Gehirnläsionen verbundenen Schlaganfall, dass die Personen nicht mehr zu dem Niveau der Leistungen zurückkehren können, wie es vor dem Schlaganfall war. Ebenso werden häufig andere Strategien angewandt, um z.B. bestimmte kognitive Leistungen zu erbringen (Rechnen, Orientierung, sprachliches Vermögen etc.).

brains are quite similar to our own. And we should expect the experiences of octopi or aliens to be different from ours to the extent that their brains are quite different from our own" (ebd.).

Nummehr kann auch die These kritisch betrachtet werden, die besagt, dass es Fälle „siliziumbasierter“ oder sonstiger künstlicher geistiger Zustände geben könnte, die als problematische Fälle multipler Realisierungen mentaler Eigenschaften anzusehen wären. Handelt es sich um phänomenale Zustände, mögen diese eine ähnliche Qualität haben wie „kohlenstoffbasierte“ mentale Zustände. Silizium und Kohlenstoff würden in dem Fall bestimmte Eigenschaften gemeinsam haben, die es ermöglichen, Schmerzen hervorzubringen. Dass es aber Realisierungen *derselben* faktisch vorhandenen Art mentaler Zustände sind, kann bezweifelt werden. So mögen siliziumbasierte Schmerzen – wenn es sie denn überhaupt geben kann – zwar taxonomisch unter die Kategorie „Schmerzen“ fallen. Hierbei handelt es sich aber nur um eine Idealisierung einer natürlichen Art. Vielmehr müsste ein solcher Schmerz als ein bestimmter Subtyp der abstrakten Kategorie „Schmerz“ betrachtet werden und stellt insofern keinen problematischen Fall einer kompositorischen Variabilität dar. Schwieriger könnte es bei bestimmten kognitiven Zuständen aussehen, wie z.B. das Wissen um den Fakt, dass Paris die Hauptstadt von Frankreich ist. Wenn man es für möglich erachtet, dass ein künstliches System diesen Fakt *genauso* repräsentiert wie ein Mensch, dürfte man einen guten Grund haben, von einer multiplen Realisierung dieses kognitiven Zustandes zu sprechen. Unterschiedliche kognitive Subtypen, die diesen Fakt jeweils irgendwie anders repräsentieren, scheinen kurios zu sein. Allerdings ließe sich überlegen, ob es bei eingehenderer Betrachtung nicht unter bestimmten Umständen doch sinnvoll sein kann anzunehmen, dass es verschiedene *Arten des Wissens* eines solchen Fakt es gibt. Eine Maschine mag eine Input-Output-Relation aufweisen, die die Zu-

schreibung der Überzeugung, dass Paris die Hauptstadt von Frankreich ist, nahelegt. Es muss aber nicht zwangsläufig ein Zustand sein, der dem entspricht, den wir als Menschen aufweisen. So haben wir als Menschen z.B. bestimmte Assoziationen, die wir mit Paris verbinden, vielleicht auch bestimmte emotionale Einfärbungen unseres Bewusstseins, weil wir z.B. in Paris unsere Flitterwochen verbracht haben. Ein kognitiver Zustand ist bei uns also für gewöhnlich in ein umfassendes semantisches Netzwerk eingebunden und/oder mit phänomenalen Qualitäten verbunden und fast nie ein bloßer isolierter Fakt. Es könnte daher aus gutem Grund hinterfragt werden, ob es wirklich *derselbe* Typ eines kognitiven Zustandes ist, den diese Maschine aufweist oder ob – wenn überhaupt – hierfür eine bestimmte Subtypisierung durchgeführt werden müsste. Vielleicht müsste man hier sogar so weit gehen, dass auch bei zwei Menschen genau genommen nur von dem Vorhandensein eines kognitiven Zustandes *ein und derselben* Kategorie bzw. Granularitätsebene gesprochen werden kann, wenn auch die (semantischen) Einbettungsbedingungen vergleichbar sind (was faktisch wohl nur äußerst selten gegeben sein dürfte). Diese große Unschärfe bei der Bestimmung genuiner mentaler Typen böte so jedenfalls auch eine Möglichkeit zur Entgegnung auf das von Block und Fodor (1972) bzw. Lycan (1987/2007) vorgetragene Argument der Neuroprothesen als Fälle multipler Realisierungen mentaler Zustände. Wenn es heißt, dass auch siliziumbasierte (statt kohlenstoffbasierte) Schmerzen möglich sein könnten, ließe sich außerdem entgegenhalten, dass die hier spezifizizierte Identitätstheorie nicht unbedingt annehmen muss, dass es für Schmerzen wesentlich ist, aus einem ganz bestimmten *Material* (Kohlenstoff vs. Silizium) geformt zu sein, sondern vielmehr eine spezifische (mikrostrukturelle) *Konfiguration* aufzuweisen (beide Materialarten werden aber bestimmte physische Eigenschaften gemeinsam haben, wenn beide es vermögen, Schmerzen hervorzubringen). C-

Faser-Reizungen sollten insofern keineswegs als eine rein *stoffliche* Art verstanden werden, sondern vielmehr als ein spezieller Typ der *Gestaltung* oder eben Konfiguration von Materie (vgl. Kap. 7.3.2). So spielt bei den kompositorischen Ebenen, wie sie im mechanistischen Ansatz propagiert werden, die Konfiguration der agierenden Entitäten der verschiedenen Ebenen eines Mechanismus die entscheidende Rolle, während die konkrete stoffliche Beschaffenheit zu vernachlässigen ist. In diesem Sinne sagt Boyd (1980), dass „there are certain configurations such that whenever they are realized by a physical system, whatever substances compose it, the qualitative feeling of pain is manifested“ (S. 96).

7.5.2.3. Die scheinbare Kontingenz psychophysischer Identitätsaussagen

Wenn man das Subtypisierungsmo­dell adäquat anwendet und auf beiden Seiten der psychophysischen Identitätsannahme mit dem gleichen Granularitätsmaß arbeitet, dürfte man eine Zuordnung erhalten, die über bloße heuristische Annahmen hinaus, einen *wesentlichen* oder natürlichen Zusammenhang herausgreift. Wenn sich also im Zuge dieses Verfahrens z.B. herausstellen sollte, dass die „klassischen“ Identitätstheoretiker in der Hinsicht doch recht haben sollten, dass menschliche Schmerzen immer mit C-Faser-Reizungen (auf der neuronalen Ebene) assoziiert sind, spricht viel dafür, dass wir hier eine Identitätsbeziehung vorliegen haben, für die das Kripkesche Diktum *notwendiger* Identitäten *a posteriori* gilt (vgl. Kap. 6.1.4.3). Diese Art der psychophysischen Relation wäre demnach eine Identität, wie sie offensichtlich zwischen Wasser und H₂O besteht: Für Wasser ist es wesentlich, dass es aus H₂O besteht, denn Wasser *ist* H₂O (auf der molekularen Ebene) und das unabhängig von jeder sprachlichen oder begrifflichen Konvention bzw. unserer wissenschaftlichen Entdeckung dieses Zusammenhangs (vgl. Boyd, 1980). So

mit müsste dann auch für menschliche Schmerzen gelten, dass es sich bei ihrer Identität mit C-Faser-Reizungen um keine Sache von Bedeutungszuschreibungen oder des Sprachgebrauchs handelt. Aus C-Faser-Reizungen (auf der neuronalen Ebene) zu bestehen, müsste demnach eine essentielle Eigenschaft von (menschlichen) Schmerzen *selbst* sein und das unabhängig von irgendwelchen Modalitäten der Beschreibung. Wenn dem so ist, dann muss eine Lösung für das von Kripke ausgemachte Problem gefunden werden, dass die – trotz des *de facto* notwendigen Zusammenhangs – *scheinbare* Kontingenz zwischen Schmerzen und C-Faser-Reizungen (wir können uns das eine ohne das andere problemlos vorstellen) vermeintlich nicht auf diese Weise ausgeräumt werden kann, wie es im Falle der ebenfalls scheinbaren Kontingenz von Wasser und H₂O möglich ist (wir mögen uns zwar „Welten“ mit Wasser vorstellen können, das nicht H₂O ist (und umgekehrt), orientieren uns bei dieser Vorstellung aber zu Unrecht an den – tatsächlich kontingenten – *qualitativen* Eigenschaften von Wasser, anhand derer wir Wasser für gewöhnlich erkennen). Ist die Folgerung Kripkes also wirklich richtig, dass eine solche Ausräumung bei psychophysischen Identitätsannahmen zwangsläufig scheitern muss, da alles, was sich anfühlt wie Schmerz, in jeder möglichen Welt Schmerz ist und jedes natürliche Phänomen, das Schmerz ist, sich so anfühlen muss wie Schmerz in dieser Welt?

Man kann diese Argumentationsweise Kripkes direkt in *psychologischer* Hinsicht infrage stellen. Denn kann man aus der reinen Vorstellbarkeit oder Denkbarkeit psychophysischer Dissoziationen, die sich offenbar nicht so gut wegerklären lassen wie bei „rein physischen“ Identitäten wie der von Wasser und H₂O, wirklich so weitreichende (ontologische) Schlüsse ziehen, wie Kripke dies tut? Wie in Kapitel 3 im Zusammenhang der Besprechung des psychologischen bzw. „subjektivistischen“ Ursprungs des Leib-Seele-Problems aufgezeigt wurde, bestehen zwischen

der perzeptuellen bzw. kognitiven Verarbeitung von mentalen und „gewöhnlichen“ physischen Größen erhebliche Unterschiede. Obwohl Gehirn und Geist keine substantiell verschiedenen Wirklichkeitsbereiche konstituieren, sondern koextensiv-identische Ebenen in einer mereologisch-mechanistischen Hierarchie, nehmen wir sie aus *psychologischen* Gründen mehr oder weniger zwangsläufig völlig unterschiedlich wahr. Die zentralen Variablen hierfür sind die unterschiedliche Aufmerksamkeitsausrichtung bei der Wahrnehmung externer physischer Größen gegenüber der Wahrnehmung mentaler Phänomene *als* mentaler Phänomene, die psychischen Strukturmerkmale der phänomenalen Transparenz und Perspektivität sowie die divergierende Verarbeitung durch – wahrscheinlich phylogenetisch verankerte – konzeptuelle Formen des Wahrnehmungssystems bzw. höherstufige konzeptuelle Systeme (Theory-of-Mind-Modul vs. Theory-of-Matter-Modul). Im Zuge dessen werden jeweils eine Reihe weiterer – stark divergierender – bedeutungshaltiger Kategorien mit aktiviert und in unser mentales vs. physisches Perzept integriert (räumlich lokalisierbar vs. räumlich nicht-lokalisierbar; sichtbar vs. unsichtbar; „greifbar/fest“ vs. „flüchtig“ usw.).

Aufgrund dieser kognitiven bzw. perzeptuellen Verarbeitungsdifferenzen werden daher auch bei der Vorstellung oder beim Nachdenken über mentale Phänomene wie Schmerzen ganz andere kognitive Kategorien aktiviert als bei der Vorstellung bzw. beim Nachdenken über C-Faser-Reizungen, wobei die Inhalte des Theory-of-Matter- und Theory-of-Mind-Moduls (sinngemäß) nach Mausfeld (2006) nicht in ein gemeinsames Datenformat überführt werden können. Wenn wir also die starke Intuition haben, dass (trotz der hier postulierten notwendigen Identität) Schmerzen ohne C-Faser-Reizungen oder C-Faser-Reizungen ohne Schmerzen existieren könnten, liegt das vor allem daran, dass wir über kein höherstufiges kognitives Modul verfügen, mittels dessen eine „Aussöhnung“ unserer per-

zeptuellen und konzeptuellen Vorstellungen stattfinden könnte. Wir haben insofern sehr unterschiedliche Begriffe in Bezug auf phänomenale versus physische Zustände und Prozesse ausgebildet, die unser Denken und Vorstellen leiten, sodass der Anschein erweckt wird, dass auch die Bezugsgegenstände dieser Begriffe ontologisch nicht zusammenfallen können (vgl. Loar, 1997/2007).

Es ist interessant, dass sich gerade in einer Fußnote in Thomas Nagels Aufsatz *Wie ist es, eine Fledermaus zu sein?* eine ganz ähnliche Argumentation findet. Nagel meint zwar, dass Kripkes Kritik am Materialismus seiner eigenen in gewisser Weise nahekomme, da er es genau wie Kripke unverstündlich fände, dass ein bestimmter Zustand des Gehirns *notwendig* einen subjektiven Charakter haben soll, ohne dass dies weiter erklärt wird (vgl. ebd., 1974/1997, S. 274). Allerdings hält er es *im Gegensatz* zu Kripke für durchaus möglich, dass der Anschein von Kontingenz als notwendig wahr zu erachtender psychophysischer Identitätsaussagen ausgeräumt werden kann, wenn man berücksichtigt, dass bei der Vorstellung von Gehirnprozessen ganz andere psychische bzw. kognitive Mechanismen greifen als bei der Vorstellung mentaler Prozesse:

Eine Theorie, die erklärte, weswegen die Beziehung zwischen dem Mentalen und dem Gehirn notwendig ist, würde uns immer noch mit Kripkes Problem zurücklassen zu erklären, warum es trotzdem kontingent zu sein scheint. Diese Schwierigkeit scheint mir in der folgenden Weise überwindbar²⁶⁹ zu sein. Wir können uns etwas vorstellen, indem wir es uns entweder sinnlich, einfühlend oder symbolisch präsentieren. Ich möchte nicht versuchen zu sagen, wie symbolische Phantasie funktioniert; aber ein Teil von dem, was in den beiden anderen Fällen geschieht, ist das Folgende: Um uns etwas sinnlich vorzustellen, versetzen wir uns selbst in einen

²⁶⁹ Hier überliegt dem Übersetzer (Ulrich Diehl) ein inhaltlich nicht unerheblicher Fehler, wenn er „surmountable“ mit der Übersetzung „unüberwindlich“ in sein Gegenteil verkehrt. Hier wird das Wort auch im Zusammenhang der folgenden Argumentationsweise Nagels korrekt mit „überwindbar“ wiedergegeben.

bewußten Zustand, der dem Zustand gleicht, in dem wir uns befänden, wenn wir es wahrnähmen. Um uns etwas einfühlend vorzustellen, versetzen wir uns selbst in einen Zustand, der der Sache selbst gleicht. (Diese Methode kann nur angewendet werden, um sich mentale Ereignisse oder Zustände vorzustellen – unsere eigenen oder die eines anderen.) Wenn wir versuchen, uns einen mentalen Zustand vorzustellen, der ohne den dazugehörigen Zustand des Gehirns auftritt, stellen wir uns zuerst durch Einfühlung das Auftreten des mentalen Zustandes vor, d.h. wir versetzen uns selbst in einen Zustand, der ihm mental gleicht. Zur gleichen Zeit versuchen wir uns das Nicht-Auftreten des zugehörigen physischen Zustandes vorzustellen, indem wir uns in einen anderen Zustand versetzen, der mit dem ersten nicht verbunden ist, – in einen Zustand, der dem gleicht, in dem wir uns befänden, wenn wir das Nicht-Auftreten des physischen Zustandes wahrnähmen. Wo die Vorstellung von physischen Eigenschaften sinnlich ist und die Vorstellung von mentalen Eigenschaften einfühlend, hat es für uns den Anschein, daß wir das Auftreten jedes Erlebnisses ohne den dazugehörigen Gehirnzustand vorstellen können und umgekehrt. Die Relation zwischen ihnen wird als kontingent erscheinen, selbst wenn sie notwendig ist, weil die verschiedenen Typen von Vorstellung voneinander unabhängig sind. (ebd.)

Es ist demnach also die Diskrepanz zwischen einer sinnlichen und einfühlenden Vorstellungsweise, die der vermeintlichen Gehirn-Geist-Kontingenz zugrunde liegt. Diese Art von Dissoziation zwischen mentalen und physischen Eigenschaften geht daher nur auf uns selbst und nicht auf eine reale Differenz (im Sinne unterschiedlicher Wirklichkeitsbereiche oder Substanzen) zurück, die Kripke hieraus abzuleiten können meint.

Aufgegriffen und einer weitergehenden philosophischen Präzisierung unterzogen hat diesen Ansatz Nagels vor allem Christopher Hill (1997/2002). Im Hinblick auf die *Vorstellbarkeit* psychophysischer Dissoziationen gilt nach Hill, dass wir uns einer sympathetischen (einfühlenden) Vorstellungskraft beim Imaginieren von geistigen Zuständen und einer perzeptuellen Vorstellungskraft beim Imaginieren von Gehirnzuständen bedienen. Die Ursache der psychophysischen Dissoziation liegt demnach darin, dass es zwei unterschiedliche psychische Mechanismen, zwei Ar-

ten der Imagination gibt, die relativ unabhängig voneinander operieren können. Die Vorstellung „körperloser Geister“ berufe sich somit darauf, dass wir „eine sympathetische Vorstellung einer Situation, die einen geistigen Zustand enthält, mit der perzeptuellen Vorstellung einer Situation, die keinen Gehirnprozess enthält, [zusammenfügen]“ (ebd., S. 189). Aus der reinen Möglichkeit, dass wir so etwas widerspruchsfrei tun können, folgt nun aber keineswegs, dass es für den vorgestellten geistigen Zustand bzw. für geistige Zustände überhaupt faktisch möglich ist, ohne einen hiermit assoziierten Hirnzustand zu existieren.

Im Hinblick auf die *Denkbarkeit* psychophysischer Dissoziationen geht Hill (ähnlich wie Loar) davon aus, dass der Begriff des Schmerzes und der Begriff der C-Faser-Reizung ganz unterschiedlichen Klassifikationszwecken dienen. Mittels des Schmerz Begriffes wird die Klassifikation von Phänomenen aufgrund von introspektiven Informationen ermöglicht, während mittels des Begriffes der C-Faser-Reizung eine Klassifikation von Phänomenen ermöglicht wird, die wir über unsere Sinne erfassen. Daher würden es ganz unterschiedliche Erfahrungen sein, die den Gebrauch dieser beiden Begriffsklassen leiten würden (vgl. ebd., S. 204). „Aber wenn nun der Gebrauch eines Begriffes von Erfahrungen geleitet wird, die völlig unabhängig von denen sind, die den Gebrauch eines zweiten Begriffes leiten, dann ist es für uns generell unmöglich, allein durch unsere Kompetenz im Gebrauch dieser Begriffe festzustellen, ob sie koreferentiell sind“ (ebd.). Insofern sei es nicht verwunderlich, dass unser korrekter Gebrauch von solchen Begriffen nicht allein zu der Einsicht führen kann, dass die Bezugsgegenstände dieser Begriffe miteinander identisch sind. Hill schließt zwar nicht aus, dass es unter Umständen möglich ist, korrekte modale *a posteriori* Intuitionen hervorzubringen (also dass eine vorgestellte oder gedachte Dissoziation in bestimmten Fällen durchaus den Tatsachen entsprechen kann), grundsätzlich gelte aber,

dass solche Intuitionen vor allem dann falsch sind, „wenn wir sie ohne vollständige Kenntnis der relevanten empirischen Tatsachen ausbilden“ (S. 206). Aus der bloßen gedanklichen oder imaginativen Möglichkeit zu psychophysischen Dissoziationen sollten jedenfalls keinerlei Rückschlüsse auf die *reale* Möglichkeit solcher Dissoziationen gezogen werden.

Berücksichtigt man die Artspezifität und Granularitätsparität bei der Formulierung psychophysischer Typen-Identitäten zeigt sich, dass die Kripkesche Argumentation ebenso schnell in sich zusammenfällt, wenn man sie nicht bereits in psychologischer Hinsicht infrage stellt, sondern sich auf den Weg der Überprüfung der *begrifflichen* Kohärenz ihrer Annahmen macht. Offenbar greift Kripkes These nur unter der Voraussetzung, dass man annimmt, dass mittels identitätstheoretischer Aussagen eine exklusive Wesensbestimmung des Mentalen vorgenommen werden soll (was – wie aufgezeigt – aber keineswegs der Fall sein muss). Gemäß der hier entwickelten Identitätskonzeption ist die von Kripke angegriffene Identitätsaussage: „Schmerzen sind (nichts anderes als bzw. in Wirklichkeit) C-Faser-Reizungen“ nämlich tatsächlich falsch. Denn es kann zunächst nur von einer *artspezifischen* Identität ausgegangen werden, sodass gilt, dass (zumindest) *menschliche* Schmerzen (auf der neuronalen Ebene) C-Faser-Reizungen sind. Dass Schmerzen im ganzen Universum und in jeder möglichen Welt an C-Faser-Reizungen gebunden sein müssen, wird somit gar nicht behauptet, was Kripkes Argumentation bereits weitestgehend den Boden entzieht. Des Weiteren müsste man auch menschliche Schmerzzustände genau genommen weiter subtypisieren, bis man auf neuronaler und mentaler Ebene dasselbe Granularitätsmaß vorliegen hat (der Einfachheit halber sei hier aber eben angenommen, dass menschliche Schmerzen tatsächlich immer mit C-Faser-Reizungen assoziiert sind). Berücksichtigt man diese Artspezifität als besonderen Fall des Subtypisierungsansatzes, steht der Ausräumung der vermeintlichen Kontin-

genz zwischen Schmerzen und C-Faser-Reizungen nichts mehr im Weg: Natürlich kann es Bewusstseinszustände bzw. phänomenale Eigenschaften geben, die sich „oberflächlich“ so anfühlen könnten wie menschliche Schmerzen, aber *de facto* keine solchen Schmerzen sind. So könnten „Putnams Außerirdische“ oder Tiere durchaus Bewusstseinszustände aufweisen, die menschlichen Schmerzzuständen ähnlich sind, trotzdem müssen sie nicht von *derselben* phänomenalen Qualität sein. Ein Tier kann daher potentiell über Schmerzzustände verfügen, die nicht an C-Faser-Reizungen gebunden sind, ohne dass die Notwendigkeit der Annahme, dass menschliche Schmerzen (auf der neuronalen Ebene) C-Faser-Reizungen sind, hierdurch gefährdet wird. Solche „tierischen Schmerzen“ stellen einfach einen anderen Subtyp von Schmerzen dar. Ganz ähnlich argumentiert Boyd (1980) im Hinblick auf die Möglichkeit, dass man sich C-Faser-Reizungen vorstellen könnte, die nicht von Schmerzen „begleitet“ sind: „For any physiological or anatomical description like the imaginary ‘C-fiber-firings’ there is certainly a possible world in which something has the qualitative properties typically associated with the term in the actual world, but really is not, in this case, an instance of ‘C-fiber-firings’. Some other sort of nerve cell might, in a different possible world, look just the way C-fibers do in the actual world” (S. 84). Der Typ “C-Faser-Reizungen” hat (wie Schmerzen) keine eindeutige qualitative Bestimmung. Insofern könnte es durchaus Vorgänge geben, die wie C-Faser-Reizungen aussehen, ohne dass sie aber *de facto* diesem Typ bzw. dieser Eigenschaftsklasse angehören. Insofern würden wir uns bei der gedanklichen Dissoziation von Schmerzen und C-Faser-Reizungen Dinge vorstellen, die oberflächlich aussehen mögen wie C-Faser-Reizungen, tatsächlich aber nicht zu dieser Klasse von Entitäten gehören (und daher auch nicht von Schmerzen „begleitet“ sein müssen). Genauso wie wir anderen Geschöpfen aus guten Gründen Schmerzen zuschreiben mögen,

die aber phänomenal keineswegs so beschaffen sein müssen wie unsere eigenen menschlichen Schmerzen²⁷⁰ (und daher auch nicht mit C-Faser-Reizungen assoziiert sein müssen). Formal gesehen besteht die von Boyd beschriebene Möglichkeit der Ausräumung der vermeintlichen Kontingenz also darin, dass man entweder den Ausdruck „Schmerz“ oder den Ausdruck „C-Faser-Reizungen“ mit einer qualitativen Beschreibung ersetzt, die kein *rigider Designator* ist (vgl. ebd., S. 84). Es ist diese qualitative Beschreibung, auf die wir uns in unseren irregeleiteten Vorstellungen und Denkwegen bezüglich psychophysischer Dissoziationen beziehen, obgleich der Gültigkeit dieser (im Zusammenhang der fraglichen Identitätsaussage) eben keine Notwendigkeit zukommt.

Soweit also die Analyse der Möglichkeit zur Ausdehnung der vorgestellten psychophysischen Konzeption auf typentheoretische Verhältnisse. Ob man dies nun als vollkommene Ausräumung des Argumentes der multiplen Realisierbarkeit akzeptieren mag oder auch nicht, ist für die vorliegende Arbeit allerdings weniger relevant, da die hier verteidigte Identitätskonzeption keineswegs von der Möglichkeit *notwendiger typentheoretischer* Annahmen *a posteriori* abhängt.

7.6. Was bleibt von der „Erklärungslücke“?

An dieser Stelle soll die Frage nach einer etwaigen explanatorischen Lücke (Levine, 1983) bzw. des schwierigen Problems des Bewusstseins (Chalmers, 1995, 1996) aufgegriffen werden (vgl. Kap. 6.1.4.4 bzw.

²⁷⁰ Jedenfalls sollte damit auch klar sein, dass über eine Verlagerung der Kripkesche Argumentation zu *menschlichen* Schmerzen keine Rehabilitation dieses Gedankengangs erfolgen kann. Es würde hierbei lediglich zu einer dementsprechenden Verlagerung der Erwiderung auf dieses Argument kommen (was den Subtypisierungs- bzw. Granularitätsgrad anbelangt).

6.5.4.1). Sie betrifft den Umstand, dass man auch aus einem perfekten Wissen über neuronale (oder überhaupt genuin physische) Tatsachen nicht ableiten kann, warum es genau ebendiese Tatsachen sind, die einem mentalen Ereignis zugrunde liegen. Es scheint also trotz der vorgestellten schlüssigen Identitätskonzeption die Frage im Raum zu bleiben, warum es gerade C-Faser-Reizungen (auf der neuronalen Ebene) sind, die mit Schmerzen identisch sind und nicht A-, B- oder D-Faser-Reizungen bzw. warum es gerade diese phänomenale Qualität des Erlebens von Schmerzen ist und nicht eine andere, wenn C-Fasern „feuern“. Um nochmals Du Bois-Reymond (1880/1974) zu Wort kommen zu lassen, lässt sich vermeintlich durch „keine zu ersinnende Anordnung oder Bewegung materieller Teilchen [...] eine Brücke ins Reich des Bewusstseins schlagen“ (S. 70).

Ist eine solche strenge Erklärungslücke, eine solche epistemische Kluft zwischen dem Psychischen und seinen physischen Grundlagen aber wirklich gegeben? Wie aufgezeigt wurde, bietet bereits der Ansatz der Funktionalisierung bzw. funktionalen Reduktion mentaler Eigenschaften eine Möglichkeit, zumindest für kognitive Eigenschaften diese Kluft weitestgehend zu schließen (vgl. Kap. 6.5.1). Wenn man denn die Prämisse zu akzeptieren bereit ist, dass eine mentale Eigenschaft nichts anderes als ihr neuronaler Realisierer ist, dann gilt, dass die mentale Eigenschaft eben mit genau diesem Realisierer identisch ist und nicht mit irgendeiner anderen physischen Größe, weil er genau das *kausale Profil* erfüllt, durch das die mentale Eigenschaft charakterisiert ist. Dass also gerade C-Faser-Reizungen die Realisierer menschlicher Schmerzen sind, liegt einfach daran, dass es diese neuronale Aktivität ist, die das funktionale Profil von Schmerzen innehat (z.B. vorausgehende Gewebeverletzungen, nachfolgendes Verhalten wie z.B. Schonung, Einnahme von Schmerztabletten etc.).

Ein noch umfassenderer Erklärungsansatz bestimmter psychischer Leistungen wird unzweifelhaft durch die hier favorisierten mechanistischen Erklärungsprinzipien bereitgestellt. Es wird dabei keineswegs einfach eine psychoneurale Identität (auf dem Niveau von Eigenschaftsinstantiierungen) behauptet, sondern eine mentale Leistung über mereologisch verfasste Mechanismen unterschiedlicher Granularitätsstufen *erklärt*. Wie immer wieder als Beispiel herangezogen, kann die Existenz des Phänomens des räumlichen Gedächtnisses eines Organismus über die agierenden Entitäten der verschiedenen, hier zum Tragen kommenden mechanistischen Ebenen hinreichend erklärt werden. So ist eine mentale Leistung mit einem ganz bestimmten, mehrere natürliche Organisations-ebenen umgreifenden Mechanismus verbunden, der im Falle des räumlichen Gedächtnisses vor allem die (eine räumliche Karte generierende) Hippocampus-Formation, die Langzeit-Potenzierung im synaptischen Spalt sowie die NMDA-Rezeptor-Aktivierung umfasst (vgl. Craver, 2008, S. 165ff). Dabei gilt (vereinfacht gesagt), dass die verschiedenen Ebenen des Mechanismus miteinander identisch sind, da sie verschiedene (mikro-) strukturelle Konfigurationen ein und derselben Eigenschaft sind (vgl. Kap. 7.1). Da die Konfiguration bzw. die Aktivität der Entitäten einer Ebene zugleich über die Konfigurationen und Aktivitäten der niedrigeren Ebene erklärt werden kann, liegt insofern keine beliebige Verbindung zwischen den einzelnen Ebenen vor, sondern eine, deren Sosein explanatorisch recht abgeschlossen ist. Wenn nun die Frage gestellt würde, warum es z.B. gerade die Hippocampus-Formation ist, die auf neuronaler Ebene eine entscheidende Rolle spielt und nicht eine andere Hirnregion, dann kann darauf verwiesen werden, dass es nun mal das kausale *und* strukturelle Profil ebendieser neuronalen Formation ist, das hierzu offensichtlich benötigt wird. Denn natürlich könnte man zwar auch fragen, warum es gerade dem Herzen und nicht der Leber zukommt, für die Blut-

zirkulation im Körper zu sorgen, aber offensichtlich handelt es sich hierbei um eine absurde Frage, da es nun mal die kausale und strukturelle Verfasstheit des Herzens ist, die seiner Funktionstüchtigkeit zugrunde liegt, was dann aber auch in gleicher Weise für den Hippocampus und die Fähigkeit zur räumlichen Orientierung gilt.

Aber in der Tat könnte es sich im Fall phänomenaler bzw. bewusster Eigenschaften schwieriger verhalten als im Falle kognitiver Eigenschaften wie dem räumlichen Gedächtnis. Es wurde zwar aufgezeigt, dass es prinzipiell möglich ist, auch für phänomenale Eigenschaften wie Schmerzen oder Farbempfindungen vergleichbare mechanistische Erklärungsprinzipien zu formulieren, allerdings scheint hier – vor allem in Bezug auf sinnliche Empfindungen – eine epistemische Kluft bestehen zu bleiben. Denn ungeachtet der Möglichkeit der Etablierung eines mechanistischen (oder auch eines „klassischen“ funktionalistischen) Modells für Farbempfindungen, scheint eine gewisse Beliebigkeit zwischen den physischen Basisstrukturen und der konkreten phänomenalen „Struktur“ oder Qualität eines solchen Ereignisses zu bestehen. Es bleibt die Frage im Raum, warum es auf neuronaler Ebene z.B. gerade V4 bzw. das neuronale Netzwerk V_{CORF} ist, das mit Farbempfindungen verbunden ist. Während es (zumindest vom Grundsatz her) nachvollziehbar ist, warum gerade der Hippocampus auf neuronaler Ebene mit der Fähigkeit zu räumlicher Orientierung verbunden ist (weil hier eben bestimmte materielle Mechanismen lokalisiert sind, die eine räumliche Repräsentation der Umgebung ermöglichen, ganz ähnlich wie es der materielle Mechanismus einer Taschenuhr ermöglicht, den Zeitfluss zu „repräsentieren“), scheint es bei phänomenalen Qualitäten bzw. bewussten Ereignissen nicht einsichtig zu sein, *warum* oder überhaupt *wie genau* es die physischen Basisstrukturen ermöglichen, diese hervorzubringen und warum sie gerade mit einer ganz bestimmten Qualität (wie ganz spezifische Farbempfin-

dungen) verbunden sind (und nicht vielleicht mit auditiven oder olfaktorischen Eindrücken). Man kann zwar den Weg von der retinalen Aufnahme von Farbinformationen bis hin zur kortikalen Verarbeitung im Okzipitallappen nachverfolgen und im Zuge einer sogar unterstellten „perfekten“ Neurowissenschaft alle auf diesem Weg installierten Mechanismen studiert haben; dennoch scheint es uneinsichtig zu bleiben, warum, wie und wo genau plötzlich eine „phänomenale Welt“, ein bewusstes Erleben solcher farblichen Eindrücke einsetzt. In dieser epistemisch-explanatorischen Hinsicht erlebt das in ontologischer Hinsicht bereits als harmlos abgehandelte Leibnizsche Mühlengleichnis bzw. Bieris Rundgang durch das Gehirn, der kein Bewusstsein auffinden lässt (vgl. Kap. 2.1.4), also offenbar ein „Comeback“.

Hat eine psychophysische Identitätstheorie daher tatsächlich mit einem schwerwiegenden epistemischen Defizit im Hinblick auf phänomenale Eigenschaften zu kämpfen? Vielleicht lässt sich diese Erklärungslücke tatsächlich nicht vollkommen schließen. Das unterstellte Ausmaß eines explanatorischen Defizits und dessen Bedeutsamkeit für die Gültigkeit einer psychophysischen Identitätstheorie werden jedoch stark überschätzt. Zunächst kann man hinterfragen, ob das hier geforderte Ausmaß an explanatorischem Gehalt überhaupt von einer begrifflich-konzeptuellen „Theorie des Geistes“ ernsthaft erwartet werden *kann*. Sind es also überhaupt prinzipiell einlösbare Forderungen an den explanatorischen Gehalt einer psychophysischen Identitätstheorie? Dies wird von Autoren wie David Papineau (1995, 2002) vehement bestritten und darauf verwiesen, dass Identitätsbehauptungen grundsätzlich keiner Erklärung bedürfen:

Wenn bewußte Eigenschaften identisch mit materiellen Eigenschaften sind, dann, so behaupte ich, gibt es kein Rätsel hinsichtlich der Frage, weshalb materielle

Eigenschaften bewußte Eigenschaften „hervorbringen“. Das ist so, weil Identitäten keiner Erklärung bedürfen. Wenn die „zwei“ Eigenschaften in Wirklichkeit eine sind, dann ist es nicht so, daß die materielle Eigenschaft die bewußte Eigenschaft „hervorbringt“ – sie ist die bewußte Eigenschaft. Und wenn sie es ist, dann gibt es kein Rätsel, warum sie ist, was sie ist. (ebd., 2002, S. 231)

Nach Ansicht Papineaus ist es unter den Voraussetzungen der Identitätstheorie also eine sprachliche Verwirrung, wenn man sagt, dass ein neuronales Geschehen ein (bewusstes) mentales Geschehen *hervorbringt*. Wenn es sich um ein und dasselbe Geschehen handelt bzw. um ein und dieselbe Eigenschaft eines Systems, dann ist es natürlich unsinnig zu sagen, dass das eine das andere hervorgebracht hat. Der Begriff der „Hervorbringung“ einer mentalen Eigenschaft impliziert genau genommen eine Kausalbeziehung, die mit der hier verteidigten Identitätskonzeption unverträglich sein dürfte. Wenn man sich die Identität von Wasser und H₂O anschaut, sei letztendlich nur die Begründung der Identität darüber möglich, dass sie „stets an derselben Stelle in der kausalen Ordnung der Dinge auftreten“ (ebd., S. 231f). Eine weitergehende Erklärung, *warum* sie denselben Stoff bezeichnen, sei hingegen unsinnig. Daher gelte auch für phänomenale Eigenschaften wie Schmerzen:

Wir können nach Aufklärung darüber verlangen, was denn *zeigt*, daß Schmerzen identisch mit dem Feuern von C-Fasern sind (erneut lautet die Antwort, daß sie an derselben Stelle in der kausalen Ordnung der Dinge erscheinen). Doch es macht keinen Sinn, nach einer weiteren Erklärung dafür zu fragen, *warum* das Feuern von C-Fasern Schmerzen hervorbringt. Wenn sie das Feuern von C-Fasern sind, werden sie nicht durch diese „hervorgebracht“. Sie sind diese, und das ist alles. (ebd., S. 232; Hervorh. im Orig.)

Nach Ansicht Papineaus sind Identitäten also im Allgemeinen bloße, nicht näher zu erklärende Fakten. Man kann sie feststellen oder entdecken, aber ihr Vorhandensein nicht weiter begründen. C-Fasern *sind* (auf

neuronaler Ebene) einfach *nichts anderes als* C-Faser-Reizungen. Farberpfindungen *sind* (auf neuronaler Ebene) *nichts anderes als* Aktivitäten in der V4-Region.

Jedoch räumt Papineau ein, dass bei Identitäten wie der von Wasser und H₂O eine gewisse Erklärung der Identitätsbeziehung vorliegen mag. Denn es ist hier durchaus möglich und sinnvoll, nach einer Erklärung dafür zu verlangen, warum Wasser H₂O ist, wenn man hiermit meint, dass eine Erklärung dafür abgegeben werden soll, warum H₂O die qualitativen Eigenschaften aufweist, die wir für gewöhnlich Wasser zuschreiben (also eine geschmacks- und geruchslose, durchsichtige Flüssigkeit zu sein). Hierauf könnte man dann mithilfe der Erkenntnisse der Chemie eine Erklärung für die behauptete Identität abliefern (vgl. ebd., S. 233). Dagegen scheint eine solche Erklärung der Identität von Schmerzen und C-Faser-Reizungen nicht möglich zu sein, denn offenbar sind wir eben nicht in der Lage zu erklären, „warum gewisse physische Zustände die Beschreibungen erfüllen, die *a priori* mit unseren alltäglichen Begriffen von Bewußtseinszuständen verknüpft sind“ (ebd.; Hervorh. im Orig.).

Papineau versucht diese Diskrepanz dadurch aufzulösen, dass phänomenale Begriffe einen anderen Bezug aufweisen als nicht-phänomenale Begriffe und daher in diesen beiden Fällen nicht die gleichen Kriterien an eine Erklärung angelegt werden können (vgl. ebd., S. 236ff). So bestreitet er, dass es eine Referenz phänomenaler Begriffe gibt, die durch eine apriorische Verknüpfung mit bestimmten Beschreibungen konstituiert wird (wie dies bei Wasser der Fall ist). Vielmehr hänge die Verwendung phänomenaler Begriffe von der Bezugnahme auf die *Simulation* der phänomenalen Zustände ab, auf die die phänomenalen Begriffe referieren (in Kürze: Menschen machen Erfahrungen, dabei werden bestimmte Hirnregionen aktiviert; Menschen können nun gewissermaßen diese Hirnareale reaktivieren und so die entsprechenden phänomenalen Zu-

stände *simulieren*, auf die sie sich bei der Verwendung phänomenaler Begriffe beziehen; ebenso kämen solche Simulationen zum Tragen, wenn wir phänomenale Eigenschaften in andere „hineinprojizieren“). Phänomenale Begriffe referieren infolge dessen durch eine solche Simulation auf ihre Gegenstände, ohne dabei irgendwelche Beschreibungen zu berücksichtigen. Insofern seien die explanatorischen Ansprüche an die Identitäten von Wasser und H₂O auf der einen und Schmerzen und C-Faser-Reizungen auf der anderen Seite nicht miteinander vergleichbar. „Da der Ausdruck ‚Schmerz‘ einen phänomenalen Begriff bezeichnet, ist er nicht mit irgendwelchen Beschreibungen *apriorisch* verknüpft; daher gibt es keine Möglichkeit zu zeigen, wie verknüpfte Beschreibungen einen physischen Zustand herausgreifen, wie es im Fall von ‚Wasser‘ möglich ist“ (ebd., S. 237; Hervorh. im Orig.). Wenn also – wie Pauen (2002a) es zusammenfasst – der „Gegenstandsbezug phänomenaler Begriffe nicht über bestimmte Kriterien vermittelt [ist], die in einer funktionalen Beschreibung erfasst werden können“ und „[...] funktionale Beschreibungen damit bereits aus begrifflichen Gründen ausgeschlossen sind, dann ist der Vergleich mit reduktiven Erklärungen physischer Entitäten, die solche funktionalen Beschreibungen erlauben, von vornherein verfehlt“ (S. 180).

Bleibt dennoch der Anschein einer Erklärungslücke bestehen, liege das daran, dass Menschen nun mal grundsätzlich mit dualistischen Intuitionen an das Gehirn-Bewusstsein-Verhältnis herantreten (psychische Ursachen hierfür wurden in Kapitel 3 bereits besprochen). Dass man nicht bereit sei, phänomenale Eigenschaften als eine besondere Art von physischen Eigenschaften zu akzeptieren bzw. dass phänomenale Begriffe letztendlich auf physische Eigenschaften referieren, geht nach Papineau (2002) vor allem auf einen „antipathetischen Fehlschluss“ (ebd., S. 239) zurück. Einen solchen Fehlschluss stellt die (Alltags-)Überzeugung dar,

dass physische Prozesse keine phänomenalen Eigenschaften aufweisen können²⁷¹. Die Ursache dessen sei, dass neurobiologische (oder allgemein physische) Begriffe nicht mit einer Simulation phänomenaler Zustände verbunden sind, wie dies bei phänomenalen Eigenschaften der Fall ist (vgl. ebd., S. 239ff). Wenn behauptet wird: Schmerzen sind C-Faser-Reizungen, würde man dies als unplausibel zurückweisen, weil man Schmerzen phänomenal simulieren (und dabei ihre „Schmerzhaftigkeit“ erleben) kann, C-Faser-Reizungen hingegen nicht (jedenfalls nicht so, dass hierbei die qualitativen Eigenschaften von Schmerzen offenbart werden). Auch wenn Papineau das selbst nicht so formuliert, spielt hier offensichtlich das psychologische Strukturmerkmal der *Perspektivität* eine entscheidende Rolle. Aus der Dritte-Person-Perspektive des Neurowissenschaftlers ist nicht ersehbar, dass C-Faser-Reizungen in Perspektive der ersten Person des betrachteten Individuums die phänomenale Beschaffenheit aufweisen, die für Schmerz charakteristisch ist. Hieraus kann aber keineswegs abgeleitet werden, dass zwischen Schmerzen und C-Faser-Reizungen kein Identitätsverhältnis besteht. Nur weil „phänomenale Merkmale nicht zu den Verwendungskriterien neurobiologischer Begriffe gehören“ (Pauen, 2002a, S. 181), lässt sich an der Gültigkeit einer psychophysischen Identitätstheorie nicht ernsthaft zweifeln. Trotz einiger stichhaltiger Ideen hat Papineau das Argument der Erklärungslücke hiermit dennoch vollständig ausgeräumt. So wurde im Hinblick auf vor allem Farbempfindungen aufgezeigt, dass diese den Verteidigern funktionalistischer Reduktionen größte Schwierigkeiten bereiten, weil es im Fall von unterschiedlichen Farbeindrücken nicht

²⁷¹ Es ist gewissermaßen das Pendant zum bereits erwähnten *phänomenologischen Fehlschluss*, den Place (1956) als entscheidend für die Existenz des intuitiven Dualismus ansieht. Er besagt, dass wir dazu neigen, mentale Zustände auf unangemessene Weise zu verdinglichen (vgl. ebd., S. 48 bzw. Kap. 6.1.1 dieser Arbeit).

zwangsläufig einen Unterschied beim hiermit zusammenhängenden kausalen (Verhaltens-)Profil gibt (vgl. Kap. 6.5.4.1), womit der Ansatz der Funktionalisierung phänomenaler Eigenschaften zum Scheitern verurteilt ist bzw. nur eine partielle Gültigkeit haben kann. Der alleinige Verweis darauf, dass Identitäten keiner Erklärung bedürfen, erscheint an dieser Stelle unbefriedigend.

Diese explanatorische Lücke kann jedoch unter Rekurs auf die hier konzipierte, auf mereologisch-mechanistischen Prinzipien fußende Identitätstheorie zumindest minimiert werden. Zunächst kann nämlich festgestellt werden, dass die Gültigkeit der hier verteidigten „Theorie des Geistes“ nicht von der Etablierbarkeit kausaler Profile in Form von alltags-sprachlichen Verhaltensmustern abhängig ist. Denn sie basiert nicht auf der „Kausale-Rollen-Ersetzungsannahme“, bei der eine mentale Eigenschaft nur über ihre kausale Rolle auf der neuronalen Ebene bestimmt wird und die Existenz einer mentalen Eigenschaft ausschließlich darin besteht, ein neuronaler Realisierer zu sein, der das funktionale (Verhaltens-)Profil dieser Eigenschaft aufweist. Nach Maßgabe des hier verteidigten Ansatzes sind phänomenale Eigenschaften mit einer (potentiell unendlichen) Abfolge mikrostruktureller Konfigurationen auf den verschiedenen mereologisch verfassten Ebenen des diese Eigenschaft realisierenden Mechanismus identisch. Daher gilt notwendigerweise das Prinzip der mereologischen Supervenienz, das besagt, dass mit einer phänomenalen Differenz auf der psychischen Ebene zwangsläufig ein Unterschied auf allen niedrigeren Organisationsebenen und damit auch auf der neuronalen Ebene verbunden ist. Jede Veränderung im Hinblick auf die phänomenale Qualität von Farbeindrücken steht also notwendigerweise mit einer Veränderung auf allen niedrigeren Ebenen (und damit auch der neuronalen Ebene) des „hervorbringenden“ Mechanismus in Verbindung. In der Tat findet man bereits bei den herkömmlichen neuronalen

Wahrnehmungsmechanismen Unterschiede bei der Verarbeitung der verschiedenen farblichen Qualitäten (siehe Abb. 34).

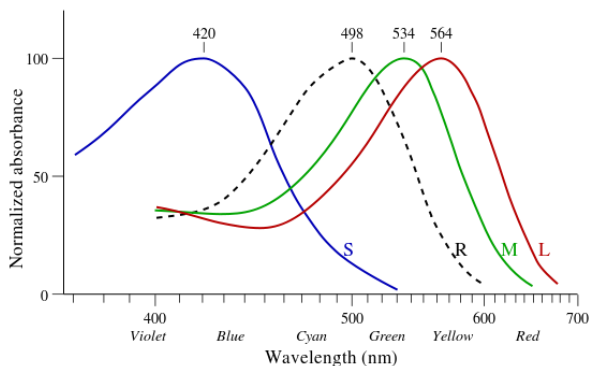


Abb. 34: Absorptionskurven des Lichts der verschiedenen Zapfentypen

So findet bereits auf der Verarbeitungsstufe der Lichtresorption durch die in den Augen befindlichen Zapfen eine Aufspaltung des Lichtspektrums statt, in dem es von drei verschiedenen Zapfentypen aufgenommen wird, die jeweils für andere Wellenlängenbereiche sensibel sind: L-Zapfen für längere Wellenlängen mit einem Absorptionsmaximum bei ca. 560 nm („Rot-Zapfen“, obwohl es genau genommen der gelblich-grüne Bereich ist); M-Zapfen für mittlere Wellenlängen mit einem Absorptionsmaximum von ca. 530 nm („Grün-Zapfen“) sowie S-Zapfen für kürzere Wellenlängen mit einem Absorptionsmaximum bei etwa 420 nm („Blau-Zapfen“). Die Absorptionskurven aller drei Zapfenarten sind verhältnismäßig breit und überlappen sich zudem stark. Deswegen liefert ein Zapfentyp allein keine eindeutige Information über den Farbeindruck, den eine Person aufweist. Dies wird erst durch eine *Verrechnung* von mindestens zwei verschiedenen Zapfentypen möglich.

Es findet also bereits auf der Ebene der Rezeptoren eine unterschiedliche Aufnahme verschiedener Wellenlängen statt, die offensichtlich von der spezifischen Beschaffenheit der einzelnen Zapfentypen abhängt. Auch in den weiteren Verarbeitungsstufen der nachgeschalteten retinalen Ganglienzellen, die über bestimmte rezeptive Felder verfügen, im Corpus geniculatum laterale (CGL) und schließlich der kortikalen Verarbeitung finden sich Unterschiede in der Verarbeitung der verschiedenen Wellenlängen (resp. Farbeindrücke). Um sich nicht in neurobiologischen Details zu verstricken, kann man sich auf die für den hiesigen Kontext relevante Tatsache beschränken und feststellen, dass die Informationen, die über die verschiedenen Zapfentypen aufgenommen werden, in verschiedenen Schichten des CGL und (okzipitalen) Kortex verarbeitet werden.

Es ist daher zu erwarten, dass sich jede Veränderung unserer phänomenalen Farbeindrücke in der physischen bzw. neuronalen Struktur ihren Niederschlag findet. Dies zeigt sich bei den verschiedenen Formen von Farbfehlsichtigkeit: Bei so genannten Rot-Blinden hat man festgestellt, dass sie tatsächlich über keine Rot-Rezeptoren verfügen, bei so genannten Grün-Blinden fehlen entsprechend die „Grün-Zapfen“. In beiden Fällen spricht man vom Phänomen der *Dichromasie*, da die Betroffenen lediglich über zwei Zapfentypen Wellenlängen des Lichts (resp. Farben) verarbeiten. Ebenso gibt es Personen, die lediglich Helligkeitsschattierungen wahrnehmen, weshalb man von Farbenblindheit oder *Monochromasie* spricht. Bei ihnen fehlen Zapfen gänzlich (vgl. Goldstein, 2008, S. 167).

Es mag daher zwar sein, dass es beim Vorhandensein unterschiedlicher Farbeindrücke nicht zwangsläufig einen detektierbaren Unterschied im Verhalten gibt, also einen solchen, der sich im kausalen Profil einer Person bzw. in einer Differenz zwischen den kausalen Profilen zweier Personen niederschlägt. Ein explanatorischer Ansatz, der sich mechanistischer Prinzipien bedient, ist aber nun keineswegs an die Etablierbarkeit von

Verhaltensprofilen phänomenaler Eigenschaften gebunden. Entgegen dem „klassischen“ Funktionalismus besteht hier keineswegs die Annahme, dass die konkrete physische Basis einer mentalen Eigenschaft mehr oder weniger irrelevant bzw. austauschbar ist, solange es eine irgendwie beschaffene physische Basis gibt, die diesen funktionalen Prozess realisiert. Funktion und (materielle) Struktur sind im Modell der agierenden Entitäten eines Mechanismus untrennbar miteinander verknüpft (vgl. Kap. 5.1.2.3 und 7.3). Alle Unterschiede in Bezug auf die agierenden Entitäten des Farbmechanismus müssen daher auf (u.a.) der neuronalen Ebene feststellbar sein (wenn hierbei zwei Personen verglichen werden, muss natürlich vorausgesetzt werden, dass man den Subtypisierungsansatz adäquat verfolgt hat und auf der korrekten Granularitätsstufe eine „physische Basis“ der verschiedenen *Farbtönen* ausfindig gemacht hat): Daher muss eine Rot-Empfindung eine divergierende physische (und somit auch neuronale) Basis haben als eine Grün-Empfindung. Mit „Basis“ muss natürlich nicht gemeint sein, dass sie zwangsläufig eine andere abgegrenzte Region ist. Natürlich kann es bestimmte überlappende Konfigurationen der agierenden Entitäten dieser Ebene geben. Entscheidend ist, dass es aber nicht ein und dieselbe Konfiguration sein kann, die einmal mit einer grünen und einmal mit einer roten Farbeempfindung verbunden ist. Wenn es dennoch möglich scheint anzunehmen, dass jemand bei exakt gleicher physischer Basis einen anderen Farbeindruck haben kann, geschieht dies nur unter Ausblendung der bestehenden neurobiologischen Kenntnisse über die relevanten Verarbeitungsmechanismen bzw. indem man Farbeindrücke als atomistische Zustände ansieht, ohne ihre vorausgehenden Verarbeitungswege und -stufen zu berücksichtigen. Letztendlich haut man hierbei in die gleiche Kerbe, wie bei der in der Philosophie teilweise immer noch geläufigen Annahme, dass die erlebte „Einheit des Bewusstseins“ eine Segregation auf den zugrunde

liegenden physischen Ebenen ausschließen würde. Aus der reinen Art des Erlebens mentaler Zustände sollte man nicht auf den „wahren Charakter“ des Psychischen oder dergleichen schließen.

Bleibt es nicht aber trotzdem rätselhaft, warum phänomenale Eigenschaften in Abhängigkeit von ihren physischen Grundlagen (mit denen sie koextensiv-identisch sind) *genau diese* phänomenale Qualität haben, die sie haben? So könnte man weiterhin kritisieren, dass das, was über die bekannten (neuronalen) Mechanismen der Wahrnehmung genau genommen aufgezeigt wird, nur der Umstand ist, dass es bestimmte Differenzen in der Verarbeitung bestimmter *Wellenlängen* des Lichts gibt. Dabei gelten zwischen diesen Wellenlängen und unserer phänomenalen Gegebenheitsweise dieser Wellenlängen (in unserer Welt und speziesspezifisch) zwar gesetzesartige Zusammenhänge, sodass wir Objekte, die aufgrund ihrer Oberflächenbeschaffenheit vor allem das Licht einer Wellenlänge von 530 nm reflektieren als „grün“ wahrnehmen. Die Verarbeitung bestimmter Wellenlängen des Lichts durchläuft dabei spezifische Verarbeitungswege und ist an spezifische strukturelle Gegebenheiten gebunden (allein die verschiedenen Zapfentypen weisen eine unterschiedliche Größe auf), die neurobiologisch voneinander unterschieden werden können, sodass es keinen einfachen „Austausch“ phänomenaler Qualitäten bei gleichbleibender Physis geben kann, wie es philosophische Gedankenexperimente suggerieren. Was aber fraglich bleibt, sind Aspekte wie diese: Warum stellt sich „Grün“ aber genauso für uns dar, wie es sich darstellt? Inwiefern geht es auf die physische Basis dieses Erlebnisses zurück, dass es genau diese phänomenale Qualität hat? Warum fühlen sich Schmerzen auf genau diese Art und Weise als schmerzhaft an und warum ist es gerade die Reizung von C-Fasern, die mit genau dieser Erlebnisqualität verbunden ist? Wäre nicht eine Welt denkbar, in der Wellenlängen von 530 nm (was in unserer Welt mit einem grünen Farb-

eindruck verbunden wäre) zwar durch den gleichen Zapfentyp verarbeitet würden, grundsätzlich aber so erlebt würden, wie wir in unserer Welt „Rot“ erleben? Könnte es nicht eine Welt geben, in der es zwar C-Faser-Reizungen gibt, diese aber mit einer anderen Erlebnisqualität verbunden sind als sie Schmerzen in unserer Welt hat? Einerseits liegt hier offensichtlich wiederum die bereits besprochene Kontingenzvermutung zwischen phänomenalen Zuständen und den mit ihnen identischen physischen Zuständen vor, deren Zustandekommen und Verfehltheit bereits hinlänglich besprochen wurde (vgl. Kap. 7.5.2.3). Andererseits bleibt es aber vielleicht wirklich ein Mysterium, warum in unserer Welt bzw. durch unseren Organismus genau diese phänomenalen Qualitäten hervorgebracht werden und keine anderen. Was also in der Tat schwierig zu erklären ist, ist, warum unser Organismus dazu übergegangen ist, eine Wellenlänge von 530 nm als phänomenales „Grün“ oder Gewebeverletzungen, die zu C-Faser-Reizungen (und bestimmten zentralnervösen Aktivierungsmustern) führen, als phänomenalen Schmerz zu interpretieren. Allerdings handelt es sich hierbei um kein explanatorisches Defizit, das ein spezielles Manko des hier entwickelten Identitätstheoretischen Ansatzes wäre. Auch jede andere Theorie des Geistes (inklusive dualistischer Positionen) hat hier keine weitergehende Erklärung anzubieten, sodass über den etwaigen Fortbestand der Existenz dieser Facette der Erklärungslückenproblematik kein Grund zum Zweifeln an der Gültigkeit des hier verteidigten Ansatzes besteht. Obgleich es sich hierbei um ein durchaus interessantes Phänomen handelt, stellt es außerdem nur eine Randerscheinung dar, die das Vorhaben der Formulierung einer tragfähigen Theorie zur begrifflich-konzeptuellen Fundierung der Psychologie nahezu unberührt lässt.

Vielleicht hat Kim (2005) nicht ganz unrecht, wenn er sinngemäß sagt, dass phänomenale Eigenschaften so etwas wie Platzhalter sind, über die

wir Ähnlichkeiten und Unterscheide in unserer Umwelt bzw. bezüglich der Repräsentation unseres eigenen Körpers darstellen (wie unterschiedliche Wellenlängen des Lichts, die von Objekten reflektiert werden oder Verletzungen des eigenen Gewebes). Über eine rein kognitive (kausal-funktionale) Repräsentation gäbe es nur eine strukturelle oder formale Abbildung unserer Umwelt bzw. unserer selbst, die als Bewusstseinserscheinung „leer“ bleiben müsste, womit diese Unterschiede offenbar gar nicht auf eine Art dargestellt werden könnten, mittels der wir geistig – gerade im Angesicht komplizierter Problemstellungen – operieren könnten. Da die interne und bewusste Repräsentation der Umwelt aber einen Überlebensvorteil bedeutet (sonst hätte sie sich nicht in diesem Ausmaß durchgesetzt), haben die höheren Organismen dieser Welt im Zuge der Evolution die Fähigkeit immer weiter verfeinert, internale und externale Differenzen auf bestimmte Weise innerlich darzustellen und so sukzessive den phänomenalen Reichtum geschaffen, den unsere Bewusstseinswelt heute hat. Dabei muss es offenbar von Vorteil sein, dass wir z.B. in Bezug auf die Wellenlänge, die von Blut reflektiert wird, den phänomenalen Eindruck „Rot“ haben, weil eine solche phänomenale Qualität aktivierend oder eben sogar alarmierend wirkt, was bei einer Interpretation in Form eines grünen Farbeindrucks wohl weniger der Fall wäre. Sicherlich gehören individuelle Erfahrungen dazu, die „Rot“ zu einer solchen Signalfarbe machen (Verletzungen mit Blutaustritt), dennoch ist eine phänomenale Darstellung als „Rot“ offensichtlich effektiver als andere mögliche Varianten (phänomenales Rot scheint ein anderes intrinsisches Potential als Signalfarbe zu besitzen als phänomenales Grün). Entsprechendes gilt für die anderen Farben und phänomenalen Qualitäten. Nimmt man die Grundannahmen der Evolutionstheorie ernst, muss daher davon ausgegangen werden, dass sich diejenigen „Farbinterpretationen“, über die wir in unserer phänomenalen Welt für gewöhnlich verfügen, aus gutem

Grund durchgesetzt haben, obgleich es sicherlich viele andere „Interpretationsweisen“ unserer Umwelt und unseres Körpers in unserer stammesgeschichtlichen Entwicklung gegeben haben wird (und in Form von – von uns als „Fehlwahrnehmungen“ klassifizierten – Phänomenen wie Dichromasie auch weiterhin gibt).

Vielleicht bleiben phänomenale Eigenschaften in gewisser Hinsicht auch zukünftig synchron emergente Eigenschaften, vielleicht wird im Fortgang der Forschung aber auch für die verbleibende explanatorische Lücke eine hinreichende Lösung gefunden. Jedenfalls sollte man vorsichtig damit sein, zu behaupten, es würde eine *prinzipielle* explanatorische Kluft bestehen, wie es die Vertreter des Erklärungslückenargumentes gerne tun. Denn im Laufe der Wissenschaftsgeschichte hat sich immer wieder gezeigt, dass solche als prinzipiell unerklärbar vermuteten Zusammenhänge im Fortgang der wissenschaftlichen Erkenntnis doch erklärt werden konnten. So war z.B. John Locke unter Verallgemeinerung des wissenschaftlichen Kenntnisstandes des 17. Jahrhunderts davon überzeugt, dass die Mechanismen, die die einzelnen Atome einer Substanz aneinander binden und dabei Veränderungen des Aggregatzustandes erlauben, nicht nur unbekannt seien, sondern sich prinzipiell dem menschlichen Verständnis entziehen würden (vgl. Locke, 1690/2000, Essay Buch II, Kap. 23, § 26; vgl. Pauen, 2001/ 2005, S. 201). Heute weiß man durchaus, wie dies zustande kommt. Die Existenz einer prinzipiellen Erklärungslücke würde jedenfalls wohl kein Naturwissenschaftler mehr akzeptieren.

Eine grundlegende Idee dafür, wie eine solche explanatorische Theorie des Geistes aussehen könnte, liefert Dörners (1998/2008) *Bauplan für eine Seele*, in dem er systematisch eine künstliche Erschaffung von internen Repräsentationen, Gefühlen und Bedürfnissen vorführt. Von einzelnen einfachen Systemzuständen zu immer komplexeren wird hier über kyber-

netische Regelkreisläufe und mechanistische Modelle ein anschauliches Gesamtsystem etabliert, das auch in Bezug auf das Dasein und Sosein phänomenaler Eigenschaften wichtige Erklärungsprinzipien beinhaltet. Da das Werk Dörners ein mehr oder weniger geschlossenes Ganzes darstellt und der Schwerpunkt meiner Arbeit eher die *ontologische* (begrifflich-konzeptuelle) und weniger die epistemisch-explanatorische Fundierung der Psychologie ist, möchte ich auf weitergehende Darstellung der hier implementierten Erklärungsprinzipien verzichten. Es würde einer relativ ausführlichen, voraussetzungsvollen Darstellung bedürfen, die hier nicht geleistet werden kann.

8. Zur Unabkömmlichkeit der Psychologie: Eine Entzauberung des (Neuro-)Reduktionismus

Wenn nun im letzten Kapitel ein Modell einer allgemeinen Architektur des Psychischen etabliert wurde, was bedeutet dies für die eingangs aufgegriffene Debatte um die Biologisierung der Psychologie und die zunehmend geschürte Angst vor einem Neuroreduktionismus?

Wie herausgestellt wurde, sind es zwei Befürchtungen, die besprochen werden müssen: Die erste Befürchtung ist *epistemischer* Natur und betrifft den Status der Psychologie als Wissenschaft und die Güte psychologischer Konstrukte, Modelle oder Theorien. Kann es sein, dass die Psychologie im Zuge der derzeitigen Verschiebungen zugunsten biologischer Ansätze bei der Erforschung mentaler Phänomene zunehmend obsolet wird? Werden vielleicht bald Biologen oder Mediziner die ursprünglichen Forschungsfragen der Psychologie viel besser, exakter etc., sprich: mit einer höheren wissenschaftlichen Güte beantworten können? Gibt es also tatsächlich ein *explanatorisches Primat* der Mikroebenen? Eine wichtige Intuition, die solche Befürchtungen nährt, liegt darin, dass den Beschreibungen und Erklärungen, die auf einer niedrigeren Ebene ansetzen (z.B. der der Biologie oder Chemie), vermeintlich ein höherer Status zukommt als solchen, die auf einer Makroebene ansetzen (z.B. Psychologie, Soziologie oder Pädagogik). Die zweite Befürchtung ist *ontologischer* Natur und betrifft den Gegenstand der Psychologie, der in der Folge des Siegeszugs neurowissenschaftlicher Forschung offenbar zur Disposition steht. Wenn das Psychische, wie aufgezeigt wurde, zwar eine eigene natürliche Organisationsebene konstituiert und daher gleichzeitig mit seinen physischen Grundlagen koextensiv-identisch sein muss, ist dies dann nicht gleichbedeutend damit, dass es auf ebendiese physischen Grundlagen *reduziert*

werden kann und welche Folgen hätte das? Kommt außerdem nicht den niedrigeren Ebenen, die das Psychische konstituieren, ein *ontologischer Vorrang* zu? Denn wenn man annimmt, dass das Psychische eine natürliche Organisationsebene konstituiert und damit offensichtlich auch den Naturgesetzen unterliegt, folgt dann nicht zwangsläufig, dass es – ob nun reduzierbar oder nicht – zumindest durch die Entitäten und Prozesse der Mikroebenen vollständig festgelegt bzw. *determiniert* wird? Sind es nicht die Mikrogesetze der Physik, auf die letzten Endes alles kausale Geschehen zurückzuführen ist und – hier schließt sich der Kreis zur epistemischen Facette – sind es damit nicht auch die Naturwissenschaften, die ein solches Geschehen am besten beschreiben und erklären können?

Um zu klären, ob diese Befürchtungen einen realistischen Gehalt haben, hilft nur die in der psycho-therapeutischen Praxis bewährte Konfrontation. Es müssen daher zunächst die begrifflichen Grundlagen freigelegt werden, die eine Reduktion psychologischer Theorien oder auch des Psychischen selbst ermöglichen würden. Von zentraler Bedeutung ist an dieser Stelle die Differenzierung zwischen eher harmlosen *reduktiven* und den eigentlich subversiven *reduktionistischen* Ansätzen. Daraufhin werden die in der Wissenschaftstheorie bestehenden reduktiven und reduktionistischen Modelle einer kritischen Prüfung unterzogen. Nach dem klassischen Modell der Theoriereduktion nach Ernest Nagel ist die Frage der Reduzierbarkeit im Wesentlichen eine Frage der logischen *Ableitbarkeit* der Zieltheorie von der Basistheorie sowie der *Verknüpfbarkeit* der beiden Theorien bzw. ihrer Begrifflichkeiten. Obgleich die grundsätzliche oder durchgängige Erfüllbarkeit dieser Bedingungen im Falle psychologischer versus neurobiologischer Theorien infrage gestellt werden kann, da in vielen Fällen von wissenschaftlichen Konzepten keine gute intertheoretische Passung vorliegt, wird herausgestellt, dass selbst im Zuge einer erfolgreichen Reduktion nach dem Schema Nagels keine (abschließende)

explanatorische Reduktion der Psychologie auf die Neurobiologie herbeigeführt würde. Eine Theoriereduktion dieser Qualität hätte keineswegs die befürchteten negativen Auswirkungen auf die Psychologie und ihre Bezugsgegenstände, sondern würde ganz im Gegenteil eher zu einer gewissen „Aufwertung“ ebendieser führen, da sich die Psychologie und ihre Entitäten, Eigenschaften und Prozesse so in das wissenschaftliche Gesamtsystem einpassen lassen und ihren – für viele (Natur-)Wissenschaftler immer noch gegebenen – „nebulösen“ Status verlieren würden. Ganz ähnlich ist der bereits im Zusammenhang der Realisierungstheorie besprochene Ansatz *reduktiver Erklärungen* zu bewerten.

Hochproblematisch sind hingegen die *reduktionistischen* Ansätze, die im Hinblick auf die schlechte intertheoretische Passung zwischen Psychologie und Neurobiologie schlussfolgern, dass neurobiologische Theorien psychologische Theorien zu *ersetzen* haben. Sie stehen im Mittelpunkt der Neuroreduktionismus-Debatte. Es kann allerdings systematisch nachgewiesen werden, dass eine solche Ersetzung psychologischer Theorien aus *prinzipiellen* Gründen scheitert. Auch die vollkommenste Neurowissenschaft der Zukunft wäre bei der Erforschung des Gehirns ohne die Führung durch das begrifflich-konzeptuelle System der Psychologie blind. Erst durch die Zuordnung von psychologisch zu beschreibenden Funktionen zu bestimmten Hirnbereichen bzw. deren Verhalten simulierenden neuronalen Netzwerkmodellen wird den indifferenten, zusammenhangslosen und ziellos verlaufenden neuronalen Geschehnissen überhaupt erst eine Bedeutung zuteil.

Vor dem Hintergrund der gewonnenen Erkenntnisse soll dann aufgezeigt werden, in welche Sackgasse die Reduktionismusdebatte verläuft, wenn man – wie leider sowohl in der Philosophie als auch Psychologie sehr verbreitet – in einer psychophysischen Identitätsannahme den Kern der Gefahr für die Eigenständigkeit der Psychologie sieht. Ebendiese

Überzeugung hat zu zwei problematischen Ansätzen der Verteidigung der „Autonomie“ der Psychologie geführt, die hier näher beleuchtet werden sollen: Einerseits ist es die Annahme, dass sich die Psychologie dem Funktionalismus zu verschreiben habe, weil (nur) dieser die – einer Identitätsannahme entgegenstehende – multiple Realisierbarkeit mentaler Eigenschaften gewährleiste. Andererseits ist es die Überzeugung, dass anstatt einer Identität eine vermeintlich harmlose Mikro-Determinierung des Psychischen angenommen werden könne, die dafür sorgt, dass das Psychische nicht aus dem Kreis wissenschaftlich respektabler Phänomene aussteigt, zugleich aber auch nicht – wie bei einer Identitätsannahme – auf seine physischen Grundlagen *reduziert* werden kann. Ironischerweise folgt aber nun gerade aus diesen Prämissen, wie gezeigt werden wird, dass die Ebene neurobiologischer Prozesse gegenüber der Ebene psychischer Prozesse ein *ontologisches* Primat und zugleich die Neurobiologie gegenüber der Psychologie ein *explanatorisches* Primat hätte.

Schließlich wird in diesem Kapitel systematisch nachgewiesen, dass ein solches Primat der Mikroebenen oder Mikroerklärungen unter den Voraussetzungen der hier entwickelten mereologisch-mechanistischen Architektur des Geistes scheitert und dies nicht trotz, sondern *aufgrund* der dabei getätigten psychophysischen Identitätsannahme. Dass es dennoch den Eindruck eines Primats der Mikroebenen bzw. von Mikroerklärungen geben mag, muss auf kognitive Tendenzen bzw. Verzerrungen zurückgeführt werden, die sich aus bestimmten (unreflektierten) inkonsistenten (ontologischen) Hintergrundannahmen oder einer Fehlinterpretation der Aussagekraft wissenschaftlicher Erkenntnisse ergeben.

8.1. Entwirrung der begrifflichen Grundlagen

Wenn sich Psychologen zur derzeitigen Entwicklung in ihrem Fach äußern, kommt dabei meist die große Besorgnis zum Ausdruck, dass die Psychologie zunehmend auf die Biologie *reduziert* werden könnte. Nicht zu guter Letzt haben sich viele Hirnforscher und biologisch orientierte Psychologen tatsächlich – zusätzlich verstärkt durch eine hohe mediale Präsenz – häufig dahingehend geäußert, dass im Zuge der Revolution der Neurowissenschaften durch die neuen bildgebenden Verfahren nun die eigentliche (sprich: naturwissenschaftliche) Erklärung für viele vormals nur notdürftig oder quasiwissenschaftlich von der Psychologie erklärten Phänomene gefunden werden könnten (vgl. Mausfeld, 2010). Ihre Intention scheint also zu sein, vormals z.B. kognitionswissenschaftliche Theorien auf neurobiologische Theorien zu *reduzieren*, sodass psychologische Erklärungen eines Phänomens zunehmend obsolet würden.

Dadurch, dass wir die neuronalen Grundlagen unseres Geistes in den Blick bekommen und nun – wie es häufig heißt – „dem Gehirn beim Denken zuschauen können“, scheint eine psychophysische Reduktion in der Hinsicht möglich zu sein, dass sich auch der eigentliche Gegenstand der wissenschaftlichen Disziplin zur Erforschung des Geistes nur auf die neuronalen oder überhaupt physischen Grundlagen des Psychischen erstreckt. Beziehen wir uns also, wenn wir uns bestimmte mentale Vorgänge oder Eigenschaften vergegenwärtigen, in Wirklichkeit auf bestimmte Vorgänge oder Eigenschaften unseres Gehirns? Ist das Psychische letztendlich nicht nur etwas „Neuronales“? Zusammengefasst: Ist nicht die folgerichtige Konsequenz aus der gegenwärtigen Entwicklung ein umfassender Neuroreduktionismus, der überflüssige vorwissenschaftliche Entitäten wie das Psychische oder überflüssige vorwissenschaftliche Disziplinen wie die Psychologie aus der Welt schafft?

Soweit eine Explikation dessen, was wohl von den meisten Psychologen unter dem Vorhaben einer Reduktion psychologischer Theorien oder des Psychischen selbst verstanden wird. Und wie bereits erwähnt, ist dies ein Bild, was von vielen Neurowissenschaftlern (inklusive einiger Biopsychologen; vgl. Birbaumer, 2003) tatsächlich befeuert wird. Es ist also nicht verwunderlich, dass alles, was mit dem Etikett „Reduktion“ versehen wird oder nur annäherungsweise versehen werden könnte, von vielen mit einer großen Skepsis betrachtet oder sogar von vornherein abgelehnt wird. Jeglicher Möglichkeit einer Reduktion scheint vorgebeugt und die Annahme einer irgendwie gearteten psychophysischen Identität scheint abgestritten werden zu müssen, wenn die Psychologie ihr Existenzrecht bzw. ihre „Autonomie“ aufrechterhalten möchte.

Mausfeld (2010) zeichnet ein Bild der gegenwärtigen Psychologie, in dem neuroreduktionistische Positionen (implizit) eine ungerechtfertigt starke Dominanz hätten. Bis in die 1980er Jahre hätte die Psychologie eine friedliche Koexistenz und eine Beziehung gegenseitiger Befruchtung zur Biologie gepflegt, die sich von der Genetik, über die Ethologie bis hin zur Evolutionsbiologie erstreckte. „Erst mit den *kognitiven Neurowissenschaften* verließ die Psychologie diesen offenen und am explanatorischen Bedarf orientierten Bezug zur Biologie, wie er seit je in der Theorieentwicklung der Naturwissenschaften üblich und fruchtbar ist, und band sich in einem historischen Regress in weiten Teilen wieder an die Philosophie²⁷², indem sie sich mit dem Neuroreduktionismus eine metaphysi-

²⁷² Hier erliegt Mausfeld meines Erachtens entweder einer äußerst missverständlichen Darstellung oder gar einer schwerwiegenden Fehleinschätzung. Wenn wir tatsächlich einmal annehmen, dass sich die gegenwärtige Psychologie einer überholten metaphysischen Grundlage verschrieben hat, dann liegt das gerade nicht daran, dass sie sich in einem historischen Regress wieder an die Philosophie gebunden hat, sondern ganz im Gegenteil das Band zur Philosophie aus einem grundlegenden Missverständnis wissenschaftlicher Emanzipation und einem „Minderwertigkeitskomplex“ gegenüber naturwissenschaftlicher For-

sche Position implizit oder explizit zu eigen machte, die naturwissenschaftlich unsinnig und philosophisch fragwürdig ist“ (S. 183; Hervorh. im Orig.). Das Ziel Mausfelds ist daher der Nachweis, dass der Reduktionismus eine veraltete, unhaltbare metaphysische Position darstellt, die sich aus verschiedenen (der Sache nach nicht gerechtfertigten) Gründen aber dennoch in der Psychologie seit den 1980er Jahren durchgesetzt hätte.

Auch wenn die von Mausfeld aufgeführten Argumente eine große Schlagkraft haben und in der Tat neuroreduktionistischen Bestrebungen das Wasser abzugraben scheinen, muss jedoch zunächst festgestellt werden, dass von Mausfeld – was symptomatisch für fast alle ist, die sich zum Thema der Gefahr eines Reduktionismus für die Psychologie äußern – nicht genügend geklärt wird, was *genau* unter den kritisierten Begriffen der Reduktion oder des Reduktionismus verstanden wird. Ist tatsächlich jedwede Form reduktiver Bestrebungen unsinnig und unhaltbar oder untergräbt sie die Daseinsberechtigung der Psychologie? Es bedarf zunächst also einer systematischeren Betrachtung des Reduktionsbegriffes. Wenn sich zeigen sollte, dass es einen inhaltlich konsistenten und prinzipiell möglichen reduktiven Ansatz geben sollte, muss darüber hinaus geklärt werden, ob dieser wirklich zu den oben skizzierten negativen Folgen für die Psychologie und ihren Gegenstand führen würde.

schungspraxis durchschnitten hat und damit in weiten Teilen blind geworden ist für die eigenen begrifflich-konzeptuellen Grundlagen. Es ist die Folge einer Vernachlässigung der für jedes andere Fach (inkl. der Naturwissenschaften) üblichen Reflexion solcher Grundlagen, die für zweifelhafte wissenschaftstheoretische Fundierungen anfällig macht und nicht andersherum.

8.1.1. Reduktion und Reduktionismus

Dass der Begriff „Reduktion“ sehr viele negative Konnotationen hat, dürfte auch damit zu tun haben, dass er in bestimmten Zusammenhängen auch als „Verringern“, „Beschneiden“ oder „Herabstufen“ aufgefasst wird, wenn z.B. von einer Reduktion der Angestelltenzahl eines Betriebes oder des monatlichen Gehalts die Rede ist.

Wenn man sich die ursprüngliche Wortbedeutung von „Reduktion“ vor Augen führt, kann man allerdings feststellen, dass der Begriff unter etymologischen Gesichtspunkten viel neutraler ist, als er heute oft verwendet oder verstanden wird. Der Begriff Reduktion entstammt dem lateinischen *reducere* und heißt nicht mehr als „zurückführen“. Gemeint ist hiermit zunächst nur, dass man z.B. die Beschaffenheit eines bestimmten Phänomens x dadurch verständlich machen kann, dass man diese mit der Beschaffenheit eines anderen Phänomens y in Verbindung bringt und x auf diese Weise auf y zurückführt. Es gibt etwas an y , das dafür verantwortlich ist, dass x so ist, wie es ist.

Insofern ist es nicht verwunderlich, dass Reduktionen in der Wissenschaftstheorie als eine entscheidende Triebfeder für den wissenschaftlichen Fortschritt bezeichnet werden:

The explanatory triumphs of the resulting theoretical integrations have richly rewarded the eagerness with which scientists have pursued reductive projects over the past one hundred fifty years. Reduction has probably been the single most effective research strategy in the history of modern science, engendering more precise accounts of the mechanisms (and their operations) underlying everything from magnetic forces to organisms' inheritance of traits to the visual perception of moving objects – to note but three examples from three different levels of analysis in science and three different collections of decades in the two centuries in question. Exploring reductive possibilities opens new avenues for sharing methodological, theoretical, and evidential resources. Successful reductions reliably generate productive programs of research at the analytical levels from which the candidate theories hail, squaring the lower level, mechanical details with the upper level phenomenal pat-

terns and refining our understanding of both in the bargain. (McCauley, 2007, S. 105f)

Wie Bischof (2008) sagt, könnte man zwar meinen, dass es die zentrale Aufgabe der Wissenschaft ist, Tatsachen festzustellen. Dies greift aber viel zu kurz. Denn da der Status eines Phänomens als Tatsache fraglich ist, weil sich vermeintliche Tatsachen im Fortgang des wissenschaftlichen Verfahrens immer wieder als falsch herausgestellt haben, besteht die Aufgabe eher darin „Informationen zu produzieren“ (S. 206). Da diese Zielvorgabe aber zu einer ungeordneten Informationsflut führen kann, besteht eine wesentliche Komponente wissenschaftlicher Arbeit in einer Theoriebildung, die Regelhaftigkeiten erkennt und mittels Abstraktion die produzierte Informationsmenge wieder *reduziert*. Da die Psychologie eine empirische Wissenschaft ist und auf diese Feststellung bekanntermaßen großen Wert legt, müsste auch sie darum bemüht sein, „die in ihrem Gegenstandsfeld enthaltene Informationsfülle nach Kräften zu reduzieren. Hier regt sich aber alsbald Widerstand: Der Begriff ‚Reduktion‘ hat bei uns keinen guten Klang – wer sich auf eine solche Aktivität einlässt, läuft Gefahr, als ‚Reduktionist‘ angeprangert zu werden“ (ebd., S. 206f).

Denn das Konzept der Reduktion (bzw. reduktiven Erklärungen) wird nur selten von dem des Reduktionismus klar abgegrenzt (vgl. z.B. Kim, 1998, S. 89). Während Reduktionen oder reduktive Erklärungen nützliche, unentbehrliche und in der Regel für die „reduzierten“ Phänomene harmlose Mittel des wissenschaftlichen Methodenkanons sind, wird unter dem Begriff *Reduktionismus* in der Wissenschaftstheorie in der Regel vielmehr ein programmatisches Vorgehen oder generelles Wissenschaftsprogramm verstanden, bei dem Reduktionen nicht nur ein Mittel, sondern nahezu der Zweck selbst sind, den die Wissenschaft zu verfolgen hat.

Reduktionisten haben sich meist dem Ideal einer Einheitswissenschaft verschrieben (vgl. Oppenheim & Putnam, 1958), eine Vorstellung, nach der alle Phänomene der Welt im Prinzip durch die grundlegendste Wissenschaft (was derzeit die Mikrophysik ist) zu erklären sind.

Des Weiteren dürfte die generelle Skepsis gegenüber Reduktionen und vor allem dem Reduktionismus in dem Umstand zu suchen sein, dass der Begriff der Reduktion neben der wissenschaftstheoretischen auch eine wissenschaftssoziologische Bedeutung hat, die als bedrohlich für die akademische Exklusivität höherstufiger Wissenschaften wie der Psychologie erlebt wird (vgl. Bischof, 2008, S. 207). Wie in dieser Arbeit aufgezeigt wurde (vgl. Kap. 5), siedeln sich Wissenschaften bzw. die von ihnen untersuchten Phänomene (annäherungsweise) auf unterschiedlichen Komplexitätsebenen an. Wenn nun Reduktionen eingefordert werden, müsste sich ein Wissenschaftler auf die Erkenntnisse basalerer Nachbardisziplinen einlassen und „prinzipiell zugeben, dass er auf solche Hilfe auch angewiesen ist. Die affektive Abwehr gegen Reduktionismus wird also gespeist vom unverzichtbaren Anspruch auf Kompetenz im eigenen Tätigkeitsfeld, wenn nicht gar von wissenschaftshierarchischem Dünkel“ (ebd.). Insofern verwundert es nicht, dass Sozialwissenschaftler oder Historiker „vor einer ‚Psychologisierung‘ und Psychologen vor einer ‚Biologisierung‘ warnen zu müssen glauben. Die einzige empirische Wissenschaft, die ohne Hemmungen reduktiv vorgehen kann, ist die Physik“ (ebd.).

8.1.2. Formen der Reduktion und des Reduktionismus

Eine wichtige begriffliche Unterscheidung besteht zwischen einer ontologischen, methodologischen und epistemischen *Reduktion* (vgl. z.B. Brigandt & Love, 2012), obgleich alle drei Reduktionsformen sehr stark zu-

sammenhängen. Jede kann als ein Mittel oder Charakteristikum des wissenschaftlichen Prozesses angesehen werden. Zu einem ontologischen, methodologischen oder epistemischen *Reduktionismus* werden sie erst, wenn man sie zum (alleinigen) Maßstab, programmatischen Ziel bzw. (Selbst-)Zweck des wissenschaftlichen Agierens erhebt.

8.1.2.1. Ontologische Reduktion und ontologischer Reduktionismus

Der Begriff der ontologischen Reduktion ist auf den ersten Blick relativ unspektakulär. Er wird in der Philosophie und Wissenschaftstheorie nur selten und wenn dann meist synonym mit dem Begriff der Identität verwendet, sodass üblicherweise auch dieselben Beispiele zur Erläuterung herangezogen werden. So bezeichnen der Abendstern und der Morgenstern, die früher für zwei verschiedene Sterne gehalten wurden, in Wirklichkeit denselben Himmelskörper und zwar keinen Stern, sondern den Planeten Venus. Insofern findet durch die Identitätsannahme eine Reduktion im Hinblick auf die Existenzannahme bzgl. realer Objekte statt, dass nämlich nicht drei verschiedene Himmelskörper (Morgenstern, Abendstern, Planet Venus) existieren, sondern nur einer. Ontologische Reduktionen behandeln die Frage, ob es sich bei bestimmten Phänomenen um verschiedene oder faktisch um ein und denselben Gegenstand handelt. So sind wissenschaftliche Entdeckungen häufig von dem heuristischen Prinzip geleitet, dass die Wirklichkeit nicht notwendigerweise ebenso viele verschiedene Phänomene enthält, „wie es wahre Beschreibungen mit verschiedener Bedeutung gibt. *Empirische Entdeckung ist oft Reduktion durch Identifikation*“ (Bieri, 1981/1997, S. 38; Hervorh. im Orig.).

In Bezug auf die Biologie wäre eine ontologische Reduktion z.B. die Annahme, dass jedes biologische System bzw. jeder biologische Organis-

mus durch nichts anderes als Moleküle und ihre Interaktionen konstituiert wird (vgl. Brigandt & Love, 2012). In Bezug auf die Psychologie wäre eine ontologische Reduktion z.B. die These, dass es sich bei mentalen Prozessen und Eigenschaften um Prozesse und Eigenschaften des Gehirns handelt. Insofern wird eine ontologische Reduktion für gewöhnlich mit einem Physikalismus bzw. Materialismus gleichgesetzt, bei dem entweder eine typenbezogene oder ereignisbezogene psychophysische Identitätsannahme unterstellt wird.

Ein mit der „physikalistischen Weltansicht“ verbundener wissenschaftstheoretischer Grundsatz ist in diesem Zusammenhang, dass auch die Bestandteile von Vorgängen, die von einer höherstufigen Theorie bzw. Wissenschaft beschrieben werden, eine physische Basis haben. Demzufolge muss es auch für alle Vorgänge, die z.B. von Historikern beschrieben werden, eine physische Grundlage geben:

[...] wir müssen also nicht auf die Physik zurückgreifen, um die Entstehung des Ersten Weltkriegs zu erklären, doch dem Physikalismus zufolge muß es für alle Veränderungen, auf die sich der Historiker berufen kann, korrespondierende Veränderungen auf der Ebene physikalisch beschreibbarer Konstituentien geben. Der Historiker kann also erwarten, daß er für die Flugbahn jeder einzelnen Kugel eine physikalische Beschreibung finden würde, wenn er die entsprechenden Ausgangsdaten hätte, er kann aber nicht erwarten, daß er höherstufige Ereignisse wie etwa den Verlauf einer Schlacht aus physikalischen Erklärungen rekonstruieren könnte. (Pauen, 1999, S. 44)

Jedes Ereignis, das von höherstufigen Wissenschaften wie der Biologie, Geschichtswissenschaft, Psychologie oder auch Soziologie beschrieben wird, muss also ein physisches Gegenstück aufweisen, wenn die Beschreibung eine natürliche (beobachterunabhängige) Größe darstellen soll. Dennoch ist mit solch einer ontologischen *Reduktion* keineswegs gemeint, dass man alle biologischen, historischen, psychologischen oder

soziologischen Beschreibungen auch in solche der Physik übersetzen könnte, geschweige denn, dass eine solche Übersetzung (wenn sie denn möglich wäre) die Forschungsfragen der höherstufigen Wissenschaften hinreichend beantworten könnte.

Dagegen beinhaltet ein ontologischer *Reduktionismus* eine These über die grundsätzliche Beschaffenheit der Welt: „It says that we only need to accept into our ontology some limited set of ‚basic‘ entities, properties, events, or processes. Usually this basic set is supposed to be provided by physics (strictly speaking) or the physical sciences (broadly construed)” (Polger, 2007, S. 52f). Demnach hat nicht nur jedes psychologische oder soziologische Ereignis eine physische *Basis*, vielmehr erschöpft sich die gesamte Wirklichkeit in denjenigen Phänomenen, die von den Naturwissenschaften bzw. der Physik beschrieben werden. Mikroreduktionisten gehen dabei nicht selten so weit, dass sie als die *eigentliche* Wirklichkeit wiederum nur die Entitäten und Prozesse der Basisebene der Physik, also die Elementarteilchen, anerkennen.

8.1.2.2. Methodologische Reduktion und methodologischer Reduktionismus

Eine methodologische *Reduktion* liegt vor, wenn man annimmt, dass sich bestimmte Erhebungen wie sie z.B. die Psychologie anwendet (z.B. Fragebogenverfahren) auch mittels Methoden der niederstufigeren Wissenschaften untersuchen lassen (also z.B. über bildgebende Verfahren). Ein methodologischer *Reduktionismus* beinhaltet die Überzeugung, dass wissenschaftlich zu untersuchende Gegenstände grundsätzlich am fruchtbarsten auf der niedrigsten verfügbaren Ebene untersucht werden können. Wenn man den von Bechtel und Richardson (1993) eingeführten (bzw. wiederentdeckten) mechanistischen Ansatz heranzieht und die Zer-

legbarkeit eines komplexen Systems in Ebenen immer kleinerer Teile bzw. agierender Entitäten anerkennt, dann wäre es z.B. die Überzeugung, dass man das Verhalten eines Organismus am besten auf der Ebene der biochemischen Komponenten seiner Zellen untersuchen kann (vgl. Brigandt & Love, 2012). In Bezug auf das zuvor herangezogene Beispiel aus den Geschichtswissenschaften würde ein methodologischer Reduktionist fordern, dass auch die Erklärung für die Auslösung des ersten Weltkrieges durch eine physikalische Beschreibung zu erfolgen hat (wie auch immer man sich diese vorzustellen hat). Es dürfte jedenfalls deutlich geworden sein, dass bei methodologischen Reduktionen die Grenze zwischen Reduktion und Reduktionismus in einem besonderen Maße fließend ist. Obgleich ein methodologischer Reduktionismus zumeist darauf zurückgeht, dass man bestimmte ontologische Reduktionsannahmen akzeptiert, besteht keine zwingende Verbindung zwischen beiden Reduktionsarten, wie Brigandt und Love (2012) mit Verweis auf Wimsatt (1980) treffend darlegen:

While methodological reductionism is often motivated by the presumption of ontological reduction, this procedural recommendation does not follow directly from it. In fact, unlike the idea of ontological reduction, methodological reductionism can be quite controversial. It has been argued that *exclusively* reductionistic research strategies can be systematically biased so as to overlook salient biological features, and that for certain questions a more fruitful methodology consists in integrating the discovery of molecular causes with the investigation of higher level features. (Hervorh. im Orig.)

Ganz ähnlich kritisiert auch Mausfeld (2003) völlig zu Recht die Abwegigkeit eines methodologischen Reduktionismus. Nur weil eine Methode eher mit einer niederstufigen, vermeintlich exakteren Wissenschaft verknüpft ist, heißt es noch lange nicht, dass das, was sie an Ergebnissen hervorbringt, tatsächlich auch *aussagekräftiger* im Hinblick auf die For-

schungsfragen ist, die die höherstufigen Wissenschaften zu beantworten haben:

Nun ist es vielleicht hilfreich, daran zu erinnern, dass eine psychologische Zugangsweise nicht bereits dadurch zu einer naturwissenschaftlichen wird, dass als abhängige Variablen Cortisolwerte, EEG-Ableitungen oder Darstellungen des lokalen Stoffwechselverbrauchs im Gehirn fungieren, statt beispielsweise Werte einer Ratingskala. Solange das psychologisch-theoretische Substrat, das in solche Untersuchungen eingeht, nicht über Klassifikationen und Korrelationen hinausgeht, wird man ebensowenig von naturwissenschaftlicher Grundlagenforschung sprechen wollen wie bei einem mit Anwendungsfragen befassten Lebensmittelchemiker, der einen Gaschromatographen für seine Arbeiten heranzieht. (S. 186)

Selbst aus der Einräumung einer (epistemischen) Reduktionsmöglichkeit der Psychologie auf die Neurobiologie würde kein methodologischer Reduktionismus folgen. Denn hieraus kann nichts darüber abgeleitet werden, wie wissenschaftliche Entdeckungen in der Psychologie gemacht werden können oder gemacht werden sollten (vgl. Melnyk, 2007, S. 46).

8.1.2.3. Epistemische Reduktion und epistemischer Reduktionismus

Wenn die Frage gestellt wird, ob die Psychologie auf die Neurobiologie reduziert werden kann, ist hiermit primär die *epistemische* Beziehung zwischen zwei wissenschaftlichen Disziplinen angesprochen. Es ist die Frage, ob der Erkenntniszugang, der mit einer bestimmten Disziplin verbunden ist, auf den Erkenntniszugang einer anderen Disziplin (bezüglich des gleichen Forschungsgegenstandes) zurückgeführt werden kann. Es ist die Ansicht, dass die Kenntnisse einer wissenschaftlichen Domäne (typischerweise solche über Prozesse der höheren Ebenen) auf Kenntnisse einer anderen wissenschaftlichen Domäne reduziert werden können (typischerweise solche der niedrigeren Ebenen).

Es können wiederum zwei zentrale Formen epistemischer Reduktionen unterschieden werden: 1. Modelle einer Theoriereduktion, die behaupten, dass man eine Theorie logisch von einer anderen (niederstufigeren) Theorie *ableiten* kann sowie 2. Modelle explanatorischer Reduktionen, die versuchen, Charakteristika der höheren Ebenen über den Rekurs auf die Charakteristika der niedrigeren Ebenen zu *erklären* (vgl. Brigandt & Love, 2012).

Vertreter eines epistemischen Reduktionismus gehen demgegenüber deutlich weiter und geben physikalischen Theorien grundsätzlich einen Vorzug gegenüber solchen aus dem Bereich höherstufiger Wissenschaften. In Bezug auf das Beispiel des Ausbruchs des ersten Weltkrieges würde dies bedeuten, dass die physikalischen Theorien, die (potentiell) die damaligen Ereignisse beschreiben könnten, die adäquateren Erklärungen ebendieses Ereignis wären (als von bestimmten Machtkonstellationen zwischen Staaten, dem auf Krieg gestimmten Zeitgeist und anderen Variablen auszugehen, wie sie die Geschichtswissenschaft heranziehen würde). In diesem Sinne wird auch der Psychologie ein bloßer *heuristischer* Wert zugestanden (vgl. z.B. Bickle, 1998, 2003). So sei sie lediglich aus bestimmten pragmatischen Gründen in vielen Fällen unentbehrlich, „weil die naturwissenschaftlichen Beschreibungen neuronaler Prozesse zu komplex würden. Dennoch wird hier unterstellt, daß erst die Naturwissenschaften, in diesem Falle vor allem die Neurobiologie, die *wirklichen* Gesetzmäßigkeiten zu erkennen vermögen, die die neuronalen und damit auch die kognitiven Prozesse beherrschen“ (Pauen, 1999, S. 45; Hervorh. im Orig.). Ein epistemischer *Neuroreduktionismus* behauptet insofern, dass die *eigentlichen* Erklärungen für psychische Gegebenheiten auf der Ebene von Eigenschaften der neuronalen Organisation des Gehirns ansetzen und daher von Neurophysiologie und Hirnforschung (an-

statt der Psychologie) aufgedeckt werden müssten (vgl. Gold & Stoljar, 1999).

Die Anerkennung einer bestimmten Form epistemischer Reduktion kann dabei durchaus durch die Akzeptanz einer bestimmten ontologischen Reduktion und eines methodologischen Reduktionismus *befördert* werden. Die Möglichkeit einer epistemischen Reduktion folgt allerdings keineswegs *automatisch* aus einer Verbindung von einer ontologischen und methodologischen Reduktion (vgl. Brigandt & Love, 2012). Natürlich kann eine ontologische Reduktion, die eine psychophysische Identitätsbehauptung offensichtlich mit sich bringt, wahr sein, ohne dass sich auch die entsprechenden psychologischen Theorien auf die der Neurobiologie (oder gar Physik) zurückführen lassen (vgl. z.B. Polger, 2004, S. 194). Denn obgleich es in der Tat wohl niemals eine rein physikalische Erklärung zum Ausbruch des ersten Weltkrieges geben wird, wird wohl kaum jemand daran zweifeln, dass sich die Geschehnisse der damaligen Zeit in einer physisch „realisierten“ Form ereignet haben. Ebenso muss es nicht zwangsläufig eine physikalische oder (neuro-)physiologische Erklärung des Psychischen geben, damit man dieses (auch) als ein physisches oder (neuro-)physiologisches Geschehen betrachten kann.

Allerdings wird in der Philosophie des Geistes häufig die Annahme vertreten, dass eine ontologische Reduktion nur als Fakt betrachtet werden kann, wenn diese auch epistemisch abgesichert werden kann (z.B. Carrier & Mittelstraß, 1989; Kim, 1989/2007, 1998). Hier wird eine „direkte Verbindung zwischen dem ontologischen und dem epistemologischen Status bestimmter Entitäten hergestellt: Wenn eine höherstufige Eigenschaft physisch ist, dann muß sie auch unter dieser Beschreibung von physikalischen oder besser: naturwissenschaftlichen Theorien zu *bestimmen* und zu *erklären* sein. Scheitern solche Versuche, dann *kann* es sich dieser Argumentation zufolge nicht um eine physische Eigenschaft handeln“

(Pauen, 1999, S. 46). Dabei wird allerdings der unterschiedliche Status von ontologischen Reduktionen (bzw. Identitätsannahmen) und epistemischen Reduktionen außer Acht gelassen. „Während es sich bei einer Identitätsbehauptung um eine *ontologische* Aussage über das Verhältnis von *Objekten* oder *Eigenschaften* handelt, beinhaltet eine Theorienreduktion eine *epistemische* Aussage über die Beziehung zweier *Theorien*. Das Scheitern einer [epistemischen; M.L.] Reduktion schließt jedoch nicht das Bestehen einer Identitätsbeziehung aus“ (ebd., S. 51; Hervorh. im Orig.). An anderer Stelle geht Pauen aber durchaus von einer wechselseitigen Abhängigkeit von einer ontologischen und epistemischen Reduktion aus. Wenn nämlich das Modell reduktiver Erklärungen mentaler Eigenschaften als Grundsatz zur Auflösung der Erklärungslücke annimmt, dass mentale Eigenschaften über Rückgriff auf neurophysiologische Theorien über Gehirnprozesse erklärt werden können, dann sei hiermit eine wichtige Voraussetzung zur Rechtfertigung von Identitätsbehauptungen erfüllt: „Wenn bestimmte mentale Eigenschaften aus dem Wissen über neuronale Prozesse erklärt werden können, dann wird damit gleichzeitig die *Notwendigkeit* der Identitätsbeziehung deutlich, die die Voraussetzung dieser Erklärung ist“ (ebd., 1999, S. 15; Hervorh. im Orig.). Nach Pauen müssen bei Vorliegen einer psychophysischen Identität Theorien über Gehirnprozesse auf Dauer als intuitiv plausible Erklärungen der korrespondierenden mentalen Zustände anerkannt werden. Andererseits würde seines Erachtens ein dauerhaftes Scheitern solcher Erklärungsversuche „letztlich die Frage aufwerfen, ob die behauptete Identitätsbeziehung denn wirklich existiert. Wenn Identität nämlich die Übertragbarkeit der zugehörigen Erklärungen impliziert, dann müsste ein solcher Fehlschlag zu dem Schluß führen, daß es sich hier doch um zwei unterschiedliche Entitäten handelt“ (ebd.).

8.2. Reduktion psychologischer auf (neuro-)biologische Theorien?

Obgleich sich die verschiedenen Reduktionsformen in der „Praxis“ nur selten vollkommen voneinander trennen lassen, kann man doch konstatieren, dass die Frage nach der Reduktion der (Theorien der) Psychologie auf die (Theorien der) Neurobiologie vor allem eine Frage nach einer epistemischen Relation ist. Es geht also primär um die Beziehung zweier Wissensarten oder Erkenntnisarten und weniger um die Beziehungen zwischen den Entitäten selbst, wie sie von den Wissenschaften untersucht werden. Allerdings wird sich im Folgenden auch zeigen, dass diese – in der Wissenschaftstheorie verbreitete – Annahme durchaus problematisch ist. Eine vollkommene Ausblendung der Ontologie oder ein Begriff der Wissenschaft als rein epistemisches Manöver führt schnell zu schwerwiegenden Fehlannahmen.

8.2.1. Die „klassische“ Theoriereduktion nach Ernest Nagel

Es war Ernest Nagel, der in seinem 1961 erschienenen Werk *The Structure of Science* das klassische Modell der Theoriereduktion formuliert und hiermit über viele Jahrzehnte die Debatte in der Wissenschaftstheorie und auch Philosophie des Geistes dominiert hat. Für die Möglichkeit einer solchen Reduktion hat er relativ strenge Kriterien aufgestellt: „a reduction is effected when the experimental laws of the secondary science (if it has an adequate theory, its theory as well) are shown to be the logical consequences of the theoretical assumptions (inclusive of the coordinating definitions) of the primary science“ (ebd., S. 352).

Eine so verstandene Reduktion beinhaltet also die Beziehung zwischen zwei wissenschaftlichen Theorien: eine sekundäre bzw. Zieltheorie und

eine primäre oder Basistheorie. Es sind nach Ansicht Nagels dabei zwei Bedingungen, die erfüllt sein müssen, damit eine Theoriereduktion erfolgen kann: Die erste Bedingung ist *Ableitbarkeit*: Demnach ist die Frage der Reduzierbarkeit im Wesentlichen eine Frage der logischen Ableitbarkeit der Zieltheorie von der Basistheorie. Nach Ansicht Nagels liegt nur dann eine erfolgreiche Reduktion einer wissenschaftlichen Theorie A auf eine wissenschaftliche Theorie B vor, wenn sich alle Gesetze von A aus den Gesetzen von B ableiten lassen. Die zweite Bedingung ist die *Verknüpfbarkeit* der beiden Theorien bzw. ihrer Begrifflichkeiten (vgl. ebd., S. 354), die besonders zum Tragen kommt, wenn man Theorien vorliegen hat, deren Begrifflichkeiten sich stark unterscheiden bzw. zumindest nicht hinreichend überlappen (vgl. Schouten & Looren de Jong, 2007, S. 4). Ein Beispiel wäre die Reduktion unserer (Alltags-)Theorien über Wasser auf die chemische Theorie von H₂O. Hierzu müsste man nun alle alltagstheoretischen Charakteristika von Wasser (kocht bei 100 °C...) aus den chemischen Gesetzen zu H₂O ableiten können. Da aber der Begriff des Kochens in der chemischen Theorie zu H₂O gar nicht vorkommt, bedarf es unter solchen Umständen so genannter *Brückenprinzipien* (*bridge laws*), die das in den beiden Theorien verwendete Vokabular miteinander verbinden. In dem geschilderten Fall müsste daher das Alltagskonzept des Kochens beispielsweise mit bestimmten molekularen Bewegungen *verknüpft* werden. „Establishing such connections between scientific theories motivates *programs* for unifying science via ‘microreductions’, in which the entities the reducing theory discusses constitute the components of the entities that the reduced theory endorses” (McCauley, 2007, S. 108; Hervorh. im Orig.)

Wenn beide Bedingungen erfüllt sind, ist die alte Theorie insofern ein Bestandteil der neuen Theorie, als dass sie nunmehr als Spezialfall der neuen Theorie fungiert, der unter ganz bestimmten Umständen zum

Tragen kommt. Ein Beispiel wäre die Reduktion der Keplerschen Gesetze zu den Umlaufbahnen der Planeten auf die Newtonschen Gesetze der Bewegung und Gravitation, wobei sie nunmehr als Spezialfälle der Newtonschen Gesetze fungieren (vgl. Schouten & Looren de Jong, 2007, S. 4f). Das Beispiel, das das Konzept der Reduktion, die Nagel vorschwebte, aber wohl am besten illustriert, ist die logische Ableitbarkeit des Galileischen Fallgesetzes aus der Newtonschen Mechanik bzw. dem Newtonschen Gravitationsgesetz (vgl. Beckermann, 1999, S. 109f):

(G) Galileisches Fallgesetz:
$$s = \frac{1}{2} \cdot g \cdot t^2 \quad (g = 9,81 \cdot \frac{m}{s^2})$$

(N1) Zweites Gesetz der Newtonschen Mechanik

$$F = m \cdot a \text{ (Kraft = Masse} \cdot \text{Beschleunigung)}$$

(N2) Newtonsches Gravitationsgesetz

$$F = f \cdot \frac{m_1 \cdot m_2}{r^2} \quad (f = 6,67 \cdot 10^{-11} \text{ Nm}^2\text{kg}^{-2})$$

Die Ableitung sieht dann wie folgt aus:

(1) Erdmasse: $m_E = 5,97 \cdot 10^{24} \text{ kg}$

(2) Erdradius: $r = 6.370.000 \text{ m}$

Dabei gilt, dass sich aus (1) und (2) aufgrund von (N2) die Kraft ergibt, die auf einen Körper mit der Masse m_1 an der Erdoberfläche wirkt:

(3)
$$F = \frac{6,67}{10^{11}} \cdot \frac{m_1 \cdot 5,97 \cdot 10^{24}}{6.370.000^2} = 9,81 \cdot m_1$$

Aus dieser Kraft resultiert nun nach dem Grundsatz N1 die Beschleunigung eines Körpers mit der Masse m_1 an der Erdoberfläche:

(4)
$$a = \frac{F}{m_1} = 9,81 \frac{m}{s^2}$$

Aus der so verstandenen Beschleunigung lässt sich nun die Geschwindigkeit berechnen, die dieser Körper nach t Sekunden erreicht (bei einer Anfangsgeschwindigkeit 0):

$$(5) \quad v = \int 9,81 \frac{\text{m}}{\text{s}^2} dt = 9,81 \cdot t \frac{\text{m}}{\text{s}}$$

Aus dieser Bedingung ergibt sich wiederum die Strecke s , die der Körper nach t Sekunden zurückgelegt hat, womit zugleich eine erfolgreiche Ableitung von (G) aus (N₁) und (N₂) evident zu sein scheint:

$$(6) \quad s = \int 9,81 t \frac{\text{m}}{\text{s}} dt = \frac{1}{2} \cdot 9,81 \cdot t^2 \text{m}$$

Somit kann das Galileische Fallgesetz offenbar als ein Spezialfall der Newtonschen Gesetze angesehen werden. Obgleich diese Ableitung sicherlich nicht einer gewissen Eleganz entbehrt, ergeben sich dennoch zumindest zwei Probleme: Erstens ist genau genommen keine exakte Ableitung vorgenommen worden, sondern nur eine *näherungsweise* Lösung. So hat bereits Popper (1957) darauf hingewiesen, dass zwischen den beiden Theorien Inkompatibilitäten und Diskrepanzen bestehen, sodass „from a logical point of view, Newton’s theory, strictly speaking, contradicts both Galileo’s and Kepler’s“ (S. 29f; zit. n. Schouten & Looren de Jong, 2007, S. 8). So würde z.B. ein geworfener Stein nach Galileo einer parabolischen Bahn folgen, während Galileo von einer elliptischen Flugbahn ausgeht. Sie unterscheidet sich allerdings kaum von einer parabolischen Flugbahn, wenn man lediglich geringe Distanzen betrachtet. Dennoch ist die Ableitbarkeit des Galileischen Fallgesetzes aus den Newtonschen Gesetzen nur annäherungsweise gegeben:

Denn in der angeführten Form gelingt sie ja nur, wenn man unberücksichtigt lässt, daß sich die Gravitationskraft, die die Erde auf den fallenden Körper ausübt, wäh-

rend des Falls – wenn auch nur geringfügig – verändert. Streng genommen lässt sich aus der Newtonschen Mechanik und dem Newtonschen Gravitationsgesetz also nur ein Gesetz ableiten, das mit dem Galileischen Fallgesetz *fast* identisch ist und das für fast alle praktischen Zwecke mit Galileis Gesetz gleichgesetzt werden kann. (Beckermann, 1999, S. 110; Hervorh. im Orig.)

Selbst wenn man also eng aneinandergrenzende Naturwissenschaften betrachtet bzw. sogar die Beziehung zwischen Theorien innerhalb der Physik, bleibt es unklar, ob es – streng genommen – in der Wissenschaftsgeschichte überhaupt erfolgreiche Fälle einer strengen Ableitung der Gesetze bzw. Theorien nach dem Modell nach Nagel gegeben hat (vgl. Wimsatt, 1979; Hooker, 1981; McCauley, 1981, 2007; Suppes, 1993). Wie Kim (1998) aufzeigt, kann der Nagelsche Ansatz als eine Erweiterung des Hempel-Oppenheimerschen deduktiv-nomologischen Modells wissenschaftlicher Erklärungen auf intertheoretische Kontexte aufgefasst werden²⁷³ (vgl. ebd., S. 26). Allerdings konstatiert Carl Gustav Hempel (1969) interessanterweise selbst, das Nagelsche Reduktionsmodell sei eine „untenable oversimplification which has no strict application in science and which, moreover, conceals some highly important aspects of the relationship to be analyzed“ (S. 197; zit n. Schouten & Looren de Jong, 2007, S. 8).

Hiermit ist zugleich die Kritik an der zweiten Bedingung, der Verknüpfbarkeit der beiden Theorien über Brückenprinzipien angesprochen. Sie ist zugleich der Punkt, der im Zusammenhang der Frage nach der Mög-

273 „Just as Hempelian explanation consists in the derivation of the statement describing the phenomenon to be explained from laws taken together with auxiliary premises describing relevant initial conditions, Nagelian reduction is accomplished in the derivation of the target theory from the base theory taken in conjunction with bridge laws as auxiliary premises. It is therefore more than a little the D-N model of explanation has had few committed adherents for over three decades, Nagel 's derivational model of reduction is still serving as the dominant standard in discussions of reduction and reductionism“ (Kim, 1998, S. 26).

lichkeit der Anwendung des Nagelschen Reduktionsmodells auf das Verhältnis von Psychologie und Neurobiologie vor allem diskutiert wurde. Obwohl die Annahme solcher Brückenprinzipien eine entscheidende Voraussetzung für die Durchführbarkeit des Nagelschen Modells darstellt, bleibt ihr Status bei Nagel selbst im Vagen. Nach Nagel (1961) selbst sind sie universell gültige bikonditionale („wenn, und nur wenn x dann y“) oder vielleicht sogar nur konditionale („wenn x dann y“) Beziehungen zwischen den Termen der beiden Theorien (vgl. ebd., S. 355f). Dies läuft auf eine bloße Korrelation hinaus und wird von den meisten Wissenschaftstheoretikern als eine zu schwache Anforderung an die Beziehung abgelehnt (vgl. z.B. Causey, 1972; Hooker, 1981). Es bedarf mehr als rein korrelativer Beziehungen, weil – selbst wenn man eine Ableitbarkeit der höherstufigen Theorie aus der Basistheorie als gegeben ansieht – eine Erklärung für die Existenz der postulierten Korrelation abgegeben werden muss (vgl. Schouten & Looren de Jong, 2007, S. 5).

Obleich Nagel auch und vielleicht sogar vor allem bloße (bi-)konditionale Beziehungen zwischen den Termen der beiden Theorien favorisiert hat, um irgendwelchen ontologischen Vereinnahmungen aus dem Weg zu gehen, stellt sich die Frage, ob die Einnahme eines solchen ontologisch „neutralen“ Standpunktes unter seinen Voraussetzungen überhaupt *möglich* ist (vgl. Schouten & Looren de Jong, 2007, S. 6). In der Philosophie des Geistes und in weiten Teilen der Wissenschaftstheorie hat sich jedenfalls zunehmend die Ansicht durchgesetzt, dass die Brückenprinzipien als Identitätsbeziehungen aufgefasst werden müssen, wenn das Nagelsche Reduktionsmodell tatsächlich greifen können soll (ebd.; Schaffner, 1967; Fodor, 1974). Insofern gibt es hier offenbar eine notwendige Verschränkung der epistemischen und ontologischen Reduktionsform. Durch diese Art der Verknüpfung wird also nunmehr nicht nur eine bloße Relation der Konzepte zweier Theorien behauptet, sondern eine

(Identitäts-)Beziehung zwischen den Entitäten und Eigenschaften, die von den Theorien postuliert werden. Daher besteht eine (von Nagel vielleicht ungewollte) Verbindung zum mikroreduktionistischen Ansatz von Oppenheim und Putnam (1958), die in ihrem Aufsatz *The Unity of Science as a Working Hypothesis* eine Vereinheitlichung der Wissenschaften als eine zentrale Zielsetzung der Wissenschaftsentwicklung ansehen. Auf der Basis ihrer – hier bereits in einem anderen Zusammenhang thematisierten (vgl. Kap. 4.1.1) – Annahme einer mereologischen Strukturgliederung der Welt, die von sozialen Gruppen, über Lebewesen, Zellen, Molekülen, Atomen bis hin zu Elementarteilchen reicht, soll nach Oppenheim und Putnam eine Reduktion *aller* Einzelwissenschaften auf *eine einzige* grundlegende Wissenschaft erreicht werden (was wohl nur der Physik zufallen kann). Durch diese Einheitswissenschaft soll nach Ansicht der Autoren dann eine umfassende und einheitliche Sicht auf die Wirklichkeit ermöglicht werden. Gemäß der Teil-Ganzes-Beziehungen würden die Eigenschaften der höherstufigen Entitäten nämlich auf die Eigenschaften und Konfigurationen der Basisebene zurückgehen, sodass der wissenschaftliche Anspruch sein müsste, letzten Endes jedes wissenschaftlich zu erklärende Phänomen hinreichend mikrophysikalisch erklären zu können (allerdings räumen sie durchaus ein, dass eine solche Mikroreduktion aufgrund der Unzulänglichkeit des theoretischen Systems der reduzierenden Wissenschaft nicht immer durchgeführt werden kann, weil der niederstufigen Wissenschaft bestimmte Begriffe zur hinreichenden Beschreibung (der Eigenschaften) eines Objektes fehlen würden, sodass es sich hierbei eher um ein Ideal oder eben eine *Arbeitshypothese* handelt; vgl. ebd., S. 8, 13).

8.2.2. Chancen und Grenzen der Nagel-Reduktion psychologischer Theorien

Die Frage, die es nun zu beantworten gilt, ist, ob unter den Voraussetzungen, wie sie Nagel selbst an eine Theoriereduktion bindet, irgendeine Gefahr für die „Autonomie“ bzw. den wissenschaftlichen Status der Psychologie ausgeht? Wird sie also in ontologischer oder explanatorischer Hinsicht bei einer unterstellten erfolgreichen Theoriereduktion nach dem Schema Nagels wirklich obsolet?

Wie bereits erläutert, kann nach Nagel ein Gesetz oder ein Konzept der reduzierten Theorie, wenn es nicht direkt über die bikonditionalen Brückengesetze aus der Basistheorie ableitbar ist, mit Hilfe der Brückengesetze in das Vokabular der Basistheorie umgeschrieben werden und als zusätzliches Gesetz der Basistheorie aufgefasst werden. Es ist der oben erwähnte Sonderfall der Basistheorie, auf die ein Gesetz oder ein Begriff reduziert wurde. Im Falle rein bikonditionaler oder gar bloßer konditionaler Beziehungen, bei denen die Begriffe der beiden Theorien nur in eine bestimmte (kontingente) Abhängigkeitsbeziehung gesetzt, aber nicht miteinander identifiziert werden, wird letztendlich gar keine begriffliche oder ontologische Vereinfachung vorgenommen. Bei Einbezug von Brückenprinzipien ist das Nagelsche Modell also keineswegs durch eine *ontologische Parsimonität* („Sparsamkeit“) gekennzeichnet (vgl. Schouten & Looren de Jong, 2007, S. 6). Denn die Reduktionsbasis wird durch dieses Prinzip essentiell in der mit ihr verknüpften „Weltanschauung“ (neue Prädikate/Begriffe) und Ontologie (neue Eigenschaften) erweitert, indem die Brückengesetze zu den reduktiven Ressourcen als Hilfsprämissen hinzugefügt werden (vgl. Kim, 2005, S. 99). Um tatsächlich eine Reduktion mit Konsequenzen für unsere ontologischen Annahmen über die Welt vornehmen zu können, müssten die reduzierten Gesetze *ausschließlich* mittels der in die Basistheorie verfügbaren Erklä-

rungsressourcen abgeleitet werden können, was Kim am Beispiel der Korrelation von Schmerzen und C-Faser-Stimulationen (Cfs) erläutert: „On a Nagel reduction of pain theory to neurophysiology, then, the pain-Cfs bridge law is an unexplained premise of the derivation of pain laws from laws of neurophysiology. This means that the pain-Cfs correlation law is now one of the axioms of the expanded base theory, and that this new base theory no longer is a pure neurophysiological theory – its primitive predicates include ‘pain’, and among its axioms is a law about pain” (vgl. ebd., S. 100). Letztendlich sind damit selbst alle denkbaren dualistischen Leib-Seele-Theorien mit einer solchen „Reduktion“ der Psychologie auf die Neurobiologie vereinbar (vgl. ebd., 1998, S. 97). Von einer Art Auflösung psychologischer Theorien oder Einverleibung durch die Neurobiologie kann also gar keine Rede sein.

Ebenso wird auch im Zuge einer erfolgreichen Reduktion nach dem Schema Nagels keine (abschließende) *explanatorische* Reduktion der Psychologie auf die Neurobiologie herbeigeführt. So wird von Kim (1998, 2005) bemängelt, dass die Nagel-Reduktion keinen Beitrag zu dem von ihm favorisierten Ansatz der *reduktiven Erklärung* des Psychischen leisten würde. Wenn also eine Erklärungslücke (vgl. Kap. 6.1.4.4) z.B. darin bestehen mag, dass es unklar bleibt, warum Schmerzen mit C-Fasern und nicht A-Fasern korreliert sind, dann müsste eine explanatorisch gehaltvolle Reduktion nachvollziehbar machen, wie bestimmte Phänomene aus grundlegenden Phänomenen hervorgehen. Unter diesen Voraussetzungen gilt nun aber, dass die Nagelsche Reduktion der Psychologie, die mit Brückengesetzen als unerklärten Hilfsprämissen arbeitet, keinen Beitrag hierzu leistet (vgl. Kim, 1998, S. 95f). Stattdessen bräuchte man nach Kim eine Erklärung, warum es diese Körper-Geist-Korrelationen gibt. „Nagel reduction of psychology is like taking mind-body supervenience as an unexplained brute fact“ (ebd., S. 96).

By adding the bridge laws to the reductive resources as auxiliary premises, *Nagel reduction essentially extends the reduction base*. If we take reduction to be an explanatory process which yields an explanation of the laws and phenomena being reduced on the basis of the laws of the base theory, Nagel reduction fails to generate such explanations. For, to do so, the reductive derivation must derive the laws being reduced *solely from the explanatory resources available in the base domain*. (Kim, 2005, S. 99f; Hervorh. im Orig.)

Da die Brückenprinzipien beanspruchen, universell gültige Verbindungen zwischen den Termen der beiden Theorien herzustellen, ist – im Angesicht der aufgezeigten Schwierigkeiten von bloßen konditionalen Beziehungen – offensichtlich eine *Identität auf Typen-Niveau* impliziert. Für eine Körper-Geist-Reduktion setzt das Nagelsche Modell somit voraus, dass jede mentale Eigenschaft mit einer nomologisch koextensiven physischen Eigenschaft versehen ist und das über alle Spezies und strukturellen Typen hinweg (vgl. Kim (1998, S. 26), was im Angesicht des Argumentes der multiplen Realisierbarkeit unplausibel erscheint. Wie aufgezeigt wurde, ist eine Integration kompositorischer Variabilität zwar durchaus möglich, allerdings nur unter den Voraussetzungen der Art-spezifität und Granularitätsparität (vgl. Kap. 7.5.2.2). Wenn Identitätsannahmen auf Typen-Niveau gültig sein sollen, müssen sie sowohl mit Bezug auf eine bestimmte Spezies als auch auf einem Niveau formuliert werden, das keine Differenz in der Körnung zwischen den (funktional oder mechanistisch konzipierten) mentalen Typen und neurobiologischen Typen beinhaltet. Um die Möglichkeit zur Anwendung des Nagelschen Ansatzes auf die Psychologie aufrechterhalten zu können, muss der Anspruch an universell gültigen Verbindungen also zwangsläufig eingeschränkt werden, da es sich hierbei um lediglich *lokale Reduktionen im Sinne spezies- bzw. strukturspezifischer Brückengesetze* handeln kann (vgl. Kim, 1998, S. 93).

Autoren wie Esfeld und Sachse (2007) bzw. Soom, Sachse und Esfeld (2010) sind nicht nur davon überzeugt, dass so eine befriedigende Antwort *prinzipieller Natur* auf das Argument der multiplen Realisierbarkeit gegeben werden kann, sondern dass die klassische (Theorie-)Reduktion der Psychologie auf die Neurowissenschaften mittels des Nagelschen Ansatzes so tatsächlich *durchführbar* würde, da über die vorgeschlagene Subtypisierung die Formulierung von bikonditionalen Beziehungen zwischen psychologischen und neurobiologischen Ausdrücken (in Form von typentheoretischen Identitätsaussagen) ermöglicht würde. Da mentale Typen Abstraktionen oder Idealisierungen echter mentaler Subtypen sind, müsse der Subtypisierungsansatz solange verfolgt werden, bis an irgendeinem Punkt keine weitere multiple Realisierbarkeit bzw. kompositorische Variabilität mehr auftritt (vgl. Soom, Sachse & Esfeld, 2010, S. 12). Die Konsequenz ist aber, dass es sich im Zuge dieser Prozedur in vielen Fällen um äußerst feinkörnige mentale Typen handeln würde. In den wenigsten Fällen dürften sie bereits existierende Begriffe oder Elemente bestehender psychologischer Theorien ausmachen, sondern sind vielmehr alltagssprachlich als auch in der wissenschaftlichen Psychologie unbekannte Größen, die vielleicht von einer zukünftigen Psychologie aufgegriffen werden könnten, vielleicht aber auch nicht. Wenn man z.B. davon ausgeht, dass es gemäß den neurobiologischen Parametern verschiedene Subtypen von Zahnschmerzen gibt (wie z.B. Zahnschmerz der Kategorie k_{xg1} vs. Zahnschmerz der Kategorie k_{xg2} usw.), bleibt es fragwürdig, ob wir überhaupt in der Lage sind, diese in unserer Selbstwahrnehmung voneinander abzugrenzen, obgleich es eine entsprechende neuronale Differenz geben mag, die über hochsensible wissenschaftliche Methoden wie fMRI oder PET erfasst werden können. Wenn ein solcher Typ auf psychischer Seite aber nun gar nicht oder kaum detektierbar ist bzw. zwei ähnliche Typen nicht voneinander unterschieden werden

können (obgleich die Differenz faktisch gegeben sein mag), stellt sich die Frage, ob wir jemals einen solchen mentalen Typ als *psychologisches* Konstrukt einführen können werden bzw. ob es überhaupt sinnvoll wäre, einen solchen Typ als *mentalen* Typ einführen zu wollen.

Wenn man es also auch *prinzipiell* für möglich erachtet, dass bestimmte Theorien der Psychologie auf bestimmte Theorien der Neurobiologie reduziert werden könnten, sagt dies deshalb noch nichts über eine *faktische* Reduktionsmöglichkeit der *aktuell gegebenen* Theorien der Psychologie aus. Wäre eine solche Reduktion nun aber auch faktisch möglich, würde auch dies keineswegs zu einer Eliminierung psychologischer Theorien führen, wie es von den Gegnern jeglicher reduktiver Bestrebungen unterstellt wird. Um dies zu verdeutlichen, zunächst zurück zu einem Beispiel aus den Naturwissenschaften: Einer der wenigen Fälle, in denen man für gewöhnlich von einer weitestgehend erfolgreichen Theoriereduktion spricht, ist die der Quantenchemie auf die Quantenmechanik (vgl. Suppes, 1993; Wünsch, 2000). Trotz dieser Reduktion ist es hierdurch keineswegs zu einer Eliminierung der quantenchemischen Theoriebildung und Forschung gekommen. Dies ist aus methodologischen und erkenntnistheoretischen Gründen unmöglich. Denn obgleich die Quantenchemie *vom Grundsatz her* tatsächlich auf die Quantenmechanik zurückgeführt werden kann, gilt es mittlerweile als unmöglich, die *Forschungsfragen und -probleme* der Quantenchemie durch Anwendung der grundlegenderen Gesetze der Quantenmechanik zu lösen: „The combinatorial explosion is so drastic and so overwhelming that theoretical arguments can be given that not only now but also in the future it will be impossible by direct computation to reduce the problems of quantum chemistry to problems of ordinary quantum mechanics“ (Suppes, 1993, S. 46).

Auf sehr anschauliche Weise erläutert auch Gerald Wünsch (2000) diesen Zusammenhang:

Man muss zumindest wissen, was Chemie ist, wo man hin will, ehe man eine solche Entwicklung zielgerichtet einleiten kann. Die Lösung der Schrödinger-Gleichung macht bereits für ein Zwei-Elektronen-System Schwierigkeiten. Die Folge ist, dass allein auf der Grundlage der Physik, mit rein quantenmechanischen Mitteln, kein den Chemiker befriedigendes Verständnis zu erzielen ist für das, was man eine chemische Bindung nennt. Quanten-Chemie unterscheidet sich von Quanten-Physik gerade dadurch, dass empirisch gewonnenes chemisches Vorwissen zu Grunde gelegt wird. Unter dergestalt stofflich definierten Randbedingungen macht man eine Abfolge geeigneter Näherungen, bis man zu einem befriedigenden Ergebnis kommt. [...] Man kann also bei geeigneter Kombination beider Fächer chemisch interessante Phänomene vom quantenmechanischen Standpunkt her verständlich und beschreibbar machen. Im Ergebnis hat man einen wichtigen Beitrag zur Chemie geleistet, man hat sie aber nicht aus der Physik deduziert, sondern man hat originär chemische Fragen mit Konzepten der Physik behandelt. Nach den wissenschaftstheoretischen Kriterien liegt keine starke Reduktion vor. Erfreulicher Weise hat man damit aber schwache Reduktion geleistet: Man hat gezeigt, dass es in den beiden Fächern anerkannt wichtige Strukturen gibt, die miteinander logisch verträglich sind. Man kann sich also nicht nur in der Praxis, sondern auch in der Theorie gegenseitig helfen. Es ist dies der Nachweis, dass die kühne Vision des Reduktions-Programms erfolgreich betrieben werden kann, dass also die Wissenschaften von ihrem Selbstverständnis her wieder stärker aneinander gebunden werden. (S. 202)

Wenn also schon keine Reduktion chemischer Theorien auf physikalische Theorien in einer Weise möglich zu sein scheint, die die Bedeutung der chemischen Forschungsebene irgendwie schmälern würde, muss der Versuch umso fragwürdiger erscheinen, die (Theorien der) Psychologie mittels des Nagelschen Ansatzes derart auf die (Theorien der) Neurowissenschaft oder gar die (Theorien der) Physik reduzieren zu wollen, dass die Theorien der Psychologie bzw. die Psychologie als Wissenschaft überhaupt überflüssig würde(n).

Keinesfalls würde es jedenfalls – wie auch Wunsch es für die Chemie darlegt – dazu kommen, dass es nur eine einseitige Beeinflussung der Neurobiologie auf die Psychologie geben könnte. Wenn es eine erfolgreiche Reduktion in der von Wunsch beschriebenen wünschenswerten

Form geben können sollte, dann ist im Regelfall eine *wechselseitige Beeinflussung* zu erwarten (vgl. McCauley, 1986; Pauen, 1999).

Eine weitere wichtige allgemeine Kritik an Nagels Modell der Reduktion formuliert McCauley (2007), indem er darauf verweist, dass „the reducing theory’s resources do not merely encompass those of the reduced theory. If, for no other reason, on the basis of its added precision alone, the reducing theory usually appears to *improve* upon the reduced theory’s account of things. Not infrequently, it corrects it” (S. 110). Ein Beispiel ist, dass durch die neurowissenschaftlichen Arbeiten von van Essen und Gallant (1994) festgestellt wurde, dass zwischen den beiden Verarbeitungspfaden visueller Informationen (Was- vs. Wo-Pfad) eine Vielzahl an Verbindungen besteht, die einen Informationsaustausch zwischen ihnen ermöglichen. Dies stellt eine Korrektur der vorhergehenden Annahme von Ungerleider und Mishkin dar, die ein voneinander unabhängiges Agieren dieser beiden Subsysteme postuliert hatten (vgl. McCauley, 2007, S. 110).

Wie bereits Wimsatt (1976, S. 218) festgestellt hat und wie es später auch von P. M. Churchland (1989, S. 48) aufgegriffen wurde, weist die Einräumung einer solchen Korrekturmöglichkeit nun aber eine gewisse Selbstwidersprüchlichkeit des Nagelschen Ansatzes nach (wenn er behauptet, dass falsche Theorien im Zuge einer solchen Reduktion auf korrektere Theorien reduziert würden):

If reducing theories often *correct* reduced theories in intertheoretic reductions, then, on the standard model of reduction, the reduced theories’ laws should *not* follow with deductive validity from premises about the laws of the reducing theories in conjunction with the bridge principles. In dealing with some of the most impressive reductions in the history of science, advocates of the standard model find themselves faced with the embarrassing dilemma of having to repudiate the D-N model of explanation unless they will accept reduction functions that leave enough semantic slack to render the putative derivation guilty of equivocation. *After all, false reduced*

theories cannot be validly deduced from true reducing theories, and they cannot even appear to do so unless the argument involves an equivocation. (McCauley, 2007, S. 110, Hervorh. im Orig.)

8.3. Ersetzung psychologischer durch (neuro-)biologische Theorien?

Wenn aufgrund der (teilweisen) Inkommensurabilität bzw. Inkompatibilität der psychologischen und (neuro-)biologischen Begriffsbildung die (zumindest durchgehende) Anwendbarkeit einer Nagel-Reduktion bezweifelt werden kann, dann stellt sich die Frage, ob es nicht vielmehr der Fall sein kann, dass die Neurobiologie einfach zunehmend die Psychologie verdrängen wird, wenn es um die Erklärung des menschlichen Erlebens und Verhaltens geht, ohne dass es hierfür überhaupt einer konkreten Verbindung psychologischer und (neuro-)biologischer Begrifflichkeiten und Theorien bedarf²⁷⁴.

8.3.1. Die „neue Welle“ des Reduktionismus

Von den Vertretern der so genannten *neuen Welle* des Reduktionismus („*New Wave Reductionism*“) in der Wissenschaftstheorie und Philosophie des Geistes wie Paul M. Churchland (1979, 1989), Patricia S. Churchland (1986), Clifford Hooker (1981) oder John Bickle (1998, 2003) wird nun behauptet, dass es keineswegs der normale Verlauf der Wissenschaft sei, dass eine Theorie nach den Nagelschen Prinzipien auf eine andere (basalere) Theorie reduziert wird, indem zwischen diesen beiden Theorien und

²⁷⁴ Im Folgenden werde ich mich primär auf den brillanten Beitrag von Robert N. McCauley (2007) beziehen.

den wissenschaftlichen Konzepten der verschiedenen Ebenen systematische Verbindungen hergestellt werden. Das, was Nagel beschreibt, sei lediglich ein äußerst selten vorkommender Sonderfall der Wissenschaftsentwicklung, ein Idealmodell der Reduktion, das von dem realen Wissenschaftsbetrieb viel zu stark abstrahiert.

Die New-Wave-Reduktionisten scheinen außerdem im Recht zu sein, wenn sie sich auf Thomas Kuhns (1962/1976) Modell der *Struktur wissenschaftlicher Revolutionen* stützen. Nach Kuhn (aber ganz ähnlich Paul Feyerabend, 1962) kommt es im Zuge der Phase einer „außerordentlichen“ Wissenschaft, die einen Paradigmenwechsel einleitet, zu einer Reihe von Inkommensurabilitäten zwischen dem alten Paradigma und seinem aufkommenden Nachfolger. So ändern sich mitunter die Auswahl der zu bearbeitenden Forschungsprobleme und eben auch die Begriffe, die Wissenschaftler verwenden. Auch wenn das Vokabular – oberflächlich betrachtet – dasselbe bleiben mag, ändern sich die hiermit bezeichneten *Begriffe* zumeist stark. Manche Begriffe werden eliminiert, manche neu eingeführt. Gewissermaßen agieren die Wissenschaftler der konkurrierenden Paradigmata sogar *in verschiedenen Welten*, denn wenn man z.B. ein geozentrisches Weltbild besitzt, führt dies zu einer grundlegend anderen Betrachtungsweise der Welt, als wenn man „durch die Brille“ eines heliozentrischen Weltbildes in ebendiese blickt. Weitere Beispiele, die Kuhn diskutiert, sind die Ablösung der Phlogiston-Theorie durch Lavoisiers Sauerstoff-Chemie oder die Ablösung der klassischen Newtonschen Physik durch Einsteins Relativitätstheorie.

Entscheidend für die New-Wave-Reduktionisten ist aber nun der Umstand, dass es nach Kuhn nicht zu einer allmählichen, kumulativen Weiterentwicklung wissenschaftlicher Theorien (z.B. durch die Anwendung des Popperschen Falsifikationsprinzips) kommt, sondern ein Paradigma wird dadurch aufgegeben, indem es durch ein anderes *ersetzt* wird. Nach

den wissenschaftssoziologischen Vorstellungen Kuhns ist dies vor allem der Fall, wenn hinreichend viele (Nachwuchs-)Wissenschaftler das neue Paradigma akzeptiert haben und auf seiner Grundlage neue (erfolgreiche) Theorien formulieren oder Experimente durchführen. Keinesfalls müssen hierzu zwischen den Theorien oder Konzepten der beiden Paradigmata Verknüpfungen gegeben sein, wie sie Nagel fordert. Die Theorie, die auf dem neuen Paradigma fußt, hat eine größere explanatorische Kraft und kann die Probleme (besser) lösen als die vorhergehende Theorie, die auf dem alten Paradigma basierte und die Wissenschaft von ihrem Normalbetrieb in die Phase der außerordentlichen Wissenschaft geführt hat. Die neue Theorie tritt daher einfach an die Stelle der unzureichenden Vorgängertheorie.

Statt einer Theoriereduktion im Sinne einer strengen logischen Ableitbarkeit wird hier nur eine reine *Ersetzung* („replacement reduction“, Sklar, 1967, S. 4) angenommen. Nach Bickle (1998) ist Reduktion „a proof of displacement (in principle), showing that a typically more comprehensive theory contains explanatory and predictive resources that parallel those of the reduced theory“ (S. 28). Als Beispiel aus den Naturwissenschaften wird zumeist aufgeführt, dass die klassische Theorie der Thermodynamik einfach empirisch und konzeptuell falsch ist und daher nicht über das Nagel-Schema auf die statistische Mechanik nach James Clerk Maxwell und Ludwig Boltzmann reduziert werden kann, sondern stattdessen durch jene ersetzt werden muss (vgl. Hooker, 1981, S. 49).

Diese Ersetzungsdoktrin findet sich in weiten Teilen der Neurowissenschaften bzw. in vielen Aussagen ihrer prominentesten Vertreter wieder (vgl. Gold & Stoljar, 1999). So schreibt Francis Crick (1994): „The scientific belief is that our minds – the behavior of our brains – can be explained by the interactions of nerve cells (and other cells) and the molecules asso-

ciated with them” (S. 7). Nach Gerald Edelman (1989) bedarf es einer umfassenden Theorie des Geistes, deren Ziel sein muss, „[to] construct a scientific theory of the mind based directly on the structure and workings of the brain. By ‚scientific‘ in this context, I mean a description based on the neuronal and phenotypic organization of an individual and formulated solely in terms of physical and chemical mechanisms giving rise to that organization” S. 8f). Auch bei Semir Zeki (1993) findet sich die Überzeugung vor, dass nur die Neurobiologie einen substantiellen Beitrag zum Verständnis des Geistes leisten könne und wir uns über kurz oder lang von den verstaubten psychologischen bzw. philosophischen Begriffen alter Tage (wie „unbewusst“, „Wissen“ oder „Information“) verabschieden müssten (vgl. ebd., S. 7).

Wenn ein zentrales Problem der Nagel-Reduktion war, dass sie – wenn sie das allgemeingültige Modell der Theorie-Reduktion sein will – annehmen muss, dass falsche reduzierte Theorien von korrekten reduzierenden Theorien abgeleitet werden können (was aufgrund der dem Nagelschen System inhärenten begrifflichen Äquivokationen unvermeidbar zu sein scheint), gehen die New-Wave-Reduktionisten (z.B. Bickle, 1998) und ihre Wegbereiter (z.B. Schaffner, 1967) nicht mehr von einem formal eindeutigen, idealisierten Modell intertheoretischer Reduktion aus, sondern behaupten, dass das, was von der reduzierenden Theorie abgeleitet werden kann, nicht die reduzierte Theorie selbst ist, sondern lediglich ein *Analogon* hierzu (vgl. McCauley, 2007).

On this “New Wave” model the reducing theory does not explain the reduced theory, so the dilemma disappears. Instead, it explains an analogue of the reduced theory constructed within the conceptual framework of the reducing theory. Thus, on the New Wave view the analogical relationship between the reduced theory and its re-

construction in terms of the reducing theory's conceptual resources enables the reducing theory both to correct the reduced theory and to explain at least something very much like it at the same time. (ebd., S. 111)

Dieses vergleichbar aussagekräftige Bild („equipotent image“, P. M. Churchland, 1989, S. 49) soll nun die zuvor von der reduzierten Theorie behandelten Phänomene erklären und vorhersagen. Nach Bickle (1998) ahmt das Analogon die Strukturen der reduzierten Theorie nach (vgl. ebd., S. 65). Hierbei gibt es nach Hooker (1981) ein Spektrum der Stärke der Analogie, die von einer „erhaltenden“ Reduktion bis hin zu einer reinen Ersetzungsreduktion reicht. Diesen Gedanken präzisiert wiederum vor allem Bickle (1998) dahingehend, dass für die Qualität einer Reduktion die *Passung* zwischen den Modellvorstellungen der beiden Theorien und ihren Anwendungen entscheidend ist. Dieser modelltheoretische Ansatz der Reduktion Bickles soll nun ermöglichen, eine Bestimmung der Güte der intertheoretischen *Passung* vorzunehmen. Die Annahme ist, dass es ein Kontinuum gibt, das auf der einen Seite eine sehr gute intertheoretische *Passung* aufweist, was die Ideal-Reduktion nach Nagel zulassen würde, bis hin zu einem äußerst schlechten „Mapping“, das zu einer Theorie-Eliminierung führen würde (vgl. Abb. 35).

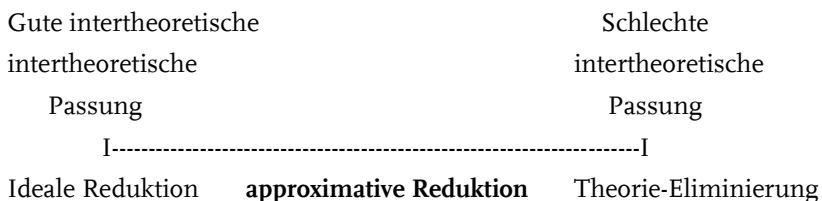


Abb. 35: Kontinuumsmodell der approximativen Theoriereduktion (Aus: McCauley, 2007, S. 113)

Bickle (1998, 2003) weist nun darauf hin, dass die meisten (historischen) physikalischen Reduktionen nur approximative Reduktionen waren, bei denen es keine gute intertheoretische Passung gab. Wenn aber nun schon rein physikalische Reduktionen keine gute Passung aufweisen, so die Argumentation, wird dies im Fall der Beziehung psychologischer und neurobiologischer Theorien noch viel geringer ausfallen.

Wie McCauley (2007) aufzeigt, sind sich daher sowohl Verteidiger einer Autonomie der Psychologie wie Fodor als auch deren Gegner wie die Churchlands oder Bickle darin einig, dass sie die Psychologie am rechten Pol des Spektrums der intertheoretischen Passung verorten, womit eine Nagel-Reduktion ausgeschlossen wäre. Während Fodor (1974) hieraus die Nicht-Reduzierbarkeit der Psychologie ableitet, ziehen die New-Wave-Reduktionisten auf derselben Basis ganz andere Schlussfolgerungen:

New Wave reductionists hold that such situations make not for theory reduction but for the “historical theory succession” that marks scientific revolutions [...]. The superior theory simply displaces its inferior counterpart. If their intertheoretic mappings are as tenuous as those in uncontroversial historical cases such as, say, those between Stahl’s account of combustion and Lavoisier’s or those between Gall’s phrenological hypotheses and modern cognitive neuroscience, we are, presumably, justified in speaking of the complete *elimination* of the inferior theory. Of course, it appears that the theories that risk elimination in the case at hand are, at the very least, those of folk psychology and, presumably, those in other areas of psychology that appeal to similar notions. (McCauley, 2007, S. 112)

Insbesondere die Churchlands haben immer wieder darauf hingewiesen, dass zwischen unseren alltagspsychologischen Konzepten, die auch die Theoriebildung in der wissenschaftlichen Psychologie weitestgehend dominieren, und der neurobiologischen Begriffs- und Theoriebildung eine unüberbrückbare Kluft liegt (P.S. Churchland, 1986; P.M. Churchland, 1989). Es sei von vornherein aussichtslos, eine Verbindung zwischen den psychologischen und neurobiologischen Theorien bzw. den in

ihnen enthaltenen Konstrukten vornehmen zu wollen, denn bei der (alltags-)psychologischen Theoriebildung handele es sich um eine überholte, wissenschaftlich unrespektable Konstruktion. Es könne daher keine Rede davon sein, zwischen einer fehlgeleiteten Alltagstheorie und einer „echten“ wissenschaftlichen Theorie (aus dem Spektrum der Neurowissenschaften) systematische Verbindungen in Form von Brückenprinzipien herstellen zu wollen. Denn wenn die Alltagspsychologie nach wissenschaftlichen Kriterien falsch ist, dann kann sie auch nicht über eine Nagel-Reduktion in den Kanon der wissenschaftlichen Theorien eingereiht werden (wie aufgezeigt wurde, führt eine Nagel-Reduktion keineswegs zu einer Auslöschung der reduzierten Theorie, sondern vielmehr zu einer Integration in das begriffliche Netzwerk der reduzierenden Theorie). Wie bereits an anderer Stelle besprochen, behauptet der vor allem von den Churchlands propagierte *Eliminative Materialismus* vielmehr, dass es sich bei psychologischen Theorien um überholte vorwissenschaftliche Theorien handelt, die durch neurowissenschaftliche Theorien zu *ersetzen* sind (gleichzeitig sind auch die von diesen Theorien hypostasierten psychologischen Konstrukte lediglich Relikte archaischer alltagstheoretischer Vorstellungen, die – zumindest im wissenschaftlichen Kontext – ebenfalls durch neurobiologische Konstrukte zu ersetzen sind; vgl. Kap. 3.1.3). Auch nach Bickle (2003) haben psychologische Theoriebildung und Forschung lediglich einen unabhkömmlichen *heuristischen* Wert (vgl. ebd., S. 114, 130).

8.3.2. Zur Persistenz der psychologischen Begriffs- und Theoriebildung

Ist es also wirklich so, dass die psychologische Theoriebildung ein vorwissenschaftliches Relikt ist, das – im Zuge der gegenwärtigen „(neuro-)wissenschaftlichen Revolution“ – durch die Neurobiologie zu ersetzen ist

bzw. dass psychologische Begriffe wie Selbstwirksamkeit, Leistungszielorientierung oder kognitive Dissonanz letztendlich den Status von Phlogiston in der Chemie innehaben? Haben sie also den Status überholter wirklichkeitsfremder Konstrukte, die vielleicht ein gewisses explanatorisches Potential besitzen mögen, *de facto* aber durch aussagekräftigere und ontologisch korrektere Größen (aus den Neurowissenschaften) abgelöst werden können? Hieran gibt es nicht nur starke Gründe zu zweifeln; der New-Wave-Reduktionismus unterliegt bei genauerer Betrachtung einer ganzen Reihe von grundsätzlichen Fehlannahmen und irreführenden Schlussfolgerungen.

Wie McCauley (2007) darlegt, sind es zwei problematische Bestimmungen, die die New-Wave-Reduktionisten relativ unbedarft von dem Konzept der Reduktion übernommen haben, wie es bereits der *Logische Empirismus* formuliert hat: Erstens nimmt der New-Wave-Reduktionismus (ebenso wie der „klassische“ Ansatz nach Nagel) an, dass man eine Wissenschaft letztendlich als ein bloßes Gefüge von Theorien betrachten kann. Eine Reduktion im wissenschaftlichen Kontext ist daher nichts anderes als ein rein *epistemisches* Manöver, bei dem es um die Beziehung von bloßen theoretischen Konstruktionen geht²⁷⁵ (vgl. z.B. P.M. Churchland, 1989, S. 48ff). Im Kern bedarf es hierbei keiner metaphysischen bzw. ontologischen Betrachtungen; Ebenen sind lediglich theoretische Konstruktionen oder Analyseebenen eines Systems. Daher sind auch Theorien und die Fachbegriffe der Wissenschaften einfach die zu einem Zeitpunkt besten *Beschreibungsformen* der Wirklichkeit, ohne dass man sich hiermit auf eine Beschaffenheit der Wirklichkeit an sich festlegen

²⁷⁵ Dabei unterscheiden sich die verschiedenen Vertreter des New Wave Reduktionismus in der Frage, ob es sich um eine (approximative) Reduktion von Gesetzen handelt (P. M. Churchland) oder ob der Einbezug von Gesetzen für eine psychoneurale Reduktion irrelevant ist (Bickle, 2003 oder Schaffner, 1992; vgl. McCauley, 2007, S. 121f).

müsste. Eine solche Position lässt sich als *epistemischer Relativismus* der Wissenschaft bezeichnen. Zweitens nehmen die Vertreter des New Wave Reduktionismus (wie auch die Verfechter Nagels) an, dass es ein *einziges* Modell für intertheoretische Reduktionen mit weitreichenden Implikationen gibt, mittels dessen *alle* Verbindungsformen zwischen wissenschaftlichen Theorien erfasst werden können.

Selbst wenn man die fragliche Festlegung, dass Wissenschaften letztendlich nichts anderes als eine Menge von wissenschaftlichen Theorien und Konzepten sind, teilt, kann hieraus keineswegs abgeleitet werden, dass die Eliminierung der Psychologie ein wissenschaftstheoretisch notwendiger, vernünftiger oder überhaupt möglicher Schritt wäre. Zunächst gilt, dass wenn die Psychologie nur einen – wenn überhaupt – heuristischen Wert hat, weil sie eine ungenaue höherstufige Beschreibungsebene ist, dann muss man bei der Anwendung dieses Argumentes auf andere Kontexte annehmen, dass auch die subzelluläre oder gar molekulare Forschungsebene der Neurowissenschaften (für die sich die New-Wave-Reduktionisten bekanntlich vehement einsetzen) nur einen heuristischen Status hat, da dieselben Themen auch auf noch basaleren Ebenen der entsprechenden Forschungsfelder der Chemie oder gar der Physik thematisiert werden könnten. In der Tat äußert sich Bickle (2003) dann auch dahingehend (vgl. ebd., S. 115, 157; McCauley, 2007, S. 122).

Wenn Wissenschaft letztendlich nur ein System von Beschreibungen, Konstruktionen und Programmen ist, dann ist unter diesen Voraussetzungen die Neurowissenschaft und selbst die Physik „nur“ ein solches. Auch die Physik kann unter diesen Voraussetzungen aber nicht *der Wirklichkeit selbst* habhaft werden. Wenn dem so ist, dann fehlt ein klares Kriterium für die Bestimmung des Primats einer Wissenschaft gegenüber einer anderen. Die reine Güte der Vorhersage des zukünftigen Verhaltens eines Systems dürfte jedenfalls kein Kriterium abgeben, das (zu-

mindest zum gegenwärtigen Zeitpunkt) zugunsten neurobiologischer oder gar physikalischer Theorien gegenüber psychologischen Theorien (des Geistes) sprechen könnte. Oder wie Pauen (1999) sagt:

Wollte man einem dieser Programme einen Primat gegenüber einem anderen zuzubilligen, das auf seiner Ebene erfolgreich ist, dann müsste man von einem unabhängigen Standpunkt aus beurteilen können, wie nah diese Programme ‚der‘ Wahrheit kommen, in deren Besitz man selbst in diesem Moment sein müsste. Soweit bekannt, hat bislang weder ein Philosoph des Geistes noch ein Wissenschaftstheoretiker diesen beneidenswerten Standpunkt erreicht, und so fehlt die Basis dafür, der Physik, der Physiologie oder irgendeiner anderen Wissenschaft einen Primat gegenüber anderen Bereichstheorien zuzubilligen. (S. 47)

Ebenso scheint eine neurobiologische Erforschung menschlichen Verhaltens und Erlebens ohne den Rückgriff auf psychologische Theorien und Begrifflichkeiten prinzipiell gar nicht möglich zu sein. Obgleich die New-Wave-Reduktionisten dies vordergründig abstreiten mögen, sind sie auch bei ihren Vorschlägen für eine adäquate Erforschung des Geistes immer auf diese „archaischen“ Größen angewiesen. Gerade die Modelle neuronaler Netzwerke, die vor allem von den Churchlands lange Zeit als Revolution in der Erforschung menschlichen Denkens gepriesen haben, kommen nicht ohne Bezug auf psychologische Begrifflichkeiten aus, wenn sie irgendeine gehaltvolle Aussage treffen wollen:

They do so even though the neuroscientists and neural network modelers, whose work they discuss, often appeal to findings from higher level psychological sciences to support the neuroscientific and neurocomputational models they prefer. Crucially, those researchers look to these findings (and, by implication, to the methods and techniques by which they were generated) not merely for guidance [...] but for *support* for their favorite hypotheses. [...] Both standard and New Wave models' spotlight on the *reduction* of theories in cross-scientific contexts renders them largely insensitive to the contributions higher level sciences regularly make to the *justification* of the very scientific theories – particularly some in neuro-science – that they champion. Neuroscientists and neurocomputational modelers regularly cite psy-

chological evidence in support of their proposals, yet New Wave reductionists often lose sight of all of this in their explicit philosophical analyses and especially in their examinations of the relations between psychology and neuroscience. (McCauley, 2007, S. 122; Hervorh. im Orig.)

Das Paradebeispiel der New-Wave-Reduktionisten ist die Erforschung des menschlichen Gedächtnisses. Schaut man sich hier z.B. die neurocomputationalen Modelle von Sejnowski und Rosenberg (1988) an, sieht man allerdings, dass auch diese auf Befunden der kognitiven Psychologie basieren und auch, dass die Ergebnisse der Modellierung meist nur Replikationen von Befunden der klassischen psychologischen Gedächtnisforschung sind (vgl. McCauley, 2007, S. 122f). Dies muss nicht verwundern, denn wie auch Dörner (2004) sagt, bedarf es für eine Erklärung des Umstandes, dass neuronale Netzwerke psychische Prozesse hervorbringen bzw. mit solchen identisch sind, die richtigen „Suchbilder“, ohne die eine Neurowissenschaft völlig blind wäre: „Man muss wissen, wonach man sucht. Wer in riesigen Neuronenverbänden nicht ziellos nach den Grundlagen von Denken, Fühlen, Wahrnehmen stöbern will, braucht eine klare Vorstellung davon, was damit eigentlich gemeint ist – was also die mentalen Prozesse ausmacht. Die Modelle für all diese Prozesse liefert aber nicht die Hirnforschung, sondern die Psychologie. Ohne deren Beitrag stochern selbst Neurobiologen nur mit der langen Stange ihrer ausgefeilten Methoden im Nebel – und die Wahrscheinlichkeit, auf diese Weise etwas herauszufinden, ist denkbar gering“ (S. 38). Dass neurobiologische Theorien und ihre Konzepte also einfach psychologische Theorien und ihre Konzepte komplett ersetzen werden, ist aufgrund der „semantische[n] Blindheit“ (Laucken, 2002) der Neurowissenschaften *prinzipiell* unmöglich, denn diese bedürfen – um eine inhaltliche Aussagekraft zu haben – der Führung durch die Psychologie. Ohne deren begriffliches Gerüst offenbart der Blick ins Gehirn lediglich eine graue glitschige und

relativ undifferenzierte Masse (vgl. McGinn, 1989). Erst durch die Zuordnung von psychologischen Funktionen zu bestimmten Hirnbereichen bzw. deren Verhalten simulierenden neuronalen Netzwerkmodellen wird diesen indifferenten, zusammenhangslosen und ziellos verlaufenden Geschehnissen gewissermaßen „Leben“ oder überhaupt eine Bedeutung eingehaucht.

Des Weiteren gibt es guten Grund daran zu zweifeln, dass es überhaupt einen solchen grundsätzlichen epistemischen Relativismus der Wissenschaften gibt, wie ihn die New-Wave-Reduktionisten unterstellen. Zunächst weist eine Wissenschaft neben den Konzepten und Theorien, die eher nur die Produkte einer Wissenschaft darstellen, auch eine Reihe von charakteristischen Methoden, wissenschaftlichen Praktiken und Experimentalkonzeptionen auf (McCauley, 2007, S. 124, 128). Darüber hinaus ist die Annahme eines solchen epistemischen – Ontologie-freien – Status der Wissenschaften abwegig. Wie in dieser Arbeit herausgestellt wurde, sind Ebenen keineswegs generell als rein analytische oder theoretische (beobachterabhängige) Konstruktionen aufzufassen, sondern sie spiegeln zumindest immer dann eine *natürliche* (beobachterunabhängige) Strukturgliederung eines Objektes wider, wenn Teil-Ganzes-Beziehungen zwischen den (agierenden) Entitäten ausfindig gemacht werden können, die sich in Abhängigkeit von den verschiedenen Perspektiven auf dieses Objekt manifestieren (vgl. Kap. 5.3). Auch bei den von Churchland und Sejnowski (1994/1997) herausgestellten Organisationsebenen des Nervensystems, die als Forschungsebenen des Gehirns fungieren, also Moleküle, Synapsen, Neurone, neuronale Netzwerke usw. handelt es sich offenbar auch nach ihrer eigenen Ansicht nicht um bloße analytische Ebenen, denn diese werden von den Autoren klar von Organisationsebenen abgegrenzt (die Ebenen des psychologischen Modells der Wahrnehmung nach Marr werden hier als Prototyp analytischer Ebenen heraus-

und den genannten Organisationsebenen gegenübergestellt; vgl. Kap. 4.2.1.1). Daher hat McCauley (2007) vollkommen recht, wenn er schlussfolgert, dass „our most successful theories provide significant clues about the furniture of the universe. This suggests that levels of analysis in science correspond to levels of organization in nature” (S. 124).

Wenn nun aber gilt, dass die Neurobiologie ohne psychologisches Geleit blind ist bei der Erforschung des Gehirns, dann verdeutlicht dies zugleich die hier vertretene Ansicht, dass es sich auch beim Psychischen um eine eigenständige, unverzichtbare Forschungsebene handelt: „If one entity contains others as its parts and if explanations of (some of) its behaviors require appeal to further organizing principles beyond those concerned with those parts, *then it occurs at a higher level of organization and likely points to a distinguishable analytical level*“ (ebd., S. 125; Hervorh.: M.L.).

8.3.3. Die fundamentale Differenz zwischen Intra- und Inter-Level-Kontexten

Kommen wir nun zur zweiten fundamentalen Fehlannahme des New Wave Reduktionismus, der Überzeugung, dass es nur *eine* mögliche Interpretation ihres Modells intertheoretischer Relationen gibt. Die Verfechter dieses Ansatzes unterliegen einer folgenschweren Konfundierung zwischen Reduktionen *innerhalb einer* Forschungsebene (*intralevel*) oder *zwischen verschiedenen* Forschungs-ebenen (*interlevel*). Dabei hat bereits Wimsatt (1976) überzeugend dargelegt, dass es eine grundlegende Unterscheidung zwischen Inter-Level- und Intra-Level-Reduktionen gibt (vgl. S. 215ff). Die von Kuhn beschriebenen Fälle wissenschaftlicher Revolutionen, bei denen ein grundlegendes Forschungsparadigma ein anderes ablöst, beziehen sich grundsätzlich auf Fälle von Reduktionen *innerhalb* einer Forschungsebene. Bei dem Vorgang, dass die Phlogiston-Theorie

durch die Sauerstoff-Theorie ersetzt wurde, handelt es sich um einen Vorgang zur Erklärung des Phänomens der Verbrennung, der auf *derselben* Forschungsebene innerhalb der Chemie angesiedelt ist. Entsprechendes gilt für die sich in der Physik ereignende Ablösung der Newtonschen Mechanik durch die Einsteinsche Relativitätstheorie usw. Intra-Level-Reduktionen beschreiben eine (zeitliche) Abfolge von Theorien („theory succession“; McCauley, 2007, S. 130), bei der es sich um konkurrierende Theorien in einem zeitlichen Verlauf innerhalb einer bestimmten Wissenschaft auf einer ganz bestimmten Analyse- bzw. Organisationsebene handelt (vgl. ebd.).

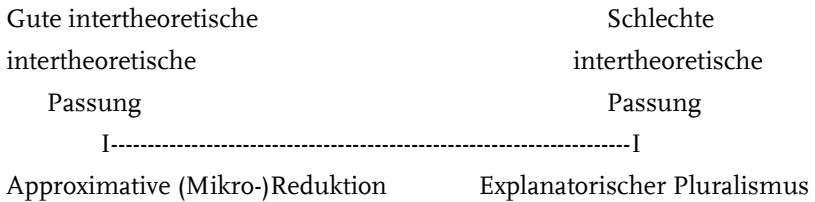
Bei Inter-Level-Reduktionen handelt es sich nach Wimsatt (1976) um wissenschaftsübergreifende Beziehungen zwischen Theorien²⁷⁶, die zur selben Zeit auf verschiedenen (analytischen) Ebenen in den Wissenschaften vorherrschen. Ein Beispiel wäre die Beziehung zwischen der aus psychologischen Befunden abgeleiteten Vermutung, dass es eine besondere Fähigkeit der visuellen Informationsverarbeitung von Gesichtern gibt, wie sie Farah, Wilson, Drain und Tanaka (1998) formuliert haben, und dem neurowissenschaftlichen Ansatz, der postuliert, dass der fusiforme Gyrus im Temporallappen die Region im Gehirn ist, die ebendiese Funktion erbringt (Kanwisher et al., 1997; vgl. McCauley, 2007, S. 130).

Entscheidend ist nun, dass sich Intra-Level-Reduktionen und Inter-Level-Reduktionen fundamental darin unterscheiden, welche Folgen der unterschiedliche Grad an intertheoretischer Passung bezüglich des zuvor be-

²⁷⁶ Genau genommen, geht es bei Inter-Level-Reduktionen nach Wimsatt nicht einmal primär um eine Reduktion von Theorien verschiedener Ebenen, wie McCauley es darstellt, sondern vielmehr um eine Reduktion der beteiligten *Phänomene* selbst (er spricht hier deshalb etwas missverständlich von *phänomenalen* Reduktionen). Reduktionen, die verschiedene Forschungsebenen betreffen, sind seines Erachtens also vor allem *ontologische* und weniger epistemische Reduktionen.

schriebenen Kontinuumsmodells für die reduzierte Theorie nach sich zieht. Statt der von den New-Wave-Reduktionisten als allgemeingültig betrachteten eindimensionalen Interpretation des Kontinuums, muss – wie McCauley (2007) darlegt – eine weitere Dimension berücksichtigt werden und zwar die Verortung der jeweiligen Reduktion *innerhalb* einer Forschungsebene versus *zwischen* verschiedenen Forschungsebenen (siehe Abb. 36).

a) Inter-Level (oder wissenschaftsübergreifende) Kontexte



b) Intra-Level (oder „fortlaufende“) Kontexte

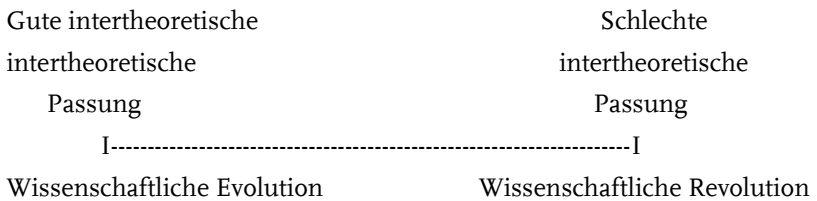


Abb. 36: Kontextabhängiges Modell der Theoriereduktion (Aus: McCauley, 2007, S. 133)

So haben wir bei den unten in der Abbildung abgetragenen *Intra-Level-Kontexten* ein Kontinuum, das von einer *wissenschaftlichen Evolution* (im Sinne der von Thomas Kuhn als „normale Wissenschaft“ deklarierten Phase) bis hin zur *wissenschaftlichen Revolution* (der „außerordentlichen“

Wissenschaftsphase) reicht, bei der aufgrund der schlechten intertheoretischen Passung tatsächlich eine neue Theorie bzw. ein neues Paradigma auftritt, die bzw. das im Extremfall dazu führt, dass die alte Theorie bzw. das alte Paradigma und die durch sie bzw. durch es postulierte Ontologie verworfen werden.

Dagegen gibt es im Fall des *Inter-Level-Kontextes* bei einer guten intertheoretischen Passung die Möglichkeit zu einer (Mikro) Reduktion (im Sinne Nagels), während aus einer schlechten Passung keinesfalls die Eliminierung der fraglichen Theorie folgt, sondern ganz im Gegenteil: die Anerkennung eines *explanatorischen Pluralismus*. „Explanatory pluralism underscores the ongoing interaction of scientific enterprises carried out at the various analytical levels. All scientific explanation is partial explanation from the perspective of some analytical level or other“ (McCauley, 2007, S. 137; Hervorh. im Orig.). Wie Polger (2004) sagt, sind wissenschaftliche Theorien und die durch sie angeregten Erklärungen immer selektiv; sie erklären nie alles oder etwas vollkommen (vgl. ebd., S. 203). Insofern bestreitet der u.a. von McCauley (2007, 2007a), Polger (2004) oder auch Bechtel (2007) verfochtene explanatorische Pluralismus, dass intertheoretische Beziehungen „exhaust all of the selection pressures in the resulting co-evolutionary process and that those selection pressures are exerted exclusively from the bottom up“ (McCauley, 2007, S. 137). Hierzu an späterer Stelle mehr (vgl. Kap. 8.7 und Kap. 9.4.).

Zunächst kann festgestellt werden, dass die von den New-Wave-Reduktionisten postulierte Theorie-Eliminierung bei einer schlechten intertheoretischen Passung nur auf den Intra-Level-Kontext zutrifft:

In fact, these sorts of large-scale, fell-swoop eliminations are an accurate prognosis in intralevel settings only. Where theories are substantially discontinuous in intralevel, successive contexts, the result, when the new theory triumphs, is something like a Kuhnian scientific revolution. [...] Reliably, the examples of theory elimination

to which New Wave reductionists point – the theories of the crystalline spheres, alchemical essences, phlogiston, caloric fluid, the aether, phrenological faculties, etc. – concern a new theory superseding an older theory at the *same* level of analysis. They concern, in short, *intra*level relations. Because they often fail to distinguish them sharply, New Wave reductionists attribute the profile and outcomes characteristic of intertheoretic relations in intralevel settings to *both* sorts of contexts. (McCauley, 2007, S. 134; Hervorh. im Orig.)

Aufgrund dieser fehlenden Differenzierungsleistung erwarten die New-Wave-Reduktionisten nun fälschlicherweise eine Theorie-Eliminierung, wenn Wissenschaften *verschiedener* Analyseebenen divergierende Ansichten bezüglich ein und desselben Phänomens haben. Wenn es nicht möglich ist, eine Analogie der höherstufigen Theorie innerhalb des Rahmens der niederstufigeren Theorie herzustellen, signalisiere das daher zwangsläufig die Eliminierung der höherstufigen Theorie. Weil New-Wave-Reduktionisten gleichzeitig annehmen, dass Wissenschaften letztendlich aus nichts anderem als ihren Theorien bestehen – kommen sie zu der abwegigen Annahme, dass die Eliminierung einer Theorie in solchen Inter-Level-Kontexten gleichbedeutend ist mit der Eliminierung einer ganzen Wissenschaft wie der Psychologie (vgl. ebd.).

Wie aber nun klar geworden sein sollte, kann aus dem Scheitern von Reduktionsversuchen nach dem Modell der klassischen Theorie-Reduktion nur dann Zweifel an der Existenzberechtigung einer Theorie abgeleitet werden, wenn zwischen den Theorien ein direktes Konkurrenzverhältnis besteht, wie es zwischen der Sauerstofftheorie der Verbrennung der modernen Chemie und der vormaligen Phlogiston-Theorie vorliegt (vgl. McCauley, 1986, S. 77). Und dies ist eben nur in Intra-Level-Kontexten der Fall. Wenn Theorien auf unterschiedlichen Ebenen verortet sind (wie diejenigen der Psychologie und Neurobiologie), dann sind solche Folgen nicht zu fürchten. Wie Pauen (1999) sagt: „Gegeneinander abwägen kann man die Theorien nur dann, wenn sie in ihren Prognosen oder

Erklärungen bezüglich einer Entität direkt miteinander konkurrieren. Fehlt hingegen eine solche Vergleichsmöglichkeit, weil wir es mit zwei unterschiedlichen Ebenen der Beschreibung zu tun haben, dann haben wir keine Basis dafür, einer Theorie hier einen Primat zuzuschreiben“ (S. 46f).

Im Allgemeinen ist die von einigen New-Wave-Reduktionisten vorgetragene Idee, dass die gesamte Psychologie als Wissenschaft auf die Neurobiologie als Wissenschaft reduziert werden sollte (und im Zuge dessen gänzlich eliminiert wird), begrifflich und pragmatisch vollkommen absurd:

Fears about a *science's* elimination or enervation or dispensability on the basis of its alleged (standard or New Wave) reduction are unfounded. Reductions of psychological theories to theories in neuroscience are conceptual bridges that permit intellectual traffic to flow between these two sciences. Constructing such integrative relationships amounts to establishing the sort of infrastructure that facilitates important forms of scientific progress at both levels. Nor does endorsing either the power of reductive strategies in science generally or individual reductive proposals at the border of psychology and neuroscience specifically require any expectations about sciences that operate at higher analytical levels eventually being forced to conduct going-out-of-business sales. It never has been nor is there any reason to think that it ever will be the case that up-and-running sciences collapse on the basis of successful theoretical reductions, or that they could so collapse, or, especially, that they *should* so collapse. *Sciences* are not the sorts of things that get reduced or eliminated. Theories are. (McCauley, 2007, S. 140; Hervorh. im Orig.)

Schaut man sich die Ausführungen der New-Wave-Reduktionisten genauer an, räumen sie selbst –Zähne knirschend mit Verweis auf einen angeblich „lediglich heuristischen Stellwert“ – zumeist ein, dass die Psychologie als eigenständige und unabkömmliche Wissenschaft anerkannt werden muss. So hat gerade Bickle, der programmatisch einen *rücksichtslosen Reduktionismus* vertritt, an einigen Stellen die Einsicht erteilt, dass es eine *Ko-Evolution* von Psychologie und Neurowissenschaft

gibt, die einer Eliminierung der Psychologie entgegensteht, und die Psychologie einen wichtigen Beitrag zur Charakterisierung kognitiver Funktionen höherer Ebenen leistet (vgl. ebd., 1998, S. 141; 2003, S. 128ff). Trotz allen Fortschritts in den Neurowissenschaften wird es auch zukünftig einer ebenenübergreifenden Forschung bedürfen, sodass „psychological causal explanations still play *important heuristic roles* in generating and testing neurobiological hypotheses“ (Bickle, 2003, S. 115; Hervorh. im Orig.).

8.4. Die reduktive Erklärung des Psychischen durch die Neurobiologie

Was bisher besprochen wurde, ist die Frage, ob es möglich ist, psychologische *Theorien* auf (neuro-)biologische *Theorien* zu reduzieren oder erstere durch letztere zu ersetzen. Allerdings ist dies keineswegs die einzige Beziehung, die zwischen der Psychologie und der Neurobiologie und ihren Forschungsgegenständen besteht. Statt dass man die Theorien der Psychologie in Bezug zu biologischen Theorien setzt, scheint ein vielsprechender Ansatz zu sein auszuloten, ob nicht mentale *Phänomene*, wie sie die Psychologie für gewöhnlich zum Gegenstand hat, hinreichend durch Theorien oder Phänomene der Neurobiologie *erklärt* werden können.

Es handelt sich hierbei um den Ansatz funktionaler Erklärungen mentaler Eigenschaften, wie er bereits im Zusammenhang der Besprechung der physischen Realisierungstheorie vorgestellt wurde (vgl. Kap. 6.5.1) und in der Gegenwartsdebatte in der Philosophie des Geistes vor allem mit Levine (1993) und Kim (1998, 2005) verbunden wird. Hier noch einmal eine Formulierung des „Schlachtplans“ des funktionalen Modells der

Reduktion einer mentalen Eigenschaft *M* auf eine physische (neuronal) Eigenschaft *P*, wie ihn Kim (1998) beschreibt:

[...] to reduce a property *M* to a domain of base properties, we must first 'prime' *M* for reduction by construing, or reconstructing, it *relationally* or *extrinsically*. This turns *M* into a relational/extrinsic property. For functional reduction we construe *M* as a second-order property defined by its causal role – that is, by a causal specification *H* describing its (typical) causes and effects. So *M* is now the property of having a property with such-and-such causal potentials, and it turns out that property *P* is exactly the property that fits the causal specification. And this grounds the identification of *M* with *P*. *M* is the property of having some property that meets specification *H*, and *P* is the property that meets *H*. So *M* is the property of having *P*. (S. 98f; Hervorh. im Orig.)

Kim (2005) präzisiert dies später noch dahingehend, dass es drei Schritte sind, die eine erfolgreiche funktionale Reduktion umfassen sollte: Erstens die Funktionalisierung der Zieleigenschaft *M* (über konzeptuelle Analysen die verhaltensbezogenen Eigenschaften von *M* aufdecken, die für seine kausale Rolle *C* charakteristisch sind), zweitens die Identifikation der physischen (neuronalen) Realisierer von *M* und drittens die Entwicklung einer erklärenden Theorie, die zum Gegenstand hat, wie die Realisierer von *M* die kausale Rolle *C* bewerkstelligen (ebd., S. 101f). Solche funktionalen Reduktionen würden nun im Gegensatz zum Nagelschen Modell der Reduktion eine tatsächliche *Erklärung* der fraglichen Phänomene liefern (vgl. S. 102). Darüber hinaus seien Instantiierungen der reduzierten Eigenschaften ausschließlich auf der Basis von Informationen über Phänomene auf der Basisebene vorhersagbar.

Was bedeutet aber *explanatorisches Reduzieren* im Sinne dieses Ansatzes nun genau und welche Konsequenzen sind hiermit für den Wissenschaftlichkeitsstatus der Psychologie verbunden? Kim erklärt die Harmlosigkeit von Reduktionen dieses Typs für die höherstufigen Wissenschaften am Beispiel des Gens: Mendel hatte ein theoretisches Konzept eines Gens im

Sinne eines Mechanismus in einem Organismus postuliert, mittels dessen genetische Informationen verschlüsselt und übertragen werden. „[T]he concept of a gene is a ‘functional’ concept, and the property of being a gene is a functional property defined by a ‘job description’” (S. 163). Nun habe die Wissenschaft gezeigt, dass DNA-Moleküle die Aufgabe erfüllen, die Mendel beschrieben hat, und sie damit die Realisierer von Genen sind. „When all this is in, we can say that the gene has been physically reduced, and that we now have a reductive explanation of how the process of heredity works at the molecular level“ (ebd.). Das höherstufige Konzept des Gens ist aber nun seit der Entdeckung, dass DNA der (materielle) Funktionsträger der Vererbung ist, nicht aus dem Kreis der Wissenschaft verbannt worden. Ganz im Gegenteil gehört es zum festen Kanon der biologischen Begriffe, auf den selbst die „basaleren“ biologischen oder biochemischen Subdisziplinen ganz selbstverständlich zugreifen. Eine funktionale Reduktion liefert auch keinerlei Anlass für die Eliminierung der funktional „reduzierten“ Phänomene.

Dasselbe gilt für die Erklärung psychischer Phänomene durch neurobiologische Aspekte. So führt die funktionale Reduktion von (menschlichen) Schmerzen auf C-Faser-Reizungen keineswegs dazu, dass wir das Konzept des Schmerzes oder die psychologische Theorie zur Entstehung, zum Verlauf oder zur Behandlung von Schmerzen nicht mehr bräuchten. Wie Pauen (1999) sagt, ist bei der Erklärung mentaler Phänomene „ein Rückgriff auf naturwissenschaftliche Erkenntnisse möglich, ohne daß es zu einer Reduktionsbeziehung zwischen Psychologie und Neurophysiologie kommt“ (S. 52). Eine reduktive Erklärung hat eben keineswegs den Status einer vollständigen Theoriereduktion, wie sie Nagel mit seinem Ansatz anvisiert. Denn dafür, dass die Psychologie zur Erklärung mentaler Phänomene auch Erkenntnisse der klassischen Naturwissenschaften wie der Biologie oder Chemie heranziehen kann, ist es überhaupt nicht

notwendig, dass die psychologischen Gesetze vollständig aus denen der Neurophysiologie abgeleitet werden können. Diese Art von Ableitbarkeit wird durch das funktionale Modell der Reduktion gar nicht berührt.

Im Prinzip wird hier also für das Verhältnis von Psychologie und Neurophysiologie nur in Anspruch genommen, was in anderen Bereichen selbstverständliche Praxis ist. Physikalische Theorien sprechen im allgemeinen weder von Wirtschaftskrisen noch von Fußballspielen und auch nicht von Mozart-Sinfonien. Ein Wissenschaftler, der ein „vollständiges“ Wissen über die physischen Prozesse etwa in Fußballstadien besitzt, aber noch niemals etwas von der dort betriebenen Sportart gehört hat, würde hieraus kaum ein konkretes Wissen über Fußballspiele ableiten können; zweifellos blieben ihm entscheidende Wissenslücken bezüglich der Regeln, Gepflogenheiten und Ziele des Spiels. Dennoch geht – abgesehen vielleicht von einigen sehr leidenschaftlichen Anhängern dieses Sports – wohl kaum jemand davon aus, daß Fußballspiele etwas Übernatürliches seien. Zu diesem Schluß würde man erst kommen, wenn man zeigen könnte, daß die Hauptakteure tatsächlich die ihnen zuweilen nachgesagte Fähigkeit haben, die Gesetze der Physik außer Kraft zu setzen. (ebd., S. 53)

Auch wenn man also das Fußballspielen funktional auf seine physischen Realisierer reduzieren kann, weil es diese sind, die die kausale Rolle einnehmen, über die sich Fußballspielen definiert, erschöpft sich der *epistemische* Gehalt dieser Sportart nicht in seiner „Materialität“ (sprich: es gibt mehr zu *wissen*, als das, was sich aus einer rein physikalischen Betrachtungsweise heraus hier zu erschließen möglich ist). Trotzdem bleibt es dabei, dass höherstufige Eigenschaften wie die Regeln eines Fußballspiels und alle anderen höherstufigen Eigenschaften *substantiell* oder *ontologisch* nichts sind, was über ihre stofflich-funktionale Seinsweise hinausreichen würde.

Die Reduzierbarkeit mentaler Eigenschaften hängt nunmehr von der *Funktionalisierbarkeit* ab und nicht von der Verfügbarkeit von Brückengesetzen. Brückengesetze sind nach Ansicht Kims weder notwendig noch hinreichend für eine erfolgreiche Reduktion (vgl. ebd., 1998, S. 27). Natur-

lich könne es auch andere Realisierer von Genen geben, aber diese multiple Realisierbarkeit ist kein Hindernis für Reduktionen oder reduktive Erklärungen, die lediglich lokal zu bestimmen sind: „Functional reduction, as I call it, can focus on the reduction of a mental property, or a group of them, for a specific population – that is, neural research on pain will aim at *local* reductions, not a one-shot *global* reduction [...]. Neural bases may differ for different instances of pain, but individual pains must nonetheless reduce to their respective neural/physical realizers” (ebd., 2005, S. 25; Hervorh. im Orig.).

So könnte es – in der modallogischen Sprechweise ausgedrückt – in bestimmten möglichen Welten andere Realisierer von Genen als DNA-Moleküle geben, die die kausalen Funktionen zeigen, die für Gene definitiv sind (vgl. ebd., 1998, S. 25). Dennoch bleibt die funktionale oder explanatorische Reduktion von Genen im hier betrachteten Fall des Menschen gültig (die Voraussetzung hierfür sei lediglich, dass die grundlegenden Naturgesetze beibehalten werden, es sich also bei den möglichen Welten – in Relation zu unserer Welt – um *nomologisch mögliche* Welten handelt; vgl. ebd.).

Kim räumt ein, dass es unwahrscheinlich ist, dass wir alle gegenwärtigen und nomologisch möglichen Realisierer von mentalen Eigenschaften ausfindig machen können. Aber auch hieraus rühre kein Grund zum Zweifel an der Gültigkeit des funktionalen Modells der Reduktion: „Even if we have not identified the actual realizer – perhaps we never will – it would make not much difference philosophically; we know that there must be a lower-level physical realizer, even if we don’t have a perspicuous description of it in an underlying theory, and we know the phenomenon involved to be reducible to its physical realizer, whatever it is” (ebd., 2005, S. 164).

Auch wenn Kim davon überzeugt ist, dass diese Art der Reduktion dem Paradigma der Reduktion in der Wissenschaft im Allgemeinen entspreche (vgl. ebd., 1998, S. 24f), dürfte bereits der Verweis auf die vorangegangenen Abschnitte dieses Kapitels aufgezeigt haben, dass von solch einem „Standardmodell“ nicht die Rede sein kann (vgl. McCauley, 2007). Auch Pauen (2007) kritisiert, dass das Modell der Reduktion nach Kim bzw. Levine keineswegs das einzige bzw. *das* naturwissenschaftliche „Standardmodell“ der Reduktion ist; eher sei dies ein Maximalmodell, das sehr hohe Hürden aufstellt. „Tatsächlich herrscht weder in der Physik noch in der Philosophie Einigkeit über die Kriterien für erfolgreiche reduktive Erklärungen. Bei diesem Modell handelt es sich keineswegs um ‚das‘ Standardmodell, sondern um eine Variante, auf die sich vor allem die Vertreter des Erklärungslückenargumentes berufen“ (S. 136f). Denn wie bereits in Kapitel 6.5.4.1 erläutert wurde, ist ein großes Problem, dass die hier geforderten konzeptuellen Analysen zur Funktionalisierung bzw. zur Erstellung eines kausalen Profils mentaler oder überhaupt höherstufiger Eigenschaften nicht immer so leicht zu erbringen sind, wie dies im Fall der Beziehung zwischen Gen und DNA der Fall sein mag. Daher überhöhen die Vertreter des Erklärungslückenargumentes dieses Modell der funktionalen Reduktion gerne selbst zu *dem* Reduktionsmodell der Wissenschaften, um hiermit zugleich der Möglichkeit einer funktionalen Reduktion bzw. überhaupt irgendeiner Reduktion mentaler Eigenschaften zu widersprechen, weil wahrscheinlich nicht für alle phänomenalen Eigenschaften hinlängliche, eindeutig voneinander abgrenzbare kausale Profile erstellt werden könnten. Dabei fügen sich nicht einmal die vielen erfolgreichen naturwissenschaftlichen Reduktionen, wie sie im Vorfeld diskutiert wurden, nahtlos in das Schema der funktionalen Reduktion ein, denn nicht immer kann das kausale Profil höherstufiger Eigenschaf-

ten über konzeptuelle Analysen ermittelt werden (vgl. ebd., S. 137). Dies ist für eine erfolgreiche Reduktion aber auch gar nicht nötig.

8.5. Irrwege der „Autonomisierung“ der Psychologie

In den letzten 40 Jahren hat es eine Reihe einflussreicher Arbeiten zu der Frage gegeben, wie die Psychologie sich gegenüber einer Vereinnahmung durch die (Neuro-)Biologie erwehren kann. Viele Philosophen sind hierbei als Verteidiger der „Autonomie“ der Psychologie aufgetreten und haben diese vor der in ihren Augen gefährlichen Nagel-Reduktion und der sie offenbar ermöglichenden Annahme einer psychophysischen Identität bewahren wollen. Der so genannte *nicht-reduktive Physikalismus* ist in der Philosophie immer noch eine sehr verbreitete Strömung, die sich (u.a.) genau dieser Angelegenheit verschrieben hat und statt einer Identität „lediglich“ eine gewisse *Abhängigkeit* des Geistes von seinen physischen Grundlagen postuliert.

Wie sich zeigen wird, ist es aber gerade diese Mikroabhängigkeit, die einen Mikrophysikalismus und damit ein ontologisches Primat der Biologie gegenüber der Psychologie mit sich bringen würde. Nicht die Akzeptanz einer psychophysischen Identität gefährdet die Eigenständigkeit und Relevanz der Psychologie, es ist der Mikrophysikalismus, von dem diese Gefahr in der Tat ausgeht.

8.5.1. Der (metaphysische) Funktionalismus als Garant von „Autonomie“

Es war vor allem Jerry Fodor (1974), der in seiner wirkmächtigen Schrift *Special Sciences* den seines Erachtens mit einer erfolgreichen Nagel-Reduktion verbundenen Mikroreduktionismus, wie er von Oppenheim

und Putnam (1958) postuliert wurde, zu bekämpfen suchte. Als Untertitel seines Aufsatzes hat Fodor daher augenfällig die Aussage *The disunity of science as a working hypothesis* gewählt, was den Charakter seines Aufsatzes als einer Replik auf Oppenheim und Putnam verdeutlicht.

Wie bereits in Kapitel 6.1.4.2 besprochen, kritisiert Fodor (1974) die (klassische) psychophysische Identitätstheorie nicht nur mithilfe des Argumentes der multiplen Realisierbarkeit mentaler Eigenschaften, sondern verknüpft die Möglichkeit zu multiplen Realisierungen mentaler (oder anderer höherstufiger) Eigenschaften grundsätzlich mit der Existenzberechtigung und Nicht-Reduzierbarkeit der höherstufigen Spezialwissenschaften wie der Psychologie.

If only physical particles weren't so small (if only brains were on the *outside*, where one can get a look at them), *then* we would do physics instead of paleontology (neurology instead of psychology; psychology instead of economics; and so on down). There is an epistemological reply; namely, that even if brains were out where they can be looked *at*, as things now stand, we wouldn't know what to look *for*: we lack the appropriate theoretical apparatus for the psychological taxonomy of neurological events. If it turns out that the functional decomposition of the nervous system corresponds to its neurological (anatomical, biochemical, physical) decomposition, then there are only epistemological reasons for studying the former instead of the latter. But suppose there is no such correspondence? Suppose the functional organization of the nervous system crosscuts its neurological organization (so that quite different neurological structures can subservise identical psychological functions across times or across organisms). Then the existence of psychology depends not on the fact that neurons are so sadly small, but rather on the fact that neurology does not posit the natural kinds that psychology requires. (ebd., S. 113; Hervorh. im Orig.)

Wenn sich Fodor gegen mikroreduktionistische Bestrebungen wendet, kämpft er hierbei hintergründig vor allem gegen die Möglichkeit einer Nagel-Reduktion der Psychologie auf die Neurobiologie oder Physik, indem er darauf verweist, dass höherstufige Eigenschaften wie mentale

Eigenschaften durch verschiedene physische Eigenschaften realisierbar sind und daher die benötigten Brückenprinzipien nicht etablierbar sind. Statt einer klaren Zuordnung von neurobiologischen oder physischen Typen zu psychologischen Typen, würde ein jeder mentale Typ immer mit unsystematischen, wissenschaftlich und ontologisch fragwürdigen Disjunktionen neurobiologischer (resp. physischer) Typen verbunden sein: „The problem all along has been that there is an open empirical possibility that what corresponds to the natural kind predicates of a reduced science may be a heterogeneous and unsystematic disjunction of predicates in the reducing science, and we do not want the unity of science to be prejudiced by this possibility“ (ebd., S. 108). Die multiple Realisierbarkeit höherstufiger Eigenschaften, die als (Kernstück der) Negation einer psychophysischen Identitätsannahme aufgefasst wird, wird somit auf verhängnisvolle Weise zu einem Garant der Existenzberechtigung der höherstufigen bzw. Spezialwissenschaften. Wer identitätstheoretische Annahmen bejaht, läuft unter diesen Voraussetzungen daher Gefahr, die Psychologie als eigenständige Wissenschaft preiszugeben.

Was nach Fodor entscheidend ist, ist nicht der „Stoff“, aus dem etwas gemacht ist, sondern vielmehr die Legitimität oder *Autonomie* einer Wissenschaft bei der Erklärung von Phänomenen. Wobei es sich um eine Auffassung handelt, die sich nicht nur in der Philosophie (der Psychologie) bzw. Wissenschaftstheorie wiederfindet, sondern auch in der Psychologie selbst. Mausfeld (2010) z.B. sagt im Zusammenhang der Ablehnung eines Erklärungsprimats der Neurobiologie gegenüber der Psychologie: „Die Psychologie ist explanatorisch autonom (wie auch in der Biologie und den Naturwissenschaften allgemein Erklärungen auf höherer Analyseebene explanatorisch autonom sind)“ (S. 183).

Wie Polger (2004) in Anlehnung an Antony und Levine (1997) herausstellt, ist eine Erklärung oder Theorie nach diesem Verständnis nur dann

autonom, „if it characterizes a unique regularity, that is, a regularity that cannot be captured in other terms“ (S. 191). Um die Realität mentaler Zustände und Eigenschaften zu belegen, muss man dann zeigen, dass psychologische Erklärungen nicht nur nützlich sind, sondern zwingend notwendig. Dabei wurde der Funktionalismus als *das* Paradigma auserkoren, das nicht nur der multiplen Realisierbarkeit mentaler Eigenschaften Rechnung tragen soll, sondern auch die „Autonomie“ der Psychologie abzusichern vermag: „If we are confident in the reality of mental states as described by psychology then we conclude that psychological explanation is autonomous. And if psychology is autonomous then we can use that fact to argue for functionalism“ (Polger, 2004., S. 192).

Im Hintergrund dieser Umklammerung des Funktionalismus steht die Annahme, dass die Identitätstheorie (die als dessen einzig ernstzunehmende Alternative angesehen wird) mit der Konsequenz verbunden sei, dass alle psychischen Fakten ohne expliziten Rückgriff auf psychologische Begrifflichkeiten erklärt werden könnten: „If science is to be unified, then all such taxonomies must apply *to the same things*. If physics is to be basic science, then each of these things had better be a physical thing. But it is not further required that the taxonomies which the special sciences employ must themselves reduce to the taxonomy of physics“ (Fodor, 1974, S. 114; Hervorh. im Orig.). Nach dieser Denkweise folgt aus einer ontologischen Reduktion offensichtlich automatisch eine epistemische Reduktion. Obgleich bereits hinreichend gezeigt wurde, dass dies keineswegs der Fall ist, lässt sich dieser Einschätzung etwas abgewinnen, wenn man sie auf das Nagelsche Reduktionsmodell bezieht. Denn mit psychophysischen Identitätsaussagen als Brückenprinzipien erscheint eine direkte Verbindung zwischen psychologischen Theorien und ihren Begrifflichkeiten auf der einen und neurobiologischen (oder physikalischen) Theorien und ihren Begrifflichkeiten *prinzipiell möglich* zu sein. Aus solch

einer Möglichkeit wird nun geschlussfolgert, dass die Identitätstheorie das Vorhaben einer Eliminierung des Mentalen begünstigen würde. Will man eine solche unsägliche Konsequenz vermeiden, müsse man sich dieser Denkweis zufolge auf den Funktionalismus berufen: „If psychological kinds are functional kinds then there is hope for the autonomy of the mental. Because functionalist psychology construes mental states as biologically abstract (multiple realizable) it purports to characterize regularities that cannot be captured by explanations in neuroscientific terms. Functionalism would thus salvage the autonomy of psychology and the reality of mental states and properties“ (Polger, 2004, S. 192). Nach Ansicht Polgers fungiert der Funktionalismus hier nicht einfach als ein (epistemisches) wissenschaftliches Forschungsparadigma, sondern hat vielmehr den Status einer *metaphysischen* Hintergrundtheorie. Wie erläutert, bedeutet dies, dass das eigentliche *Wesen* mentaler Eigenschaften letztendlich darin bestehen würde, funktionaler Natur *und nicht* materieller/stofflicher Natur zu sein (vgl. Kap. 6.5 sowie Kap. 7.5.2.1).

In Anbetracht dieser Argumentationskette zeigt sich, dass es höchstproblematisch ist, wenn Reduktionismus und Identitätstheorie als mehr oder weniger austauschbare Konzepte verstanden werden. Faktisch ist mit einem Reduktionismus etwas ganz anderes intendiert als mit der Formulierung einer psychophysischen Identitätstheorie. Reduktionismus ist ein umfassendes programmatisches Vorgehen, bei dem die Autonomie höherstufiger Wissenschaften wie der Psychologie von vornherein negiert wird. Die Autonomie-These wird wie von Fodor (1997) aber häufig als Zurückweisung einer jeglichen Form von Reduktion formuliert, wodurch die Unterscheidung zwischen einer in der Regel harmlosen (ontologischen) Reduktion und einem kruden Reduktionismus völlig verwischt wird:

Simply to have a convenient way of talking, I will say that a law or theory that figures in bona fide empirical explanations, but that is not reducible to a law or theory of physics, is ipso facto *autonomous*; and that the states whose behavior such laws or theories specify are *functional* states. (In fact; I don't know whether autonomous states are ipso facto functional. For present purposes all that matter is whether functional states are ipso facto autonomous). (S. 149; Hervorh. im Orig.)

Wie Polger (2004) nun schlüssig darlegt, zieht Fodor hier zwar augenscheinlich eine ontologische Schlussfolgerung, obwohl die „Autonomie“ der Psychologie *de facto* primär eines epistemischen bzw. explanatorischen Nachweises bedarf. Dagegen ist die Identitätsthese eine rein ontologische oder metaphysische Aussage. Ihre Gültigkeit hängt keineswegs davon ab, mit welchen anderen explanatorischen Ressourcen sie verknüpft werden kann, obgleich dieser Fehlschluss eine gewisse Verbreitung gefunden haben mag (vgl. ebd., S. 194; Kap. 6.1.4.4 dieser Arbeit). Nur wenn man annimmt, dass die Reduzierbarkeit der Psychologie auf die Neurobiologie die Identitätstheorie mit sich bringt und das Zutreffen der Identitätstheorie die Reduzierbarkeit der Psychologie auf die Neurobiologie sicherstellt, ergibt sich die Konsequenz, dass die multiple Realisierbarkeit mentaler Eigenschaften das letzte Bollwerk gegenüber solch einer Vereinnahmung der Psychologie durch die Biologie und einer völligen Auflösung des Psychischen im Gehirn darstellt. Daher sichert das Argument der multiplen Realisierbarkeit dieser Denkweise zufolge die „Autonomie“ der Psychologie ab, indem es die Möglichkeit einer (Nagel) Reduktion verhindert. Diese Verwirrung führt nun dazu, dass viele Identitätstheoretiker glaubten, argumentieren zu müssen, dass die Psychologie nicht „autonom“ sein kann (denn die Korrektheit der Identitätsaussage ist für sie gesetzt und unter den hier beschriebenen Voraussetzungen mit einer „Autonomie“ der Psychologie unvereinbar), während viele Funktionalisten ganz im Gegenteil glaubten, argumentieren zu müssen,

dass sie „autonom“ ist, weil die Identitätsthese aufgrund der multiplen Realisierbarkeit mentaler Eigenschaften falsch ist. Dabei hängt die Gültigkeit der Identitätsthese als einer *ontologischen* Theorie faktisch keineswegs von der Reduzierbarkeit psychologischer Theorien oder Erklärungen (auf solche der Neurobiologie oder gar Physik) ab, denn hierbei handelt es sich um ein *epistemisches* Manöver (vgl. ebd.).

Durch diese Form der Autonomisierung der Psychologie ergibt sich nun jedenfalls eine vermeintlich vorteilhafte Lage der Psychologie. Sie ist in der Tat *autonom*, weil ihre Theorien und Konzepte bewusst von der naturwissenschaftlichen Forschung auf den basaleren Ebenen abweichen können. Oder wie McCauley (2007) sagt:

The bleak reductive prospects Fodor foresees, in effect, insulate psychological theorizing and research. Psychology enjoys great theoretical and methodological leeway, essentially unencumbered by what might appear to be divergent evidence arising from research at lower levels of analysis. Inquiries at those levels merely concern the details of implementation [...] They will not bear in any integral way on the theories and principles at stake at the higher level. (S. 118)

Eine solche „Selbst-Gesetzgebung“ ist nun aber keineswegs ein wünschenswerter Zustand für eine wissenschaftliche Disziplin. Die Psychologie läuft damit Gefahr, von den anderen wissenschaftlichen Disziplinen isoliert zu werden und bietet – entgegen dem hier zugrunde liegenden Bestreben – Raum für die von den New-Wave-Reduktionisten geforderte Ersetzung psychologischer Theorien durch solche der Neurobiologie oder noch basalerer Wissenschaften.

8.5.2. Mikrodetermination statt psychophysischer Identität

Wie in Kapitel 6.1 erläutert wurde, hat die Identitätsrelation in der Philosophie des Geistes und auch über diese hinaus kein gutes „Image“, da sie

infolge der ungünstigen Interpretation, die sie durch vor allem Jack Smart erfahren hat, mit einem relativ kruden reduktiven Physikalismus gleichgesetzt wurde. Eine psychophysische Identität zu behaupten, heißt demnach, dass das Psychische *nichts anderes als* ein Gehirnprozess ist, dass mentale Eigenschaften *nichts anderes als* neuronale Eigenschaften sind. Wenn nun auch noch Fodors Analyse hinzugezogen wird, die besagt, dass die Autonomie der Psychologie durch die Identitätstheorie infrage gestellt wird, würde die Akzeptanz einer jeder psychophysischen Identitätsannahme daher nicht nur bedeuten, den Geist auf das Gehirn zu reduzieren, sondern auch die Psychologie auf die Neurobiologie. Dass beide Annahmen unberechtigt sind, wurde gezeigt.

Wenn man den Einschätzungen Fodors dennoch folgt und eine psychophysische Identität ablehnt, um vermeintlich die „ontologische Eigenständigkeit“ und kausale Wirksamkeit mentaler Eigenschaften sowie die „Autonomie“ der Psychologie abzusichern, gleichzeitig aber das Fundament naturwissenschaftlicher Grundprinzipien wie das Prinzip der kausalen Geschlossenheit der physischen Welt nicht verlassen möchte, scheint der einzige Ausweg darin zu bestehen, statt einer Identität „lediglich“ eine gewisse *Abhängigkeit* des Psychischen von seinen physischen Grundlagen zu akzeptieren. So kann der so genannte *nicht-reduktive Physikalismus* als Versuch betrachtet werden, die grundsätzliche Position des Physikalismus anzuerkennen, ohne damit gleichzeitig eine Reduktion des Psychischen auf neuronale (oder allgemein: physische) Aspekte vertreten zu müssen. Sein Anliegen ist also die Klärung der Frage, ob es möglich ist, „eine Antwort auf das Leib-Seele-Problem zu formulieren, die zwar einerseits das Primat des Physikalischen und einen ontologischen Monismus bewahrt, die aber unseren antimaterialistischen Intuitionen so weit wie möglich entgegenkommt, indem sie zeigt, warum das

Mentale auf begrifflicher Ebene aus prinzipiellen Gründen nicht auf das Physikalische reduziert werden kann“ (Metzinger, 2007d, S. 219).

Es sind die Theorien der Supervenienz, Emergenz und einer *nicht-reduktiven* Realisierung (wie sie z.B. von Shoemaker, 2007, oder auch Gillett, 2002, 2003, 2003a, vertreten werden; vgl. Kap. 6.5), die auserwählt wurden, diese Funktion einer minimalen Abhängigkeit des Psychischen von seinen physischen Grundlagen zu erfüllen, die auf der einen Seite kein vollkommen losgelöstes, dem Naturgeschehen gegenüberstehendes Phänomen begründen und auf der anderen Seite aber doch die gefürchtete Annahme einer psychophysischen Identität verhindern soll.

Die Supervenienzrelation beschreibt eine Kovarianz von Eigenschaften, bei der die Makroeigenschaften eines Systems von dessen Mikroeigenschaften abhängen, gibt aber selbst keine Erklärung dafür ab, warum es zu einer solchen Kovarianz kommt (vgl. Kap. 6.3). Dabei impliziert das – über eine bloße systemische Emergenz hinausgehende – Konzept einer starken synchronen Emergenz ebenfalls die Gültigkeit der Supervenienz-These und liefert darüber hinaus nur eine negative Kennzeichnung der psychophysischen Relation ab (vgl. Kap. 6.4). Daher kann zumindest auf der Basis der Konzepte der Supervenienz und Emergenz allein keine hinlängliche Theorie des Geistes entwickelt werden. Denn obgleich diese Konzepte eine gewisse Beschreibung der psychophysischen Relation abliefern, die von vielen Autoren (auch in der Psychologie) herangezogen wird, und ebenfalls im Hinblick auf die mereologische Supervenienz und systemische Emergenz zur Kennzeichnung der hier entwickelten mereologisch-mechanistisch basierten Identitätstheorie hilfreich ist, ermöglicht der alleinige Rekurs auf diese Konzepte keineswegs die Formulierung einer begrifflich konsistenten und gleichzeitig inhaltlich gehaltvollen Theorie zur Bestimmung der ontologischen Beziehung zwischen den natürlichen Organisations-ebenen des „Gehirn-Geist-Komplexes“.

Daher sahen sich die nicht-reduktiven Physikalisten dazu genötigt, das Konzept der (nicht-reduktiven) Realisierung hinzuzuziehen, um eine solche weitergehende Bestimmung abliefern zu können. Einer solchen Realisierungstheorie zufolge sind mentale Eigenschaften nicht mit physischen Eigenschaften identisch, um vermeintlich hierdurch ihre ontologische „Eigenständigkeit“ zu bewahren, aber doch durch sie realisiert, wodurch sie in den Bereich naturwissenschaftlich erklärbarer Phänomene eingereiht werden sollen und zugleich ihre kausale Wirksamkeit – trotz der Akzeptanz der kausalen Geschlossenheitsannahme der physischen Welt – gerettet werden soll.

Wenn die Realisierungsrelation das Problem mentaler Verursachung lösen soll, muss sie so expliziert werden, dass verständlich wird, dass mentale Eigenschaften ungeachtet ihrer Eigenständigkeit ontologisch so eng an ihre physikalischen Realisierer gebunden sind, dass die Intuition einer ‚Konkurrenz‘ nicht greift. Eine Theorie der Realisierungsrelation muss daher deutlich machen, dass eine mentale Eigenschaft und ihr physikalischer Realisierer zwar *zwei* Eigenschaften sind, aber eben nicht zwei Eigenschaften, die ontologisch so verschieden sind, dass sie überhaupt in irgendeinem Sinn miteinander konkurrieren können. (Walter, 2009, S. 696; Hervorh. im Orig.)

Während im Zusammenhang der Supervenienzrelation und damit auch im Hinblick auf das Konzept der Emergenz potentiell eine *wechselseitige Abhängigkeit* zwischen den Makro- und Mikroeigenschaften eines Systems angenommen wird (denn hier wird eben nur eine *Kovariation* zweier Ereignisreihen postuliert), die nur durch Zusatzannahmen auf eine reine Mikro-Makro-Determination festgelegt werden kann (und in ebendieser Weise auch für gewöhnlich festgelegt wird, da eine wechselseitige Abhängigkeit dem allgemeinen Wissenschaftsverständnis zuwider zu laufen scheint), impliziert eine (nicht-reduktive) Realisierungsrelation bereits mit begrifflicher Notwendigkeit eine *einseitige Abhängigkeit* des Mentalen

von seinen physischen Grundlagen: „Das Realisierungsideom bringt dabei eine *notwendige und asymmetrische Abhängigkeitsbeziehung* zwischen mentalen und physikalischen Eigenschaften zum Ausdruck: die Instantiierung einer physikalischen Eigenschaft impliziert (nomologisch oder metaphysisch) notwendig die Instantiierung einer von ihr realisierten mentalen Eigenschaft, nicht aber umgekehrt“ (ebd., 2009a; S. 52; Hervorh.: M.L.).

Wenn so eine psychophysische Identität von vornherein ausgeschlossen wird und trotzdem keine „Konkurrenz“ zwischen mentalen Eigenschaften und ihren Realisierern bestehen soll, ist der einzige Ausweg daher in der Tat, eine *einseitige Abhängigkeit* der zu realisierenden Makroeigenschaften eines Systems von den Mikroeigenschaften des Systems als dessen physischen Realisierern anzunehmen. Denn unter gleichzeitigem Ausschluss einer psychophysischen Identität *und* einer Mikro-Makrodetermination würde es nichts mehr geben, was den Geist an seine materiellen Grundlagen binden oder die Möglichkeit mentaler Verursachung erklären könnte. Man müsste dann notgedrungen, wie dies Leibniz tat, so etwas wie eine prästabilisierte Harmonie der mentalen und physischen Ereignisreihe annehmen, um einen solchen Standpunkt aufrechtzuerhalten.

Was im Folgenden beleuchtet werden soll, ist, dass der nicht-reduktive Physikalismus in einer ganz bestimmten Hinsicht problematisch ist, nämlich, dass er statt der beabsichtigten „Rettung“ des Psychischen vor allem im Hinblick auf dessen kausale Wirkmächtigkeit mit einer immensen Schwächung oder gar Auflösung ebendieser Funktion verbunden ist. Der Kern des Problems ist, dass der nicht-reduktive Physikalismus einen so genannten *Mikrophysikalismus* impliziert und zwar in dem Sinne, dass es die Mikroebenen sind, die die Makroebenen determinieren, ohne dass die gegenteilige „Wirkrichtung“ angenommen werden kann.

Zunächst muss man einräumen, dass der Mikrophysikalismus in der Tat so etwas wie eine natürliche Konsequenz unseres wissenschaftlichen Weltbildes zu sein scheint, nach dem sich zusammengesetzte Dinge immer nach Maßgabe ihrer Teile verhalten. Alle makroskopischen Entitäten der Welt bestehen demnach aus mikrophysikalischen Entitäten und werden durch mikrophysikalische Gesetze beherrscht (vgl. Pettit, 1993, S. 253). Oder wie Kim (1993) sagt: „The broad metaphysical conviction [...] is the belief that ultimately the world – at least, the physical world – is the way it is because the micro-world is the way it is“ (S. 100).

Wie der Physiker und Wissenschaftstheoretiker Andreas Hüttemann (2004) in seinem brillanten und im Folgenden primär herangezogenen Werk *What Is Wrong With Microphysicalism?* darlegt, geht diese Überzeugung einer Mikrodetermination auf die Vergegenwärtigung der (natur-)wissenschaftlichen Erklärungspraxis zurück, nach der das Sosein von Makroeigenschaften eines Systems für gewöhnlich mit Bezug auf das Sosein der Mikroeigenschaften dieses Systems erklärt wird (also ganz im Sinne des Modells reduktiver Erklärung, vgl. Kap. 8.4): „The sciences are highly successful at providing micro-explanations. We explain why chalk is white in terms of the possible energy states of the molecules that chalk consists of. The specific heat of mental can be understood on the basis of the behaviour of the ions and the electrons, i.e. on the basis of the behaviour of its parts. The success of micro-explanation seems to be evidence for the truth of microphysicalism“ (S. 1). Insofern beziehen sich die nicht-reduktiven Physikalisten bei der Formulierung einer physikalischen Theorie mit großer Selbstverständlichkeit auf den – ihrer Einschätzung nach – minimalen Nenner eines jeden Physikalismus, die mikrophysikalische Determination. Dies erklärt die obige Aussage Metzingers (2007d), der genau in diesem Sinne den Physikalismus des nicht-reduktiven Physikalismus darin begründet sieht, dass er „den *Primat des Physikalischen*

und einen ontologischen Monismus bewahrt“ (S. 219; Hervorh.: M.L.). Dieser Ansicht nach sollte jede Theorie, die ein höherstufiges Phänomen zu erklären beabsichtigt und dabei beansprucht, eine *physikalistische* Theorie zu sein, einen *Vorrang der Mikroebenen* (z.B. in Form einer einseitigen Mikro-Determinationsbeziehung) als vermeintlich selbstevidenten Eckpfeiler unseres naturwissenschaftlichen Weltbildes beinhalten.

Man kann solche mikrophysikalistischen Überzeugungen in zumindest drei verschiedene Arten einteilen: Es ist die bereits genannte *Mikrodetermination*, nach der das Verhalten oder die Eigenschaften von zusammengesetzten Systemen durch das Verhalten oder die Eigenschaften ihrer Konstituenten bzw. deren struktureller Organisation determiniert wird. Darüber hinaus gibt es ein Verständnis der Beziehung zwischen höher- und niederstufigen Ebenen, das sich als *Mikrobeherrschung* (*micro-government*) bezeichnen lässt, nach der die Gesetze der Mikroebenen auch die Makroebenen von Systemen beherrschen; und es gibt die These der *Mikroverursachung*, nach der jedes Kausalgeschehen kraft der Verursachung auf der Ebene der (kleinsten) Teile, also auf den Mikroebenen, stattfindet (vgl. Hüttemann, 2004, S. 7).

Allerdings gesteht man mit der Akzeptanz eines solchen Mikrophysikalismus den mikrophysikalischen Entitäten unweigerlich eine *ontologische Priorität* zu. Sie sind es, die die Makroeigenschaften determinieren oder beherrschen, während letztere keine solche Funktion aufweisen. Besonders prägnant bringt dies Robert Klee (1984) zum Ausdruck:

Micro-explanation is powerful in virtue of the fact that when a level of organization within a system can be explained in terms of lower-levels of organization this must be because the lower-levels (i.e. the microproperties) determine the higher-levels (i.e. the macro-properties). This is why micro-explanation makes sense — *the direction of explanation recapitulates the direction of determination*. (ebd., S. 59f, zit. n. Hüttemann, 2004, S. 9; Hervorh.: M.L.)

Es wird hier eine ontologische Konsequenz aus dem Vorliegen einer Vielzahl von erfolgreichen Mikroerklärungen gezogen, die Hüttemann (2004) als *explanatorischer Realismus* bezeichnet: „Explanatory realism is the view that some ontological relation has to obtain in nature in order to make explanations work“ (S. 14).

Hiermit ist zugleich die bitterböse Ironie eines *nicht-reduktiven Physikalismus* freigelegt, der von einer einseitigen Mikro-Makro-Determination ausgeht. Angetreten, um eine „Auflösung“ des Geistigen in der materiellen Seinsweise der Welt zu verhindern, die die Vertreter vor allem in einem Reduktionismus ausfindig zu machen glaubten, der unweigerlich von einer psychophysischen Identitätstheorie ausgeht, aber doch bestrebt, grundlegende naturwissenschaftliche Constraints mitzutragen, sind die Verfechter einer solche Position wohl oder übel zu der Annahme gezwungen, den Entitäten und Eigenschaften der Mikroebenen ein ontologisches Primat gegenüber denjenigen der Makroebenen wie eben Geist oder Bewusstsein zuzuordnen. Es ist also gerade ein nicht-reduktiver Physikalismus, der statt einer Identitätsannahme „lediglich“ von einer asymmetrischen Mikro-Makro-Determination (respektive -Beherrschung, respektive -Verursachung) ausgeht, auf dessen Basis die Neurobiologie gegenüber der Psychologie in einer privilegiierteren Position stände, da alle psychischen Prozesse letztendlich vollkommen durch neurobiologische Prozesse determiniert sein würden.

Und auch die programmatisch zum eigenen Anspruch erhobene „Rettung“ psychophysischer Kausalität im Sinne der Möglichkeit zu einer mentalen Verursachung physischer Ereignisse ist mit dem nicht-reduktiven Physikalismus nur schwer vereinbar, denn wenn das Psychische eben nicht identisch mit seinen physischen Grundlagen ist, bleibt es unklar, wie es kausal wirkungsvoll sein soll, ohne das Prinzip der kausalen Geschlossenheit der physischen Welt bzw. den Energieerhaltungssatz zu

verletzen. Wie Walter (2008, 2009a) schlüssig dargelegt hat und auch im Rahmen dieser Arbeit in Bezug auf die nicht-reduktive Realisierungstheorie nach Schoemaker und Gillett erläutert wurde (vgl. Kap. 6.5.3), bleibt die Möglichkeit einer mentalen Verursachung unter diesen Voraussetzungen nicht nur äußerst mysteriös, sondern es droht der gänzliche *Ausschluss* der kausalen Wirksamkeit mentaler Eigenschaften. Es ist das Kausalgeschehen im Gehirn, das maßgeblich für das (erlebte) Kausalgeschehen in unserem Geist ist. Insofern kommt auch der Erforschung neuronaler oder biochemischer Prozesse ein Vorrang zu, da hier das psychische Geschehen determiniert wird und das Psychische selbst keinen Einfluss auf die Vorgänge im Gehirn hat. Wie klar zu sehen ist, führt dieser Weg sehr nah an eine Sichtweise, die bereits unter dem Stichwort *Epiphenomenalismus* besprochen wurde (vgl. Kap. 3.1.2), nach der alles Geistige letztendlich nur Begleiterscheinungen von Gehirnvorgängen sind. Die bittere Ironie des nicht-reduktiven Physikalismus besteht also darin, dass über die Verweigerung einer Identitätsannahme gerade das Gegenteil dessen erreicht wird, was diese Position eigentlich erreichen wollte: Statt eines „autonomen“, ontologisch eigenständigen und kausal wirksamen Phänomens wird das Psychische zu etwas, was durch seine physischen Grundlagen restlos determiniert und dem die Möglichkeit einer kausalen Wirkmächtigkeit gänzlich genommen wird.

8.6. Warum ein ontologisches Primat der Mikroebenen scheidet

Im Folgenden wird es nicht einfach darum gehen aufzuzeigen, dass eine einseitige Mikroabhängigkeit etwa für die Psychologie nicht Wünschenswertes ist und daher vermieden werden sollte, sondern dass eine solche

Interpretation der Mikro-Makro-Beziehung schlichtweg falsch ist. Es mag in der Tat so etwas wie eine zumindest implizite Gewissheit sein, dass den agierenden Entitäten bzw. Eigenschaften niedrigerer Ebenen in gewisser Hinsicht ein höherer Wirklichkeitsstatus zukommt als denen höherer Ebenen. Diese Intuition dürfte sich vor allem aus einem Weltverständnis speisen, das vermeintlich durch die Wissenschaft nahegelegt wird. So scheint es der Art und Weise der Erkenntnisgewinnung der Wissenschaften zu entsprechen, dass sie das Auftreten und die Beschaffenheit höherstufiger (Alltags-)Phänomene verständlich macht, indem sie auf mikrostrukturelle Gegebenheiten verweist. Wenn wir – wie in den von Hüttemann (2004) genannten Beispielen – die Farbeigenschaften von Kalk mittels der energetischen Zustände seiner Moleküle erklären können, die Wärmeeigenschaft von bestimmten Metallen mittels des Verhaltens ihrer Ionen und Elektronen und des Weiteren auch die Aggregatzustände von Stoffen (fest, flüssig, gasförmig) mittels der spezifischen Lage- und Bewegungseigenschaften ihrer Teilchen (Moleküle oder Atome), dann scheint klar zu sein, dass ebendiese Mikroeigenschaften eines Systems dessen Makroeigenschaften *determinieren*. Dass es sich hierbei um eine *asymmetrische* Beziehung in Form einer einseitigen Mikro-Makro-Abhängigkeit handelt, wird offensichtlich dadurch verbürgt, dass es in den genannten Fällen absurd wäre anzunehmen, dass es die Farbe des Kalks ist, die die energetischen Zustände der Kalkmoleküle determiniert oder auch der Aggregatzustand eines Stoffes die Lage- und Bewegungseigenschaften seiner Teile. Somit hätten die agierenden Entitäten bzw. Eigenschaften der Mikroebenen gegenüber denen der Makroebenen und damit wohl auch die Biologie gegenüber der Psychologie, die Chemie gegenüber der Biologie und schließlich auch die (Mikro-)Physik gegenüber der Chemie einen ontologischen Vorrang.

Dabei ist – wie Hüttemann (2004, 2009) eindrücklich und sehr detailliert nachweist und im Folgenden in wesentlichen Zügen nachgezeichnet wird – eine solche ontologische Priorität der Mikroebenen bei genauerer Betrachtung nicht zu halten. Der Erfolg der Mikroerklärungen hängt zwar in der Tat davon ab, dass es eine Mikrodetermination zwischen den Teilen und dem Ganzen gibt. Aber es ist eben keine ausschließliche Determination des Ganzen durch seine Teile, sondern eine *wechselseitige Determination* zwischen den Teilen und dem Ganzen. „Under the assumption that physics provide a complete description of the relation of parts and wholes we are forced to the conclusion that a *mutual dependence* of parts and wholes is all there is. *A fortiori*, there is no ontological priority of either the micro- or the macro-level“ (ebd., S. 2; Hervorh.: M.L.). Obwohl es potentiell durchgehend möglich sein könnte, Mikroerklärungen für alle höherstufigen Ereignisse oder Eigenschaften zu erhalten, können hieraus keine ontologischen Schlussfolgerungen gezogen werden.

8.6.1. Zur Vermengung von Physikalismus und Mikrophysikalismus

Ein grundlegender Irrtum ist bereits die Annahme, dass ein „Physikalismus“ bzw. identitätstheoretische Positionen immer auch ein „Mikrophysikalismus“ sein muss. Tatsächlich wird in der philosophischen Literatur zwischen den Konzepten „Physikalismus“ und „Mikrophysikalismus“ kaum differenziert, wie sich bereits im Kontext der Besprechung der multidimensionalen Realisierungstheorie Gilletts (2003a) gezeigt hat (vgl. Kap. 6.5.3 dieser Arbeit) und auch durch das folgende Zitat von Campbell und Bickhard (2011) verdeutlicht wird:

The physicalist claim is that all facts obtain *in virtue of* the distribution of the fundamental entities and properties – whatever they turn out to be – of completed *fundamental physics*. These fundamental entities are basic particulars of some sort: bits of

matter, concrete events, or whatever, which are the fundamental constituents out of which everything in the world is composed, and whose properties and relations are sufficient to determine everything that is true about the world. (S. 45; Hervorh. im Orig.)

Physikalismus bzw. eine psychophysische Identität und ein Mikrophysikalismus werden hier offensichtlich als voneinander untrennbare Konzepte aufgefasst. Eine Bejahung einer psychophysischen Identität bringt demnach notwendigerweise einen ontologischen Vorrang der Mikroebenen mit sich. Eine Ansicht, die wiederum auf die Interpretation der „klassischen“ Identitätstheorie zurückgeht, gemäß der mentale Eigenschaften eben *nichts anderes als* neuronale Eigenschaften sind.

The reason why the identification of mental with physical properties is perceived to have far-reaching consequences for the mental is our belief that the mental somehow *falls prey* to the physical. By identifying a mental property with a physical property the mental becomes subject to the dictatorship of the proletariat. Suddenly, the mental no longer lives a life of its own, it is governed by physical laws. It seems to me that these claims express the core materialist or physicalist intuitions. (Hüttemann, 2004, S. 61; Hervorh. im Orig.)

Dabei sind die Annahmen eines Mikrophysikalismus versus einer psychophysischen Identität *de facto* unabhängig voneinander: „You can embrace identity-physicalism without subscribing to micro-physicalism. The relation of parts and wholes and the relation between different kinds of properties of wholes are simply two separate issues. Therefore, our rejection of microphysicalism does not extend to identity-physicalism“ (Hüttemann, 2004, S. 2). Die Identitätstheorie, auch in ihrer „klassischen“ Formulierung, ist vollkommen unabhängig von der Annahme eines Primats der Mikroebenen. Man muss an dieser Stelle allerdings beachten, dass sich Hüttemann, wenn er sich auf das Verhältnis von Gehirn und Geist bezieht, primär an den begrifflichen Bestimmungen Jaegwon

Kims orientiert, der zwischen einerseits mikrobasierten Eigenschaften wie den neuronalen Realisierern einer mentalen Eigenschaft und andererseits Makroeigenschaften (zweiter Ordnung) unterscheidet, die selbst nicht mikrobasierte Eigenschaften sind und für die die Frage der Identität mit einer mikrobasierten Eigenschaft (auf dem Niveau von Eigenschaftsinstantiierungen) geklärt werden müsste.

Gemäß der Argumentationsweise der vorliegenden Arbeit sind mentale Eigenschaften dagegen *selbst* (und nicht nur über ihre Beziehung zu ihren neuronalen Realisierern) mikrobasierte Eigenschaften, die auf einer eigenen höherstufigen Organisationsebene auftreten, die durch neuronale Netzwerke konstituiert wird. Trotz dieser Diskrepanz ist der Befund Hüttemanns, dass aus einer psychophysischen Identitätsannahme keineswegs ein Primat der Mikroebenen erfolgt bzw. Physikalismus und Mikrophysikalismus begrifflich voneinander unabhängig sind, weiterhin korrekt. Denn über die Annahme einer Identität zwischen einer mentalen Eigenschaftsinstantiierung mit ihrem neuronalen Realisierer wird die mentale Eigenschaft – obgleich auf indirektem Wege – auch nach dem Modell Kims zu einer mikrobasierten Eigenschaft. Der Unterschied bestünde dann nur darin, dass mentale Eigenschaften als Eigenschaften zweiter Ordnung auf derselben Ebene wie neuronale Eigenschaften angesiedelt sind und keine eigene natürliche Organisationsebene bilden (wobei es mysteriös bleibt, was es überhaupt heißen soll, eine Eigenschaft zweiter Ordnung zu sein; vgl. Kap. 6.5.4.2).

Auch wenn gemäß der hiesigen Konzeption mentale Eigenschaften mit den (mikro-)strukturellen Konfigurationen der agierenden Entitäten einer mereologisch-mechanistischen Hierarchie identisch sind, ist auch hiermit noch kein Mikrophysikalismus begründet. Vielmehr müsste hierfür nachgewiesen werden, dass zwischen den verschiedenen Ebenen eines Systems (eines Mechanismus), die durch eine Teil-Ganzes-Bezie-

hung charakterisiert sind, eine *asymmetrische* Abhängigkeit der Makroebenen von den Mikroebenen besteht.

The identification of mental (or biological or solid-state physical) with (micro-) physical properties is at best a necessary condition for such a claim. It furthermore depends crucially on whether or not micro-determination, micro-government and microcausation turn out to be true. It depends on whether the behaviour or property of the complicated compound system that the mental, biological or solid-state physical property has been identified with is indeed determined by the behaviour of the constituents but not vice versa, and on whether the real causal agents - the hidden springs - are the constituents rather than the compound itself. It is one thing to identify mental properties (or biological properties or the properties of many-particle physics) with micro-based properties of the same system. The really interesting question is what we may legitimately assert about the properties of this system on the basis of micro-explanation [...] and whether that forces us to revise our conception of the mental, the biological, etc. (Hüttemann, 2004, S. 61)

Es bedarf nunmehr also der expliziten Betrachtung, ob man auf der Basis der potentiell durchgehenden Möglichkeit von Mikroerklärungen den ontologischen Schluss ziehen kann, dass zwischen dem Ganzen und seinen Teilen eines (mechanistischen) Systems, zwischen der „Ebene des Geistes“ und den verschiedenen Organisationsebenen, die unser Gehirn konstituieren, tatsächlich ein Verhältnis der Mikrodetermination, Mikrobeherrschung oder der Mikroverursachung besteht. Wenn sich herausstellen sollte, dass dies nicht der Fall ist, hat die Identität von mentalen Eigenschaften mit mikrobasierten Eigenschaften des Gehirns keineswegs die Auswirkungen, die für gewöhnlich angenommen werden (vgl. ebd., S. 61f). Das heißt, selbst wenn mentale Eigenschaften (qua ihrer selbst oder qua ihrer Realisierung durch neuronale Eigenschaften) mikrobasierte Eigenschaften sind, brächte dies kein Primat der Mikroebenen mit sich.

8.6.2. Zum Scheitern einer unidirektionalen Mikrodeterrmination

Häufig wird die Gültigkeit einer Mikrodeterrmination bereits mit der Gültigkeit der (mereologischen) Supervenienz begründet. Dabei ist jede Supervenienzrelation, und so auch die mereologische Supervenienz, eine *symmetrische* Relation (vgl. Kap. 6.3.1). Wenn keine Veränderung eines Ganzen erfolgen kann, ohne eine Veränderung seiner Teile, dann kann es auch keine Veränderung ebendieser Teile geben, ohne dass sich zugleich das Ganze verändert. Eine Vorherrschaft des Mikrophysikalischen lässt sich hierüber nicht begründen (vgl. Hüttemann, S. 63).

Als zweiter vermeintlicher Beleg für die Gültigkeit der Mikrodeterrmination wird der bereits erwähnte *explanatorische Realismus* herangezogen. Da es vor allem Klee (1984) war, der ganz explizit davon sprach, dass die Richtung der Erklärung eines Systems zugleich die Richtung der Deterrmination dieses Systems rekapitulieren würde (vgl. ebd., S. 59f), nennt Hüttemann (2004) diese Überzeugung *Klees Diktum* (vgl. ebd., S. 71).

Obgleich die Annahme, dass der Möglichkeit einer korrekten Mikroerklärung immer auch eine Mikrodeterrmination des fraglichen Objektes zugrunde liegt, nicht völlig abwegig ist, folgt hieraus keineswegs, dass es sich hierbei um eine asymmetrische bzw. unidirektionale Mikro-Makro-Deterrmination handelt. Hüttemann (2004) erläutert dies unter anderem am Beispiel des Boyle-Charles-Gesetz idealer Gase (vgl. ebd., S. 77f). Es beinhaltet die Gleichung $pV = RT$. Es besagt, dass für jeden beliebigen explanatorischen Kontext gilt, dass die Temperatur eines bestimmten Gases einen bestimmten Wert hat und zwar hat es diesen auf der Basis der Werte des in ihm herrschenden Drucks und seines Volumens. Damit basiert eine solche Erklärung einer Deterrminationsbeziehung zwischen Druck und Volumen auf der einen Seite und der Temperatur auf der anderen, wobei die Druck- und Volumenausprägungen die Temperatureigenschaft dieses Gases deterrminieren. Nach Klees Diktum würde es

sich hierbei um eine asymmetrische Abhängigkeit handeln, da die Richtung der Erklärung lediglich von „unten nach oben“ verläuft. Wenn dies der Fall wäre, könnte man in der Tat sagen, dass Druck und Volumen die Temperatur des Gases *hegemonisch* determinieren. Tatsächlich ist dies schlichtweg falsch, was sich anhand der Beziehung zwischen Druck, Volumen und Temperatur, wie sie durch die mathematische Gleichung beschrieben wird, leicht aufzeigen lässt:

Thanks to the equation $pV = RT$ we know that it is also the case that the values of temperature and volume determine that of pressure, or the values of temperature and pressure that of volume. [...] The determination relation does not give rise to hegemony of pressure and volume over temperature simply because the same kind of determination relation obtains between any two of the free variables and the third. The ideal gas law gives rise to all kinds of determination not just to the one that the explanation of temperature in terms of volume and pressure relied on. The determination relation is non-hegemonic. (ebd., S. 78)

Das gleiche Prinzip gilt nun auch für kompositorische Systeme, bei denen eine Teil-Ganzes-Mikroerklärung vorgenommen wird. Genau wie beim Beispiel des idealen Gases liegt hier eine bidirektionale Beziehung vor, was Hüttemann (2009) am Beispiel eines zusammengesetzten Systems mit der Masse M erläutert (vgl. S. 72f). Dieses bestünde aus drei Teilsystemen mit den Massen m_1 , m_2 und m_3 , womit sich – dank der rein aggregativen Beziehung der Masse-Eigenschaft (vgl. Kap. 5.1.2.2) – ein einfaches „Zusammensetzungsgesetz“ ergibt: $M = m_1 + m_2 + m_3$. M ist dabei unweigerlich festgelegt, wenn für die Teilmassen bestimmte Werte angenommen werden. Daher determinieren die Teilmassen m_1 bis m_3 in der Tat die Gesamtmasse M . „Es gilt aber gleichfalls, dass M , m_1 und m_2 zusammen m_3 determinieren, den Massewert für das dritte Teilsystem. Beliebige drei Werte legen den jeweils vierten fest. Dies verdankt sich dem Zusammensetzungsgesetz [...]. Die vier Werte determinieren sich

wechselseitig. Kein Wert ist ausgezeichnet“ (Hüttemann, 2009, S. 72f; Hervorh.: M.L.). Insofern kann aus dem Vorliegen einer Teil-Ganzes-Beziehung keineswegs geschlossen werden, dass den Teilen die „eigentlichen“ Kausalkräfte zukämen. Dasselbe gilt für alle Mikroerklärungen in der klassischen als auch Quantenmechanik und überhaupt mereologisch verfasste Systeme, auch wenn es sich hierbei nicht um einfache aggregative Verhältnisse zwischen den Ebenen handelt (vgl. ebd., 2004, S. 79ff). Denn obgleich wir bei komplexeren („chaotischen“) Systemen nicht immer ein konkretes „Zusammensetzungsgesetz“ zur Verfügung haben werden oder vielleicht auch nur haben können, bleibt der aus diesen einfachen Systemen ableitbare Fakt bestehen, dass sich das Ganze und die Teile eines Systems *wechselseitig* determinieren. Die Grenzen unseres Erkenntnisvermögens sind keine Bruchstellen der Wirklichkeit, sie schaffen keine neuen ontologischen Fakten.

Obgleich die Bedeutung mereologischer Prinzipien zur Charakterisierung des Gehirn-Geist-Verhältnisses bei ihm eher unklar bleibt, hat im Grunde dennoch bereits Fechner (1851/1922) festgestellt, dass aus einer psychophysischen Identitätsannahme keine einseitige Mikrodetermination abgeleitet werden kann, wenn er – hier zunächst vielleicht noch etwas missverständlich – darlegt: „Man darf [...] den Zusammenhang der Erscheinungen nicht einseitig so darstellen, als müßten die Empfindungen, Gedanken einfach immer so verlaufen, wie es ihnen die Bewegungen im Gehirn, in den Nerven vorgeschrieben wird, woran sie – infolge der Identität des Grundwesens – nun einmal gebunden sind. Das könnte den Anschein erwecken, als wollten wir den Geist aus der Materie herleiten“ (S. 257f). Darauf präzisiert er diesen Gedanken insofern, dass er zwar durchaus einräumt, dass die Abfolge von Gedanken nur so verlaufen kann, wie es die Gehirnprozesse gestatten, jedoch gilt auch, dass ebenso wenig „die Bewegungen im Gehirn anders verlaufen [können], als es die

Gedanken gestatten, woran sie gebunden sind“ (ebd., S. 258). Es ist demnach zwar richtig, dass man, wenn man die neuronale Ebene der „psychophysischen Einheit“ betrachtet, einen (potentiell) mittels naturwissenschaftlichen Gesetzen beschreibbaren Prozess vor sich hat. Dieselbe Aktivität wäre allerdings aus der Perspektive desjenigen Menschen, der diesen Hirnprozess aufweist, beispielsweise eine geistig verfasste Entscheidung zu einer bestimmten Handlung (auf die neuronalen, molekularen Ebenen hat derjenige aufgrund der phänomenalen Transparenz mentaler Zustände keinen repräsentationalen – d.h. intern darstellbaren – Zugriff). Wie in Kapitel 7.4 aufgezeigt, vollziehen sich die Aktivitäten auf beiden bzw. allen ontologisch gleichwertigen Ebenen in *Gleichzeitigkeit*. Es ist ein und dasselbe Kausalgeschehen, das alle Ebenen der mereologisch-mechanistischen Hierarchie durchdringt, so dass man auch allen beteiligten Ebenen dieselbe Wirkmächtigkeit zuschreiben muss.

Zusammenfassend kann man daher sagen, dass Klees Diktum, dass die Richtung einer Erklärung die Richtung der Determination eines Objektes rekapituliert, im Falle von Mikroerklärungen kompositorisch verfasster Systeme falsch ist. Die zugrunde liegende Determinationsrelation ist *ungerichtet*. Insofern sind das Vorhandensein und der wissenschaftliche Stellenwert von Mikroerklärungen kein Beleg für die Gültigkeit einer Mikrodeterminationsbeziehung im Sinne einer hegemonischen Mikro-Makro-Relation. Vielmehr scheint es sich bei dieser Überzeugung um eine gewisse kognitive Verzerrung zu handeln: „The view that the behaviour of the compound 'results' from the behaviour of the parts can be seen as a reflection of how we usually explain the behaviour of compound systems but not as an adequate representation of an underlying ontological quasi-causal relation between parts and wholes. The fact that a mutual determination relation obtains between parts and wholes implies that micro-determination is false“ (Hüttemann, 2004, S. 86). Auf diesen

Aspekt einer kognitiven Verzerrung als Ursache der Annahme eines Primats der Mikroebenen werde ich an späterer Stelle noch ausführlicher eingehen (vgl. Kap. 8.7.2).

8.6.3. Zum Scheitern des Primats mikrophysikalischer Gesetze

Auch wenn aus der Verfügbarkeit und Aussagekraft von Mikroerklärungen keine einseitige Mikrodetmination abgeleitet werden kann, könnte man – wie Hüttemann (2004) darlegt – immer noch versucht sein, eine Vorherrschaft des Mikrophysikalischen abzuleiten, indem man darauf insistiert, dass es die *Gesetze* der Mikroebenen sind, die letztendlich alles beherrschen, während die Gesetze, Regelmäßigkeiten oder Zusammenhänge, die makroskopische Spezialwissenschaften wie die Psychologie behandeln, die Gesetze der Mikroebenen weder ergänzen noch unabhängig von diesen sind. Wie Pettit (1993) sagt, haben sie nicht das Potential „to conflict with them and they do not serve to reinforce them, representing an extra booster for sequences of events that are established in accordance with those laws“ (S. 217).

Hüttemann (2004) erläutert die hier aufgeworfene Frage am System eines physikalischen Halbleitersystems:

A particular semi-conductor, for instance, is realized by a compound system consisting of silicon atoms. This system is a realization of a semi-conductor because it can be shown that the characteristic features of the semi-conductor are due to the microstructure of the system, which consists of silicon atoms. Does this identification imply that the laws that pertain to the silicon atoms govern the semi-conductor's behaviour? Does it imply that the atoms are the real causal agents rather than the semi-conductor as a whole? (S. 59; Hervorh.: M.L.)

Wie sich im Folgenden zeigen wird, bestehen nicht nur gewichtige Gründe, an einer solchen Vorherrschaft des Mikrophysikalischen zu zweifeln,

sondern es lässt sich sogar infrage stellen, dass mikrophysikalische Gesetze in komplexen Systemen überhaupt instantiiert sind.

8.6.3.1. Kein ontologischer Vorrang für Mikrogesetze

Wenn man nun versucht genauer auszubuchstabieren, wie eine solche Vorherrschaft des Mikrophysikalischen in Form von alles beherrschenden Mikrogesetzen aussehen könnte, finden sich hierbei zumindest vier Interpretationsmöglichkeiten.

Eine erste wäre, dass die Aussage, dass Mikrogesetze alles beherrschen, so gemeint ist, dass Mikro-gesetze alles *determinieren* (vgl. Hüttemann, 2004, S. 89). Einschätzungen dieser Art finden sich in der philosophischen Literatur tatsächlich häufig vor. So räumt z.B. Pauen (1999) der Determinationsrelation denselben wissenschaftlichen Status ein wie dem Prinzip der kausalen Geschlossenheit der physischen Welt. Das Prinzip der physischen Determination würde demnach besagen, dass „*alle positiven Fakten* in unserer Welt *durch physische Fakten determiniert* werden, daß es hier also keine Veränderungen gibt, die nicht auch eine Entsprechung in *physikalisch beschreibbaren* Prozessen haben“ (ebd., S. 44; Hervorh.: M.L.). Allerdings kann dieser Ansicht schnell widersprochen werden, insofern sie als eine gerichtete Mikro-Makro-Determination aufgefasst wird. Denn wie bereits gezeigt wurde, ist die Determinationsrelation in mereologisch verfassten Systemen notwendigerweise von einer *wechselseitigen* Natur.

Eine zweite Möglichkeit zur Verteidigung des Mikrophysikalismus im Sinne einer Mikrobeherrschung wäre darauf zu insistieren, dass Mikrogesetze die *einzigsten* Gesetze sind, die überhaupt von Systemen instantiiert werden. Makrogesetze (womit im Folgenden auch die gesetzesartigen Zusammenhänge, Abhängigkeiten etc. bezeichnet werden, wie sie von

der Psychologie postuliert werden) seien dagegen faktisch nicht instantiiert. In der Beantwortung dieses Arguments verdeutlicht Hüttemann zunächst den simplen Grundsatz, dass ein Gesetz immer dann als instantiiert betrachtet werden kann, wenn das Verhalten, das es beschreibt, von einem beliebigen System gezeigt wird. Dabei gelten Gesetze genau genommen nur für völlig *isolierte* Systeme, also unter (kontrafaktischen) Idealbedingungen ohne irgendwelche „Störfaktoren“ (vgl. ebd., 2004, S. 16ff). Wenn nun die zu erklärende Größe von Mikroerklärungen das (gesetzesartige) Verhalten von kompositorischen Systemen ist, dann erklären diese das Verhalten eines Gesamtsystems unter Rückgriff auf deren Teile. Mikroerklärungen setzen dabei eine Abstraktion von „Störfaktoren“ voraus, denn sie blenden bei der Erklärung des Verhaltens des Ganzen externale Faktoren oder Einbettungsbedingungen des jeweiligen Systems aus. Zugleich wird zu erklärende Verhalten des makroskopischen Gesamtsystems als eine *manifeste* Größe betrachtet, was dann aber nach dem oben genannten Grundsatz auch bedeutet, dass das System die jeweiligen *makroskopischen* „Gesetze“ instantiiert haben muss, die genau dieses Verhalten beschreiben. Damit kann aus der explanatorischen Praxis der Wissenschaften nicht mehr geschlossen werden, dass generell keine Makrogesetze in irgendwelchen Systemen instantiiert werden (vgl. ebd., S. 90). Wie sich noch zeigen wird, muss sogar eher daran gezweifelt werden, dass gerade *mikrophysikalische* Gesetze in komplexen kompositorischen Systemen überhaupt instantiiert sind.

Die dritte Interpretationsmöglichkeit der Mikrobeherrschung bezieht sich auf *Dispositionen* eines Systems. Demnach bedeutet die Behauptung, dass Mikrogesetze alles beherrschen, dass die einzig wirklich existierenden Dispositionen die der (mikroskopischen) Teile eines Systems sind und sich die Teile daher gemäß diesen mikrogesetzlich beschreibbaren Dispositionen verhalten. Wenn dem so ist, dann gibt es keine Disposition

des Ganzen, sich gemäß makroskopischen Gesetzen zu verhalten. Wenn man dies nun untersucht, muss zunächst in Erinnerung gerufen werden, dass im Zuge von Mikroerklärungen angenommen werden muss, dass das (gesetzesartige) Verhalten des kompositorischen Systems tatsächlich beobachtbar ist. Wenn es sich hierbei also um ein manifestes Verhalten handelt, dann gibt es einen guten Grund für die Annahme, dass das Ganze die Disposition hat, sich nach Maßgabe des *makroskopischen* Gesetzes zu verhalten: „Successful micro-explanations presuppose two things: the disposition of the subsystems to accord with the micro-laws if the subsystems are isolated, and the disposition of the compound to accord with the macro-law. (The latter disposition is manifest.)” (ebd.). Damit ist dann aber auch aufgezeigt, dass die Interpretation der Mikrobeherrschungsmetapher in Form von bestimmten Dispositionen der fraglichen Systeme das Ziel verfehlt, den Mikroebenen eine ontologische Priorität zuordnen zu können.

Die vierte Interpretationsweise ist etwas komplexer bzw. voraussetzungsreicher. Sie besagt, dass die einzigen überhaupt existierenden Gesetze mereologisch basierte Mikrogesetze sind. Um das Argument zu entfalten, bezieht sich Hüttemann zunächst auf zwei Definitionen und Axiome von David Lewis, die als Grundannahmen der Mereologie als eines rein *logisch-mathematischen* Systems anzusehen sind:

Definition: x and y *overlap* iff they have some common part. Iff not, they are (*entirely*) *distinct*.

Definition: Something is a *fusion* of some things iff it has all of them as parts and has no part that is distinct from each of them. [...]

Transitivity: If x is a part of some part of y, then x is a part of y.

Unrestricted Composition: Whenever there are some things, then there exists a fusion of those things.

Uniqueness of Composition: It never happens that the same things have two different fusions. (Lewis, 1991, S. 73f, zit. n. Hüttemann, 2004, S. 90f; Hervorh. im Orig.)

Hierzu ist anzumerken, dass eine Fusion in diesem Zusammenhang als eine bloße *Aggregation* angesehen wird (vgl. Kap. 5.1.2.2). Zum Problem werden diese Bestimmungen nun, wenn Lewis ableitet, dass wir, wenn wir auf die Existenz aller möglichen Arten von mereologischen Fusionen festgelegt sind, offenbar folgende Konsequenz ziehen müssten: „But given a prior commitment to cats, say, a commitment to cat-fusions is not a *further* commitment. The fusion is nothing over and above the cats that compose it“ (Lewis, 1991, S. 81; zit. n. Hüttemann, 2004, S. 91; Hervorh. im Orig.). Insofern könnte man ableiten, dass wenn man sich zur Annahme der Existenz des Verhaltens eines zusammengesetzten Systems verpflichtet, sich letztendlich auf nichts anderes verpflichtet als auf das Verhalten von dessen Teilen. Die Aussage „Mikrogesetze beherrschen alles“ besagt somit in diesem Fall, dass „a commitment to macro-laws is no further commitment“ (Hüttemann, 2004, S. 91).

Hüttemann stellt heraus, dass sich an dieser Stelle zwei verschiedene Lesarten dieses mereologischen Argumentes ergeben. Nach der ersten Lesart gehören „Zusammensetzungsgesetze“ nicht zur Gruppe der Mikrogesetze, denn Mikrogesetze werden als die Gesetze definiert, die beschreiben, wie sich Subsysteme kompositorischer Systeme unter Isolationsbedingungen verhalten würden. Allerdings ist unter diesen Voraussetzungen die Behauptung, dass Makrogesetze keine weitere Verpflichtung über die Annahme dieser Mikrogesetze hinaus darstellen würden, falsch, denn ein Makrogesetz ist nicht einfach eine Menge von Mikrogesetzen. Nach Hüttemann beinhaltet ein solches auch die „Zusammensetzungsgesetze“, also Informationen darüber, wie sich die verschiedenen Subsysteme einbringen und im Gesamtsystem aufgehen. Insofern gilt, „there is more to a physical compound than the subsystems or constituents merely sitting next to one another. Physical composition is more intimate, so to speak, and that is the reason why we need laws of compo-

sition that tell us how these systems' behaviour combine. Thus, in considering a system as a compound physical system we are not merely conceiving the behaviour of the parts, we *furthermore* assert how their behaviour combines" (ebd.; Hervorh. im Orig.). Dass es sich so verhält, wird vor allem durch den mechanistischen Erklärungsansatz verdeutlicht. Denn hier handelt es sich auch niemals um bloße mereologische Komponenten, die isoliert nebeneinander stehen, sondern immer um *agierende* Entitäten, die in einen ebenenspezifischen bzw. Gesamtmechanismus eingebettet sind. Erst aus der Kenntnis der spezifischen Konfiguration bzw. des eine bestimmte Organisationsstruktur aufweisenden Geflechts dieser agierenden Entitäten kann die Leistung des Mechanismus, eine bestimmte (mentale) Funktion hervorzubringen, erklärt werden (vgl. Kap. 5.1.2.3). Dies ist ein entscheidender Unterschied zwischen mereologischen Betrachtungen in rein logisch-mathematischer Hinsicht und mechanistischen Ebenen zur Beschreibung bzw. Erklärung *empirischer* Phänomene, die zwar auch durch Teil-Ganzes-Beziehungen gekennzeichnet sind, deren „Träger“ aber Lebewesen mit einem viel komplexeren Aufbau sind und bei dem Interaktionen der *agierenden* (und eben nicht statisch-abstrakten) Entitäten von entscheidender Bedeutung sind.

Nach der zweiten Leseart akzeptiert der Mikrophysikalist dieses Argument, insistiert nun aber darauf, dass die Mikrogesetze *und* die Kompositionsgesetze zusammen genommen *alle* Gesetze ausmachen, die es gibt. Weiterhin würde es daher keinen Grund zur Annahme der Gültigkeit von Makrogesetzen geben, da diese bestenfalls eine Begleiterscheinung der Mikro- und Zusammensetzungsgesetze wären. Allerdings kann man Mikrophysikalisten dann fragen, warum sie bei der Annahme von Mikrogesetzen plus Kompositionsgesetzen als unhinterfragten Bezugspunkten mit ihrer Argumentation beginnen. Dies ist genau genommen recht willkürlich, denn „[t]hey could just as well start elsewhere and the conclusion

would be that certain micro-laws are no further commitment or that they come for free" (ebd. S. 92). Denn die Zusammensetzungsgesetze erzeugen lediglich eine wechselseitige Determinationsbeziehung zwischen den Teilen und dem Ganzen und können keine Vorherrschaft der Mikroebenen begründen.

Mikrophysikalisten tendieren nun dazu, einen ganz bestimmten Startpunkt ihrer ontologischen Weltbeschreibung heranzuziehen, was nur schlüssig wäre, wenn es sich hierbei um den einzig möglichen Standpunkt handeln würde. Allerdings spricht nichts dagegen, die Makroebenen bzw. Makrogesetze eines Systems als Startpunkt heranzuziehen. „If the micro-laws (plus the laws of composition and those of interaction) are given, the macro-laws are entailed, determined or involve further commitment. But why start at this point? The determination, entailment and commitment relation fails to distinguish the micro-level as something special" (ebd.).

8.6.3.2. Sind Mikrogesetze in komplexen Systemen überhaupt instantiiert?

Soweit nun also die Widerlegung der These, dass es eine (nomologische) Vorherrschaft des Mikrophysikalischen gibt. Hüttemann geht aber noch einen Schritt weiter und stellt in einer Art Gegenangriff heraus, dass der Rekurs auf Mikroerklärungen sogar die Behauptung rechtfertigt, dass es gerade die *Mikrogesetze* (und nicht die Makrogesetze) sind, die innerhalb eines mikroexplanatorischen Kontexts *nicht* instantiiert sind (vgl. ebd., S. 92ff). Denn wie zuvor erläutert wurde, beschreiben Naturgesetze, wie sich ein physikalisches System unter Idealbedingungen, das heißt, unter Abwesenheit jeglicher „Störfaktoren“ verhalten würde. Innerhalb eines kompositorischen Systems sind solche Bedingungen aber niemals für

dessen Subsysteme gegeben, sodass sich das (gesetzesartige) Verhalten der Subsysteme genau genommen gar nicht in der vom Mikroggesetz beschriebenen Form manifestiert.

The composition laws we use in micro-explanations connect the behaviour of the compound system and that of its parts. Therefore the relations of the micro-macro relation are, first, the (lawful) behaviour of the compound system and, second, the (lawful) behaviour of the parts. What is referred to in micro-explanations is how the parts would behave if they were isolated, i.e. to the parts' behaviour under counterfactual conditions. The actual behaviour of parts is not referred to. It is the behaviour of the compound that is actual. Therefore, the law that describes the behaviour of the compound is instantiated whereas the laws that describe how the parts would behave given the right circumstances fail to be instantiated. (ebd., S. 94)

Wenn also das Verhalten komplexer Systeme beschrieben werden soll, wie es die Psychologie zur Aufgabe hat, dann liegt hier gerade *keine* Erfassungsmöglichkeit mittels mikrophysikalischer Gesetze vor, da die Einbettungsbedingungen solcher Systeme viel zu komplex sind, als dass Mikroggesetze hier als instantiiert betrachtet werden könnten. Das System verhält sich nicht nach Maßgabe von Mikroggesetzen, sodass diese auch nicht das zukünftige Verhalten eines solchen Systems prognostizieren könnten. Hüttemann zeigt, dass dies bereits bei einfachen physikalischen Phänomenen greift, wie bei der Wärme eines Metalls (vgl. ebd., S. 94f). Ein jedes Metall kann als kompositorisches System verstanden werden, welches aus einem kristallinen und einem elektronischen Subsystem besteht. Beide leisten einen Beitrag zur gegebenen Wärme eines Metalls. Wenn man die Interaktionen an dieser Stelle ausblendet, kann das Verhalten des zusammengesetzten Systems darüber „mikroerklärt“ werden, wie sich die Subsysteme in Isolation verhalten würden. „Lediglich“ bestimmte Zusammensetzungsgesetze müssten zusätzlich herangezogen werden. Die Folge ist dann allerdings, dass wiederum die Zusammenset-

zungsgesetze instantiiert sind, während es die Gesetze für die Teile nicht sind. Nur die spezifische Wärme des Metalls kann direkt gemessen werden, wohingegen die Beiträge der einzelnen Subsysteme unklar bleiben. Das zentrale Argument Hüttemanns ist also, dass Mikrogesetze in kompositorischen Systemen faktisch nicht instantiiert werden, weil solche Systeme keine geschlossenen bzw. isolierten Systeme sind.

Therefore, the conclusion has to be drawn that even though the law for the compound is instantiated and even though this law can be explained by relying on the laws for the subsystems' behaviour under counterfactual conditions these latter laws are not instantiated. Micro-explanation does provide evidence, first, for macro-laws to be instantiated and, second, for the micro-laws, i.e. the laws for the parts, not to be instantiated. (ibd.)

Wenn man nun geneigt sein mag, nicht den Mikroebenen, sondern dem Ganzen eine (ontologische) Priorität gegenüber seinen Teilen zuzuschreiben, muss man sich allerdings bewusst machen, dass auch die Entscheidung, was man überhaupt als ein zusammengesetztes System aus der Menge des Existierenden herausgreift, auch eine Frage bestimmter pragmatischer und keine in sich selbst begründete Tatsache ist. Daher ist auch der Umstand, dass wir ein Gesetz als instantiiert betrachten, immer von einer vorhergehenden Festlegung abhängig. Hüttemann (2004) spricht in diesem Zusammenhang daher von einer *Relativität der Instantiierung* (vgl. S. 96) und erläutert dies wiederum an Beispielen aus der Physik.

Wenn zuvor gesagt wurde, dass für Gase das Gesetz gilt: $pV = RT$, dann muss eingeräumt werden, dass dieses simple Gesetz eigentlich nur unter idealen, praktisch nie vorhandenen Bedingungen gültig ist, sodass es durch das Van-der-Waals-Gesetz ersetzt wurde, das besagt: $(p + a/V^2)(V - b) = RT$, wobei a und b die Interaktion der Moleküle und ihre begrenzte Ausdehnung in Rechnung stellen sollen. Allerdings ist auch dieses Ge-

setz wieder durch die so genannte Dieterice-Gleichung ersetzt worden, die eine noch komplexere Struktur aufweist (vgl. ebd.). So hat man in der Physik zunächst angenommen, dass sich Druck, Volumen und Temperatur gegenseitig bedingen, bis sich herausstellte, dass andere Faktoren ebenso wichtig sind. Das sich abzeichnende Problem ist nun, dass dieses Beispiel die generelle Tendenz verdeutlicht, dass je mehr wir über ein System wissen, desto mehr Verbindungen lassen sich unter den Konstanten und Variablen aufdecken, die für die zu messenden Größen relevant sind.

In a world in which forces have infinite ranges, such as ours, a lot of physical systems depend on a lot of other physical systems, even far away ones. According to Newtonian physics the redistribution of massive particles (the crossing of my legs, for instance) simultaneously affects all other massive objects in the universe. (In relativistic physics it takes some time and affects objects in the forward light-cone only.) Given this degree of interconnectedness it is hard to see how a faithful description of a part of the universe can take into account less than the universe as a whole. The interconnectedness of the universe forces us to the conclusion that the universe as a whole is the only proper object that can be said to instantiate a law. (ebd., S. 96f)

Diese Schlussfolgerung mag drastisch und überzogen klingen, können wir doch aufgrund unserer idealisierten Gesetze recht gute Vorhersagen des Verhaltens von Objekten machen. Wenn es aber nicht um solche praxisbezogenen Erwägungen im Sinne einer lebensweltlichen Bewährung geht, sondern um eine größtmögliche *ontologische Exaktheit*, dann ist diese Einschätzung Hüttemanns durchaus korrekt. Das Problem liegt dann aber auf der Hand: Wenn es nämlich so viele (potentielle) Faktoren gibt, die auch die zusammengesetzte Entität beeinflussen, dann ist die Aussage, dass sich in solchen Fällen das Verhalten des Ganzen und nicht das der Teile manifestiert, offenbar falsch. Damit würde es sich dann aber sowohl beim Verhalten der Subsysteme als auch beim Verhalten des Gan-

zen um etwas Kontrafaktisches handeln, solange nicht die Einbettung in das gesamte Universum berücksichtigt wird (vgl. ebd., S. 97). Hier handelt es sich um einen Aspekt, der allgemein für den Kontext von Mikroerklärungen charakteristisch ist. Denn wenn ein kompositorisches System über den Rekurs auf seine Bestandteile bzw. die hier vorherrschenden Mikrogesetze erklärt wird, dann wird genau genommen immer eine *kontrafaktische* Welt und niemals die tatsächliche Welt erklärt. Wenn also das Verhalten des Ganzen, das als gegeben betrachtet wird, mittels des Verhaltens seiner Teile erklärt wird, dann ist dieses Verhalten – aufgrund der hiermit verbundenen Ausblendung von Einfluss- und Einbettungsbedingungen – notwendigerweise kontrafaktisch *in Relation* zu dem Verhalten des Ganzen. Insofern gilt, dass die Frage, ob ein Gesetz instantiiert ist oder nicht, immer relativ zu einer vorhergehenden Entscheidung darüber ist, ob ein System als ein isoliertes zusammengesetztes System angesehen wird. Für diesen mikroexplanatorischen Kontext gilt nun in der Tat, dass das Gesetz für das Ganze instantiiert ist, während dies nicht für die Mikrogesetze gilt (vgl. ebd., S. 97f).

The essential point is that the above considerations do nothing to undermine the argument in the previous section. [...] If a system is treated as a compound as in micro-explanation – the subsystems fail to be isolated. A fortiori the behaviour of the subsystems is not manifest. Furthermore, if a system is treated as a compound its behaviour, the explanandum has to be taken to be manifest. Therefore, whenever a system is micro-explained such an explanation provides evidence for the instantiation of the macrolaw and against the instantiation of the micro-laws. (ibd., S. 98)

8.6.4. Zum Scheitern des Ansatzes einer Mikroverursachung

Unter anderem Kim (1993) hat das Vorhandensein einer Mikroverursachung zum Problem erhoben, weil hiermit jede makroskopische Verursachung und damit auch die Möglichkeit einer mentalen Verursachung

untergraben würde: „In this global microdeterministic picture there is no place for irreducible macrocausal relations. We expect any causal relation between two macroevents [...] to be micro-reductively explainable in terms of more fundamental causal processes, like any other facts involving macroproperties and events” (S. 97).

Dass eine Mikroverursachung im Sinne einer Inter-Level-Kausalität bei genauerer Betrachtung mereologisch-mechanistischer Systeme aber eine nicht zu haltende Annahme ist, wurde bereits hinreichend aufgezeigt (vgl. Kap. 7.4.1). So kann es – bezogen auf einen bestimmten Zeitpunkt – keine echte ebenenübergreifende Verursachung (egal welcher Richtung) innerhalb eines mereologisch verfassten Gesamtmechanismus geben, weil es unsinnig ist anzunehmen, dass sich eine agierende Struktur gewissermaßen selbst verursacht. Denn wenn die agierenden Entitäten der verschiedenen Ebenen eines Mechanismus in einer Identitätsbeziehung stehen, dann ist es absurd zu sagen, dass das Verhalten der Teile das Verhalten des Ganzen *verursacht* hat oder andersherum.

Daher kann man mit Bezugnahme auf Craver und Bechtel (2007) eher von einem *mechanistisch vermittelten Effekt* sprechen, der keinen Fall einer „herkömmlichen“ Verursachung darstellt. Denn hierbei handelt es sich um eine symmetrische Relation, weil der Mechanismus als ein Ganzes vollkommen durch die organisierte Aktivität seiner Teile konstituiert wird, sodass gilt „components act as they do because of factor acting on mechanism, and mechanisms act as they do because of the activities of their lower-level components“ (ebd., S. 553). Dagegen ist Kausalität (im Sinne der *causa efficiens*) eine *asymmetrische* Relation (A bewirkt B, während B aber nicht zugleich A bewirkt).

Gemäß des Prinzips der *mereologischen Supervenienz* schlägt sich jede psychische Veränderung bis zur Ebene mikrophysikalischer Entitäten herab nieder. Kommt es z.B. zu einer Veränderung des semantischen Netzwer-

kes einer Person, dann muss auch eine Modifikation der Konfiguration aller niedrigeren Ebenen der entsprechenden Mechanismen stattgefunden haben. Denn eine mentale Eigenschaft ist etwas, das *alle* Ebenen des sie hervorbringenden Mechanismus „durchdringt“. Sie ist genauso ein mentales Erlebnis mit einer phänomenalen Qualität oder einem kognitiven Gehalt wie auch etwas, das gewissermaßen in den Atomen dieser mechanistischen Hierarchie „steckt“. Der psychische Kausalnexus *ist* daher der physische Kausalnexus in einer anderen Perspektive bzw. auf einer anderen natürlichen Organisationsebene einer mit sich selbst identischen mechanistischen Hierarchie (vgl. Lüdmann, 2009, S. 124). Dass es in diesem Sinne keine Differenz oder gar kausale *Konkurrenz* zwischen dem Ganzen und seinen Teilen geben kann, verdeutlicht auch Hüttemann (2009) am Beispiel der Eigenschaft eines Billardballs, einen bestimmten Impuls zu haben:

Diese Eigenschaft wird durch die Eigenschaften der Teile des Billardballes, die ebenfalls jeweils einen bestimmten Impuls besitzen, realisiert. Mit dem Impuls des Billardballes sind nun bestimmte kausale Fähigkeiten verknüpft, z. B. die Fähigkeit, einen ruhenden Billardball zu bewegen. Aber auch die Teile des Billardballs haben Impulse, mit denen kausale Fähigkeiten verknüpft sind. Kommt es hier zu einer Konkurrenz zwischen den kausalen Fähigkeiten der realisierten Eigenschaft einerseits und den realisierenden Eigenschaften andererseits? Das scheint nicht der Fall zu sein. Hier gilt in der Tat, dass eine kausale Konkurrenz von Teilen und Ganzem unplausibel ist. (S. 72)

Somit kann es für kompositorisch bzw. mereologisch verfasste Systeme keine Inter-Level-Kausalität geben, sodass auch die Annahme einer Mikro-Makro-Abhängigkeit im Sinne des Epiphänomenalismus verworfen werden muss:

The part-whole relation is not a causal relation. If it was a causal relation one would expect the time para-meter to play a role in the exact description of the relation.

However, no such time parameter is mentioned in the laws of composition [...] Therefore, the behaviour of the parts cannot be said to cause the behaviour of the compound. Thus, micro-explanation does not support the epiphenomenalist reading of the claim that macro-causal relations hold in virtue of microcausal relations. (ebd., 2004, S. 116)

Allerdings verdeutlicht Hüttemann zu Recht, dass es genau genommen falsch ist, dass das kausale Agieren des Ganzen *nichts anderes als* das kausale Agieren seiner Teile ist, denn hier müssen zusätzlich noch die – wie er sie nennt – „Zusammensetzungsgesetze“ berücksichtigt werden: „The behaviour of the parts *plus* the laws of composition determine the behavior of the compound“ (ebd., 2004, S. 117; Hervorh. im Orig.). Dies stellt die hier vertretene These, dass es nur *ein* Kausalgeschehen gibt, das sich auf allen beteiligten mechanistischen Ebenen *zugleich* manifestiert, allerdings keineswegs infrage. Es ist nur sein terminologischer Rahmen der Besprechung, bei dem er sich einerseits auf kompositorische im Allgemeinen (d.h. *nicht* speziell auf mechanistische Systeme) bezieht und andererseits zumeist anhand der Frage nach der Instantiierung von Gesetzen argumentiert, was es in der Tat erforderlich macht, auch das strukturelle Geflecht der Teile, die *Organisationsstruktur* der kompositorischen Ebene, explizit zu erwähnen. Mechanistischen Ebenen, die der hier entwickelten psychophysischen Identitätskonzeption zugrunde liegen, ist dieser strukturelle Aspekt bereits inhärent. Ein Mechanismus definiert sich über eine bestimmte Organisationsstruktur seiner agierenden Entitäten. Es ist die Identität der verschiedenen *mikrostrukturellen Konfigurationen* einer mentalen Eigenschaft, die hier behauptet wird (vgl. Kap. 7.1 und 7.2) und nicht einfach eine Identität eines Ganzen und seiner beliebig angeordneten Teile. Allerdings würde selbst die völlige Unkenntnis der Zusammensetzungsgesetze nichts daran ändern, dass es sich bei einem Kausalvorgang um einen *ebenendurchdringenden* Prozess handelt.

Epistemische Begrenzungen liefern kein brauchbares Kriterium für ontologische Feststellungen.

Des Weiteren wurde in Kapitel 7.4 aufgezeigt, dass die Annahme einer Inter-Level-Kausalität – und damit auch die einer Mikroverursachung – nicht nur scheitert, sondern dass im Hinblick auf mechanistisch vermittelte Effekte sogar eher das Gegenteil gilt, nämlich dass es die höchste in einer mechanistischen Hierarchie realisierte Ebene ist, die für den ebenendurchdringenden Kausalitätsvollzug ausschlaggebend ist. Denn obgleich es zwar keine Ebene gibt, auf der *allein* sich der Kausalprozess ereignet, ist die Ebene des Mentalen bei einem Vorgang, bei dem z.B. ein bestimmtes Ziel verfolgt wird, insofern maßgeblich, als dass erst hier die eigentlichen Gründe für diese Handlung gegeben sind. Würden die mentalen Eigenschaften nicht vorliegen, würde die gesamte mechanistische Hierarchie aufgrund der mereologischen Supervenienz-Relation zwischen den Ebenen nicht die Struktur aufweisen können, die sie bei Vorhandensein der mentalen Eigenschaft hat. Wenn ein Mensch zu bestimmten mentalen Leistungen fähig ist, dann kann er dies nur *kraft* der psychischen Eigenschaften, die hierfür notwendig sind, auch wenn sich die kausale Aktivität bei diesen mentalen Leistungen auf allen hierzu installierten Mechanismen vollzieht. Die mereologisch-mechanistischen Hierarchien dieser Vorgänge sind nach Maßgabe der mentalen Eigenschaften gestaltet. Ohne die mentalen Eigenschaften gäbe es auch die jeweiligen „genuin physischen“ Ereignisketten nicht in dieser Form.

Wenn die Ausgangsfrage dieses Abschnitts war zu klären, ob aus der verbreiteten Praxis von Mikroerklärungen in den Wissenschaften ein *ontologisches* Primat der Mikroebenen abgeleitet werden kann, dann kann dies nun eindeutig verneint werden. Alle drei infrage kommenden Arten eines ontologischen Vorrangs der Mikroebenen, die Mikrodetermination, die

Mikrobeherrschung als auch die Mikroverursachung, scheitern. Im Gegenteil scheint es sogar einen gewissen Vorrang oder eine Maßgeblichkeit der höherstufigen Eigenschaften, Prozesse und Gesetze (bzw. Regularitäten) in kompositorisch verfassten Systemen zu geben. Wenn also die nicht-reduktiven Physikalisten annehmen, dass eine asymmetrische Bottom-Up-Abhängigkeitsbeziehung zwischen den Teilen und dem Ganzen bzw. die damit einhergehende Vorrangstellung der Mikroebenen ein konstitutives Element unseres wissenschaftlichen Weltbildes ist, auf das sie sich bei ihrem (vermeintlich) „minimalen“ Physikalismus stützen möchten, dann irren sie sich bereits an dieser Stelle. Hier herrscht ein – trotz seiner immensen Verbreitung – fundamentaler Irrtum vor.

8.7. Warum ein explanatorisches Primat der Mikroebenen scheidert

Wenn zwar nun eine wechselseitige Determination zwischen dem Ganzen und seinen Teilen und damit kein *ontologisches* Primat der Mikroebenen vorliegt, könnte man die Gültigkeit des Mikrophysikalismus dann nicht trotzdem darüber absichern, dass den Mikroebenen bzw. „niederstufigen“ Wissenschaften ein zumindest *explanatorisches* Primat zukommt? Ist es nicht vor allem die *Erklärung* von Makrophänomenen durch Mikrophänomene im Sinne der Aufdeckung der Modalität einer Mikrodetermination von Makroeigenschaften eines Systems, die den eigentlichen wissenschaftlichen Erkenntnisfortschritt begründet, während die gegenteilige Richtung der Erklärung zwar faktisch möglich sein mag, explanatorisch jedoch zu vernachlässigen ist? So könnte der Mikrophysikalist zwar einräumen, dass es in *ontologischer* Hinsicht korrekt sein mag, dass zwischen den Mikro- und Makroebenen eine wechselseitige

Determination vorliegt, aber zugleich bestreiten, dass eine Erklärung in Top-down-Ausrichtung oder eine solche, die auf ein und derselben Erklärungsebene angesiedelt ist, einen echten Beitrag zum Verständnis eines bestimmten Phänomens leisten könnte. Oder bezogen auf den Psychologie-Biologie-Kontext könnte ein überzeugter neurobiologischer Reduktionist behaupten, dass mit einer erfolgreichen vollständigen Identifizierung der neuronalen Grundlagen einer mentalen Eigenschaft alles wissenschaftlich Relevante zur Erklärung ebendieser Eigenschaft gesagt sei und die vorherigen psychologischen und vielleicht auch soziologischen, pädagogischen oder ökonomischen „Behelfserklärungen“ obsolet würden.

Im Folgenden kann gezeigt werden, dass auch ein solches *explanatorisches* Primat nicht existiert, denn ohne die Organisations- und Einbettungsbedingungen eines (mechanistischen) Phänomens kann dieses in der Regel nicht hinlänglich erklärt und es können auch keine adäquaten Vorhersagen getätigt werden²⁷⁷. Es stellt sich dann aber die Frage, warum ein solches Primat dennoch eine so hartnäckige Hintergrundannahme in der Wissenschaftsgemeinschaft oder überhaupt im menschlichen Denken darstellt.

8.7.1. Das explanatorische Defizit von Mikroerklärungen

Insbesondere mit Verweis auf den mechanistischen Erklärungsansatz kann aufgezeigt werden, dass ein Erklärungsprimat der Mikroebenen

²⁷⁷ In *ontologischer* Hinsicht wurde hierzu bereits im letzten Abschnitt in Bezug auf die Frage, ob Mikrogesetze in makroskopischen Systemen überhaupt instantiiert sind, einiges ausgesagt. Hier wird die Analyse nun auf die *epistemische* Ebene gehoben sowie anhand der Psychologie-(Neuro-)Biologie-Relation verdeutlicht, dass der bloße Rekurs auf Mikroerklärungen *explanatorisch defizitär* ist.

nicht gegeben ist, denn wie bereits Wimsatt (1976) betont und gegenwärtig vor allem Bechtel (2007, 2008, 2009) in systematischer Weise nachgewiesen hat, ist neben dem „Blick nach unten“, zumeist auch ein „Blick zur Seite“ oder gar ein „Blick nach oben“ notwendig, um ein Phänomen tatsächlich adäquat verstehen und erklären zu können.

The quest to explain phenomena by identifying responsible mechanisms involves an inherent reductionistic²⁷⁸ commitment such research decomposes the mechanism into its parts and their operations. But unlike more traditional philosophical accounts of reduction, the mechanistic perspective is not exclusively reductionistic, for it requires also recomposing the mechanism by taking into account the organization among the parts of the mechanism and situating the mechanism in its environment. Studying the parts and operations, organization, and situatedness of a mechanism requires different sets of investigatory tools. Understanding mechanisms requires a pluralistic approach that looks not just down but around and up. (ebd., 2009, S. 559)

Es sind also die Organisations- und Einbettungsbedingungen eines Mechanismus, die zu seiner bloßen Zerlegung in Teilstrukturen bzw. agierende Entitäten verschiedener Komplexitätsstufen hinzukommen müssen, damit ein Phänomen adäquat erklärt werden kann bzw. die überhaupt gegeben sein müssen, damit das Phänomen tatsächlich realisiert wird.

Zunächst ist es natürlich vollkommen korrekt, dass eine Mikroerklärung eines Phänomens durchaus einen wichtigen Beitrag zu dessen Verständnis leistet. Denn wie bereits erwähnt, können wir z.B. die Wärmeeigenschaft von bestimmten Metallen mittels des Verhaltens ihrer Ionen und Elektronen oder auch die Aggregatzustände von Stoffen (fest, flüssig, gasförmig) mittels der spezifischen Lage- und Bewegungseigenschaften

²⁷⁸ Leider differenziert auch Bechtel hier nicht zwischen *reduktiv* und *reduktionistisch*, was zu unnötigen Missverständnissen führen kann.

ihrer Teilchen (Moleküle oder Atome) sehr gut erklären. Ebenso ist die zunehmend feinkörnigere Identifikation von Gehirnbereichen, die bestimmten mentalen Leistungen zugrunde liegen und die voranschreitende Aufdeckung der hier installierten mikrostrukturellen und neurophysiologischen Mechanismen ein Meilenstein der (neuro-)biologischen Forschung.

Dabei ist die simple Lokalisierung von bestimmten Prozessen jedoch noch nicht sehr aussagekräftig. Ein viel größerer Erkenntnisfortschritt ist dadurch zu erzielen, dass man ein Phänomen in Operationen seiner Komponenten zerlegt, die mit verschiedenen Teilen des Systems verbunden sind. Bechtel (2009) spricht in diesem Zusammenhang von einer *komplexen Lokalisierung* (S. 547). Ein Beispiel liefert wiederum die Erforschung des visuellen Systems, wie sie in Kapitel 7.5.1 beschrieben wurde. So legten neuroanatomische Pfade und auch Läsionsstudien an Tieren und Menschen schon frühzeitig den Okzipitalkortex als „Ort des Sehens“ nahe. Aufgrund der Anwendung verhaltensbezogener und neurophysiologischer Methoden wurden eine Vielzahl an hochspezifizierten Hirnregionen als auch der implementierten Informationsverarbeitungsmechanismen lokalisiert und dies auf immer feinkörnigeren Ebenen. Nunmehr sind über 30 verschiedene Areale und Pfade und eine Vielfalt an neurophysiologischen Mechanismen identifiziert worden, die an der Verarbeitung visueller Informationen beteiligt sind (Felleman & van Essen, 1991; van Essen & Gallant, 1994; vgl. McCauley & Bechtel, S. 748; Bechtel, 2009, S. 547ff). Dabei ist die simple (abwärts gerichtete) Lokalisierungsannahme, dass ein bestimmter Teil eines Systems für die Hervorbringung eines Phänomens verantwortlich ist, ein explanatorischer Kurzschluss. Die bloße Behauptung, dass der striäre Kortex jeglicher visueller Informationsverarbeitung zugrunde liegt, erklärt für sich genommen wenig bis überhaupt nichts. Was betrachtet werden muss, ist die genaue Organisa-

tionsstruktur des Mechanismus. „Beyond distinguishing phenomena and decomposing the responsible system functionally and structurally, mechanistic explanation requires researchers to determine how the various component parts are organized such that the operations are coordinated appropriately to realize the overall phenomenon” (Bechtel, 2007, S. 177). Hierfür ist entscheidend, dass es sich bei höherstufigen Eigenschaften eines Systems nicht einfach nur um eine bloße Aggregation niederstufiger Eigenschaften des Systems bzw. von dessen Teilen handelt. Denn wie nun bereits mehrfach herausgestellt wurde, stellen aggregative Ebenen eine extrem seltene Ausnahme im Hinblick auf die Organisationsstruktur kompositorisch verfasster Systeme dar. Schon gar nicht spielen sie im Fall komplexer informationsverarbeitender Systeme wie dem Menschen eine bedeutende Rolle. Denn wie ein Mechanismus als ein Ganzes agiert, ist häufig völlig verschieden von der konkreten Ausgestaltungsweise der Operationen, die seine Komponenten ausführen.

Neurons in different parts of the visual system generate action potentials in response to release of transmitters by cells on which they synapse, for example, while the (ventral) visual system as a whole identifies what object is presented to the person and makes that information available to other cognitive systems (those engaged, for example, in encoding memories or making decisions). [...] Organization itself is not something inherent in the parts (even selforganizing systems only organize themselves under appropriate conditions). Accordingly, investigators who already understand in detail how the parts behave are often surprised by what happens when they are organized in particular ways. (Bechtel, 2007, S. 185f)

Daher wurden auch in der Forschung und Theoriebildung zur visuellen Informationsverarbeitung Modelle konzipiert, bei denen einige Teile des Systems von den vorhergehenden Operationen anderer Teile des Systems abhängen, um ihre eigenen Operationen auszuführen. In einer linear aufgebauten Organisation „the product of the operation of one part is

operated on by the next in a sequence until an output product is produced” (ebd., S. 187). So haben z.B. Hubel und Wiesel (1962) festgestellt, dass die Antwortmuster der Zellen des im Thalamus befindlichen LGN (lateral geniculate nucleus bzw. im Deutschen gebräuchlicher: Corpus geniculatum laterale, CGL) durch eine Zentrum-Umfeld-Struktur gekennzeichnet sind, während einfache Zellen in V1 auf Kanten an ganz spezifischen Orten und komplexe Zellen auf Ecken an verschiedenen Orten im rezeptiven Feld oder auch auf bewegende Kanten hin reagieren (vgl. Kandel & Mason, 1996, S. 438ff). Sie nahmen daher eine lineare Verarbeitungsprozedur an, bei der „a simple cell would sum inputs from several LGN cells whose center lay along a line with a given orientation and respond when these inputs exceeded a threshold [...]. They further proposed that a complex cell that fired whenever a line with an appropriate orientation anywhere in its visual field did so by responding whenever it received inputs from the simple cells to which it was connected“ (Bechtel, 2009, S. 551). Ein so aufgebautes lineares System kann andere Leistungen erbringen als seine Komponenten. Während bei einem Aggregat jeder Teil durch einen anderen ausgetauscht werden kann (ohne dass es zu einer Veränderung der herangezogenen Eigenschaft des Ganzen kommt), hängen lineare Systeme von der richtigen Reihenfolge der Teile ab, die es ermöglicht, dass die Produkte einer Operation als Inputs für die nachfolgende Operation zur Verfügung stehen. Trotzdem ist die hier angenommene Organisation des Systems weiterhin eine rein additive, bei der die Operation einer Komponente einfach zur Operation einer anderen hinzugefügt wird (vgl. Bechtel, 2007, S. 187). Wie sich gerade in der Forschung der letzten 30 Jahre gezeigt hat, reichen auch solche linearen Modelle daher keineswegs aus um die tatsächlichen Zusammenhänge adäquat abzubilden bzw. korrekte Vorhersagen zu machen.

So entsteht im Hinblick auf das Modell von Hubel und Wiesel das Problem falscher Vorhersagen. Während faktisch zu beobachten ist, dass sich die Reaktion gesättigter einfacher Zellen abschwächt, wenn der Kontrast zwischen Zentrum und Umgebung des rezeptiven Felds groß wird, sagt ihr lineares Modell hingegen voraus, dass es einen kontinuierlichen Zuwachs des Antwortverhaltens über die jeweilige Zeitspanne gibt. „Further, when two patterns are overlaid, one representing the cell's preferred stimulus, a simple cell's response is diminished, whereas a linear model predicts it would respond normally to the preferred stimulus” (Bechtel, 2009, S. 551).

Heeger, Simoncelli und Movshon (1996) haben daher ein nicht-lineares Modell entwickelt, das das Antwortverhalten des Systems korrekter vorhersagt, indem es den summierten Input durch einen Faktor aufteilt, der die zusammengefasste Antwort der *umliegenden* Neuronen repräsentiert. Wie Bechtel (2009) sagt: „They then extended the application of these non-linear models to MT cells which respond not just to individual components of motion, but to complex patterns. According to their model, a given MT cell sums inputs from V1 cells compatible with the composite motion and normalizes its response by dividing the summed input by the pooled responses of other MT cells” (S. 551f). Mit einem Voranschreiten der Verarbeitungsstufen visueller Reize im Gehirn werden Informationen also dadurch extrahiert, dass die jeweiligen Gehirnregionen adäquat mit anderen Gehirnbereichen verbunden sind, die auf einer einfacheren Stufe der Informationsextraktion lokalisiert sind. Weil die Aussagekraft der Verarbeitungsprozesse in einer ganz bestimmten Gehirnregion daher sehr begrenzt ist, ist es notwendig, viele weitere Stufen der Verarbeitung visueller Informationen ausfindig zu machen (vgl. ebd., S. 552). Dabei gibt es sowohl vorwärts gerichtete Prozesse, wie Ungerleider und Mishkin (1982) den dorsalen und ventralen Verarbeitungspfad visueller Informa-

tionen zunächst konzipierten, als auch – wie sich in der Forschung zunehmend herausstellt – eine Vielzahl rekurrenter und rückgekoppelter Projektionen (vgl. Bechtel, 2009, S. 553ff; 2007, S. 187ff). Wie Bechtel (2009) mit Verweis auf Ballard (1991) und Buzsáki (2006) herausstellt: „The extensive interconnection of brain areas with feedforward, recurrent, and collateral projections suggests that it too might be a system that is inherently active, generating patterns of neural activity on its own [...]. If it is an active system, then it may be better to think of sensory inputs as perturbing the brain's internal dynamics rather than as providing inputs to which the brain passively responds” (S. 555). Wenn ein Mechanismus rekurrente und Rückkopplungsschleifen aufweist, hängt die Operationsweise vieler Komponenten von dem Agieren anderer Komponenten ab und ein Theoretiker kann nicht einfach das Verhalten des Gesamtsystems dadurch erklären, dass er das Verhalten der Teile des Systems einfach „aufaddiert“. Vielmehr muss er berücksichtigen, wie die Komponenten durch Bedingungen festgelegt sind, die sich aus der Operationsweise anderer Teile des Systems ergeben (vgl. ebd., 2007, S. 189).

Ebenso spielen die konkreten Einbettungs- bzw. Umgebungsbedingungen eines Mechanismus eine entscheidende Rolle zur Erklärung des von ihm hervorgebrachten Phänomens. So sind Mechanismen keine isolierten Systeme, sondern werden durch Umwelteinflüsse moduliert: „This means that understanding the behavior of a mechanism requires focusing not just on its lower level constitution and its organization, but also on the specific character of the inputs it receives from its environment” (vgl. ebd., 2009, S. 557). So ist aus der Wahrnehmungspsychologie bekannt, dass es eine Vielzahl von Top-down-Kontexteffekten gibt, was Abbildung 37 illustrieren soll. Hier werden die dargestellten Zeichen als DAS und OHR bzw. THE und CAT wahrgenommen, obwohl die Zeichen für H und A in der Abbildung völlig identisch dargestellt sind. „Der allgemeine Kontext,

den die Wörter liefern, führt zur richtigen Deutung“ (Anderson, 2001, S. 63).

DAS OHR
TAE CAT

Abb. 37: Kontexteffekt der Wahrnehmung (nach Selfridge, 1955; vgl. Anderson, 2001, S. 63).

Ein weiteres Beispiel ist *Ames Raum*, der von einem bestimmten Blickpunkt aus wie ein gewöhnliches Zimmer wirkt, bei dem die Wände zueinander sowie zu Boden und Decke rechtwinklig stehen (vgl. Ames, 1951). Faktisch ist der Raum jedoch trapezförmig verzerrt. Aufgrund unserer Vorerfahrungen mit solchen Wahrnehmungsgebilden erscheint uns die uns gegenüberliegende Wand parallel, obwohl die eine der beiden Ecken weiter entfernt ist als die andere. Der interessante Effekt ist nun, dass wenn sich zwei gleich große Personen in diese Ecken stellen, die weiter entfernte Person kleiner wahrgenommen wird als die andere. Auch für diesen Effekt sind ganz bestimmte Kontextbedingungen (Interpretation des Wahrnehmungsobjektes nach dem Schema für „herkömmliche“ rechteckig angelegte Räume) entscheidend. Insofern hat Bechtel (2009) völlig recht, wenn er sagt, dass „[behind] providing input to a mechanism, the environment often has important systemic properties that are missed if one simply looks at inputs to the mechanism sequentially. That is, the environment is not just a collection of stimuli, but of objects behaving in systematic ways“ (S. 558). Nur weil man z.B. ein Merkmal unter Rückgriff auf den Pfad, der von den molekularen genetischen Mechanismen bis hin zu seinem faktischen Auftreten führt, nachzeichnen kann, hat man nicht

die Funktion, die das Merkmal in einem noch größeren System leistet, erklärt. Die Erklärung eines Phänomens mittels des begrifflich-konzeptuellen Systems derselben Ebene, auf der es angesiedelt ist oder sogar höherstufiger Ebenen, hat eine mindestens ebenso große Erklärungskraft (vgl. Wimsatt, 1976, S. 249). Zu verstehen, wie das visuelle System organisiert und mit anderen physiologischen Systemen koordiniert ist sowie auf externale Stimuli reagiert, bedarf daher eines grundlegenden Wissens, das jenseits von der Identifikation der Komponenten des visuellen Systems und ihrer Operationen liegt.

Problematisch ist allerdings, dass Bechtel – wie sich auch schon zuvor immer wieder gezeigt hat – in diesem Zusammenhang nicht immer klar zwischen Reduktion und Reduktionismus unterscheidet. Denn wenn er sagt, dass solche Organisations- und Einbettungsbedingungen eines Mechanismus einer bloßen *reduktionistischen* Interpretation des mechanistischen Erklärungsansatzes zuwiderlaufen würden (vgl. z.B. ebd., 2009, S. 559), ist dies nur insofern richtig, als dass man die hier herausgestellte und über eine bloße Reduktivität weit hinausgehende Interpretation des Begriffes des *Reduktionismus* zugrunde legt, nach der es die Mikroebenen sind, die die *eigentliche* Wirklichkeit ausmachen, auf der die *eigentlichen* Kausalprozesse ablaufen etc. (vgl. Kap. 8.1.2.1). Falsch wäre es dagegen anzunehmen, dass aus der Kontextabhängigkeit bei solchen Wahrnehmungsphänomenen oder auch optischen Täuschungen eine *Nicht-Reduzierbarkeit* der höherstufigen „ganzheitlichen“ Phänomene folgt. Wie Dörner (1979) aufzeigt, lassen sich natürlich auch interaktionale Verknüpfungen von Kausalketten in reduktiven Modellen berücksichtigen. Daher lässt sich das Verhalten des Ganzen durchaus „aus der Kenntnis der Teile und ihrer Interaktion erklären. Allerdings bedeutet die Existenz von ganzheitlichen Phänomenen wohl, daß ein *radikaler* Reduktionismus nicht möglich ist“ (ebd., S. 127; Hervorh. im Orig.). Ein solcher

würde annehmen, dass man allein durch die Analyse der Entitäten der niedrigeren Ebenen eine vollkommene Erkenntnis der Entitäten der höheren Ebenen erhalten könne. Dies ist im Hinblick auf das Verhältnis verschiedener Wissenschaften unsinnig:

Der Neurophysiologe wird nicht ‚automatisch‘ die Gesetze der Psychologie mitentdecken, genauso wenig, wie der Psychologe die Gesetze der Soziologie automatisch mitentdecken wird, wenn er sich mit den psychischen Erscheinungen bei Einzelmenschen befaßt. Ganzheitlichkeit ist ein Interaktionsphänomen, dessen Gesetzmäßigkeiten sich nur dann enthüllen, wenn man die Teile in Interaktion miteinander betrachtet. Man entdeckt die für die Interaktion wesentlichen Eigenschaften von Dingen erst, wenn man die Dinge interagieren lässt. (ebd.)

In diesem Sinne ist durch das Studium *einzelner* Atome keineswegs feststellbar, dass die Anzahl von Elektronen auf der äußersten Schale wichtig für die chemische Reagibilität des entsprechenden Elements ist (vgl. ebd., S. 128). Trotzdem kann hieraus nicht abgeleitet werden, dass es sich um ein nicht-reduktives Phänomen handelt. Somit gilt einerseits, dass die Organisations- und Einbettungsbedingungen von Mechanismen keineswegs die Möglichkeit einer Reduktion der höherstufigen Phänomene ausschließen. Andererseits wird einem kruden Reduktionismus à la Bickle oder Churchland, der höherstufigen Wissenschaften das Existenzrecht abzusprechen sucht, aber durchaus das Wasser abgegraben. Insofern hat Bechtel (2007) dann doch recht, wenn er sagt: „As a result, organized mechanisms become the focus of relatively autonomous disciplines – disciplines which deploy their own tools of investigation and develop their own distinctive accounts of the phenomena associated with these mechanisms. Thus, to understand the autonomy achieved by higher levels within a mechanistic framework, we need to focus on the sorts of organization that figure in mechanisms” (S. 186).

Damit ist das mechanistische Modell der Erklärung gewissermaßen ein Hybrid aus dem klassischen reduktiven Erklärungsansatz, wie ihn Nagel konzipierte, und dem funktionalistischen Ansatz. Während Ersterer Bottom-up-Erklärungen als konstitutives Element mechanistischer Integrationen betrachtet und kontextuelle Erklärungen ausblendet, betont Letzterer „outward- and upward-looking contextual explanations and neglects constitutive explanation. Both, therefore, are incomplete from the point of view of mechanism“ (Polger, 2004, S. 205). Oder wie Bechtel (2009) es ausdrückt:

One of the sharpest distinctions between the mechanistic perspective on reduction and more traditional accounts involves the attitude about how complete the knowledge available at any given level is. In nomological accounts of reduction, the lowest level is often viewed as itself providing a complete account of everything that happens: if one provides the appropriate initial conditions, then one can derive from lower-level laws all the consequences. Appealing to higher-level explanations then appears problematic since any causal processes at higher levels would overdetermine events already determined by lower-level processes. In particular, appealing to mental causes for events determined by brain processes appears problematic (Kim, 1998). From a mechanist perspective, there is no basis for assuming that one could provide a complete account of the functioning of the mechanism in terms of the parts alone. The behavior of the mechanism depends not just on the parts but how they are organized and the context in which they are situated. (S. 559f)

8.7.2. Der Vorrang der Mikroebenen als kognitive Heuristik/Verzerrung

Wenn Mikroerklärungen also in den meisten Fällen nicht ausreichend sind, um ein komplexes Phänomen adäquat zu erklären, stellt sich die Frage, warum sie dennoch allem Anschein nach in den Augen der meisten Forscher und wahrscheinlich auch in der Allgemeinbevölkerung einen Vorrang gegenüber Erklärungen auf derselben oder höheren Ebenen haben. Warum hat man also z.B. den Eindruck, dass die Information, dass meiner Orientierungsleistung in Situation XYZ eine Aktivität in

bestimmten kortikalen Zentren zugrunde liegt, eine größere wissenschaftliche bzw. explanatorische Aussagekraft hat, als wenn man die konkreten psychischen (funktionalen) Mechanismen und Vorbedingungen sowie die (sozialen, kulturellen...) Umweltbedingungen freilegt. Oder warum würden wohl viele Menschen – zumindest im ersten Moment – einer (neuro-)biologischen Erklärung des Aufkommens des Nationalsozialismus oder der Psychopathie und des Größenwahns Hitlers (wenn sie denn verfügbar wären²⁷⁹) einen größeren wissenschaftlichen Stellenwert beimessen als der feinmaschigsten historischen Analyse der politischen und gesellschaftlichen Umstände bzw. der abstrusen Psychodynamik des „Führers“?

Wimsatt (1976, 1979, 1980, 1986, 2006) spricht in diesem Zusammenhang vom Phänomen der (subjektiv empfundenen) *Erklärungspriorität* der Mikroebenen oder Mikrowissenschaften. Er zeigt dabei zunächst auf, dass die Annahme eines ontologischen auf der einen und eines explanatorischen Primats auf der anderen Seite stark zusammenhängen, und ist überzeugt, dass die Annahme eines ontologischen Primats abwegig ist: „... entities and things [...] at different levels are equally real, and none is secondary, in its *reality*, to another“ (Wimsatt 1976, S. 242).

Nach Wimsatt (1976) geht die trotzdem häufig vorgenommene Zuschreibung eines höheren Realitätsgrades zu den Mikroebenen darauf zurück, dass angenommen wird, dass kompositorische Systeme asymmetrisch verfasst sind, weil die Entitäten der höheren Ebenen in gewisser Weise aus Entitäten der niederen Ebenen zusammengesetzt sind und insofern vermeintlich etwas Sekundäres, Abgeleitetes und entwicklungsgeschicht-

²⁷⁹ Rein fiktiv ließe sich hier annehmen, dass es bei diesem Personenkreis eine Häufung neuronaler Aktivität im Hinzentrum x_{Gz1} oder eine verminderte Ausschüttung eines bestimmten Transmitters y gibt, wobei die Ausprägung dieser Variablen vom Überzeugungsgrad bzgl. der nationalsozialistischen Ideologie abhängt.

lich Späteres darstellen (in die letztgenannte Richtung argumentieren in der Tat z.B. Oppenheim & Putnam, 1958). Neben der Verfügbarkeit von Mikroerklärungen liegt nach Wimsatt (1976) ein Grund darin, dass man sich auf die folgende Logik kompositorischer Systeme bezieht: Wenn Xs aus Ys zusammengesetzt sind, sind Ys nicht aus Xs zusammengesetzt. Auch wenn das korrekt ist, ist hieraus natürlich kein ontologischer Vorrang der Mikroebenen abzuleiten, denn die Inversion der Kompositionsbeziehung ist die Partitionsbeziehung (Ys sind Teile von X), welche ebenso asymmetrisch ist (vgl. ebd., S. 243). Ein weiterer Grund könnte die Annahme sein, dass die niedrigeren Ebenen allgemeiner/unbestimmter sind, während die höheren Ebenen nur eine spezielle Teilmenge dessen freilegen, was tatsächlich aus den Entitäten der niederen Ebenen „gemacht“ werden könnte. Aber auch diese Überzeugung ist nach Wimsatt abwegig, denn man könnte auch sagen, dass es viele Teile gibt, die man aus Ganzheiten herauslösen könnte, was zeigt, dass auch hier genauso gut eine vergleichbare asymmetrische Interpretation in umgekehrter Richtung vorgenommen werden kann (vgl. ebd., S. 243f). Das dies *de facto* nicht geschieht, hat offenbar nichts mit ontologischen Gesichtspunkten bzw. den betrachteten wissenschaftlichen Objekten zu tun, sondern vielmehr mit uns selbst. Eine weitere Quelle für die wahrgenommene Asymmetrie könnte darin liegen, sich zu vergegenwärtigen, wann wir uns dazu genötigt sehen, eine Bedeutungsveränderung eines bestimmten wissenschaftlichen Ausdrucks anzuerkennen. Mit Verweis auf Roth (1974) verdeutlicht Wimsatt, dass Bedeutungsveränderungen ein zentrales Kriterium zur Etablierung verschiedener Beschreibungsebenen ist: Die Entität oder Eigenschaft, auf die sich der Ausdruck bezieht, wird in dem Fall sozusagen auf einer niedrigeren Beschreibungsebene „umgeschrieben“. Wenn wir z.B. wissen, dass Gene aus DNA bestehen, verändert dies die Bedeutung des Ausdrucks „Gen“. Wenn wir dagegen aber z.B. entdecken,

dass ein Eisenatom in der Hämgruppe auftaucht, die mit dem Hämoglobin assoziiert ist, gibt es keine Bedeutungsänderung für den Ausdruck „Eisen“ (vgl. ebd., 1976, S. 244). Daraus zieht Wimsatt (1976) folgende aufschlussreiche Konsequenz:

This shows how deeply reductionism is embedded in our conceptual scheme, but I suspect that the direction of meaning change is a consequence of explanatory primacy rather than a possible cause of it. If the source of explanations is generally to be found at lower levels, then more relevant information about the behavior of an entity will be picked up by discovering a lower level description of it than by looking for descriptions of systems of which it is a part. The most important information about an entity becomes stabilized as part of the meaning of terms that refer to it, so lower-level redescrptions are in most cases more likely to lead to changes of meaning than upper-level redescrptions. (ebd.; Hervorh.: M.L.)

Nach Wimsatt geht die Annahme, dass Dinge niedriger Ebenen *realer* sind als Dinge höherer Ebenen auf Intuitionen zurück, die mit einem (vermeintlichen) explanatorischen Vorrang der Mikroebenen bzw. ihrer Entitäten und Eigenschaften zu tun haben. Da wir es (allein schon aus den Bereichen der Technik, der Ingenieurwissenschaften etc.) gewohnt sind, höherstufige Eigenschaften durch niederstufige zu erklären, denken wir, dass niederstufige *realer* sind: „The direction of explanatory primacy is thus, on this account, simply determined by the fact that in our conceptual account of our world, more things are explained by things which are lower than them in level than by things which are higher than them in level“ (S. 245f). Nach Wimsatt ist der Irrglaube an ein Primat der Mikroebenen – ganz ähnlich wie bei Hüttemann – also eine Folge der Intuition, dass Bottom-up-Erklärungen gegenüber Top-down-Erklärungen höherwertig sind.

Bei der Annahme eines Erklärungsprimats der Mikroebenen respektive eines kruden Mikroreduktionismus scheint es sich demnach um eine *kognitive Verzerrung* im Sinne einer gewissen *Heuristik* zu handeln, wie

sie z.B. Tversky und Kahnemann (1974) in Bezug auf Urteile ausfindig gemacht haben (vgl. Wimsatt, 1980). Nach Maßgabe von Kosten-Nutzen-Abwägungen mögen Heuristiken in vielen Fällen hilfreich sein, führen aber doch zu systematischen Fehlern, also solchen, die nicht völlig zufällig verlaufen, sondern einem ganz bestimmten (Verzerrungs-)Muster folgen (vgl. ebd., S. 220). Wimsatt (1986) arbeitet dabei eine Vielzahl von Folgen heraus, die sich aus der Anwendung reduktionistischer Heuristiken bei der Begriffsbildung, bei der Konzeption von Modellen, der Theoriekonstruktion oder auch beim Beobachten und Experimentieren ergeben können.

Ein Aspekt ist die *beschreibende Lokalisierung*, bei der eine relationale Eigenschaft so beschrieben wird, als handele es sich dabei um eine einstellige Relation, also eine einfache Eigenschaft bzw. um relationale Eigenschaften einer niedrigeren Ebene. Dabei ergibt sich allerdings das Problem „... that if a property is a function of system properties and environment properties, keeping the environment constant will make the property look as if it is a function only of system properties“ (Wimsatt 1986, S. 301). Eine weitere Facette ist die auch von Bechtel (2007, 2009) kritisierte Überbetonung der *Lokalisierung* eines Phänomens, die die konkreten Organisations- und Einbettungsbedingungen eines Mechanismus vernachlässigt. Der Kontext eines Mechanismus wird hierdurch in der Theorie- und Modellbildung viel zu sehr vereinfacht. Wichtige Daten, die außerhalb des Systems liegen, werden von vornherein gar nicht erst erhoben oder diese Umweltfaktoren treten lediglich als möglichst konstant zu haltende „Störfaktoren“ auf, obgleich sie in natürlichen Umgebungen stark schwanken und das Verhalten des Systems entscheidend beeinflussen mögen (vgl. Wimsatt, 1986, S. 301). Man kann also sagen, dass die zentrale Konsequenz einer reduktionistischen Heuristik darin besteht, dass Effekte des Kontextes eines zu erklärenden Phänomens ignoriert

oder unterschätzt werden. Neben dem vielleicht zu vermeiden wollenden Aufwand, der bei einer Berücksichtigung komplexer Systeme vonnöten wäre, mag ein weiterer Grund in der bereits thematisierten *Linearität* des menschlichen Denkens liegen, die auch zu der Bevorzugung linearer Modelle gegenüber komplexeren rekurrenten und rückgekoppelten Modellen führt und somit der Erfassung der vollen Realität komplexer Systeme im Wege steht (vgl. Bechtel, 2007, S. 187; Kap. 8.7.1 dieser Arbeit). Das „Perfide“ am vermeintlichen Erklärungsvorrang ist nun, dass es sich hierbei um eine sich selbst verstärkende Idee – ganz im Sinne einer *selbsterfüllenden Prophezeiung* – handelt. Denn aufgrund der Überzeugung, dass eine Mikroerklärung einen höherwertigen Status hat, wird auch in zukünftigen Untersuchungen eines Phänomens eher nach einer solchen Erklärung gesucht werden als nach einer abwärts gerichteten Erklärung oder einer, die auf der gleichen Ebene ansetzt. Wimsatt sagt, dass der Erklärungsvorrang sogar in Fällen greift, in denen wir eine Top-down-Erklärung, aber keine Bottom-up-Erklärung besitzen. In so einem Fall würde es eine starke Tendenz geben, die Erklärung als abgeleitet, heuristisch oder nur als eine Art Lückenbüßer zu betrachten, bis wir eine „richtige“ Erklärung gefunden haben (vgl. ebd., 1976, S. 246f). Damit benennt er genau die Vorurteile, die reduktionistische Forscher und Philosophen ins Feld führen, um psychologische Erklärungsmodelle gegenüber biologischen, chemischen oder physikalischen zu diskreditieren (vgl. Kap. 2.1.2).

Wimsatt (2006) meint allerdings, dass der Versuch, reduktionistische Heuristiken ganz zu vermeiden, nicht sinnvoll ist. Es wäre wohl ein aussichtsloser Kampf gegen Windmühlen, denn wie auch im Hinblick auf die phänomenale Transparenz mentaler Zustände, die entscheidend zur Annahme eines Substanzdualismus beiträgt (vgl. Kap. 3.1), kann diese auch wider besseren Wissens nicht transzendiert werden. Stattdessen

sollten wir vielmehr den Ursprung und das Wesen dieser Verzerrungen untersuchen, sodass wir diese verstehen und auf ihre Auswirkungen hin korrigieren können (S. 466f). Interessant ist Wimsatts Ansatz besonders, weil er aufzeigt, dass bestimmte erkenntnis- und wissenschaftstheoretische Fragen stark von psychologischen Vorannahmen beeinflusst sind. Insofern ist die Debatte um den Reduktionismus keine Angelegenheit, die eine formale philosophische Analyse allein beantworten könnte, sondern eine, die von psychologischen und methodischen Belangen abhängt und offenbar auch über gezielte empirische Untersuchungen zu kognitiven Verzerrungen im wissenschaftlichen Erkenntnisprozess vorangebracht werden kann. So könnte man untersuchen, unter welchen Bedingungen eine reduktionistische Strategie wahrscheinlicher angewendet wird, wann sie Erfolg versprechend ist und wann weniger (vgl. Wimsatt 1979, S. 367).

9. Leitlinien und Konsequenzen der begrifflich-konzeptuellen Grundlegung der Psychologie

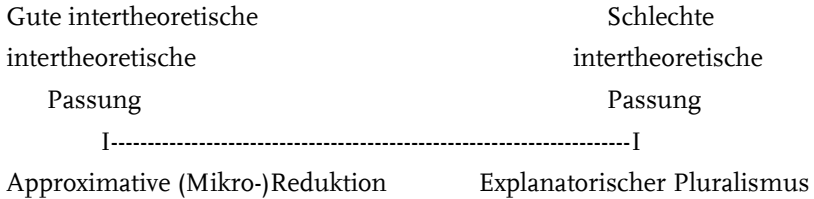
Im Folgenden sollen nun die zentralen Leitlinien und die Konsequenzen der erfolgten begrifflich-konzeptuellen Grundlegung der Psychologie zusammengetragen werden. Entgegen vielen unbegründeten Befürchtungen in der Psychologie und Philosophie stellt die Reduktivität psychologischer Theorien und Konzepte bzw. die Akzeptanz einer (mereologisch-mechanistisch fundierten) Identitätstheorie hierbei keinen Autoritäts- oder Autonomieverlust der Psychologie dar, sondern führt im Gegenteil zu einer Festigung ihrer Position im Kanon der Wissenschaften und einer größeren Respektabilität und „ontologischen Güte“ ihrer postulierten Entitäten, Eigenschaften und Prozesse. Das Psychische kann über den mechanistischen Ansatz klar in der Welt als ein reales und kausal wirksames Phänomen verortet werden. Die ontologische *Einheit* des Psychischen ist dabei zugleich mit einer Pluralität der Erklärungsweise psychischer Phänomene verbunden, was jeden Ansatz Lügen straft, der den Mikroebenen ein ontologisches oder epistemisches Primat zuzusprechen beabsichtigt. Der (identitätstheoretisch interpretierte) mechanistische Erklärungsansatz kann insgesamt als eine sinnvolle Erweiterung und Korrektur des – die gegenwärtige Psychologie immer noch sehr prägenden – funktionalistischen Paradigmas angesehen werden und sollte dieses daher als neues Leitbild psychologischer Forschung, Konzept- und Theoriebildung ablösen.

9.1. Reduktivität als Eckpfeiler einer „autonom“-integren Psychologie

Am Anfang der Arbeit stand u.a. die Frage nach dem Stellenwert von Reduktion und Reduktionismus für die Psychologie. Wie sieht es daher mit der Befürchtung aus, dass eine psychophysische Identitätsannahme oder ontologische Reduktion, wie sie hier herausgestellt wurde, Tür und Tor für eine Reduktion der Psychologie auf die Biologie bzw. psychologischer Begriffe und Theorien auf solche der (Neuro-)Biologie öffnet, was die Psychologie obsolet machen würde? Zur Erinnerung: Es war vor allem Fodor (1974), der genau diesen Zusammenhang sah und ihn wirkmächtig zu Wort brachte: „If science is to be unified, then all such taxonomies must apply *to the same things*. If physics is to be basic science, then each of these things had better be a physical thing. But it is not further required that the taxonomies which the special sciences employ must themselves reduce to the taxonomy of physics” (S. 114; Hervorh. im Orig.). Wobei seiner Auffassung nach nur die multiple Realisierbarkeit mentaler Eigenschaften und der hiermit assoziierte Funktionalismus die Psychologie vor einer Vereinnahmung durch die Biologie (bzw. letztendlich sogar durch die Physik) schützen kann.

Es lohnt sich an dieser Stelle nochmals auf das Modell der Theoriereduktion von McCauley (2007) zurückzukommen (vgl. Abb. 38), um hieran systematisch aufzuzeigen, wann und inwiefern eine Reduktion psychologischer Theorien und Konzepte möglich ist und ob hiermit tatsächlich die von Fodor und vielen anderen befürchteten Konsequenzen verbunden sind.

a) Inter-Level (oder wissenschaftsübergreifende) Kontexte



b) Intra-Level (oder „fortlaufende“) Kontexte

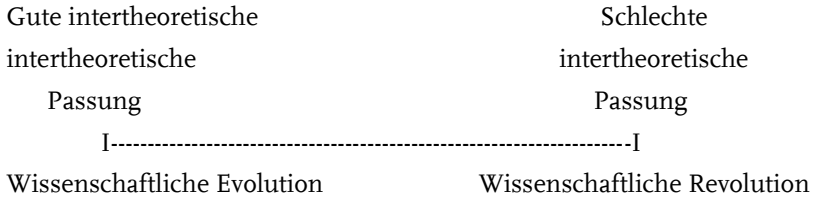


Abb. 38: Kontextabhängiges Modell der Theoriereduktion (Aus: McCauley, 2007, S. 133)

Zunächst sei daran erinnert, dass sowohl die „Verteidiger“ der Autonomie der Psychologie wie Fodor als auch die ihnen diametral entgegengesetzten New-Wave-Reduktionisten wie Bickle oder die Churchlands, die für eine weitestgehende Eliminierung der Psychologie (bzw. ihrer Konzepte und Theorien) plädieren, zur Grundlage ihrer Argumentation die schlechte intertheoretische Passung zwischen Neurobiologie und Psychologie erhoben haben. Während die einen in der multiplen Realisierbarkeit mentaler Eigenschaften eine unüberbrückbare Hürde für jegliche reduktive (und damit auch reduktionistische) Bestrebungen ansehen und funktionalistische Modelle präferieren, die möglichst unabhängig („autonom“) von den physischen respektive neuronalen Prozessen zu konzipieren seien, die mit den von ihnen betrachteten mentalen Vorgängen asso-

ziiert sind, argumentieren die anderen, dass genau hierdurch der „klassische Umstand“ für eine Ersetzung eines obsoleten (vor-)wissenschaftlichen Paradigmas vorliegen würde (da sich deren Theorien eben nicht ins naturwissenschaftliche Weltbild einpassen lassen). Im Zuge des wissenschaftlichen Fortschritts sei(en) die (Theorien der) Psychologie daher durch die (Theorien der) Neurobiologie zu ersetzen.

Wie ist eine – mit der multiplen Realisierbarkeit mentaler Eigenschaften zusammenhängende – schlechte intertheoretische Passung zwischen Neurobiologie und Psychologie aber nun tatsächlich zu bewerten? Ist sie wirklich der begrüßenswerte Garant für die „Autonomie“ der Psychologie? Entgegen der Ansicht Fodors ist eine solche Autonomie nichts, was für die Psychologie ein wünschenswerter Zustand sein kann. Schon gar nicht hängt ihre Existenzberechtigung von einer Nicht-Reduzierbarkeit ihrer Theorien und Begriffe und der Bekämpfung psychophysischer Identitätsannahmen ab. Denn aus der Überzeugung, dass wenn die Psychologie „autonom“ ist, sie dies nur aufgrund ihrer Abstraktion von den biologischen Grundlagen des Psychischen ist, ergeben sich zumindest drei gravierende Probleme:

Erstens werden durch die Art von „Autonomie“, die durch den Funktionalismus der Psychologie zugesichert bleiben soll, gerade diejenigen psychologischen Modelle und Theorien bevorzugt, die keine (neuro-)biologische Entsprechung aufweisen oder vielleicht sogar *konträr* zu neurowissenschaftlichen Befunden stehen. Denn je weniger sich die funktionalen Konzepte der Psychologie mit physischen oder (neuro-)physiologischen in Verbindung bringen lassen, desto „autonom“ wäre nach der Logik Fodors die Psychologie²⁸⁰. Mit Bezug auf die in Kapitel 4 aufgeführten

²⁸⁰ Was allerdings keineswegs heißen soll, dass funktionalistische Konzepte und Theorien des Geistes generell realitätsfern sind. Grundsätzlich können funktionalistische Konzepte

Ebenenmodelle des Geistes könnte man sagen, dass die theoretischen Ebenen, wie sie z.B. im Hinblick auf Intelligenz oder die visuelle Wahrnehmung bestehen, möglichst keine Entsprechung oder eindeutige Realisierung in Form von natürlichen Organisationsebenen haben sollten. Andernfalls bestünde die Gefahr, dass Identitätsannahmen zwischen den Termen der beiden Wissenschaften als Brückenprinzipien eine Theorie-reduktion nach Maßgabe Nagels oder auch des Ansatzes reduktiver Erklärungen ermöglichen könnten und hiermit die gesamte Psychologie (vermeintlich) obsolet würde. Hier wird der Begriff der Autonomie als *Selbstgesetzgebung* wortwörtlich genommen und „frei schwebende“ Theorien und Konzepte, denen eine ontologische Verankerung fehlt, grundsätzlich als wünschenswert angesehen. „Realitätsferne“ als Maßstab einer Wissenschaft zu erheben, ist aber nun ein völlig unsinniges Bestreben, das zugleich keineswegs notwendig ist, um die Eigenständigkeit und Relevanz der Psychologie zu begründen.

Ganz im Gegenteil ist diese Art von „Autonomie“ *zweitens* ein Türöffner für die einzig ernst zu nehmende Bedrohung der Psychologie bzw. ihrer Theorien und Konzepte: dem New-Wave-Reduktionismus. Denn wie aufgezeigt wurde, ist mit einer schlechten intertheoretischen Passung die Gefahr verbunden zu schlussfolgern, dass aufgrund der offensichtlichen Realitätsferne dieser Größen andere an ihrer Stelle treten sollten, die dieses Manko nicht haben und idealerweise eine vergleichbare oder vielleicht sogar bessere explanatorische und prognostische Güte aufweisen. Die psychologische Begriffs- und Theoriebildung mag in vielen Fällen zwar eine gewisse Unabkömmlichkeit aufweisen, die eine vollkommene Erset-

natürlich auf eine Weise verfasst werden, die eine gute Passung zu bestimmten physichen/neurobiologischen Parametern aufweist. Nur führt das weit verbreitete Diktum Fodors dazu, dass dies kein anstrebenswertes Ziel ist, sondern im Gegenteil eher vermieden werden sollte.

zung psychologischer Theorien durch neurobiologische Theorien verhindert. Dies hat zum einen den Grund, dass die propagierte Ersetzung wissenschaftlicher Theorien und Konzepte oder gar ganzer wissenschaftlicher Disziplinen auf Intra-Level-Kontexte bezogen ist. Eine *wissenschaftliche Revolution* wie die Ablösung der Phlogiston-Theorie durch Lavoisiers Sauerstoff-Chemie oder die Ablösung der klassischen Newtonschen Physik durch Einsteins Relativitätstheorie kann insofern auf das Verhältnis von Psychologie und Neurobiologie gar nicht zutreffen, da sich ihre Theorien und Konzepte auf unterschiedliche Organisationsebenen des menschlichen Organismus beziehen. Eine *ausschließlich* neurobiologische Erforschung des Denkens und Fühlens ist zum anderen auch dadurch ausgeschlossen, dass die Biologie bei der Erforschung des Gehirns ohne das semantische bzw. begriffliche Gerüst der Psychologie vollkommen blind wäre. Ohne die forschung leitende „Suchbilder“, ohne eine klare Vorstellung davon, was mentale Prozesse und Eigenschaften ausmacht, stochern Biologen mit selbst den ausgefeiltesten Methoden nur im tiefsten neuronalen Nebel herum (vgl. Dörner, 2004, S. 38). Oder wie Gardner (1987) es im Hinblick auf die Erforschung der Wahrnehmung ausdrückt:

One cannot have an adequate theory about anything the brain does unless one also has an adequate theory about the activity itself. It is not possible to study perception – even in its most fine-grained forms – without a theory of perception. [...] From this perspective, it is not possible to enter into the nervous system as a disinterested observer who is simply chronicling the facts as many neurosciences assume they are doing. (S. 286f)

Dennoch befeuert die Fodorsche Argumentationsweise krude reduktionistische Bestrebungen und verfestigt den problematischen Status der Psychologie als einer Wissenschaft, der ein ontologisches Fundament

fehlt bzw. deren Begrifflichkeiten und Erklärungsweisen ein großes Maß an Redundanz und Inkonsistenz aufweisen.

Drittens folgt aus der Ablehnung psychophysischer Identitätsannahmen und dem gleichzeitigen Bestreben, nicht völlig aus dem Kanon einer zumindest naturwissenschaftlich orientierten Wissenschaft auszusteigen (was wohl die Vereinnahmung des Substanzdualismus mit sich brächte), die Notwendigkeit der Postulierung einer (vermeintlich) „minimalen“ Abhängigkeit der psychischen Prozesse und Eigenschaften von ihren physischen Grundlagen. Wie aufgezeigt wurde, entpuppt sich der mit dieser Auffassung assoziierte *nicht-reduktive Physikalismus* als wahrer Wolf im Schafspelz. Es ist gerade die Annahme einer unidirektionalen Mikroabhängigkeit, die dazu führen würde, dass den Mikroebenen und ihren Entitäten und Eigenschaften gegenüber den Makroebenen und ihren Entitäten und Eigenschaften in Form eines Mikrophysikalismus ein *ontologisches und epistemisches Primat* zukäme. Dagegen bringen die geschmähte und vorschnell zur Gefahr für die Psychologie erklärte psychophysische Identitätsannahme (bzw. ontologische Reduktion) und selbst eine epistemische Reduktion nach dem Modell von Nagel bzw. dem der reduktiven Erklärung keineswegs einen solchen Mikrophysikalismus mit sich.

Entgegen Fodor wäre für die Psychologie die Durchführbarkeit einer (annäherungsweise) Nagelschen Reduktion aufgrund einer guten intertheoretischen Passung zwischen psychologischen und (neuro-)biologischen Termen oder die Verfügbarkeit reduktiver Erklärungen nach dem Modell von Kim (1998, 2005) etwas Wünschenswertes und keineswegs für ihre Existenz Gefährliches. Wie McCauley (2007) sagt: „[...] any putative *reductions of a science*, i.e., reductions in situations that involve *good* intertheoretic mappings in interlevel contexts, lead neither to the elimination or

the displacement of an upper level theory, but rather to its *vindication*" (S. 138; Hervorh. im Orig.).

Es ist ein großes Problem, dass ein Bemühen um eine Einheit oder Vereinheitlichung wissenschaftlicher Disziplinen, Paradigmen, Theorien oder auch nur Konzepte gegenwärtig (in sehr undifferenzierter Weise) in einem schlechten Licht steht. Verantwortlich hierfür ist primär die Annahme, dass sich eine solche Vereinheitlichung auf einen Mikroreduktivismus (bzw. Mikrophysikalismus) beziehen würde, dessen Bestreben es ist, alle Erklärungen höherstufiger Theorien auf solche der (Mikro-)Physik zurückzuführen und dass mit der Begründung, dass letztendlich alle Entitäten, Prozesse und Eigenschaften solche sind, die in den Gegenstandsbereich der (Mikro-)Physik fallen. Hierbei handelt es sich um eine Annahme, die an prominenter Stelle vor allem von Oppenheim und Putnam (1958) befeuert wurde und eine Art kollektiven Glaubenssatz aller höherstufigen Wissenschaften darzustellen scheint. So wehrt sich die Soziologie bekanntermaßen vehement gegen einen „Psychologismus“, die Psychologie gegen einen „Biologismus“, die Biologie gegen einen „Physikalismus“, womit allerdings nicht nur gemeint ist, dass man sich gegenüber einem Erklärungsprimat der basaleren Wissenschaften wehrt, sondern auch gegen alle Vereinheitlichungstendenzen, die den *Forschungsgegenstand* der verschiedenen Disziplinen betreffen. So scheint eine um ihre Existenzberechtigung fürchtende Soziologie behaupten zu müssen, dass soziale *Phänomene* sich in keiner Weise auf solche der Psychologie reduzieren lassen, während eine nach „Autonomie“ strebende Psychologie scheinbar behaupten muss, dass sich psychische *Phänomene* in keiner Weise auf solche der (Neuro-)Biologie reduzieren lassen, und die Biologen wiederum, dass biologische *Phänomene* in keiner Weise Phänomene sind, die mit solchen zusammenfallen, wie sie die Physik untersucht. Wie Bischof (2008) sagt, wird die affektive Abwehr gegen

Reduktion und Reduktionismus gespeist „vom unverzichtbaren Anspruch auf Kompetenz im eigenen Tätigkeitsfeld, wenn nicht gar von wissenschaftshierarchischem Dünkel. [...] Die einzige empirische Wissenschaft, die ohne Hemmungen reduktiv vorgehen kann, ist die Physik“ (S. 207).

Jedoch sollte nicht nur in intradisziplinärer, sondern auch interdisziplinärer Hinsicht eine zentrale Komponente wissenschaftlicher Arbeit darin bestehen, Regelmäßigkeiten zu erkennen und mittels Abstraktion die produzierte Informationsmenge (wozu auch die Anzahl postulierter wissenschaftlicher Konstrukte und Theorien gehört) zu *reduzieren* (vgl. ebd.). Oder andersherum: Es kann nicht Sinn und Zweck wissenschaftlichen Agierens sein, eine ungeordnete Informationsflut zu schaffen, die nur lose zusammenhängende oder gar völlig unverbundene empirische Einzelbefunde aufzuweisen hat. Genauso wenig wie es sinnvoll wäre, in gänzlich ungeordneter Art und Weise bzw. *ad hoc* theoretische Konstrukte zur Lösung eng umgrenzter fachlicher Fragestellungen einzuführen, ohne einen gewissen Abgleich mit dem bestehenden begrifflich-konzeptuellen Netzwerk innerhalb einer Disziplin oder auch disziplinenübergreifend vorzunehmen. Die Folgen der Nicht-Berücksichtigung dieser Grundsätze sind im Hinblick auf die Psychologie hinlänglich bekannt: eine extreme Redundanz und Uneinheitlichkeit des theoretischen Gerüsts innerhalb des Faches, eine Vielzahl von unkoordinierten Ad-hoc-Erklärungen als auch nur eine marginale intertheoretische Passung bzw. logische Verträglichkeit mit den Grundannahmen, Konzepten und Befunden angrenzender wissenschaftlicher Disziplinen. Dies sind keine Kennzeichen von „Autonomie“, sondern von einer tiefen Fehlentwicklung, die – wie einleitend in dieser Arbeit aufgezeigt wurde – mit der weitestgehend fehlenden Installation einer *Theoretischen Psychologie* zusammenhängt.

Wie Wunsch (2000) im Hinblick auf das Verhältnis von Chemie und (Quanten-)Physik sagt, mag zwar eine *starke* interdisziplinäre Reduktion (was bei ihm einem *Reduktionismus* entspricht) weder wünschenswert noch überhaupt möglich sein, denn das Betreiben von Chemie auf der Basis rein physikalischer Konzepte ist nicht möglich – hierzu bedarf es zweifelsohne *chemischer* Begrifflichkeiten. Dagegen kann eine (moderate) Reduktion in der Form der Behandlung chemischer Fragen durch Konzepte der Physik durchaus einen wichtigen Beitrag zur Chemie leisten, indem „chemisch interessante Phänomene vom quantenmechanischen Standpunkt her verständlich und beschreibbar [gemacht werden]. [...] Man hat gezeigt, dass es *in den beiden Fächern anerkannt wichtige Strukturen* gibt, die *miteinander logisch verträglich* sind. Man kann sich also nicht nur in der Praxis, sondern auch in der Theorie gegenseitig helfen“ (S. 202; Hervorh.: M.L.). Er plädiert daher aus wohl überlegten Gründen dafür, dass die kühne Vision eines Reduktionsprogramms durchaus wieder aufgegriffen werden sollte und die Wissenschaften wieder stärker aneinander gebunden werden (vgl. ebd.). Der Schlüssel hierzu ist eben, dass man sich über die von den wissenschaftlichen Disziplinen betrachteten *anerkannt wichtigen und logisch miteinander verträglichen Strukturen* verständigt. So wie hier die Chemie und Physik zu einer Verständigung darüber gekommen sind, dass sie sich auf dieselben Strukturen mit (teilweise) unterschiedlichen Begrifflichkeiten beziehen bzw. dass sie ein und denselben Gegenstand auf unterschiedlichen Analyse- bzw. Organisations-ebenen betrachten, bedarf es auch einer ebensolchen Verständigung zwischen Psychologie und (Neuro-)Biologie. Oder wie Mausfeld (2010) es sehr treffend ausdrückt:

Die Prämisse, dass psychologische Leistungen eines biologischen Organs, des Gehirns, sind, stellt zwar eine vernünftige und im Rahmen der Naturwissenschaften

konkurrenzlose Prämisse dar, doch bleibt sie umso folgenloser, je magerer das theoretische Substrat auf psychologischer und biologischer Seite ist. Je reichhaltiger hingegen eine explanatorisch angemessene Theoriebildung ist, desto größer sind auch die Chancen, sie durch biologische Befunde verfeinern und in eine übergreifende Perspektive integrieren zu können. (Mausfeld, 2010, S. 181f)

Auch die Psychologie muss als empirische Wissenschaft darum bemüht sein, die in ihrem Gegenstandsfeld enthaltene Informationsfülle, soweit es möglich ist, zu *reduzieren*. Wie in der Physik (bewerkstelligt durch die Theoretische Physik) muss in der psychologischen Theoriebildung eine Vereinheitlichung und Kohärenz der bestehenden empirischen Datelage und der mit dieser zusammenhängenden theoretischen Konstrukte geschaffen werden. Die bestehende Informationsmenge muss hierbei – idealerweise auch interdisziplinär – auf eine möglichst geringe Menge an Grundannahmen und hypothetischen Konstrukten reduziert werden. Allerdings regt sich, sobald der Begriff der Reduktion in irgendeiner Form auftaucht sofort Widerstand. „Der Begriff ‚Reduktion‘ hat bei uns keinen guten Klang – wer sich auf eine solche Aktivität einlässt, läuft Gefahr, als ‚Reduktionist‘ angeprangert zu werden“ (Bischof, 2008, S. 206f). Insofern zeigt sich abermals, wie wichtig es ist, das Konzept der Reduktion (bzw. reduktiven Erklärungen) klar von einem Reduktionismus abzugrenzen. Während Reduktionen oder reduktive Erklärungen nützliche, unentbehrliche und in der Regel für die „reduzierten“ Phänomene harmlose Mittel des wissenschaftlichen Methodenkanons sind, wird unter dem Begriff *Reduktionismus* in der Wissenschaftstheorie vielmehr ein programmatisches Vorgehen oder generelles Wissenschaftsprogramm verstanden, bei dem Reduktionen nicht nur ein Mittel, sondern nahezu ein Selbstzweck sind, den die Wissenschaft zu verfolgen hat. Keinesfalls bedeutet eine Reduktion psychischer Prozesse und Eigenschaften auf solche der (Neuro-)Biologie, dass es nur eine einseitige Be-

einflussung der Neurobiologie auf die Psychologie geben könnte. Wenn es eine erfolgreiche Reduktion in der von Wüensch beschriebenen wünschenswerten Form geben können sollte, dann ist eine *wechselseitige Beeinflussung* zu erwarten (vgl. McCauley, 1986, 2007; Pauen, 1999; Hüttemann, 2004). Eine Reduktion des Psychischen im Sinne einer psychophysischen Identitätsannahme, die auf mereologisch-mechanistischen Prinzipien fußt, schließt eine einseitige Mikroabhängigkeit aus. Es gibt weder ein ontologisches noch epistemisches Primat des Neuronalen gegenüber dem Psychischen, (neuro-)biologischer gegenüber psychologischer Erklärungen oder ein solches der (Neuro-)Biologie gegenüber der Psychologie trotz der Akzeptanz einer gewissen *Einheit* des Forschungsgegenstandes der beiden wissenschaftlichen Disziplinen.

9.2. Der Gegenstand der Psychologie bzw. psychologischer Erklärungen

Um eine Reduktion nach dem Schema Nagels²⁸¹ bzw. intertheoretische Integration der Psychologie zu bewerkstelligen, bedarf es nun in der Tat auch einer gewissen Reduktivität mentaler Eigenschaften und Prozesse selbst, damit die hierfür benötigten Brückenprinzipien, die eine Verbindung zwischen den Konstituenten beider Theorien herstellen müssen, greifen können. Es bedarf also – um noch ein-mal Wüensch (2000) sehr treffenden Worte aufzugreifen – *anerkannt wichtiger und logisch miteinander verträglicher Strukturen* im Sinne einer systematischen Verbindung

²⁸¹ Es sollen im Folgenden Nagel-Reduktionen als „Prüfstein“ herangezogen werden, um zu klären, ob hierdurch negative Folgen für die Psychologie bzw. ihren Gegenstand gegeben sind. Wenn dies nicht der Fall ist, gilt dies natürlich auch für das Modell reduktiver Erklärungen, das weniger restriktive Voraussetzungen hat.

zwischen den Konzepten (respektive deren Bezugsgegenständen), die die Psychologie und Neurobiologie in ihren Theorien zum Gegenstand haben. Die hier entwickelte, auf mereologisch-mechanistischen Prinzipien fußende psychophysische Identitätskonzeption liefert die benötigte Verbindung.

Zur Erinnerung: Nach Maßgabe der hier entwickelten Theorie des Geistes ist das Psychische ein reales Phänomen, das auf einer höherstufigen Ebene einer mereologisch-mechanistisch individuierten natürlichen Organisationsebene des menschlichen Organismus angesiedelt ist. So sind mentale Eigenschaften nicht wie in der physischen Realisierungstheorie abstrakte Eigenschaften zweiter Ordnung, sondern selbst (mikrobasierte) Kennzeichnungen der strukturellen Verfasstheit oder Konfiguration eines Systems bzw. (menschlichen) Organismus. Mentale Prozesse sind auf einer eigenen höherstufigen Komplexitätsebene angesiedelt, die durch neuronale Netzwerke konstituiert wird. Ein mentaler Prozess bildet hierbei die Spitze einer aufgespannten lokalen Teil-Ganzes-Hierarchie und ist koextensiv-identisch mit den agierenden Entitäten aller ihm zugrunde liegenden natürlichen Organisationsebenen. Eine mentale Eigenschaft ist daher eine mikrobasierte Struktureigenschaft eines (informationsverarbeitenden) Systems und ist durch eine (potentiell unendliche) Abfolge von miteinander koextensiv-identischen Mikrokonfigurationen gekennzeichnet.

Mit Bezug auf das Beispiel der *räumlichen Orientierung* bzw. des *räumlichen Gedächtnisses* lässt sich sagen, dass dieses durch einen Mechanismus bewerkstelligt wird, der sich auf verschiedenen Ebenen einer natürlichen Organisation von in einer Aktivität befindlichen Entitäten genauer untersuchen lässt. Die verschiedenen Ebenen sind durch eine Teil-Ganzes-Beziehung charakterisiert, bei der sich die agierenden Entitäten einer höheren Ebene des Mechanismus in solche niedrigerer Ebenen zerlegen

lassen (vgl. Craver & Bechtel, 2007; Craver, 2007). Die Eigenschaft „räumliche Orientierung“ (bzw. räumliches Gedächtnis) hat dabei eine bestimmte kognitive Struktur, die z.B. darin bestehen wird, sich die räumliche Umgebung (bildhaft) vorzustellen, seine eigene gegenwärtige Position und die zuvor selbst durchlaufenen Pfade mit dieser Repräsentation abzugleichen, um eine Entscheidung darüber zu treffen, welcher Weg eingeschlagen werden sollte. Es sind also Prozesse, die vor allem durch kognitionspsychologische Modelle (wie das Arbeitsgedächtnismodell Baddeleys) beschrieben werden können.

Neben dieser kognitiven Struktur, die das mentale Phänomen primär kennzeichnet, weist das Haben von räumlicher Orientierung eine Reihe weiterer *Mikrostrukturen* auf. Nach dem gegenwärtigen Kenntnisstand besteht auf der neuronalen Ebene eine komplexe Mikrostruktur, die als zentralen Bestandteil die – in einer ganz bestimmten Aktivität befindliche – Hippocampus-Formation zum Gegenstand hat. Weiterhin gilt, dass die agierenden Entitäten der neuronalen Ebene in eine Struktur zerlegt werden können, bei der die im synaptischen Spalt zwischen Neuronen stattfindende Langzeitpotenzierung eine zentrale Rolle innehat, sodass die räumliche Orientierungsleistung durch eine bestimmte Mikrostruktur auf der elektrophysiologischen Ebene gekennzeichnet ist. Schließlich weist das Haben von räumlicher Orientierung auf der molekularen Ebene eine ganz bestimmte Konfiguration auf, bei der die NMDA-Rezeptor-Aktivierung von primärer Bedeutung ist. Potentiell ließe sich diese Reihe unendlich fortsetzen oder auch durch verschiedene, hier nicht erwähnte, vielleicht in ihrer Bedeutung noch gänzlich unentdeckte oder unverstandene kompositorische Zwischenebenen (z.B. der Zellorganellen) erweitern. Hierbei handelt es sich um Ebenen, die primär von der (Neuro-) Biologie bzw. der kognitiven Neurowissenschaft untersucht, begrifflich

verfeinert und in einen theoretischen Gesamtzusammenhang gebracht werden können.

Entsprechendes gilt nun für alle mentalen – und somit auch für *phänomenale* – Eigenschaften und Prozesse, wie dies im Hinblick auf Farberfahrungen ausführlich erörtert wurde (vgl. Kap. 7.1.3). Wenn nun die Psychologie mit kognitionspsychologischen Modellen die mentale Ebene beschreibt, ist dies von keinem geringeren Stellenwert als die (neuro-)biologischen und biochemischen Modellvorstellungen der niedrigeren Ebenen der mereologisch-mechanistischen Hierarchie. Auch hier handelt es sich um *Beschreibungen* dieser Ebenen, die niemals deren volle Realität einfangen, sondern immer nur eine Reihe kausal relevanter Aspekte – im Sinne bestimmter modellhafter, abstrahierender, durch die herangezogenen Forschungsmethoden geleiteter Vorstellungen bezüglich der faktisch agierenden Entitäten – herausgreifen können.

Wichtig ist an dieser Stelle, dass eine (potentiell unendliche) Abfolge von Identitätsbeziehungen vorliegt: Die Eigenschaft der räumlichen Orientierung ist etwas, was durch verschiedene (mikro-)strukturelle Konfigurationen natürlicher – mereologisch-mechanistisch individuierter – Organisationsebenen gekennzeichnet ist. Durch die Erkenntnis, dass (beobachterunabhängige) Eigenschaften eines Systems (mikro-)strukturelle Eigenschaften sind und den Fakt, dass eine Eigenschaft auf jeder natürlichen (über die Teil-Ganzes-Beziehung individuierten) Organisationsebene eines Systems auf eine ganz bestimmte Mikrostruktur rekurriert, muss angenommen werden, dass im Hinblick auf die koextensiven Mikrostrukturen ein *Identitätsverhältnis* vorliegt. Das Haben räumlicher Orientierung besteht daher nicht nur darin, eine bestimmte kognitive Struktur zu besitzen (obgleich es sich hier um die epistemisch wichtigste Kennzeichnung handeln mag), sondern auch darin, eine Vielzahl von in einem Identitätsverhältnis stehenden mikrostrukturellen Konfigurationen in

einer mereologisch-mechanistischen Hierarchie aufzuweisen. Die realisierte psychologische Zieleigenschaft selbst ist dabei gewissermaßen der „Kit“, der aus den verschiedenen koextensiven strukturellen Gliederungsebenen eine *Einheit* bildet. Sie ist zugleich keine Eigenschaft zweiter Ordnung (im Sinne Kims), sondern eine strukturgebundene, mikrobasierte Eigenschaft (erster Ordnung), die auf einer potentiell unendlichen Abfolge natürlicher Organisationsebenen jeweils eine spezifische Konfiguration aufweist.

Insofern hat Hüttemann (2004) vollkommen recht, wenn er sagt, dass das Psychische eine (kompositorisch begründete) *ontologische Einheit* darstellt²⁸² (vgl. S. 124ff). Jede Instantiierung einer mentalen Eigenschaft weist neben ihrer psychischen immer auch eine neuronale, molekulare, atomare etc. Realität auf, denn auf jeder potentiell herausgreifbaren Ebene der mereologisch-mechanistischen Hierarchie hat die mentale Eigenschaft eine bestimmte Organisationsstruktur. Natürlich muss man aber zugleich konstatieren, dass es sich hierbei tatsächlich um einen *potentiellen* Aspekt handelt, der durch die gegenwärtig herangezogenen mechanistischen Organisationsebenen zur Erklärung von bestimmten mentalen Phänomenen keineswegs abgedeckt wird. Für viele mentale Eigenschaften und Prozesse bestehen momentan noch keine oder nur sehr unausgereifte mechanistische Erklärungsmodelle bzw. es lassen sich kaum solche mereologisch-mechanistischen Hierarchien wie im Fall des räumlichen Gedächtnisses etablieren (und selbst hier ist das Modell noch rela-

²⁸² Dabei verdeutlicht er zugleich, dass das, was wir als Einheit herausgreifen, keine *absolute* Feststellung beinhaltet, sondern dass es von unseren Vorannahmen abhängig ist, was wir als Einheit (im Sinne eines kompositorischen Systems) begreifen. So ist es eine *relative*, von bestimmten momentanen Erwägungen abhängige Feststellung, was wir als Einheit (im Sinn eines kompositorischen Systems) begreifen. Demnach ist es eine relative bzw. pragmatische Festlegung, ob wir unser Planetensystem als ein kompositorisches Ganzes oder als ein Teil der Milchstraße begreifen (vgl. Hüttemann, 2004, S. 125).

tiv lückenhaft). Obgleich mittlerweile in einer Vielzahl psychologischer Erhebungen (neuro-)biologische Parameter mit erhoben werden, findet sich ein „explanatorischer Spagat“ im Sinne der Entwicklung eines dezierten mechanistischen Erklärungsmodells bisher kaum vor.

Für die Kluft zwischen psychologischen und neurobiologischen Begrifflichkeiten, die einer (Nagel-) Reduktion im Weg steht, ist in erster Linie die multiple Realisierbarkeit bzw. kompositorische Variabilität mentaler Eigenschaften und Prozesse verantwortlich. Psychologische und (neuro-)biologische Theorien betrachten keine Einzelfälle (im Sinne einzelner Instantiierungen von Eigenschaften oder Prozessen), sondern versuchen, allgemeingültige Zusammenhänge zu erfassen. Wie aufgezeigt wurde, handelt es sich hierbei allerdings um keine prinzipielle, unüberbrückbare Hürde. Zunächst ist es eine Tatsache, dass sich die meisten (natur-)wissenschaftlichen Erkenntnisse aus der *induktiven Verallgemeinerung* bestimmter empirischer Befunde ergeben, sodass auch hier eine *heuristische* psychophysische Identitätsannahme auf Typen-Niveau eine gewinnbringende Forschungsstrategie und recht solide Basis für eine (Nagel-)Reduktion sein dürfte. Trotz der nicht zu unterschätzenden neuronalen Plastizität gibt es doch eine recht zuverlässige individuen- und sogar speziesübergreifende Übereinstimmung bei der Lokalisierung bestimmter Funktionen im Gehirn (vgl. Bechtel & McCauley, 1999; McCauley & Bechtel, 2001; McCauley, 2007, 2012; Bechtel, 2008).

Zugleich handelt es sich bei der kompositorischen Variabilität um einen Aspekt, der primär auf das unterschiedliche *Granularitätsmaß* psychischer und (neuro-)biologischer Konzepte zurückgeht (vgl. Bechtel & Mundale, 1999). Dass häufig eine schlechte Passung der Begriffe vorliegt, lässt sich vordergründig über wissenschaftshistorische Aspekte und hintergründig vor allem über eine dem Psychischen selbst inhärente „Dynamik“ erklären. So wurden psychologische und biologische (respektive

medizinische) Forschung sowie Begriffs- und Theoriebildung lange Zeit relativ unabhängig voneinander betrieben, sodass es nicht verwundern muss, dass sich das begriffliche und theoretische Instrumentarium beider Fächer relativ stark voneinander unterscheidet. Aber natürlich greift diese Erklärung zu kurz, denn es stellt sich die Frage nach den Ursachen für eine solche Tendenz, die nur eine Widerspiegelung eines fundamentalen psychischen Geschehens sein dürfte.

Entscheidend hierfür sind nun die (zumindest implizit) bereits von Gustav Theodor Fechner (1851/ 1922) herausgestellten psychischen Strukturmerkmale der Perspektivität und phänomenalen Transparenz. Vereinfacht gesagt können wir „Gehirn“ und „Geist“ als Strukturebenen in einer mit sich selbst identischen mereologisch-mechanistischen Hierarchie aufgrund unserer eigenen organismischen Dispositionen immer nur in Form bestimmter Repräsentationen auffassen. Wir haben keinen „absoluten“, sondern immer nur einen „relativen“ Zugriff auf unsere Umwelt und uns selbst. Im engeren Sinne gilt, dass wir aufgrund der Plastizität unseres Wahrnehmungssystems einen *phänomenal variablen* Zugriff auf unsere Erfahrungswirklichkeit haben (vgl. Mausfeld, 2003, 2010a). Auch „Gehirn“ und „Geist“ triggern in diesem Sinne die Einnahme bestimmter Perspektiven, die aufgrund unserer (phylogenetischen und ontogenetischen) Dispositionen inkommensurabel erscheinen. Die phänomenale Differenz zwischen Gehirn und Geist geht dabei auf die Verarbeitung des „Gehirn-Geist-Komplexes“ durch unterschiedliche konzeptuelle Formen unseres Wahrnehmungssystems zurück, die sich im Laufe unserer Phylogenese herausgebildet haben. Die konzeptuellen Formen sind wahrscheinlich integrale Bestandteile unseres perzeptuellen Systems selbst, hängen aber auch mit höherstufigen kognitiven Systemen (Theory-of-Mind- vs. Theory-of-Matter-Modul) zusammen, die sich im Laufe der Individualentwicklung ausformen und zu einer Segmentierung

unserer (perspektivengebundenen) Wahrnehmungseindrücke führen. Hinzu kommt – und dies dürfte an dieser Stelle sogar entscheidend sein – das Phänomen der *phänomenalen Transparenz*, durch das komplexe repräsentationale Systeme wie der Mensch gekennzeichnet sind (vgl. Metzinger, 2003a, 2006). Das Phänomen bezeichnet den Umstand, dass auf der Ebene des Erlebens die strukturellen (neuronalen, molekularen, atomaren etc.) Grundlagen ebendieses Erlebens selbst nicht einsichtig sind. Dass bestimmte neuronale Parameter bei der Herausbildung von Begrifflichkeiten zur Beschreibung mentaler Eigenschaften und Prozesse, die dann auch von der akademischen Psychologie aufgegriffen wurden, keine Rolle gespielt haben, muss daher keineswegs verwundern. Sie sind nichts, was das geistige Erleben in irgendeiner Form (im Regelfall) phänomenal durchdringen würde und damit (erlebnismäßig einsichtig) strukturieren könnte. Die Folge ist, dass während psychische Zustände und Prozesse weiterhin in relativ groben Klassen zusammengefasst werden, Gehirnzustände und -prozesse immer feinkörniger bestimmt bzw. voneinander abgegrenzt werden (vgl. Bechtel & Mundale, 1999). Dieser Fakt begrenzt nun tatsächlich die Reichweite der Durchführbarkeit einer Theorie-Reduktion im Sinne Nagels, denn so lassen sich die erforderlichen Brückenprinzipien zwischen den in den korrespondierenden Theorien verwendeten Konzepten in vielen Fällen nicht installieren. Es handelt sich hierbei allerdings nicht um eine prinzipielle Grenze, sondern nur um eine solche des *gegenwärtigen* Standes der Theorie- und Konzeptbildung der Psychologie und (Neuro-)Biologie.

Arbeitet man in Zukunft verstärkt daran, das Granularitätsmaß auf beiden (bzw. allen) Ebenen einer mereologisch-mechanistischen Hierarchie konstant zu halten (*Subtypisierungsansatz*), verflüchtigt sich das Problem der multiplen Realisierbarkeit bzw. kompositorischen Variabilität mentaler Eigenschaften und Prozesse. Es wird eine stabile Grundlage für die

Formulierung psychophysischer Identitätsannahmen auf Typen- bzw. Eigenschaftsniveau geschaffen, die für eine (Nagel-)Reduktion psychologischer Theorien (und ihrer Begrifflichkeiten) erforderlich sind. Bei einer grobkörnigen Analyse sind hierbei Konzepte für psychische Zustände zu fokussieren, die sich über verschiedene Individuen und Arten hinweg zeigen. Bei gleichem Granularitätsmaß erhält man so eine entsprechende (relativ grob gefasste) Aktivität in verschiedenen Hirnarealen, was die Möglichkeit einer Eins-zu-eins-Korrespondenz bewahrt (vgl. Bechtel & Mundale, 1999). Wenn man dagegen eine feinkörnige Granularität heranzieht und zwischen verschiedenen psychologischen Zuständen über Individuen hinweg differenziert oder sogar bei einem Individuum über verschiedene Zeitpunkte hinweg, dann muss man auch bei der Analyse des Gehirns ein feines Granularitätsmaß anwenden und auch hier werden sich so die psychologischen Differenzen in neuronalen Differenzen und die neuronalen Differenzen in psychologischen Differenzen widerspiegeln (vgl. ebd.).

Obleich Mausfeld in diesem Zusammenhang konstatiert, dass ihm bisher kein einziger Fall bekannt sei, bei dem bei einer psychologischen Grundlagenfrage ein Schlusswort durch die Neurowissenschaften erfolgt sei, stellt er dennoch fest: „Je reicher und tiefer eine psychologische Theorie über einen spezifischen Phänomenbereich ist, umso größer sind natürlich auch die Aussichten, dass neuropsychologische Befunde das Spektrum möglicher Theorievorstellungen weiter einzugrenzen helfen“ (in Fiedler et al., 2008, S. 36). Vielleicht sollte man in der Tat kein Schlusswort durch neurowissenschaftliche Forschung erwarten. Eine gewisse Orientierungsfunktion kann sie aber schon bieten, insbesondere wenn es um die Schaffung eines möglichst einheitlich-kohärenten, ontologisch respektablen, nicht-redundanten sowie metatheoretisch kompatiblen (in intra- und interdisziplinärer Hinsicht) begrifflich-konzeptuellen Funda-

ments der Psychologie geht. Denn wenn es in der Psychologie bisher primär nur die „Kreuzvalidierung“ (vgl. Cronbach & Meehl, 1955) ist, die zur Berechtigung der Gültigkeit bestimmter psychologischer Konstrukte herangezogen wird, bietet die Eingliederung mentaler Eigenschaften und Prozesse in mereologisch-mechanistische Hierarchiemodelle eine ganz neue Chance zur Etablierung ontologisch respektablem Größen, die solchen der Biologie, Chemie oder Physik in nichts nachstehen.

Natürlich stellt die Konstanzhaltung der Granularität psychologischer und neurobiologischer Konzepte ein in vielen Fällen sicher nur schwer zu realisierendes Ideal dar. Für das hier verfolgte Vorhaben einer begrifflich-konzeptuellen bzw. ontologischen Fundierung der Psychologie ist es aber keineswegs notwendig, dass alle psychologischen Konzepte eine solche Passung aufweisen und sich dem Subtypisierungsansatz vollends stellen müssen. Wenn die psychophysische Identitätstheorie in der hier entwickelten mereologisch-mechanistisch begründeten Form die Basis einer solchen Grundlegung darstellt, dann bedeutet es nicht, dass sich auch *faktisch* in jedem Fall bestimmte „Eins-zu-Eins-Korrespondenzen“ ableiten lassen müssen. Wie McCauley (2007) sagt: „A perfectly interesting version of the identity theory need only require that humans' mental states are identical to some of their brain states. It does not require that all mental states, especially those cast at coarse grains within psychology, are identical to one another or that all identities of psychological and neural states are cast at a single grain“ (S. 147). Die Vielfältigkeit und Vielgestaltigkeit unserer mentalen Eigenschaften mögen – auf sehr feinmaschigen Granularitätsniveaus – ihrer *perfekten* Analysierbarkeit (oder gar ihrer bloßen Klassifizierbarkeit bzw. „Individuierbarkeit“) und damit auch ihrer weitergehenden Zuordenbarkeit in einer mereologisch-mechanistischen Hierarchie entgegenstehen. Diese *epistemische* Schwierigkeit beschneidet die grundsätzliche Gültigkeit der identitätstheoretischen Kon-

zeption als *ontologische* Basis der Psychologie dabei nicht. Denn obgleich sich auch nicht jeder biologische Begriff eindeutig auf Typen-Niveau auf ein physikochemisches Pendant beziehen lässt (man denke z.B. an ökologische Nischen, „Hackordnungen“ und dergleichen), besteht in der Biologie kein Zweifel daran, dass jedes biologische Phänomen eine physische Basis oder „Realität“ hat. Selbst wenn also die Prämissen wahr wären, die der (Fodorschen) Umgarnung des Funktionalismus als Garant der „Autonomie“ der Psychologie zugrunde liegen und Typen-Identitäten in epistemischer Hinsicht nicht immer nachweisbar sind, stellt dies weder die Gültigkeit einer psychophysischen Identitätskonzeption, noch die reduktive Integration eines großen Teils des begrifflichen Instrumentariums und theoretischen Fundus der Psychologie und (Neuro-)Biologie infrage (vgl. ebd.).

Des Weiteren führt der Fakt, dass jede psychische Eigenschaft bzw. jeder psychische Prozess neben ihrer/seiner psychischen immer auch eine neuronale, molekulare, atomare etc. Realität aufweist, in keiner Weise zu der Annahme, dass eine dieser Ebenen dazu ausgezeichnet ist, die *eigentliche* Realität des Mentalen auszumachen. Es gibt in ontologischer Hinsicht kein Kriterium, das die eine Ebene gegenüber der anderen Ebene als etwas „Realeres“ herausstellen könnte (vgl. Pauen, 1999; Hüttemann, 2004; McCauley, 2007). Wenn Mausfeld (2010) problematisiert, dass wir nicht wissen „auf welcher physikalischen Ebene der Gehirnorganisation die relevanten Prinzipien der Erzeugung ‚perzeptueller Objekte‘ auf der Basis eines physico-chemischen Inputs zu suchen sind, etwa auf der Ebene von Neuronen oder komplexer dynamischer Systeme von Neuronen, auf der Ebene subzellulärer Strukturen von interagierenden Proteinen, auf der Ebene quantentheoretischer Vorgänge etc.“ (S. 181), dann liegt hier eine gewisse Verwirrung vor. Es gibt nicht eine Ebene, die die „Herberge“ des Geistes oder die von Wahrnehmungsprozessen darstellt,

sondern es handelt sich um ein „ebenendurchdringendes“ Phänomen. Allerdings ist es in der Tat die Ebene – wie Mausfeld es bezeichnet – komplexer dynamischer Systeme von Neuronen (also die Ebene neuronaler Netze), der eine gewisse Maßgeblichkeit im Hinblick auf die Existenz mentaler Prozesse und Eigenschaften bzw. für die Strukturgliederung aller niederstufigen Ebenen zukommen mag.

9.3. Die Realität und kausale Wirkmächtigkeit des Psychischen

Generell gilt also, dass ein mentaler Prozess alle Ebenen der mit ihm identischen mereologisch-mechanistischen Hierarchie agierender Entitäten „durchdringt“. Dabei gibt es – bezogen auf einen bestimmten herausgegriffenen Zeitpunkt – keine Inter-Level-Kausalität (im Sinne der *causa efficiens*) zwischen den Ebenen einer solchen Hierarchie, sondern nur ein Phänomen, dass man als *mechanistisch vermittelten Effekt* bezeichnen kann (vgl. Craver & Bechtel, 2007). Denn wenn die agierenden Entitäten der verschiedenen Ebenen eines Mechanismus in einer Identitätsbeziehung stehen, dann ist es absurd zu sagen, dass das Verhalten der Teile das Verhalten des Ganzen *verursacht* hat oder andersherum.

Obleich es sich also hier nicht um Kausalität im gewöhnlichen Sinne handelt, schlägt sich jede Veränderung auf der psychischen Ebene über die neuronale Ebene bis zur Ebene der Moleküle oder Atome nieder, weil eine *Verschränkung* der Organisationsebenen dadurch begründet ist, dass die (kausal agierenden) Entitäten der verschiedenen Ebenen in einer Teil-Ganzes-Relation stehen, für die ein Verhältnis der Koextensivität bzw. Identität besteht und somit notwendigerweise das Prinzip der *mereologischen Supervenienz* greift. Wenn etwas Teil eines Ganzen ist, dann muss

es notwendigerweise in demselben räumlichen Feld und zum selben Zeitpunkt existieren wie das Ganze. Da ein Molekül aus Atomen besteht, muss eine Veränderung auf der molekularen Ebene immer auch ein Pendant auf der atomaren Ebene haben. Verändert sich irgendein Parameter des Moleküls, muss auch eine Veränderung bei den Atomen vorliegen, aus denen es besteht. Jede Veränderung eines mentalen Zustandes muss daher genauso mit einer Veränderung auf der neuronalen, mikrozellulären, molekularen, atomaren und subatomaren Ebene verbunden sein. Kommt es – z.B. im Rahmen einer Psychotherapie – zu einer kognitiven Umstrukturierung des semantischen Netzwerkes einer Person, dann muss auch eine Modifikation der Konfiguration aller niedrigeren Ebenen der entsprechenden Mechanismen stattgefunden haben. Denn eine mentale Eigenschaft ist etwas, das eine Konfiguration auf *allen* Ebenen des sie hervorbringenden Mechanismus hat. Sie ist genauso ein mentales Erlebnis mit einer phänomenalen Qualität oder einem kognitiven Gehalt wie auch etwas, das gewissermaßen in den Atomen dieser mechanistischen Hierarchie „steckt“.

Betrachtet man statt eines ganz bestimmten Zeitpunktes einen zeitlichen *Verlauf*, dann lässt sich über die Annahme mechanistisch vermittelter Effekte hinaus sagen, dass das Psychische eine kausale Wirkmächtigkeit hat. So besteht über die verschiedenen Zeitpunkte einer gegebenen Ereigniskette hinweg ein *kausales* Geschehen. Dies ist allerdings keineswegs ein Intra-Level-Phänomen, sondern eben ein „*ebenendurchdringendes*“ Phänomen, also ein Phänomen, bei dem das kausale Geschehen auf allen Ebenen des jeweiligen (Gesamt-)Mechanismus abläuft. Auch wenn also für einen ganz konkreten Zeitpunkt nicht angenommen werden kann, dass das Geschehen auf einer Ebene für das Geschehen auf der anderen Ebene kausal (im Sinne der *causa efficiens*) verantwortlich ist, ist es aber vollkommen korrekt, wenn man ableitet, dass z.B. das Geschehen in der

mechanistischen Hierarchie zu einem Zeitpunkt t_1 das Geschehen in der herausgegriffenen mechanistischen Hierarchie zu einem Zeitpunkt t_2 verursacht hat. Allerdings lässt sich dieses *eine* kausale Geschehen auf unterschiedlichen Ebenen betrachten. Zwischen dem psychischen und physischen Kausalnexus besteht daher keineswegs ein Bruch. Sie lassen sich nicht gegenüberstellen oder gegeneinander ausspielen. Es ist ein und dasselbe Kausalgeschehen, das – vermittelt über die Teil-Ganzes-Beziehung – alle Ebenen einer mit sich selbst identischen mechanistischen Hierarchie betrifft.

Zugleich gilt, dass der Zuschreibung von Kausalität zur Ebene des Psychischen hierbei keine Willkür zukommt, sondern der psychischen Ebene bei einer mereologisch-mechanistischen Hierarchie, die bis hin zum Psychischen aufgespannt ist, in vielen Fällen faktisch eine Maßgeblichkeit für das Auftreten der nachfolgenden Ereignisse zugeordnet werden muss. Denn aufgrund der systemischen Emergenz, also aufgrund der Tatsache, dass das Ganze Eigenschaften aufweist, die seine Teile entbehren, folgt, dass ein solches ebenendurchdringendes Kausalgeschehen wie z.B. eine Bewegungsinitiation *kraft* ebendieser instantiierten mentalen Eigenschaft zustande gekommen ist. Auch wenn es zwar keine Ebene gibt, auf der *allein* sich der Kausalprozess ereignet, ist die Ebene des Mentalen bei einem Vorgang, bei dem z.B. ein bestimmtes Ziel verfolgt wird, insofern maßgeblich, als dass erst hier die eigentlichen Gründe für diese Handlung gegeben sind. Würden die mentalen Eigenschaften nicht vorliegen, würde die gesamte mechanistische Hierarchie aufgrund der mereologischen Supervenienzrelation zwischen den Ebenen nicht die Struktur aufweisen können, die sie bei Vorhandensein der mentalen Eigenschaft hat. Vereinfacht gesagt: *Geist ist geformte Materie* (die auf der psychischen Ebene eine ganz bestimmte kognitive resp. phänomenale Struktur hat). Wenn ein Mensch dazu fähig ist, zu sprechen, sich räum-

lich zu orientieren, sich in seine Artgenossen emotional hineinzuversetzen, Pläne für die Zukunft und zur Lösung komplexer Probleme zu schmieden, eine feine Ironie in einer Aussage zu erkennen und sich hierüber zu erfreuen, dann kann er dies nur *kraft* der psychischen Fähigkeiten bzw. der Eigenschaften, die hierfür notwendig sind, auch wenn sich die kausale Aktivität bei diesen mentalen Leistungen auf allen Ebenen der hierzu installierten Mechanismen vollzieht. Die mereologisch-mechanistischen Hierarchien dieser Vorgänge sind nach Maßgabe der realisierten mentalen Eigenschaften gestaltet. Ohne die mentalen Eigenschaften gäbe es auch die jeweiligen „genuin physischen“ Ereignisketten nicht (in dieser Form).

Wenn dennoch der Eindruck vorhanden sein mag, dass die Akzeptanz einer psychophysischen Identitätsannahme damit verbunden sein muss, dass das mentale Phänomen hierdurch zu *nichts anderem als* einem neuronalen (biochemischen, molekularen...) Phänomen degradiert würde, hat dies zum einen mit der problematischen Interpretation der „klassischen“ Identitätstheorie (im Sinne einer strikten *Identifikation als* anstatt einer mereologisch fundierten *Identifikation mit*), wie sie von Smart (1959) propagiert wurde, zu tun. Zum anderen suggerierte aber auch die nachfolgende Diskussion der physischen Realisierungstheorie eine Vorstellung, bei der mentale Eigenschaften als bloße Ansammlung neuronaler, biochemischer oder gar mikrophysikalischer Eigenschaften anzusehen wären. Es ist die *aggregative* Interpretation der Beziehung natürlicher Organisationsebenen, bei denen höherstufige Eigenschaften *nichts anderes als* eine Ansammlung niederstufiger Eigenschaften sind.

Aggregative Ebenen, wie sie beim bloßen Aufaddieren von den Massen der Teile zur Gesamtmasse des Systems bestehen mögen, stellen eine extrem seltene Ausnahme in Bezug auf die Beziehung zwischen natürlichen Organisationsebenen dar. Schon gar nicht spielen sie im Fall kom-

plexer informationsverarbeitender Systeme wie dem Menschen eine bedeutende Rolle. Denn wie ein Mechanismus als ein Ganzes agiert, ist häufig völlig abhängig von den Operationen, die seine Komponenten ausführen. Dabei sind die agierenden Entitäten und ihre Organisationsstruktur von Ebene zu Ebene (des Mechanismus) oft völlig verschieden bzw. es sind sehr unterschiedliche Teilmechanismen auf den einzelnen Ebenen der mereologisch-mechanistischen Hierarchie installiert.

Einer aggregativen Interpretation des Psychischen hat die Diskussion in der Philosophie (des Geistes) nichtsdestotrotz starken Vorschub geleistet, wenn bei denjenigen, die sich auf eine Identitätsrelation beziehen (wie z.B. Kim, 1998, 2005; Pauen, 1999, 2001/2005), aber auch ihren Gegnern (z.B. Shoemaker, 2003, 2007; Gillett, 2002, 2003, 2010), die Möglichkeit einer Identität von *Eigenschaftsklassen* erörtert wird. Dies ist ontologisch verfehlt und keinesfalls nur ein Ausdruck einer harmlosen sprachlichen Konvention oder Ungenauigkeit (vgl. Polger & Shapiro, 2008). Mentale Eigenschaften (bzw. ihre Instantiierungen) sind nicht einfach neuronale Eigenschaften (bzw. ihre Instantiierungen). Diese Aussage ergibt genau genommen keinen Sinn. Mentale Eigenschaften haben u.a. eine neuronale *Mikrostruktur*. Mentale Eigenschaften sind identisch mit einer (potentiell) unendlichen Anzahl von Mikrostrukturen unterschiedlichster Komplexitätsstufen. Mentale Eigenschaften *sind* aber dennoch nicht einfach neuronale Eigenschaften. Selbst wenn man an dieser Stelle nun einmal die Realisierungsmetapher gebraucht (was aus dem Grund der Ablehnung der mit der Realisierungstheorie verbundenen Annahme mentaler Eigenschaften als mysteriöser Eigenschaften zweiter Ordnung in dieser Arbeit möglichst vermieden wurde), wäre es nicht so, dass eine mentale Eigenschaft durch eine neuronale Eigenschaft (oder Ansammlung neuronaler Eigenschaften) realisiert wird. Ihre Realisierung erfolgt durch eine bestimmte neuronale (biochemische, molekulare, atomare...) *Konfigura-*

tion agierender Entitäten verschiedenster Organisationsebenen. Trotz der Akzeptanz der grundsätzlichen Möglichkeit einer „Reduktion“ psychischer Phänomene besteht daher keinerlei Anlass für eine „Nichts-andere-als-Phobie“ (vgl. Wimsatt, 1976, S. 242).

9.4. Ontologische Einheit und explanatorischer Pluralismus

Was bedeutet eine (Nagel-)Reduktion psychologischer Theorien aber nun für die Psychologie und ihren explanatorischen Status? Zunächst kann konstatiert werden, dass die alleinige Reduktion wissenschaftlicher Theorien bzw. ihrer Konzepte niemals zu einer vollständigen Erklärung bestimmter (höher-stufiger) Phänomene führt. Wie Polger (2004) sagt, sind wissenschaftliche Theorien und die durch sie angeregten Erklärungen immer selektiv; sie erklären nie alles oder etwas vollkommen (vgl. ebd., S. 203). Insofern bestreitet der u.a. von McCauley (2007, 2007a), Polger (2004) oder auch Bechtel (2007, 2008, 2009) verfochtene explanatorische Pluralismus, dass intertheoretische Beziehungen „exhaust all of the selection pressures in the resulting co-evolutionary process and that those selection pressures are exerted exclusively from the bottom up“ (McCauley, 2007, S. 137).

Die Reduktion einer psychologischen Theorie auf eine solche der Biologie ist nie ein Vorgang, bei dem alle Anforderungen an die Erklärung des von der Theorie aufgegriffenen psychischen Phänomens erfüllt würden. Vielmehr führt auch das Vorliegen einer erfolgreichen Mikroerklärung grundsätzlich nicht zu einer abschließenden Erklärung eines Phänomens. Im Gegenteil: Mit Blick auf die Komplexität der strukturellen Organisation eines Mechanismus, seiner Vielgestaltigkeit und funktionalen

Einbindung in größere Kontexte bleibt stets ein *explanatorischer Pluralismus* erforderlich: „Explanatory pluralism *underscores* the ongoing interaction of scientific enterprises carried out at the various analytical levels. All scientific explanation is partial explanation from the perspective of some analytical level or other“ (ebd.; Hervorh. im Orig.).

Nehmen wir an, wir hätten unter Berücksichtigung der begrifflichen Granularitätsparität eine psychologische Gedächtnistheorie konzipiert, deren Konzepte eine sehr gute Passung zu einer (neuro-)biologischen Theorie über die Funktionsweise des Hippocampus und einiger anderer hiermit assoziierter kortikaler Hirnzentren aufweisen. Nehmen wir weiterhin an, dass die Theorie des Arbeitsgedächtnisses nach Baddeley (1986) eine solche Theorie wäre und somit die von ihm postulierten Konstrukte der zentralen Exekutive, des visuell-räumlichen Notizblockes und der phonologischen Schleife mit bestimmten neurobiologischen Konzepten (bzw. neuronalen Strukturen) der oben genannten (neuro-)biologischen Theorie so in Verbindung gebracht werden können, dass ein adäquates mechanistisches Erklärungsmodell entsteht. Unter diesen Voraussetzungen ließe sich nach dem von Nagel vorgeschlagenen Schema eine *Reduktion* der psychologischen auf eine (neuro-)biologische Gedächtnistheorie vornehmen. Aber was würde das dann genau bedeuten?

Natürlich kann eine Mikroerklärung bestimmter Gedächtnisphänomene auf diese Weise einen wichtigen Beitrag zu deren Verständnis leisten. Wenn wir also ganz genau wüssten, welche Gehirnzentren mit welchen psychischen Phänomenen, Fähig- und Fertigkeiten verbunden sind, haben wir unser Wissen um Gedächtnisphänomene insofern erweitert, als dass wir das höherstufige Phänomen (idealerweise) exakt in seine Komponenten auf der neuronalen Ebene zerlegen können. Bestünden zum gleichen psychologischen Inhaltsbereich konkurrierende Theorien, wäre auf diese Weise wahrscheinlich eine Klärung der Frage möglich, welche

Theorie die realen Verhältnisse am besten einfängt. Eine Theorie, deren Konzepte und Annahmen keinerlei Entsprechung auf der neuronalen Ebene haben, mag zwar durch bestimmte Kriterien ihre Berechtigung haben (gute Vorhersagen zukünftigen Verhaltens...), kann aber eben nicht für sich beanspruchen, eine adäquate Beschreibung der tatsächlichen Gegebenheiten vorzunehmen.

Wenn wir jedenfalls unsere – nach Baddeley durch den visuell-räumlichen Notizblock bewerkstelligte – Fähigkeit, eine „mentale Landkarte“ zur Orientierung in einer bestimmten Umgebung konstruieren zu können, durch einen ganz bestimmten im Gehirn installierten Mechanismus erklären können, hat dies einen profunden explanatorischen Gehalt. Es ist von vergleichbarer Qualität (wenn sicher auch nicht von gleicher „Tragweite“) wie die Entdeckung, dass wir die Aggregatzustände von Stoffen mittels der spezifischen Lage- und Bewegungseigenschaften ihrer Teilchen (Moleküle oder Atome) erklären können. Wir verlagern unsere (analytische) Perspektive hierbei von einer Makro- zu einer Mikroebene und können die Beschaffenheit der einen Organisationsebene mit der Beschaffenheit einer anderen (niedrigeren) Organisationsebene (einer mereologisch-mechanistischen Hierarchie) erklären.

Genau genommen müsste nun eigentlich auch eine Passung der Konzepte der neurobiologischen Theorie zu denen der elektrophysiologischen Theorie, die u.a. die Langzeitpotenzierung im synaptischen Spalt zum Gegenstand hat, als wiederum eine solche der Konzepte dieser Theorie zu denen einer molekularen Theorie hergestellt werden, auf die sich vor allem die NMDA-Rezeptor-Aktivierung bezieht. Denn dies sind die Ebenen, die nach dem mechanistischen Erklärungsmodell als kausal relevant für bestimmte Gedächtnisleistungen (wie räumliche Orientierung) angesehen werden können. Auch die Neurobiologie müsste also um eine

Granularitätsparität der von ihr verwendeten Konzepte zu denen niederstufigerer wissenschaftlicher (Sub-)Disziplinen bemüht sein.

Durch eine solche Mikroerklärung werden aber nun die psychologischen Konstrukte zur Beschreibung des mentalen Phänomens keineswegs hinfällig. Ohne den Rückgriff auf die psychologischen Konstrukte bliebe die Hirnforschung auch zukünftig blind bei der weiteren Erforschung des Gedächtnisses. Das psychologische Konzept zu einer psychischen Eigenschaft oder einem psychischen Prozess bleibt auch in diesem Fall weiterhin der „Kit“, der es ermöglicht, aus dem relativ undifferenzierten neuronalen Geflecht funktionale bzw. mechanistische Einheiten herauszugreifen. Insofern trifft McCauleys (2007) Einschätzung zu, dass eine solche Reduktion eher eine Unterstützung für den Stellenwert der psychologischen Theorien und Konzepte bedeutet, als dass dieser hierdurch beschnitten würde.

Wenn nun die Wissenschaften so geschachtelt sind, daß die jeweils nächstniedere sich mit der Binnenstruktur von Systemen befaßt, mit deren Außenstruktur sich die jeweils höhere befaßt, so ist klar, daß die nächstniedere bezüglich der Relevanz der von ihr untersuchten Merkmale im Dunkeln tappen muß. Sie braucht Leitlinien, die sie nur von der höhergeordneten Wissenschaft beziehen kann. Mit diesen Leitlinien allerdings ist es dann möglich, daß die Binnenstruktur eines Systems so erforscht wird, daß seine Außenstruktur daraus ableitbar ist. (Dörner, 1979, S. 128)

Zugleich lässt sich das erforderliche explanatorische Manöver, das sich um Gedächtnisleistungen des Menschen dreht, durch solche Mikroerklärungen keineswegs zum Abschluss bringen. Wie Bechtel (2007, 2009) herausgestellt hat, müssen neben der bloßen Zerlegung eines Mechanismus in Teilstrukturen bzw. agierende Entitäten verschiedener Komplexitätsstufen auch die Organisations- und Einbettungsbedingungen eines Mechanismus berücksichtigt werden, damit ein Phänomen adäquat verstanden werden kann. Die simple Lokalisierung von bestimmten Pro-

zessen im Gehirn allein ist nicht sehr aussagekräftig. Es bedarf einer weiteren Abklärung der konkreten Organisation des Mechanismus, die auch und *gerade* von der Psychologie geleistet werden muss:

The contribution of organization in creating mechanisms that do things their parts cannot do undergirds the need for scientists to discover the particular forms of organization realized in a mechanism. This is what higher-level disciplines, such as psychology, have the resources to provide. Their autonomous contribution is secure even if higher-level activities are not multiply realizable. At the same time, mechanistic explanation also recognizes the value of reductionistic investigations into how the components perform their operations. Higher-level inquiries and reductionistic inquiries complement each other, and often provide heuristic guidance to each other. Neither on its own suffices and neither can do the work of the other. (Bechtel, 2007, S. 193)

Organisationsprinzipien sind dabei nichts, was in den Teilen eines Systems selbst stecken würde. Es ist im Gegenteil die Konfiguration der Komponenten eines Mechanismus, die einem bestimmten Muster folgt. Obgleich also die neuronalen Grundlagen bestimmter Gedächtnisleistungen identifiziert worden sind, bedarf es einer weitergehenden Analyse, wie z.B. die einzelnen Komponenten des Arbeitsgedächtnissystems zusammenwirken, ob sich bestimmte Prozesse gegenseitig verstärken oder hemmen. Natürlich können auch hier wiederum die neuronalen, biochemischen etc. Grundlagen dieser Aspekte in Form von mechanistischen Organisationsebenen untersucht werden und z.B. bestimmte vorwärts gerichtete, rekurrente oder rückgekoppelte neuronale Netzwerke aufgespürt werden.

Über die Organisationsprinzipien hinaus spielen die konkreten Einbettungs- bzw. Umgebungsbedingungen eines Mechanismus eine entscheidende Rolle zur Erklärung des durch ihn hervorgebrachten Phänomens. So sind Mechanismen nicht einfach passiv reagierende Systeme, sondern werden durch Umwelteinflüsse moduliert. Hier stellt sich z.B. die Frage,

inwieweit ein bestimmtes Vorwissen, Wachheit/Müdigkeit, die Kenntnis bestimmter Gedächtnisstrategien oder auch die Motivation ebendiese anzuwenden, die gezeigte Gedächtnisleistung beeinflusst. Gibt es entwicklungsbedingte Veränderungen? Hängt die Gedächtnisleistung mit bestimmten sozial-emotionalen Kompetenzen wie der Emotionsregulation zusammen etc. pp.? Dies alles sind Fragen, die nur die Psychologie adäquat beantworten kann (obgleich auch hier wiederum nach den Mikrostrukturen verschiedener Ebenen all dieser Aspekte gesucht werden kann und auch gesucht werden sollte). Die Vergegenwärtigung und Berücksichtigung von Organisations- und Einbettungsbedingungen von Mechanismen revidiert dabei keineswegs die Möglichkeit der Reduktion eines höherstufigen Phänomens, sondern zeigt nur auf, dass eine solche immer nur eine (für sich genommen unvollkommene) Komponente einer hinlänglichen wissenschaftlichen Erklärung eines solchen Phänomens sein kann:

All of the complex mechanisms and systems that populate the biological, psychological, cognitive, and socio-cultural sciences are, from the standpoint of explanation, greater than the sums of their parts and demand study as isolated, as constituted, and as situated. Such multi-level study not only offers a richer account of these particular mechanisms and their operations, it also enhances our understanding of all of the systems engaged in a mechanistic hierarchy. [...] Because they countenance the full range of possible cross-scientific relationships, 'these versions of explanatory pluralism do not confine themselves to reductive analyses only. Reductive explanation is a valuable contributor, but these approaches also embrace forms of non-reductive explanation as well [...]. For the explanatory pluralist, all explanations are partial explanations, all explanations are from some perspective, and all explanations are motivated by and respond to specific problems. (McCauley, 2007, S. 150)

Trotz einer Identitätsannahme bzw. der Akzeptanz der Möglichkeit einer ontologischen Reduktion gibt es kein ontologisches bzw. epistemisches *Primat* der Mikroebenen und damit auch keinerlei privilegierte Stellung

der (Neuro-)Biologie gegenüber der Psychologie. Denn obgleich eine Mikroerklärung mentaler Phänomene bzw. (Nagel-)Reduktion psychologischer Theorien in vielen Fällen (potentiell) möglich sein mag, ist hiermit kein Mikrophysikalismus verbunden. (Psychophysische) Identität und Mikrophysikalismus sind begrifflich unabhängig voneinander bzw. eine mereologisch-mechanistisch basierte psychophysische Identitätsannahme schließt eine unidirektionale Mikrodetermination aus. Zwischen den agierenden Entitäten der verschiedenen Ebenen einer mereologisch-mechanistischen Hierarchie besteht ein *symmetrisches* Abhängigkeitsverhältnis, eine *wechselseitige* Determination (vgl. Hüttemann, 2004, 2009; Craver & Bechtel, 2007).

Auch wenn wir also z.B. räumliche Orientierung über bestimmte neuronale Mechanismen (bzw. genauer: die Mikrostruktur agierender Entitäten auf der neuronalen Ebene) in einem gewissen Umfang erklären können, bedeutet dies nicht, dass es die neuronalen Mechanismen sind, die einseitig die Eigenschaft der räumlichen Orientierung determinieren. Diese Eigenschaft besteht vielmehr selbst darin, genau diese Mikrostruktur (und eine Vielzahl anderer in der mereologisch-mechanistischen Hierarchie) aufzuweisen. Wenn hier überhaupt sinnvoll von einer Determinationsbeziehung gesprochen werden kann, ist diese also beidseitig, denn mit ihrem Sosein legt die mentale Eigenschaft das Sosein der neuronalen Ebene fest, wie deren Sosein das Sosein der mentalen Eigenschaft bedingt.

Wie aufgezeigt wurde, ist es außerdem fraglich, ob in komplexen Systemen wie dem menschlichen Organismus, der die Existenz mentaler Eigenschaften und Prozesse ermöglicht, überhaupt mikrophysikalische Gesetze instantiiert sind. Aufgrund der spezifischen Organisation (bzw. der „Zusammensetzungsgesetze“; Hüttemann, 2004, 2009) und der Einbettung in einen ganz spezifischen Kontext können die Theorien

niederstufiger Wissenschaften wie die der Physik, Chemie, aber auch die der (Neuro-)Biologie aus prinzipiellen Gründen keine bessere, aussagekräftigere, prognostisch verlässlichere Beschreibung des Psychischen vornehmen als die Psychologie, insofern sie um eine Berücksichtigung ebendieser Facetten bemüht ist. Es gilt: „Causation requires laws to be instantiated. Therefore causal agency has to be attributed to the compound system rather than to its parts“ (Hüttemann, 2004, S. 125). Natürlich könnte die Psychologie – z.B. in einer falsch verstandenen Nachahmung der Naturwissenschaften – gerade auch einfache Gesetze des Verhaltens und Erlebens aufspüren wollen (wie einst der „orthodoxe“ Behaviorismus; vgl. hierzu Bischof, 2008), würde damit aber gerade ihr Potential, ihre Überlegenheit bei der Erklärung komplexer Verhaltensweisen bzw. dem Verhalten in komplexen Situationen anhand mentaler Eigenschaften und Prozesse nicht ausschöpfen.

Auch Hüttemann (2004) spricht sich daher anstatt eines Mikrophysikalismus für einen *pragmatischen Pluralismus* aus, der mit einer ontologischen Einheit der Natur verbunden ist. Dass eine ontologische Einheit und ein (explanatorischer) Pluralismus sonst als unvereinbar erscheinen, liegt seines Erachtens nur daran, dass eine solche Einheit für gewöhnlich mit mikrophysikalistischen Erwägungen verbunden wird, so wie bei Oppenheim und Putnam (1958) letztendlich alles aus mikrophysikalistischen Teilchen besteht und alles von den mikrophysikalischen Gesetzen beherrscht wird. Dass wir eine Einheit der Natur bzw. ihrer Entitäten haben, ohne dem Mikrophysikalischen ein Primat zuzusprechen, liegt in zweierlei Hinsicht an der bidirektionalen Determinationsrelation in kompositorischen Systemen: „First, a small number of determination relations suffices to fix the relation between all kinds of compound systems and their constituents. Second, the system of determination relations makes a unified whole out of nature“ (Hüttemann, 2004, S. 124). Insofern

ist das in der Abbildung 38 dargestellte Kontinuum zwischen einer (Nagel-)Reduktion bei guter intertheoretischer Passung und einem explanatorischen Pluralismus bei geringer intertheoretischer Passung etwas irreführend. Eine (Nagel-)Reduktion wird faktisch nie alle Bedingungen für eine vollständige Erklärung eines Phänomens abdecken können.

Zusammenfassend kann man daher sagen, dass nicht jede Form von Reduktion von der Psychologie abgelehnt werden sollte, sondern nur ein kruder (epistemischer) Reduktionismus, der behauptet, dass die eigentliche Erklärung mentaler Eigenschaften und Prozesse nur von der Neurobiologie (oder noch basalerer Wissenschaften) erbracht werden kann. Mit Rückbezug auf die zentrale Frage, ob sich die Psychologie vor einem Neuroreduktionismus hüten muss, kann dies mit Verweis auf die Ausführungen Mausfelds (2003, 2010) als korrekt eingestuft werden. Problematisch ist allerdings, dass durch die Ausführungen zugleich der Eindruck erweckt wird, dass auch jede Form psychophysischer Identität abgelehnt werden müsse (was deutlich wird, wenn er z.B. über die phänomenale Differenz zwischen Gehirn und Geist für die Trennung zwischen Biologie und Psychologie argumentiert), obgleich eine solche „ontologische Reduktion“ keineswegs mit den Zielen und Konsequenzen des von ihm angegriffenen Neuroreduktionismus einhergeht. Im Gegenteil fördert die ontologische Einheit und Integration des Psychischen den Status der Psychologie bzw. ihrer Konzepte und Theorien und sollte als ein zentrales Ziel psychologischer Begriffs- und Theoriebildung erhoben werden. Die Psychologie bleibt unabkömmlich, auch wenn die Phänomene bzw. Eigenschaften, auf die sie sich bezieht, koextensiv-identisch mit den verschiedenen (mikro-)strukturellen Konfigurationen der agierenden Entitäten der sie „hervor-bringenden“ Mechanismen sind. Sie bleibt unabkömmlich, weil es keine unidirektionale Abhängigkeit des Psychischen vom Gehirn gibt, sondern nur eine wechselseitige Determination, eine

„ebenendurchdringende“ Kausalität (die sich häufig sogar nach Maßgabe der mentalen Eigenschaft vollzieht) und eine gewisse „Eigengesetzlichkeit“ höherstufiger Ebenen komplexer Systeme bzw. ihrer Eigenschaften und Prozesse.

9.5. Das mechanistische Paradigma als Leitbild der Forschung und Theoriebildung

Was im Zuge der Besprechung der Leitlinien und Konsequenzen deutlich geworden sein sollte, ist, dass das mechanistische Paradigma einen geeigneten explanatorischen Ansatz für die Psychologie als auch für ihre Einbettung in die Kognitions- und Neurowissenschaften abliefert (vgl. Polger, 2004). Es kann dabei als eine konzeptuell kohärentere als auch wissenschaftlich plausiblere Erweiterung oder Korrektur des funktionalistischen Paradigmas angesehen werden (vgl. Eronen, 2010). Während eine funktionalistische Konzeption des Geistes zumeist die mysteriöse Beziehung mentaler Eigenschaften als Eigenschaften zweiter Ordnung zu ihren physischen Realisierern thematisiert (insofern sie eine solche Frage nicht von vornherein ausblendet), kann der mechanistische Ansatz dafür Gewähr tragen, dass mentale Eigenschaften als höherstufige Eigenschaften *erster Ordnung* einen klar bestimmbareren Platz im wissenschaftlichen Gesamtsystem erhalten.

Wie in Kapitel 4 in dieser Arbeit erläutert wurde, orientieren sich die funktionalistischen Modelle des Geistes für gewöhnlich an der *Computer-Metapher* des Menschen, d.h., dass der Mensch als ein informationsverarbeitendes System wie ein Computer aufgefasst wird. Sie postulieren relativ abstrakte theoretische Modelle psychischer Abläufe, die aufgrund be-

stimmter bekannter Eingangsvariablen zumeist sehr zuverlässige Prognosen zukünftigen Verhaltens des betrachteten Systems ermöglichen. Die Modelle liefern gewissermaßen die Operatoren, mit denen die explanatorisch ausgeklammerte mentale Blackbox des Behaviorismus nach der kognitiven (funktionalistischen) Wende gefüttert werden durften (insofern wurde der Behaviorismus nie durch den Kognitivismus *abgelöst*, wie es so oft heißt, vielmehr wurde Ersterer durch Letzteren lediglich um bestimmte Variablen *erweitert*). Kennzeichnend für die funktionalistische Konzeption des Geistes ist also, dass das Denken, Entscheiden, Planen und Handeln des Menschen über recht *abstrakte* Informationsverarbeitungsprozesse beschrieben wird, die (programmatisch) unabhängig von neurowissenschaftlichen Constraints konzipiert werden können. Wenn aber doch die Frage nach der Beziehung kognitionspsychologischer Modelle zu den tatsächlichen Vorgängen in unseren Köpfen gestellt wird, wird sie zumeist nach Maßgabe der Ausführungen Marrs (1982) beantwortet, bei dem die computationalen Prozesse seines Modells der visuellen Informationsverarbeitung (z.B. die Generierung einer dreidimensionalen Tiefenstruktur) durch bestimmte *formale Algorithmen* spezifiziert werden, von denen wiederum angenommen wird, dass sie eine bestimmte *physische Implementation* (z.B. in einem menschlichen Gehirn) aufweisen. Dass funktionale Prozesse auch eine physische Realität haben, wird also zwar durchaus zur Kenntnis genommen, was aber lange Zeit eher unangetastet blieb, war ein Versuch, die genaueren Umstände, Bedingungen, Constraints etc. einer solchen Implementation in Erfahrung bringen zu wollen. Erst der *konnektionistische* Ansatz hat sich dieser Aufgabe stärker verschrieben und in Computermodellen vor allem neuronale Netze moduliert, die näher am tatsächlichen Geschehen in unseren Gehirnen orientiert sind (vgl. Rumelhart & McClelland, 1986; Smolensky, 1988; Rumelhart, 1989/1991).

Wie ausführlich erörtert, wird die physische Implementation abstrakter funktionaler Prozesse und Eigenschaften in ontologischer Hinsicht in der Philosophie des Geistes über die Theorie der *physischen Realisierung* näher ausbuchstabiert. Dabei gilt es, das kausale Profil einer abstrakten funktionalen Eigenschaft (Verhaltensinputs und -outputs) über eine konzeptuelle Analyse ihres Begriffes zu bestimmen, um daraufhin diejenigen physischen (neuronalen) Strukturen zu ermitteln, die eben jenes kausale Profil aufweisen (vgl. Levine, 1993; Kim, 1998, 2005). Die Realisierungstheorie soll dabei den explanatorischen Ansatz dafür liefern, wie abstrakte, funktional charakterisierte Eigenschaften, wie sie von der Kognitionswissenschaft postuliert werden, ein Teil unserer physischen Welt sein können („funktionale Reduktion“). Unter ihren Voraussetzungen zeichnet sich ab, dass es die *neuronale* Realität mentaler Zustände ist, durch die das (funktionalistisch interpretierte) Psychische seinen „natürlichen“ Platz in der Welt erhalten soll und über die es kausal wirksam sein kann. Ein abstraktes, theoretisches Phänomen wird in diesem Sinne durch konkrete physische Gegebenheiten *realisiert* (quasi *materialisiert*), wobei es unterschiedliche physische Systeme sein können, die es bewerkstelligen, den computationalen oder konnektionistischen Formalismus in die Welt zu heben. Mentale Eigenschaften sind demnach abstrakte *Eigenschaften zweiter Ordnung*, die man – zumindest programmatisch – mittels funktionaler Beschreibungen erfassen und deren eigentliche Existenz man ausfindig machen kann, indem man sich auf die Suche nach denjenigen physischen Realisierern als Eigenschaften erster Ordnung macht, die genau diese funktionale Rolle ausfüllen. Der Realitätsstatus einer abstrakten funktionalen Beschreibung bzw. einer Eigenschaft zweiter Ordnung bleibt dabei äußerst fraglich.

Dagegen wird durch das mechanistische Verständnis mentaler Eigenschaften und Prozesse der funktionalistische Ansatz auf eine äußerst

gewinnbringende Art und Weise erweitert und (teilweise) korrigiert. Wie Bechtel (2008) sagt, ist ein Mechanismus eine Struktur, „performing a function in virtue of its component parts, component operations, and their organization“ (S. 13). Während der Funktionalismus bzw. die mit ihm assoziierte Realisierungstheorie die Beziehung abstrakter Informationsverarbeitungsprozesse zu bestimmten Vorgängen im Gehirn thematisiert, erweitert der mechanistische Ansatz dieses Anliegen zu einer Frage nach der Einbettung eines bestimmten (mentalen) Phänomens in eine durch die Teil-Ganzes-Beziehung charakterisierte *Strukturhierarchie verschiedenster natürlicher Organisationsebenen*. Wenn wir also eine kognitive Eigenschaft (wie räumliche Orientierung) heranziehen, stellt sich hier nicht einfach die Frage, welche Hirnstruktur ebendiese Eigenschaft *realisiert*, sondern es muss geklärt werden, welche kausal relevanten Mechanismen sich auf den verschiedensten Ebenen einer mereologisch-mechanistischen Hierarchie ausfindig machen lassen. So hat diese Eigenschaft neben ihrer kognitiven Struktur, die von der Psychologie über kognitionspsychologische Modelle beschrieben wird, eine bestimmte neuronale Struktur, die u.a. in einer Hippocampus-Aktivität besteht, aber neben anderen auch eine molekulare Struktur, für die die NMDA-Rezeptor-Aktivierung kennzeichnend ist. Ein mentales Phänomen weist also eine (potentiell unendliche) Abfolge von Mikrostrukturen auf, von denen keine die *eigentliche* Realität dieses Phänomens ausmachen würde. „This suggests that to functionally reduce a property *M* amounts to providing a mechanistic explanation for *M*. The upshot is that if we want to keep the model of functional reduction close to science, it turns out that there is no functional reduction over and above mechanistic explanation“ (Eronen, 2010, S. 146).

Der Funktionalismus bzw. der Ansatz der funktionalen Reduktion kann demnach als ein Spezialfall des mechanistischen Erklärungsansatzes auf-

gefasst werden. Statt der in der Philosophie verbreiteten Annahme, dass alle mechanistischen Eigenschaften funktionale Eigenschaften sind, insbesondere wenn sie kausal wirksame Eigenschaften sein sollen (vgl. z.B. Shoemaker, 1981; Lycan, 1987/2007), verhält es sich genau andersherum: „[F]unctional entities are mechanisms per excellence“ (Polger, 2004, S. 207; vgl. auch Polger & Flanagan, 2001). Entgegen dem Prinzip der funktionalen Reduktion spielt im mechanistischen Paradigma nicht nur die *Identifizierung* eines physischen (neuronalen) Realisierers dieser Eigenschaft eine Rolle, sondern ebenso die konkrete *Organisationsstruktur* des Mechanismus sowie dessen *Einbettung* in einen ganz bestimmten (komplexeren) Kontext. Insofern hat eine auf dem mechanistischen Ansatz basierende Identitätstheorie, wie sie in dieser Arbeit entwickelt wurde, „more explanatory resources than functionalism because it makes use of both contextual and constitutive explanations“ (Polger, 2004, S. 210). Ebenso hat sich eindrücklich gezeigt, dass das (Fodorsche) Diktum, dass der Funktionalismus als einziger Garant der „Autonomie“ der Psychologie fungiert, falsch ist. Während dieses Diktum – streng befolgt – dazu führen würde, dass „Realitätsferne“ oder metatheoretische Uneinpassbarkeit zu Kriterien einer adäquaten psychologischen Theorie- und Begriffsbildung würden und die Psychologie deswegen Gefahr liefe, kruden reduktionistischen Programmen wie dem New-Wave-Reduktionismus und aufgrund der Abwehr gegenüber einer psychophysischen Identitätsannahme zugleich einem Mikrophysikalismus anheim zu fallen, kann auf der Basis des mechanistischen Ansatzes (und seiner identitätstheoretischen Interpretation) die „Autonomie“ oder vielleicht besser *Integrität* der Psychologie abgeleitet werden, ohne sich hierbei hinter einem in weiten Teilen fadenscheinigen Verweis auf die multiple Realisierbarkeit mentaler Eigenschaften verstecken zu müssen. Wie Hüttemann (2009) sagt, kann man daher die Art von ontologischer Unabhängigkeit, die

nicht-reduktive Physikalisten wie Fodor im Blick haben, viel besser dadurch explizieren, dass „man den Realisierungsbegriff durch die physikalische Teil-Ganzes-Beziehung erläutert und ontologische Unabhängigkeit durch die ‚Universalität‘ der fraglichen Eigenschaften“ (ebd., S. 73).

Wenn in letzter Zeit häufig von einer *neurowissenschaftlichen Wende* in der Psychologie und Philosophie die Rede ist (vgl. z.B. Walde, 2008), trifft dies unter diesen Voraussetzungen nicht exakt den Kern der begonnenen, aber größtenteils sicher noch ausstehenden Weiterentwicklung des Faches. Der Ausdruck *mechanistische Wende* dürfte ein viel zutreffenderer Ausdruck und für eine adäquate Weiterentwicklung des Faches ein viel anstrebenswerterer *Maßstab* für Forschung, Konzept- und Theoriebildung sein. Denn nach Maßgabe des mechanistischen Paradigmas weist jede mentale Eigenschaft neben ihrer kognitiven und/oder phänomenalen Struktur immer eine (potentiell unendliche) Abfolge von Mikrostrukturen auf. Die neuronale Ebene ist also nur eine Betrachtungs- bzw. Organisationsebene unter vielen. Es gibt daher in *ontologischer* Hinsicht nichts, was die neuronale Ebene gegenüber den anderen Ebenen in der jeweiligen mereologisch-mechanistischen Hierarchie auszeichnen würde, obgleich sie in *explanatorischer* Hinsicht gegenüber den mikroskopischeren biochemischen oder gar mikrophysikalischen Ebenen von größerer Bedeutung sein mag (was aber eher mit unserem begrenzten analytischen Vermögen zusammenhängt als mit einer in der Sache selbst begründeten Ursache). Psychologische Theorien und Konzepte verlieren durch eine (Nagel-)Reduktion oder reduktive Erklärung bzw. die Einbettung ihres Bezugsgegenstandes in eine mereologisch-mechanistische Hierarchie keineswegs an Bedeutung. Ohne sie ist eine zielgerichtete Erforschung mentaler Eigenschaften und Prozesse, egal auf welcher Ebene dieser Hierarchie, nicht möglich.

Trotzdem bleiben auch im Hinblick auf den mechanistischen Ansatz bzw. seiner identitätstheoretischen Interpretation einige Fragen offen. So scheint es im Sinne des Erklärungslückenproblems auch hier gewisse Schwierigkeit bei der Klärung der Frage zu geben, wie genau das *Sosein* phänomenaler Eigenschaften wie Farbempfindungen, Emotionen oder Schmerzen (also ihre konkrete Erlebnisqualität) mit dem *Sosein* der installierten mechanistischen Struktur (also mit der konkreten Organisation der agierenden Entitäten der verschiedenen Ebenen) zusammenhängt. Ob eine solche Schwierigkeit allerdings wirklich eine prinzipielle Hürde der Erklärung widerspiegelt, wie viele Philosophen vermuten, oder ob es sich dabei doch nur um eine gewisse Illusion handelt (weil es hierfür keine Erklärung in der geforderten Art geben *kann*), wird sich wohl erst im Laufe der weiteren Forschung in der Psychologie und Philosophie bzw. der kognitiven Neurowissenschaft zeigen.

Literaturverzeichnis

- Aizawa, K. & Gillett, C. (2009). The (Multiple) Realization of Psychological and Other Properties in the Sciences. *Mind and Language*, 24, 181-208.
- Alais, D. & Blake R. (2005) (ed.). *Binocular Rivalry*. Cambridge: MIT Press.
- Alexander, S. (1920). *Space, Time, and Deity*, Vol. 2. London: Macmillan.
- Allolio-Näcke, L. (2011). Im Widerstreit der Ontologien. *Psychologische Rundschau*, 62, 113-116.
- Ames, A. (1951). Visual perception and the rotating trapezoidal window. *Psychological Monographs*, 65, 1-32.
- Anderson, J. A. (1973). A theory for the recognition of Items from short memorized lists. *Psychological Review*, 80, 417-438.
- Anderson, J. R. (2001). *Kognitive Psychologie*. Heidelberg: Spektrum.
- Anscombe, G. E. M. (1979). Under a Description. *Noûs*, 13, 219-233.
- Antony, L. & Levine, J. (1997). Reduction with autonomy. In J. Tomberlin (ed.), *Philosophical Perspectives 11: Mind, Causation, and World* (pp. 83-105). Boston: Blackwell.
- Aristoteles (1995). *Metaphysik*. Hamburg: Meiner.
- Aristoteles (1995a). *Über die Seele*. Hamburg: Meiner.
- Armstrong, D. M. (1966/1970). The nature of mind. In C. V. Borst (ed.), *The Mind/Brain Identity Theory* (pp. 67-79). London: Macmillan.
- Armstrong, D. M. (1978). *Universals and Scientific Realism*. Cambridge: Cambridge University Press.
- Armstrong, D. M. (1997). *A World of States of Affairs*. Cambridge: Cambridge University Press.
- Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.
- Baker, Lynn Rudder (1997). Why constitution is not identity. *The Journal of Philosophy*, 94, 599-621.

- Ballard, D. H. (1991). Animate vision. *Artificial Intelligence*, 48, 57-86.
- Barondes, S. H. (1995). *Moleküle und Psychosen. Der biologische Ansatz in der Psychiatrie*. Heidelberg: Springer.
- Barsalou, L. (1993). Challenging assumptions about concepts. *Cognitive Development*, 8, 169-180.
- Bartuschat, W. (1996). *Baruch de Spinoza*. München: Beck.
- Bechtel, W. (2001). The compatibility of complex systems and reduction: A case analysis of memory research. *Minds and Machines*, 11, 483-502.
- Bechtel, W. (2007). Reducing psychology while maintaining its autonomy via mechanistic explanation. In M. Schouten & H. Looren de Jong (eds.), *The matter of the mind: Philosophical essays on psychology, neuroscience and reduction* (pp. 172–198). Oxford: Blackwell.
- Bechtel, W. (2008). *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. New York: Psychology Press.
- Bechtel, W. (2009). Looking down, around, and up: Mechanistic explanation in psychology. *Philosophical Psychology*, 22, 543-564.
- Bechtel, W. (2012). Identity, reduction, and conserved mechanisms. In S. Gozzano & C. S. Hill (eds.), *New Perspectives on Type Identity. The Mental and the Physical* (pp. 43-65). Cambridge: Cambridge University Press.
- Bechtel, W. & McCauley, R. N. (1999). Heuristic identity theory (or back to the future): The mind-body problem against the background of research strategies in cognitive neuroscience. In M. Hahn & S. C. Stones (eds.), *Proceedings of the twenty first annual conference of the Cognitive Science Society* (pp. 67-72). Mahwah: Erlbaum.
- Bechtel, W. & Mundale, J. (1999). Multiple Realizability Revisited. Linking Cognitive and Neural States. *Philosophy of Science*, 66, 175-207.

- Bechtel, W. & Richardson, R. C. (1993). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Cambridge: MIT Press.
- Beckermann, A. (1996). Eigenschafts-Physikalismus. *Zeitschrift für philosophische Forschung*, 50, 3-25.
- Beckermann, A. (1999). *Analytische Einführung in die Philosophie des Geistes*. Berlin: Walter de Gruyter.
- Bickle, J. (1998). *Psychoneural Reduction: The New Wave*. Cambridge: MIT Press.
- Bickle, J. (2003). *Philosophy and Neuroscience: A Ruthlessly Reductive Account*. Dordrecht: Kluwer.
- Bieri, P. (1981/³1997). Einleitung in Teil 1: Materialismus. In P. Bieri (Hrsg.), *Analytische Philosophie des Geistes* (S. 31-55). Weinheim: Beltz.
- Bieri, P. (1981a/³1997). Einleitung in Teil 3: Selbstbewusstsein, Privatheit und Subjektivität. In P. Bieri (Hrsg.), *Analytische Philosophie des Geistes* (S. 201-208). Weinheim: Beltz.
- Bieri, P. (1994). Was macht Bewußtsein zu einem Rätsel? In W. Singer (Hrsg.), *Gehirn und Bewusstsein* (S.172-180). Heidelberg, Spektrum.
- Bieri, P. (2004). *Das Handwerk der Freiheit. Über die Entdeckung des eigenen Willens*. Frankfurt am Main: S. Fischer.
- Birbaumer, N. (2003). Biologische Psychologie 2010. Visionen zur Zukunft des Faches in der Psychologie. *Psychologische Rundschau*, 54, 120-128.
- Birbaumer, N. & Schmidt, R. F. (2000). Motivation und Emotion. In R. F. Schmidt & H.-G. Schaible (Hrsg.), *Neuro- und Sinnesphysiologie* (S. 455-475). Berlin: Springer.
- Birke, M. (2004). Materialismus. In P. Prechtel (Hrsg.), *Grundbegriffe der analytischen Philosophie* (S. 137-138). Stuttgart: J. B. Metzler.

- Bischof, N. (1981). Aristoteles, Galilei, Kurt Lewin – und die Folgen. In W. Michaelis (Hrsg.), *Bericht über den 32. Kongreß der Deutschen Gesellschaft für Psychologie, Band 1* (S. 17-39). Hogrefe: Göttingen.
- Bischof, N. (2008). *Psychologie: Ein Grundkurs für Anspruchsvolle*. Stuttgart: Kohlhammer.
- Bitbol, M. (2012). Downward Causation without foundations. *Synthese*, 185, 233-255.
- Blake, R. & Logothetis, N. K. (2002). *Visual competition*. *Nature Reviews Neuroscience*, 3, 13–21.
- Block, N. (1990). Can the mind change the world? In G. Boolos (ed.), *Meaning and Method: Essays on Honor of Hilary Putnam* (pp. 137-170). Cambridge: Cambridge University Press.
- Block, N. (2003). Do Causal Powers Drain Away? *Philosophy and Phenomenological Research*, 67, 133-150.
- Block, N. & Fodor, J. A. (1972). What Psychological States are Not. *The Philosophical Review*, 81, 159-181.
- Bloomfield, L. (1933/2000). *Die Sprache*. Wien: Edition Praesens.
- Bohman, J. (2000). The Importance of the Second Person: Interpretation, Practical Knowledge, and Normative Attitudes. In H. H. Kögler & K. R. Stueber (eds.), *Empathy and Agency. The Problem of Understanding in the Human Sciences* (pp. 222-224). Boulder: Westview.
- Borst, C. V. (1970). Introduction. In C. V. Borst (ed.), *The Mind/Brain Identity Theory* (pp. 13-29). London: Macmillan.
- Boyd, R. (1980). Materialism without Reductionism. What Physicalism Does Not Entail. In N. Block (ed.), *Readings in Philosophy of Psychology* (pp. 67-105). Cambridge: Harvard University Press.
- Bradley, A. (unpubliziert). *Dimensioned Realization and Multiple Realization*. 1-24.

- (<https://www.yumpu.com/en/document/view/17322428/1-dimensioned-realization-and-multiple-realization-by-adam>;
Zugriff am 13.10.2013)
- Brandtstädter, J. (1982). Apriorische Elemente in psychologischen Forschungsprogrammen. *Zeitschrift für Sozialpsychologie*, 13, 267-277.
- Brandtstädter, J. (1984). Apriorische Elemente in psychologischen Forschungsprogrammen. Weiterführende Argumente und Beispiele. *Zeitschrift für Sozialpsychologie*, 15, 151-158.
- Brandtstädter, J. (1987). „A rose has no teeth“ – Zum Problem der Unterscheidung zwischen Begriffsverwirrungen und überraschenden empirischen Befunden in der Psychologie. In J. Brandtstädter (Hrsg.), *Struktur und Erfahrung in der psychologischen Forschung* (S. 1-13). Berlin: Walter de Gruyter.
- Braun, K. & Bogerts, B. (2001). Erfahrungsgesteuerte neuronale Plastizität. Bedeutung für Pathogenese und Therapie psychischer Erkrankungen. *Nervenarzt*, 72, 3-10.
- Brauns, H.-P. (2012). Ist die rationale Psychologie eine theoretische Psychologie? In A. Stock, H.-P. Brauns & U. Wolfradt (Hrsg.), *Historische Analysen theoretischer und empirischer Psychologie* (S. 75-89). Frankfurt am Main: Peter Lang.
- Brentano, F. (1874/1971). *Psychologie vom empirischen Standpunkt*. Hamburg: Meiner.
- Bressloff P. C., Cowan, J. D., Golubitsky, M., Thomas, P. J. & Wiener, M. C. (2001). Geometric visual hallucinations, Euclidean symmetry and the functional architecture of striate cortex. *Philosophical Transactions of the Royal Society London*, 356, 299-330.
- Breuer, F. (1977). *Einführung in die Wissenschaftstheorie für Psychologen*. Aschendorff/Münster: Aschendorffsche Buchdruckerei.

- Breyer T. (2012). Perspektivität als Thema der Phänomenologie und Psychopathologie. *Journal für Philosophie und Psychiatrie*, 5. (www.jfpp.org/109.html).
- Brigandt, I. & Love, A. C. (2012). Reductionism in biology. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Summer 2012 edition)*; <http://plato.stanford.edu/entries/reduction-biology>
- Broad, C. D. (1925). *The Mind and Its Place in Nature*. London: Routledge & Kegan Paul.
- Buzsáki, G. (2006). *Rhythms of the brain*. Oxford: Oxford University Press.
- Campbell, N. A. & Reece, J. B. (2011) (Hrsg.). *Biologie*. München: Pearson Studium.
- Campbell, R. J. & Bickhard, M. H. (2011). Physicalism, Emergence and Downward Causation. *Axiomathes*, 21, 33-56.
- Carey, S. (1985). *Conceptual Change in Childhood*. Cambridge: MIT Press.
- Carey, S. & Bartlett, E. J. (1978). Acquiring a single new word. *Papers and Reports on Childs Language Development*, 15, 17-29.
- Carnap, R. (1932). Die physikalische Sprache als Universalsprache der Wissenschaft. *Erkenntnis*, 2, 432-465.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Carrier, M. & Mittelstraß, J. (1989). *Geist, Gehirn, Verhalten. Das Leib-Seele-Problem und die Philosophie der Psychologie*. Berlin: De Gruyter.
- Cattell, R. B. (1987). *Intelligence: Its structure, growth and action*. Amsterdam: North-Holland.
- Causey, R. L. (1972). Attributive Identities in microreductions. *Journal of Philosophy*, 79, 407-422.
- Chalmers, D. (1995). Facing Up to the Problem of Consciousness. *Journal of Consciousness Studies*, 2, 200-219.

- Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.
- Chomsky, N. (1975). *Reflections on Language*. New York: Pantheon Books.
- Chomsky, N. (1981). On the representation of form and function. *Linguistic Review*, 1, 3-40.
- Churchland, P. M. (1979). *Scientific Realism and the Plasticity of Mind*. New York: Cambridge University Press.
- Churchland, P. M. (1982). Is "thinker" a natural kind? *Dialogue*, 21, 223-238.
- Churchland, P. M. (1989). *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. Cambridge: MIT Press.
- Churchland, P. S. (1986). *Neurophilosophy: Toward a unified science of the mind-brain*. Cambridge: MIT Press.
- Churchland, P. S. (1993). Can neurobiology teach us anything about consciousness? *Proceedings and Addresses of the American Philosophical Association*, 67, 23-40.
- Churchland, P. S. & Sejnowski, T. J. (1988). Perspectives on Cognitive Neuroscience. *Science*, 242, 741-745.
- Churchland, P. S. & Sejnowski, T. J. (1994/1997). *Grundlagen zur Neuroinformatik und Neurobiologie*. Braunschweig: Vieweg.
- Clapp, L. (2001). Disjunctive properties and multiple realizations. *Journal of Philosophy*, 98, 111–136.
- Clark, A. (1989/2007). Mikrofunktionalismus: Konnektionismus und die wissenschaftliche Erklärung mentaler Zustände. In T. Metzinger (Hrsg.), *Grundkurs Philosophie des Geistes. Band 2: Das Leib-Seele-Problem* (S. 392-426). Paderborn: Mentis.
- Cole, D. J. (1990). Cognitive Inquiry and the Philosophy of Mind. In D. J. Cole, J. H. Fetzer & T. L. Rankin (eds.), *Philosophy, Mind, and Cognitive Inquiry* (pp. 1-46). Dordrecht: Kluwer.

- Craver, C. F. (2001). Role functions, mechanisms, and hierarchy. *Philosophy of Science*, 68, 53-74.
- Craver, C. F. (2002). Interlevel experiments and multilevel mechanisms in the neuroscience of memory. *Philosophy of Science*, 69, 83-97.
- Craver, C. F. (2007). *Explaining the Brain. Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Clarendon Press.
- Craver, C. F. (2007a). Constitutive Explanatory Relevance. *Journal of Philosophical Research*, 32, 3-20.
- Craver, C. F. & Bechtel, W. (2007). Top-down causation without top-down causes. *Biology and Philosophy*, 22, 547-563.
- Crick, F. (1994/1995). *The astonishing hypothesis*. New York: Touchstone.
- Cronbach, L. J. & Meehl, P. E. (1955). Construct validity and in psychological tests. *Psychological Bulletin*, 52, 281-302.
- Cruse, H. (2004). Ich bin mein Gehirn. Nichts spricht gegen den materialistischen Monismus. In C. Geyer (Hrsg.), *Hirnforschung und Willensfreiheit. Zur Deutung der neuesten Experimente* (S. 223-228). Frankfurt am Main: Suhrkamp.
- Damasio, A. R. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace & Company.
- Damasio, A. R. (2000). A neurobiology for consciousness. In T. Metzinger (ed.), *Neural Correlates of Consciousness: Empirical and Conceptual Questions* (pp. 111-120). Cambridge: MIT Press.
- Davidson, D. (1970/2007). Mentale Ereignisse. In T. Metzinger (Hrsg.), *Grundkurs Philosophie des Geistes, Band 2: Das Leib-Seele-Problem* (S. 222-242). Paderborn: Mentis.
- Davis, M. (1958). *Computability and Unsolvability*. New York: McGraw-Hill.
- Dehaene, S. (2010). *Lesen. Die größte Erfindung der Menschheit und was dabei in unseren Köpfen passiert*. München: Knaus.

- Dehmelt, H. (1989). Triton, Electron, Cosman...: An Infinite Regression? *Proceedings of the National Academy of Sciences*, 86, 8618-8619.
- Dennett, D. C. (1991/1994). *Philosophie des menschlichen Bewusstseins*. Hamburg: Hoffmann und Campe.
- Descartes, R. (1641/1986). *Meditationes de Prima Philosophia. Meditationen über die Erste Philosophie*. Stuttgart: Reclam.
- Deutscher, M. (1967). Mental and Physical Properties. In C. F. Presley (ed.), *The Identity Theory of Mind* (pp. 65-83). Brisbane: University of Queensland Press.
- DeWitt, B. S. (1970). Quantum mechanics and reality. *Physics Today*, 9, 30.
- Dierks, T., Linden, D. E. J., Jandl, M., Formisano, E., Goebel, R., Lanfermann, H. & Singer, W. (1999). Activation of Heschls gyrus during auditory hallucinations. *Neuron*, 22, 615-621.
- Dilling, H. Mombour, W. & Schmidt, M. H. (2013). *Internationale Klassifikation psychischer Störungen: ICD-10 Kapitel V (F). Klinisch-diagnostische Leitlinien*. Bern: Huber.
- Dimitrova, V. & Lüdmann, M. (2011). Die Entwicklung sozialer Basisprozesse und sozialer Kompetenzen. In M. Limbourg & G. Steins (Hrsg.), *Sozialerziehung in der Schule* (S. 115-130). Wiesbaden: Verlag für Sozialwissenschaften.
- Dimitrova, V. & Lüdmann, M. (2014). *Sozial-emotionale Kompetenzentwicklung: Leitlinien der Entfaltung der emotionalen Welt*. Wiesbaden: Springer VS.
- Dockrell, J. & Campbell, R. (1986). Lexical acquisition strategies in the pre-school child. In S. A. Kuczaj & M. D. Barrett (eds.), *The development of word meaning* (pp. 121-154). New York: Springer.

- Dörner, D. (1974). *Die kognitive Organisation beim Problemlösen: Versuch einer kybernetischen Theorie der elementaren Informationsverarbeitung beim Denken*. Bern: Huber.
- Dörner, D. (1979). Die Psychologie und das Problem der Einheit der Wissenschaften. In B. Kanitscheider (Hrsg.), *Materie-Leben-Geist. Zum Problem der Reduktion der Wissenschaften* (S. 121-128). Berlin: Duncker & Humblot.
- Dörner, D. (1983). Empirische Psychologie und Alltagsrelevanz. In G. Jüttemann (Hrsg.), *Psychologie in der Veränderung: Perspektiven für eine gegenstandsangemessenere Forschungspraxis* (S. 13-29). Weinheim: Beltz.
- Dörner, D. (1998/2008). *Bauplan für eine Seele*. Reinbek: Rowohlt.
- Dörner, D. (2004). Man muss wissen, wonach man sucht. *Gehirn und Geist*, 7/2004, 36 – 38.
- Dörner, D. (2008). Über die Notwendigkeit einer Theoretischen Psychologie. In B. Krause (Hrsg.), *Entwicklungen in der Experimentellen Psychologie an der Humboldt-Universität zu Berlin* (S. 13-32). Berlin: ZeE-Verlag.
- Dörner, D., Lutz, W. & Meurer, K. (1967). Informationsverarbeitung beim Konzepterwerb. *Zeitschrift für Psychologie*, 174, 194-230.
- Du Bois-Reymond, E. (1880/1974). *Vorträge über Philosophie und Gesellschaft*. Hamburg: Meiner.
- Edelman, G. M. (1989). *The remembered present*. New York: Basic Books.
- Ehlert, U. & von Känel, R. (2011) (Hrsg.). *Psychoendokrinologie und Psychoimmunologie*. Heidelberg: Springer.
- Eimas, P. D. & Quinn, P. C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development*, 65, 903-917.
- Engel, S., Zhang, X. & Wandell, B. A. (1997). Color tuning in human visual cortex measured using functional magnetic resonance imaging. *Nature*, 388, 68-71.

- Ervine, S. J. (2011). Constitution and Composition: Three Approaches to their Relation. *ProtoSociology*, 27, 212-235.
- Ekman, P., Levenson, R. W. & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, 221, 1208-1210.
- Eronen, M. I. (2010). Replacing Functional Reduction with Mechanistic Explanation. *Philosophia naturalis*, 47/48, 125-153.
- Esfeld, M. & Sachse, C. (2007). Theory reduction by means of functional sub-types. *International Studies in the Philosophy of Science*, 21, 1-17.
- Everett, H. (1957). "Relative State" Formulation of Quantum Mechanics. *Reviews of Modern Physics*, 29, 454-462.
- Fahrenberg, J. (2011). *Wilhelm Wundt – Pionier der Psychologie und Außen-seiter? Leitgedanken der Wissenschaftskonzeption und deren Rezeptionsgeschichte*. (e-book; www.jochen-fahrenberg.de)
- Fahrenberg, J. (2013). *Zur Kategorienlehre der Psychologie: Komplementaritätsprinzip, Perspektiven und Perspektiven-Wechsel*. Lengerich: Pabst.
- Farah, M. J., Wilson, K. D., Drain, H. M & Tanaka, J. R. (1998). What is 'Special' about Face Perception? *Psychological Review*, 105, 482-498.
- Fazekas, P. & Kertész, G. (2011). Causation at different levels: tracking the commitments of mechanistic explanations. *Biology and Philosophy*, 26, 365-383.
- Fechner, G. T. (1861). *Ueber die Seelenfrage. Ein Gang durch die sichtbare Welt, um die unsichtbare zu finden*. Leipzig: C. F. Amelang.
- Fechner, G. T. (1879). *Die Tagesansicht gegenüber der Nachtansicht*. Leipzig: Breitkopf & Härtel.
- Fechner, G. T. (1851/1922). *Zend-Avesta. Gedanken über die Dinge des Himmels und des Jenseits vom Standpunkte der Naturbetrachtung*. Leipzig: Insel.

- Fechner, G. T. (1860/1964). *Elemente der Psychophysik*, 2 Bde. Amsterdam: E. J. Bonset.
- Feigl, H. (1958/1967). *The "Mental" and the "Physical": The Essay and a Postscript*. Minneapolis: University of Minnesota Press.
- Feigl, H. (1960/1970). Mind-body, not a pseudo-problem. In C. V. Borst (ed.), *The Mind/Brain Identity Theory* (pp. 33-41). London: Macmillan.
- Felleman, D. J. & van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1, 1-47.
- Feyerabend, P. K. (1962). Explanation, Reduction, and Empiricism. In H. Feigl & G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science. Vol. III: Scientific Explanation, Space, and Time* (pp. 28-97). Minneapolis: University of Minnesota Press.
- Fiedler, K., Kliegl, R., Lindenberger, U., Mausfeld, R., Mummendey, A. & Prinz, W. (2008). Psychologie im 21. Jahrhundert – eine Standortbestimmung. *Gehirn und Geist*, 1/2008, 26-36.
- Fodor, F. A. (1974). Special Sciences (or the Disunity of Science as a Working Hypothesis). *Synthese*, 28, 97-115.
- Fodor, F. A. (1994). *The Elm and the Expert. Mentalese and Its Semantics*. Cambridge: MIT Press.
- Fodor, F. A. (1997). Special Sciences. Still autonomous after all these years. In J. Tomberlin (ed.), *Philosophical Perspectives 11: Mind, Causation, and World* (pp. 149-164). Boston: Blackwell.
- Frege, G. (1884). *Die Grundlagen der Arithmetik. Eine logisch mathematische Untersuchung über den Begriff der Zahl*. Breslau: Wilhelm Koebner.
- Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100, 25-50.
- Gardner, H. (1987). *The mind's new science. A history of cognitive revolution*. New York: Basic Books.

- Gaser, C., Nenadic, I., Volz, H. P., Buchel, C. & Sauer, H. (2004). Neuroanatomy of „hearing voices“: A frontotemporal brain structural abnormality associated with auditory hallucinations in schizophrenia. *Cerebral Cortex*, 14, 91-96.
- Gegenfurtner, K. R. (2003). Farbwahrnehmung und ihre Störungen. In H.-O. Karnath & P. Thier (Hrsg.), *Neuropsychologie* (S. 11-20). Berlin: Springer.
- Gelman, S. & Coley, D. (1990). The importance of knowing a dodo as a bird: Categories and inferences in 2-year-old children. *Developmental Psychology*, 26, 796-804.
- Gelman, S. & Markman, E. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development*, 58, 1532-1541.
- Gillett, C. (2002). The dimensions of realization: a critique of the Standard view. *Analysis*, 62, 316-323.
- Gillett, C. (2003). The Metaphysics of Realization, Multiple Realizability, and the Special Sciences. *The Journal of Philosophy*, 100, 591-603.
- Gillett, C. (2003a). Non-Reductive Realization and Non-Reductive Identity. In S. Walter & H.-D. Heckmann (eds.), *Physicalism and Mental Causation. The Metaphysics of Mind and Action* (pp. 31-58). Exeter: Imprint.
- Gillett, C. (2007). Understanding the New Reductionism: The Metaphysics of Science and Compositional Reduction. *The Journal of Philosophy*, 104, 193-216.
- Gillett, C. (2010). Moving beyond the subset model of realization: The problem of qualitative distinctness in the metaphysics of science. *Synthese*, 177, 165-192.
- Gillett, C. (2011). Multiply realizing scientific properties and their instances. *Philosophical Psychology*, 24, 727-738.

- Gold, I. & Stoljar, D. (1999). A neuron doctrine in the philosophy of neuroscience. *Behavioral and Brain Sciences*, 22, 585-642.
- Goldstein, E. B. (2008). *Wahrnehmungspsychologie*. Heidelberg: Springer.
- Gozzano, S. & Hill, C. S. (eds.) (2012). *New Perspectives on Type Identity. The Mental and the Physical*. Cambridge: Cambridge University Press.
- Hardin, C. L. (1987). Qualia and Materialism. Closing the Explanatory Gap. *Philosophy and Phenomenological Research*, 47, 281-298.
- Hardin, C. L. (1997). Reinverting the Spectrum. In A. Byrne & D. R. Hilbert (eds.), *Readings on Color, Bd. 1* (pp. 289-302). Cambridge: MIT Press.
- Hartmann, N. (1943). *Die Anfänge des Schichtungsgedankens in der Alten Philosophie*. Berlin: Verlag der Akademie der Wissenschaften/De Gruyter.
- Hartmann, N. (1949). *Der Aufbau der realen Welt*. Berlin: De Gruyter.
- Harris, Z. S. (1951). *Methods in Structural Linguistics*. Chicago: Chicago University Press.
- Heckmann, H.-D. & Esken, F. (1998). Generelle Einführung: Bewußtsein und Repräsentation. In F. Esken & H.-D. Heckmann (Hrsg.), *Bewußtsein und Repräsentation* (S. 11-49). Paderborn: Ferdinand Schöningh.
- Heckmann, H.-D. & Walter, S. (2001/²2006). Einleitung. In H.-D. Heckmann & S. Walter (Hrsg.), *Qualia: Ausgewählte Beiträge* (S. 15-21). Paderborn: Mentis.
- Heeger, D. J., Simoncelli, E. P. & Movshon, J. A. (1996). Computational models of cortical visual processing. *Proceedings of the National Academy of Sciences*, 93, 623-627.
- Heidelberger, M. (2000). Fechner und Mach zum Leib-Seele-Problem. In A. Arndt & W. Jaeschke (Hrsg.), *Materialismus und Spiritualismus: Philosophie und Wissenschaften nach 1848* (S. 53-67). Hamburg: Meiner.

- Heidelberger, M. (2002). Wie das Leib-Seele Problem in den Logischen Empirismus kam. In M. Pauen & A. Stephan (Hrsg.), *Phänomenales Bewusstsein – Rückkehr zur Identitätstheorie?* (S. 40–72). Paderborn: Mentis.
- Heider, F. & Simmel, M. (1944). An experimental study of apparent behaviour. *American Journal of Psychology*, 57, 243-259.
- Heim, C. & Meinlschmidt, G. (2003). Biologische Grundlagen. In U. Ehlert (Hrsg.), *Verhaltensmedizin* (S. 17-93). Berlin: Springer.
- Hempel (1969). Reduction: Ontological and linguistic facets. In S. Morgenbesser, P. Suppes & M. White (eds.), *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel* (pp. 179-199). New York: St. Martin's Press.
- Hering, E. (1878). *Zur Lehre vom Lichtsinn*. Vienna: Gerold.
- Herrmann, U. (2009). Gehirnforschung und die neurodidaktische Revision schulisch organisierten Lehrens und Lernens. In U. Herrmann (Hrsg.), *Neurodidaktik. Grundlagen und Vorschläge für ein gehirngerechtes Lehren und Lernen* (S. 148-181). Weinheim: Beltz.
- Hill, C. S. (1997/2002). Vorstellbarkeit, Denkbarkeit, Möglichkeit und das Leib-Seele-Problem. In M. Pauen & A. Stephan (Hrsg.), *Phänomenales Bewusstsein – Rückkehr zur Identitätstheorie?* (S. 184-207). Paderborn: Mentis.
- Hirsch, E. (1982). *The concept of Identity*. New York: Oxford University Press.
- Hooker, C. (1981). Towards a General Theory of Reduction, *Dialogue*, 20, 38-59, 201-36, 496-529.
- Hösle, V. (2006). Encephalius. Ein Gespräch über das Leib-Seele-Problem. In F. Hermanni & T. Buchheim (Hrsg.), *Das Leib-Seele-Problem. Antwortversuche aus medizinisch-naturwissenschaftlicher, philosophischer und theologischer Sicht*. München: Wilhelm Fink Verlag.

- Hubel, D. H. & Wiesel, T. N. (1962). Receptive Fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160, 106-154.
- Husserl, E. (1976) *Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie*. Den Haag: Nijhoff.
- Hüttemann, A. (2004). *What's Wrong With Microphysicalism?* London: Routledge.
- Hüttemann, A. (2009). Physikalische Realisierung in der Physik. In M. Backmann & J. Michel (Hrsg.), *Physikalismus, Willensfreiheit, Künstliche Intelligenz* (S. 67-73). Paderborn: Mentis.
- Huxley, T. H. (1866). *Lessons in Elementary Physiology*. London: The Macmillan Co.
- Inagaki, K. (1990). The effects of raising animals on children's biological knowledge. *British Journal of Developmental Psychology*, 8, 119-129.
- Jackson, F. (1982). Epiphenomenal Qualia. *Philosophical Quarterly*, 32, 127-136.
- Jackson, F. & Petit, P. (1990). Causation and the Philosophy of Mind. *Philosophy and Phenomenological Research*, 50, 195-214.
- James, W. (1890/1981). *The Principles of Psychology*. Cambridge: Harvard University Press.
- Jäncke, L. (2010). Hirnforschung: sinnvolle Ergänzung oder überflüssiges Anhängsel der Psychologie? *Psychologische Rundschau*, 61, 191-198.
- Jäncke, L. & Petermann, F. (2010). Wie viel Biologie braucht die Psychologie. *Psychologische Rundschau*, 61, 173-174.
- Jänig, W. (2000). Vegetatives Nervensystem. In R. F. Schmidt & H.-G. Schaible (Hrsg.), *Neuro- und Sinnesphysiologie* (S. 151-202). Berlin: Springer.

- Johnson-Laird, P. N. (1983). *Mental models. Towards a cognitive science of language, inference and consciousness*. Cambridge: Cambridge University Press.
- Jones, S. & Smith, L. (1993). The place of perception in children's concepts. *Cognitive Development*, 8, 113-139.
- Kandel, E. R. (2000). Disorders of Thought and Volition: Schizophrenia. In E. R. Kandel, J. H. Schwartz, & T. M. Jessell (eds.), *Principles of Neural Science* (pp. 1188-1207). New York: McGraw-Hill.
- Kandel, E. R. & Kupfermann, I. (1996). Kognitive Neurowissenschaft: Von den Nervenzellen zur Kognition. In E. R. Kandel, J. H. Schwartz & T. M. Jessell (Hrsg.), *Neurowissenschaften* (S. 327-352). Heidelberg: Spektrum.
- Kandel, E. R. & Mason, C. (1996). Wahrnehmung von Form und Bewegung. In E. R. Kandel, J. H. Schwartz & T. M. Jessell (Hrsg.), *Neurowissenschaften* (S. 431-457). Heidelberg: Spektrum.
- Kandel, E. R., Schwartz, J. & Jessell, T. M. (1996) (Hrsg.). *Neurowissenschaften*. Heidelberg: Spektrum.
- Kant, I. (1787/1966). *Kritik der reinen Vernunft* (Auflage B). Stuttgart: Reclam.
- Kanwisher, N., McDermott, J. & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17, 4302-4311.
- Kaufmann, L. (2007) (Hrsg.). *Kognitive Entwicklungsneuropsychologie*. Göttingen: Hogrefe.
- Keil, F. (1989). *Concepts, Kinds, and Cognitive Development*. Cambridge: MIT Press.
- Keil, F. (1991). The emergence of theoretical beliefs as constraints on concepts. In S. Carey & R. Gelman (eds.), *The Epigenesis of Mind: Essays on Biology and Cognition* (pp. 237-321). Hillsdale: Erlbaum.

- Keil, F. (1992). The origins of an autonomous biology. In M. Gunnar & M. Maratsos (eds.), *Minnesota Symposium on Child Psychology* (pp. 103-137). Hillsdale: Erlbaum.
- Kerkhoff, G. (2003). Störungen der visuellen Raumorientierung. In H.-O. Karnath & P. Thier (Hrsg.), *Neuropsychologie* (S. 175-183). Berlin: Springer.
- Kim, J. (1972/1980). Physicalism and the Multiple Realizability of Mental States. In N. Block (ed.), *Readings in Philosophy of Psychology* (pp. 234-236). Cambridge: Harvard University Press.
- Kim, J. (1976). Events as Property Exemplifications. In M. Brand & D. Walton (eds.), *Action Theory* (pp. 159-177). Dordrecht: Reidel.
- Kim, J. (1989/2007). Der Mythos des nichtreduktiven Materialismus. In T. Metzinger (Hrsg.), *Grundkurs Philosophie des Geistes, Band 2: Das Leib-Seele-Problem* (S. 248-272). Paderborn: Mentis.
- Kim, J. (1993). *Supervenience and Mind*. Cambridge: Cambridge University Press.
- Kim, J. (1997). Supervenience, Emergence, and Realization in the Philosophy of Mind. In M. Carrier & P. K. Machamer (eds.), *Mindscapes: Philosophy, Science, and the Mind*. Konstanz: Universitäts-verlag.
- Kim, J. (1996/1998). *Philosophie des Geistes*. Wien: Springer.
- Kim, J. (1998). *Mind in a physical world*. Cambridge: MIT Press.
- Kim, J. (2002). The Layered Model: Metaphysical Considerations. *Philosophical explorations*, 5, 2-20.
- Kim, J. (2003). Blocking Causal Drainage and Other Maintenance Chores with Mental Causation. *Philosophy and Phenomenological Research*, 67, 151-176.
- Kim, J. (2005). *Physicalism, or something near enough*. Princeton: Princeton University Press.
- Kim, J. (2006). Emergence: Core ideas and issues. *Synthese*, 151, 547-559.

- Klee, R. (1984). Micro-determinism and the Concept of Emergence. *Philosophy of Science*, 51, 44-63.
- Klein, C. (2008). An Ideal Solution to Disputes about Multiply Realized Kinds. *Philosophical Studies*, 140, 161-177.
- Knecht, S. & Ringelstein, E.-B. (1999). Neuronale Plastizität am Beispiel des somatosensorischen Systems. *Nervenarzt*, 70, 889-898.
- Koch, C. (2005). *Bewusstsein: ein neurobiologisches Rätsel*. Heidelberg: Spektrum.
- Koch, S. (1951). Theoretical Psychology. *Psychological Review*, 58, 295-301.
- Kohlberg, L. (1976). Moral stages and moralization. The cognitive developmental approach. In T. Lickona (ed.), *Moral development and behavior* (pp. 31-53). New York: Holt, Rinehart & Winston.
- Kripke, S. (1971). Identity and Necessity. In M. K. Munitz (ed.), *Identity and Individuation* (pp. 135-164). New York: New York University Press.
- Kripke, S. (1972). *Naming and Necessity*. Cambridge: Harvard University Press.
- Kuhn, T. S. (1962/1976). *Die Struktur wissenschaftlicher Revolutionen*. Suhrkamp: Frankfurt am Main.
- Kupfermann, I. (1996). Cortex und Kognition. In E. R. Kandel, J. H. Schwartz & T. M. Jessell (Hrsg.), *Neurowissenschaften* (S. 351-369). Heidelberg: Spektrum.
- Landau, L. D. & Lifschitz, J. M. (1989). *Lehrbuch der theoretischen Physik*. Berlin: Akademie-Verlag.
- Lang, P. J. (1993). The three-system approach to emotion. In N. Birbaumer & A. Öhman (eds.), *The structure of emotion* (pp. 18-30). Seattle: Hogrefe & Huber.
- Lange, F. A. (1866/1908). *Geschichte des Materialismus und Kritik seiner Bedeutung in der Gegenwart*, Band 2. Leipzig: J. Baedeker.

- Laucken, U. (1998). *Sozialpsychologie: Geschichte – Hauptströmungen – Tendenzen*. Oldenburg: BIS-Verlag der Universität Oldenburg.
- Laucken, U. (2002). Über die semantische Blindheit einer neurowissenschaftlich gewendeten Psychologie. Oder: Was hätte uns eine so gewendete Psychologie zum "Dialog der Kulturen" zu sagen? *Berichte aus dem Institut zur Erforschung von Mensch-Umwelt-Beziehungen*, Nr. 39. Oldenburg: Universität Oldenburg.
- Laucken, U. (2003) *Theoretische Psychologie. Denkformen und Sozialpraxen*. Oldenburg: BIS-Verlag der Universität Oldenburg.
- Laucken, U. (2005). "Gibt es Willensfreiheit?" Möglichkeiten der psychologischen Vergegenständlichung von "Willens-, Entscheidungs- und Handlungsfreiheit". *Forum Qualitative Sozialforschung*, 6 (1).
- Lazarus, R. S. & Folkman, S. (1984). *Stress, Appraisal, and Coping*. New York: Springer Pub.
- Leibniz, G. W. (1704/1996). *Neue Abhandlungen über den menschlichen Verstand*. Hamburg: Meiner.
- Leibniz, G. W. (1710/1996). *Die Theodizee*. Frankfurt am Main: Suhrkamp.
- Leibniz, G. W. (1720/2002). *Monadologie und andere metaphysische Schriften*. Hamburg: Meiner.
- 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 M. Davies & G. W. Humphreys (eds.), *Consciousness. Psychological and Philosophical Essays* (S. 121-136). Oxford: Blackwell.
- Lewes, G. H. (1875). *Problems of Life and Mind*. London: Trübner.
- Lewin, K. (1931). Der Übergang von der aristotelischen zur galileischen Denkweise in Psychologie und Biologie. *Erkenntnis*, 1, 421-466.

- Lewis, D. K. (1966/2007). Eine Argumentation für die Identitätstheorie. In T. Metzinger (Hrsg.), *Grundkurs Philosophie des Geistes, Band 2: Das Leib-Seele-Problem* (S. 169-180). Paderborn: Mentis.
- Lewis, D. K. (1980). Mad pain and Martian pain. In N. Block (ed.), *Readings in the Philosophy of Psychology* (pp. 216-222). Cambridge: Harvard University Press.
- Lewis, D. K. (1991). *Parts of Classes*. Oxford: Blackwell.
- Libet, B. (1985). Unconscious cerebral initiative and the role of unconscious will on voluntary action. *Behavioral and Brain Sciences*, 8, 529-567.
- Libet, B. (2007). *Mind Time. Wie das Gehirn Bewusstsein produziert*. Frankfurt am Main: Suhrkamp.
- Lindworsky, J. (1926). *Theoretische Psychologie im Umriss*. Leipzig: J. A. Barth.
- Loar, B. (1997/2007). Phänomenale Zustände (Zweite Version). In M. Pauen, M. Schütte & A. Staudacher (Hrsg.), *Begriff, Erklärung, Bewusstsein: Neue Beiträge zum Qualia-Problem* (S. 27-60). Paderborn: Mentis.
- Locke, J. (1690/2000). *Versuch über den menschlichen Verstand*. Hamburg: Meiner.
- Lockwood, M. (1993). The Grain Problem. In H. Robinson (ed.), *Objections to Physicalism* (pp. 271-292). Oxford: Oxford University Press.
- Lüdmann, M. (2009). *Neuronen und Halluzinationen. Schizophrenie im Angesicht des psychophysischen Problems*. Berlin: Logos.
- Lüdmann, M. (2009a). Schizophrenie im Angesicht des Leib-Seele-Problems: Eine Erörterung und Bewertung der bestehenden Lösungsansätze anhand auditiver Halluzinationen. *Journal für Philosophie und Psychiatrie*, 2, 1/2009.

- Lüdmann, M. (2010). Nothing but Neurons? Examining the Ontological Dimension of Schizophrenia in the Case of Auditory Hallucinations. *Journal of Mind and Behavior*, 31, 49-64.
- Lüdmann, M. (2012). Zur Aktualität Fechners bzw. Wundts in der gegenwärtigen Debatte um den Status des Psychischen. In A. Stock, H.-P. Brauns & U. Wolfradt (Hrsg.), *Historische Analysen theoretischer und empirischer Psychologie* (S. 75-89). Frankfurt am Main: Peter Lang.
- Lüdmann, M. (2014). Der Mythos der verspäteten Wissenschaft. Oder warum wir eine Theoretische Psychologie brauchen. In W. Mack, H. E. Lück & U. Wolfradt (Hrsg.), *Behaviorismus und Erkenntnistheorie im psychologisch-historischen Kontext* (S. 85-104). Frankfurt am Main: Peter Lang.
- Lüdmann, M. (2014a). Repräsentation und Architektur des Psychischen. In E. Aschermann & M. Kaiser-El-Safti (Hrsg.), *Gestalt und Gestaltung in interdisziplinärer Perspektive* (S. 87-108). Frankfurt am Main: Peter Lang.
- Lüdmann, M. (unpubliziert). *Frühe Theorien der Identität von Körper und Geist. Ihre Entwicklungsgeschichte und Widerspiegelung in der gegenwärtigen Philosophie des Geistes* (Magisterarbeit). Universität Trier.
- Lycan, W. (1987/2007). Die Kontinuität natürlicher Ordnungsebenen. In T. Metzinger (Hrsg.), *Grundkurs Philosophie des Geistes, Band 2: Das Leib-Seele-Problem* (S. 432-459). Paderborn: Mentis.
- Mach, E. (1886/⁴1903). *Die Analyse der Empfindungen und das Verhältniss des Physischen zum Psychischen*. Jena: Gustav Fischer.
- Mach, E. (1905/⁵1968). *Erkenntnis und Irrtum*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Machamer, P., Darden, L. & Craver, C. F. (2000). Thinking about Mechanisms. *Philosophy of Science*, 67, 1-25.

- Mack, W. (2012). Seele / Psychê – problemgeschichtliche Argumentationsskizze für die Wiederaneignung eines grundlegenden Begriffes der Psychologie. In A. Stock, H.-P. Brauns & U. Wolfradt (Hrsg.), *Historische Analysen theoretischer und empirischer Psychologie* (S. 57-74). Frankfurt am Main: Peter Lang.
- Madole, K. L. & Cohen, L. B. (1995). The role of object parts in infants' attention to form-function correlations. *Developmental Psychology*, 31, 637-648.
- Marr, D. (1982). *Vision*. New York: Freeman and Company.
- Mausfeld, R. (2003). No Psychology In – No Psychology Out. *Psychologische Rundschau*, 54, 185-195.
- Mausfeld, R. (2006). ...in jedem Moment ein Schauplatz geistiger Möglichkeiten. Multiperspektivität als funktionales Designprinzip des menschlichen Geistes. *Vortragsmanuskript der Ringvorlesung „Das neue Bild vom Menschen“ im Rahmen des Forschungsverbundes „Interdisziplinäre Anthropologie“, Friedrich-Schiller-Universität Jena (6.12.2006)*.
- Mausfeld, R. (2007). Über Ziele und Grenzen einer naturwissenschaftlichen Zugangsweise zur Erforschung des Geistes. In A. Holderegger, B. Sitter-Liver, C. W. Hess & G. Rager (Hrsg.), *Hirnforschung und Menschenbild* (S. 21-40). Fribourg: Academic Press.
- Mausfeld, R. (2010). Psychologie, Biologie, kognitive Neurowissenschaften. Zur gegenwärtigen Dominanz neuroreduktionistischer Positionen und ihren stillschweigenden Grundannahmen. *Psychologische Rundschau*, 61, 180-190.
- Mausfeld, R. (2010a). Intrinsic Multiperspectivity: On the Architectural Foundations of a Distinctive Mental Capacity. In P. A. Frensch & R. Schwarzer (eds.), *Cognition and neuropsychology: International perspectives on psychological science* (pp. 95-116). Hove: Psychology Press.

- McCauley, R. N. (1981). Hypothetical Identities and Ontological Economizing: Comments on Causey's Program for the Unity of Science. *Philosophy of Science*, 48, 218-227.
- McCauley, R. N. (1986). Intertheoretic Relations and the Future of Psychology. In: S. M. Christensen & D. R. Turner (eds.), *Folk Psychology and the Philosophy of Mind* (pp. 63-81). Hillsdale: Lawrence Erlbaum.
- McCauley, R. N. (2007). Reduction: Models of Cross-Scientific Relations and Their Implications for Psychology-Neuroscience Interface. In P. Thagard (ed.), *Handbook of the Philosophy of Science: Philosophy of Psychology and Cognitive Science* (pp. 105-158). Amsterdam: Elsevier.
- McCauley, R. N. (2007a). Enriching Philosophical Models of Cross-Scientific Relations: Incorporating Diachronic Theories. In M. Schouten & H. Looren de Jong (eds.), *The Matter of the Mind. Philosophical Essays on Psychology, Neuroscience, and Reduction* (pp. 199-223). Malden: Blackwell.
- McCauley, R. N. (2012). About Face. Philosophical Naturalism, the Heuristic Identity Theory, and Recent Findings about Prosopagnosia. In S. Gozzano & C. S. Hill (eds.), *New Perspectives on Type Identity. The Mental and the Physical* (pp. 186-206). Cambridge: Cambridge University Press.
- McCauley, R. N. & Bechtel, W. (2001). Explanatory Pluralism and Heuristic Identity Theory. *Theory & Psychology*, 11, 736-760.
- McClamrock, R. (1991). Marr's Three Levels: A Re-Evaluation. *Minds and Machines*, 1, 185-196.
- McCulloch, W. S. & Pitts, W. H. (1943). Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics*, 5, 115-133.

- McCulloch, W. S. & Pitts, W. H. (1947). On how we know universals: the perception of visual and auditory. *Bulletin of Mathematical Biophysics*, 9, 127–147.
- McCulloch, W. S. (1965/2000). *Verkörperungen des Geistes*. Wien: Springer.
- McGinn, C. (1989). Can We Solve the Mind-Body Problem? In R. Warner & T. Szubka (eds.), *The Mind-Body Problem. A Guide to the Current Debate* (pp. 99-208). Oxford: Blackwell.
- McLaughlin, B. (1992). The Rise and Fall of British Emergentism. In A. Beckermann, H. Flohr & J. Kim (eds.), *Emergence or Reduction? Essays on Prospects of Nonreductive Physicalism* (pp. 49-93). Berlin: Walter de Gruyter.
- Melnyk, A. (2007). Functionalism and Psychological Reductionism: Friends, not Foes. In M. Schouten & H. Looren de Jong (eds.), *The Matter of the Mind. Philosophical Essays on Psychology, Neuroscience, and Reduction* (pp. 31-50). Malden: Blackwell.
- Metzinger, T. (1993). Die gegenwärtige Situation in der Philosophie des Geistes. *Information Philosophie*, 21, 14-24.
- Metzinger, T. (1993a). *Subjekt und Selbstmodell: Die Perspektivität phänomenalen Bewusstseins vor dem Hintergrund einer naturalistischen Theorie mentaler Repräsentation*. Paderborn: Schöningh.
- Metzinger, T. (1995/⁵2005). Einleitung: Das Problem des Bewusstseins. In T. Metzinger (Hrsg.), *Bewußtsein. Beiträge aus der Gegenwartsphilosophie* (S. 15-53). Paderborn: Mentis.
- Metzinger, T. (2000) (ed.). *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. Cambridge, MA: MIT Press.
- Metzinger, T. (2000a). The subjectivity of subjective experience: A representationalist analysis of the first-person perspective. In T. Metzinger (ed.), *Neural Correlates of Consciousness: Empirical and Conceptual Questions* (pp. 285-306). Cambridge, MA: MIT Press.

- Metzinger, T. (2003) *Being No One. The Self-Model Theory of Subjectivity*. Cambridge: MIT Press.
- Metzinger, T. (2003a). Phänomenale Transparenz und kognitive Selbstbezugnahme. In U. Haas-Spohn (Hrsg.), *Intentionalität zwischen Subjektivität und Weltbezug* (S. 411-459). Paderborn: Mentis.
- Metzinger, T. (2006). Being No One – Eine sehr kurze deutsche Zusammenfassung. T. Metzinger (Hrsg.), *Grundkurs Philosophie des Geistes, Band 1: Phänomenales Bewusstsein* (S. 424-475). Paderborn: Mentis.
- Metzinger, T. (2007). Das Leib-Seele-Problem in der zweiten Hälfte des zwanzigsten Jahrhunderts. In T. Metzinger (Hrsg.), *Grundkurs Philosophie des Geistes, Band 2: Das Leib-Seele-Problem* (S. 11-31). Paderborn: Mentis.
- Metzinger, T. (2007a). Identitätstheorie II: Die ontologisch neutrale Analyse mentalistischer Terme. In T. Metzinger (Hrsg.), *Grundkurs Philosophie des Geistes, Band 2: Das Leib-Seele-Problem* (S. 107-109). Paderborn: Mentis.
- Metzinger, T. (2007b). Identitätstheorie III: Neutraler Monismus. In T. Metzinger (Hrsg.), *Grundkurs Philosophie des Geistes, Band 2: Das Leib-Seele-Problem* (S. 129-131). Paderborn: Mentis.
- Metzinger, T. (2007c). Identitätstheorie IV: Kausale Rollen und Eigenschafts-Physikalismus. In T. Metzinger (Hrsg.), *Grundkurs Philosophie des Geistes, Band 2: Das Leib-Seele-Problem* (S. 167-168). Paderborn: Mentis.
- Metzinger, T. (2007d). Eigenschaftsdualismus I: Emergenz. In T. Metzinger (Hrsg.), *Grundkurs Philosophie des Geistes, Band 2: Das Leib-Seele-Problem* (S. 275-276). Paderborn: Mentis.
- Moll, H. (2007). Person und Perspektivität – Kooperation und soziale Kognition beim Menschen. In F. Kannetzky & H. Tegtmeyer (Hrsg.), *Leipziger Schriften zur Philosophie. Personalität – Studien zu einem*

- Schlüsselbegriff der Philosophie* (S. 37-56). Leipzig: Leipziger Universitätsverlag.
- Moore, G. E. (1903). The Refutation of Idealism. *Mind*, 12, 433-453.
- Moore, G. E. (1922). *Philosophical Studies*. London: Routledge and Paul.
- Morgan, C. L. (1923). *Emergent Evolution*. London: Williams and Norgate.
- Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt, Brace & World.
- Nagel, T. (1974/1997). Wie ist es, eine Fledermaus zu sein? In P. Bieri (Hrsg.), *Analytische Philosophie des Geistes* (S. 261-275). Weinheim: Beltz Athenäum.
- Neisser, U (1994). Multiple Systems: A new approach to cognitive theory. *European Journal of Cognitive Psychology*, 6, 225-241.
- Newell, A. & Simon, H. A. (1976). Computer science as empirical enquiring. In J. Haugeland (ed.), *Mind Design* (S. 35-66). Cambridge: MIT-Press.
- Nida-Rümelin, M. (2002). Erklärbare und nicht erklärbare Aspekte phänomenalen Erlebens. In M. Pauen & A. Stephan (Hrsg.), *Phänomenales Bewusstsein – Rückkehr zur Identitätstheorie?* (S. 330-341). Paderborn: Mentis.
- Nietzsche, F. (1887). *Zur Genealogie der Moral*. Leipzig: C. G. Naumann.
- Oppenheim, P. & Putnam, H. (1958). The Unity of Science as a Working Hypothesis. In H. Feigl, M. Scriven & G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science. Vol. II: Concepts, Theories, and the Mind-Body Problem* (pp. 3-36). Minneapolis: University of Minnesota Press.
- Papineau, D. (1995). Der antipathetische Fehlschluß und die Grenzen des Bewußtseins. In T. Metzinger (Hrsg.), *Bewußtsein. Beiträge aus der Gegenwartsphilosophie* (S. 305-319). Paderborn: Mentis.

- Papineau, D. (2002). Achtung Lücke! In M. Pauen & A. Stephan (Hrsg.), *Phänomenales Bewusstsein – Rückkehr zur Identitätstheorie?* (S. 222-242). Paderborn: Mentis.
- Parvizi, J. & Damasio, A. R. (2001). Consciousness and the brainstem. *Cognition*, 79, 135-159.
- Pauen, M. (1999). *Das Rätsel des Bewusstseins. Eine Erklärungsstrategie*. Paderborn: Mentis.
- Pauen, M. (2001/⁴2005). *Grundprobleme der Philosophie des Geistes: Eine Einführung*. Frankfurt: Fischer.
- Pauen, M. (2002). Invertierte Schmerzen. Funktionale Eigenschaften phänomenaler Zustände und das Erklärungslückenargument. In M. Pauen & A. Stephan (Hrsg.), *Phänomenales Bewusstsein – Rückkehr zur Identitätstheorie?* (S. 266-296). Paderborn: Mentis.
- Pauen, M. (2002a). Einführung. In M. Pauen & A. Stephan (Hrsg.), *Phänomenales Bewusstsein – Rückkehr zur Identitätstheorie?* (S. 178-183). Paderborn: Mentis.
- Pauen, M. (2004). Freiheit: Eine Minimalkonzeption. In F. Hermann & P. Koslowski (Hrsg.), *Der freie und der unfreie Wille – Philosophische und theologische Perspektiven* (S. 79-112). Paderborn: Fink.
- Pauen, M. (2007). *Was ist der Mensch? Die Entdeckung der Natur des Bewusstseins*. München: DVA.
- Pauen, M. (2012). The Second-Person Perspective. *Inquiry*, 55, 33-49.
- Pauen, M. & Stephan, A. (2002) (Hrsg.). *Phänomenales Bewusstsein – Rückkehr zur Identitätstheorie?* Paderborn: Mentis.
- Pauen, S. (1999). Developing ontological categories of matter: Stable dimensions and changing concepts. In W. Schnotz, S. Vosniadou, & M. Carretero (Eds.), *New perspectives on conceptual change* (pp. 15-31). Amsterdam: Elsevier.

- Paul, L. A. (2007). Constitutive overdetermination. In J. J. Campbell (ed.), *Causation and explanation* (pp. 265-290). Cambridge: MIT-Press.
- Pauli, P. & Birbaumer, N. (2000). Psychophysiologische Ansätze. In J. H. Otto, H. E. Euler & H. Mandl (Hrsg.), *Emotionspsychologie* (S. 75-84). Weinheim: Beltz.
- Pereboom, D. (2002). Robust nonreductive materialism. *Journal of Philosophy*, 99, 499-531.
- Pereboom, D. & Kornblith, H. (1991). The Metaphysics of Irreducibility. *Philosophical Studies*, 63, 125-145.
- Perner, J. (2000). RUM, PUM, and the perspectival relativity of sortals. In J. W. Astington (ed.), *Minds in the making: Essays in honour of David R. Olson* (pp. 212-232). Oxford: Blackwell.
- Perner, J., Brandl, J. L. & Garnham, A. (2003). What is a Perspective Problem? Developmental Issues in Belief Ascription and Dual Identity. *Facta Philosophica*, 5, 355-378.
- Perner, J., Stummer, S. Sprung, M. & Doherty, M. J. (2002). Theory of mind finds its Piagetian Perspective: Why alternative naming comes with understanding belief. *Cognitive Development*, 17, 1451-1472.
- Pettit, P. (1993). A Definition of Physicalism. *Analysis*, 53, 213-223.
- Pinel, J. P. J. (1997). *Biopsychologie*. Heidelberg: Spektrum.
- Place, U. T. (1956). Is consciousness a brain process? *British Journal of Psychology*, 47, 44-50.
- Place, U. T. (1960). Materialism as a Scientific Hypothesis. *Philosophical Review*, 69, 101-104.
- Platon (2004). Phaidon. In G. Eigler (Hrsg.). *Platon. Sämtliche Dialoge, Band 2: Menon – Kratylos – Phaidon – Phaidros* (S. 1-155). Hamburg: Meiner.
- Polger, T. W. (2004). *Natural Minds*. Cambridge: MIT Press.

- Polger, T. W. (2007). Some Metaphysical Anxieties of Reductionism. In M. Schouten & H. Looren de Jong (eds.), *The Matter of the Mind. Philosophical Essays on Psychology, Neuroscience, and Reduction* (pp. 51-75). Malden: Blackwell.
- Polger, T. W. (2009). Evaluating the Evidence for Multiple Realization. *Synthese*, 167, 457-472.
- Polger, T. & Flanagan, O. (2001). A decade of teleofunctionalism. *Mind and Machines*, 11, 113-126.
- Polger, T. W. & Shapiro, L. A. (2008). Understanding the Dimensions of Realization. *Journal of Philosophy*, 104, 213-222.
- Polger, T. W. & Sufka, K. (2006). Closing the Gap on Pain: Mechanism, Theory, and Fit. In M. Aydede (ed.), *New Essays on the Nature of Pain and the Methodology of its Study* (pp. 325-350). Cambridge: MIT Press.
- Popper, K. R. (1957). The aim of science. *Ratio*, 1, 24-35.
- Popper, K. R. & Eccles, J. C. (1977/²1982). *Das Ich und sein Gehirn*. München: R. Piper.
- Prechtel, P. (2004). Individuation. In P. Prechtel (Hrsg.), *Grundbegriffe der analytischen Philosophie* (S. 102). Stuttgart: J. B. Metzler.
- Prinz, W. & Müsseler, J. (2008). Einleitung: Psychologie als Wissenschaft. In J. Müsseler (Hrsg.), *Allgemeine Psychologie*. Heidelberg: Spektrum.
- Putnam, H. (1967). Psychological Predicates. In W. H. Capitan & D. D. Merrill (eds.), *Art, Mind and Religion* (pp. 37-48). Pittsburg: University of Pittsburg Press.
- Pylshyn, Z. W. (1989/³1991). Computing in Cognitive Science. In Posner, M. I. (ed.), *Foundations of Cognitive Science* (pp. 51-91). Cambridge: MIT Press.
- Rang, B. (2000). *Identität und Indifferenz. Eine Untersuchung zu Schellings Identitätsphilosophie*. Frankfurt am Main: Vittorio Klostermann.

- Rescher, N. & Oppenheim, P. (1955). Logical Analysis of Gestalt Concepts. *British Journal for the Philosophy of Science*, 6, 89-106.
- Rosenzweig, M. R., Breedlove, S. M. & Leiman, A. L. (2001). *Biological Psychology: An Introduction to Behavioral, Cognitive, and Clinical Neuroscience*. Sunderland: Sinauer.
- Roth, G. (1994). Braucht die Hirnforschung die Philosophie? In J. Fedorowicz (Hrsg.), *Neuroworlds: Gehirn – Geist – Kultur* (S. 81-92). Frankfurt am Main: Campus Verlag.
- Roth, G. (2002). Gleichtakt im Neuronennetz. *Gehirn & Geist*, 1/2002, 38-46.
- Roth, G. (2004). Worüber dürfen Hirnforscher reden und in welcher Weise? In C. Geyer (Hrsg.), *Hirnforschung und Willensfreiheit. Zur Deutung der neuesten Experimente* (S. 66-85). Frankfurt am Main: Suhrkamp.
- Roth, N. (1974). *Progress in Modern Biology: An Alternative to Reduction*. Ph. D. dissertation. Committee on Conceptual Foundations of Science. University of Chicago.
- Rothacker (1938/1966). *Die Schichten der Persönlichkeit*. Bonn: H. Bouvier u. Co.
- Rumelhart, D. E. (1989/³1991). The Architecture of Mind: A Connectionist Approach. In M. I. Posner (ed.), *Foundations of Cognitive Science* (pp. 133-159). Cambridge: MIT Press.
- Rumelhart, D. E. & McClelland, J. L. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge: MIT Press.
- Russell, B. (1921/2000). *Die Analyse des Geistes*. Hamburg: Meiner.
- Ryle, G. (1949). *The Concept of Mind*. Chicago: University of Chicago Press.
- Sagan, C. (1985). *Contact*. New York: Pocket Books.
- Schachter, S. & Singer, J. E. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 69, 379-399.

- Schäfer, F. (2012). *UDĀNA. Die Blindgeborenen. Angehörige verschiedener Schulen.* (http://www.palikanon.com/khuddaka/udana/ud_6.htm; aufgerufen am 30. Januar 2012)
- Schaffner, K. F. (1967). Approaches to reduction. *Philosophy of Science*, 34, 137-147.
- Schandry, R. (2006). *Biologische Psychologie*. Weinheim: Beltz.
- Schedlowski, M. & Tewes, U. (1996) (Hrsg.). *Psychoneuroimmunologie*. Heidelberg: Spektrum.
- Schelling, F. W. J. (1985) [1794-1852]. *Ausgewählte Schriften*, 6 Bde. Frankfurt am Main: Suhrkamp.
- Schelling, F. W. J. (1802/2005). *Bruno oder über das göttliche und natürliche Prinzip der Dinge*. Ein Gespräch. Hamburg: Meiner.
- Schlick, M. (1918/²1979). *Allgemeine Erkenntnislehre*. Suhrkamp: Berlin.
- Schopenhauer, A. (1813/2007). *Ueber die vierfache Wurzel des Satzes vom zureichenden Grunde*. Zürich: Diogenes.
- Schouten, M. & Looren de Jong, H. (2007). Mind Matters: The Roots of Reductionism. In M. Schouten & H. Looren de Jong (eds.), *The Matter of the Mind. Philosophical Essays on Psychology, Neuro-science, and Reduction* (pp. 1-28). Malden: Blackwell.
- Schumacher, R. (2009). Hirnforschung und schulisches Lernen. In U. Herrmann (Hrsg.), *Neurodidaktik. Grundlagen und Vorschläge für ein gehirngerechtes Lehren und Lernen* (S. 124-133). Weinheim: Beltz.
- Searle, J. R. (1992/1996). *Die Wiederentdeckung des Geistes*. Frankfurt am Main: Suhrkamp.
- Sejnowski, T. J. & Churchland, P. S. (1989/³1991). Brain and Cognition. In Posner, M. I. (ed.), *Foundations of Cognitive Science* (pp. 301-356). Cambridge: MIT Press.

- Selfridge, O. G. (1955). Pattern recognition and modern computers. In *Proceedings of the Western Joint Computer Conference* (pp. 94-97). New York: Institute of Electrical and Electronics Engineers.
- Sellars, W. (1963/⁴1971). *Science, Perception and Reality*. London: Routledge & Paul.
- Selman, R. L. (1971). The relation of role-taking to the development of moral judgement on children. *Child Development*, 42, 79-91.
- Selman, R. L. (1976). Social-cognitive understanding: a guide to educational and clinical practice. In T. Lickona (ed.), *Moral development and behavior* (pp. 299-316). New York: Holt, Rinehart & Winston.
- Shapiro, L. (2000). Multiple realizations. *The Journal of Philosophy*, 97, 635-654.
- Shapiro, L. (2008). How To Test for Multiple Realization. *Philosophy of Science*, 75, 514-525.
- Shapiro, L. A. & Polger, T. W. (2012). Identity, variability, and multiple realization in the special sciences. In S. Gozzano & C. S. Hill (eds.), *New Perspectives on Type Identity. The Mental and the Physical* (pp. 264-287). Cambridge: Cambridge University Press.
- Shoemaker, S. (1980). Causality and Properties. In P. van Inwagen (ed.), *Time and Cause* (pp. 109-135). Dordrecht: D. Reidel Publishing Co.
- Shoemaker, S. (1981). Some varieties of functionalism. *Philosophical Topics*, 12, 83-118.
- Shoemaker, S. (1998). Causal and metaphysical necessity. *Pacific Philosophical Quarterly*, 79, 59-77.
- Shoemaker, S. (2001). Realization and mental causation. In C. Gillett & B. B. Loewer (eds.), *Physicalism and Its Discontents* (pp. 74-98). Cambridge: Cambridge University Press.
- Shoemaker, S. (2003). Realization, micro-realization, and coincidence. *Philosophy and Phenomenological Research*, 67, 1-23.

- Shoemaker, S. (2007). *Physical Realization*. Oxford: Oxford University Press.
- Siegel, R. K. (1992/1998). *Halluzinationen. Expedition in eine andere Wirklichkeit*. Reinbek: Rowohlt.
- Siegel, R. K. & Jarvik, M. E. (1975). Drug-induced hallucinations in animals and man. In R. K. Siegel & L. J. West (eds.), *Hallucinations: Behavior, Experience and Theory* (pp. 81-161). New York: Wiley.
- Simon, H. A. (1969/³1996). The Architecture of Complexity. In H. A. Simon (ed.), *The Sciences of the Artificial* (pp. 183-216). Cambridge: MIT Press.
- Simon, H. A. & Kaplan, C. A. (1989/³1991). Foundations of Cognitive Science. In Posner, M. I. (ed.), *Foundations of Cognitive Science* (pp. 1-47). Cambridge: MIT Press.
- Simons, P. (1987). *Parts. A Study in Ontology*. Oxford: Clarendon Press.
- Singer, W. (1999). Neuronal Synchrony. A versatile code of the definition of relations? *Neuron*, 24, 49-65.
- Singer, W. (2004). Verschaltungen legen uns fest: Wir sollten aufhören von Freiheit zu sprechen. In C. Geyer (Hrsg.), *Hirnforschung und Willensfreiheit. Zur Deutung der neuesten Experimente* (S. 30-65). Frankfurt am Main: Suhrkamp.
- Sklar, L. (1967). Types of inter-theoretic reduction. *British Journal for the Philosophy of Science*, 18, 109-124.
- Slife, B. D. & Williams, R. N. (1995). *What's behind the research? Discovering hidden assumptions in the behavioral sciences*. Thousand Oaks: Sage.
- Slife, B. D. & Williams, R. N. (1997). Torward a Theoretical Psychology. Should a Subdiscipline Be Formally Recognized? *American Psychologist*, 52, 117-129.
- Smart, J. J. C. (1959). Sensations and brain processes. *Philosophical Review*, 68, 148-156.

- Smart, J. J. C. (1963). Materialism. *Journal of Philosophy*, 60, 651-662.
- Smith, L. & Heise, D. (1992). Perceptual similarity and conceptual structure. In B. Burns (ed.), *Percepts, Concepts, and Categories* (pp. 233-272). Amsterdam: North-Holland.
- Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences*, 11, 1–23.
- Soom, P., Sachse, C. & Esfeld, M. (2010). Psycho-neural Reduction Through Functional Sub-Types. *Journal of Consciousness Studies*, 17, 7-26.
- Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New York: Macmillan.
- Spelke, E. S., Breinlinger, K., Jacobson, K., & Phillips, A. (1993). Gestalt relations and object perception: A developmental study. *Perception*, 22, 1483-1501.
- Spelke, E. S., Hirst, W. & Neisser, U. (1976). Skills of divided attention. *Cognition*, 4, 215-230.
- Spinoza, B. (1677/1977). *Ethik. Nach geometrischer Methode dargestellt*. Stuttgart: Reclam.
- Staats, A. W. (1983). *Psychology's crisis of disunity*. New York: Praeger.
- Stephan, A. (2002). Phänomenaler Pessimismus. In M. Pauen & A. Stephan (Hrsg.), *Phänomenales Bewusstsein – Rückkehr zur Identitätstheorie?* (S. 342-365). Paderborn: Mentis.
- Stephan, A. (2006). Zur Rolle des Emergenzbegriffs in der Philosophie des Geistes und in der Kognitionswissenschaft. In D. Sturma (Hrsg.), *Philosophie und Neurowissenschaften* (S. 147-166). Frankfurt am Main: Suhrkamp.
- Stern, E. (2005). Pedagogy meets neuroscience. *Science*, 310, 745.

- Stern, E., Graber, R. & Schumacher, R. (2005). *Lehr-Lern-Forschung und Neurowissenschaften: Erwartungen, Befunde und Forschungsperspektiven (Bildungsforum Bd. 13)*. Berlin: BMBF.
- Sternberg, S. (1966). High-speed scanning in human memory. *Science*, 153, 652-654.
- Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction time experiments. *American Scientist*, 57, 421-451.
- Stiles, J. (2007). Entwicklung räumlicher Fähigkeiten. In L. Kaufmann, H.-C. Nuerk, K. Konrad & K. Willmes (Hrsg.), *Kognitive Entwicklungsneuropsychologie* (S. 245-270). Göttingen: Hogrefe.
- Strack, F. (2010). Wo die Liebe wohnt. Überlegungen zur Bedeutung der bildgebenden Hirnforschung für die Psychologie. *Psychologische Rundschau*, 61, 203-205.
- Strawson, P. (1959). *Individuals*. London: Methuen.
- Stubenberg, L. (1997). Austria vs. Australia: Two versions of the identity theory. In K. Lehrer & J. C. Marek (eds.), *Austrian Philosophy Past and Present. Essays in Honor of Rudolf Haller* (S. 125-146). Dordrecht: Kluwer Academic Publishers.
- Suppes, P. (1993). *Models and Methods in the Philosophy of Science*. Dordrecht: Kluwer.
- Tessier-Lavigne, M. & Gouras, P. (1996). Farbe. In E. R. Kandel, J. Schwartz & T. M. Jessell (1996). *Neurowissenschaften* (S. 459-475). Heidelberg: Spektrum.
- Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.
- Tinbergen, N. (1951). *The Study of Instinct*. Oxford: Clarendon Press.
- Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *The Behavioral and Brain Sciences*, 28, 675-735.

- Tse, P. U. (1999). Volume Completion. *Cognitive Psychology*, 39, 37–68.
- Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.
- Uhlhaas, P. J. & Singer, W. (2006). Neural Synchrony in Brain Disorders: Relevance for Cognitive Dys-functions and Pathophysiology, *Neuron*, 52, 155-168.
- Ungerleider, L. G. & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale & J. W. Mansfield (eds.), *Analysis of visual behavior* (pp. 549-586). Cambridge: MIT Press.
- Van Gulick, R. (1982). Functionalism as a theory of mind. *Philosophical Research Archives*, 8, 185-204.
- Van Gulick, R. (1988). A functionalist plea for self-consciousness. *The Philosophical Review*, 97, 149-81.
- Van Gulick, R. (2001). Reduction, emergence and other recent options on the mind-body problem: a philosophic overview. *Journal of Consciousness Studies*, 8, 1-34.
- Van Essen, D. C. & Gallant, J. L. (1994). Neural mechanisms of form and motion processing in primate visual system. *Neuron*, 13, 1-10.
- Van Hezewijk, R. (2000). The Century of Theoretical Psychology? *Theory & Psychology*, 10, 99-106.
- Varela, F. J., Thompson, E. & Rosch, E. (1992). *Der mittlere Weg der Erkenntnis. Die Beziehung von Ich und Welt in der Kognitionswissenschaft – der Brückenschlag zwischen wissenschaftlicher Theorie und menschlicher Erfahrung*. Bern: Scherz.
- Von Weizsäcker, C. F. (1971/1989). *Die Einheit der Natur. Studien*. München: Hanser.
- Von Weizsäcker, C. F. (1985). *Aufbau der Physik*. München: Hanser.
- Von Kutschera, F. (1999). Probleme der Identität. *Facta Philosophica*, 1, 3-19.

- Wahle, R. (1889). *Über das Verhältniss zwischen Substanz und Attributen in Spinoza's Ethik*. In *Sitzungsberichte der philosophisch-historischen Classe der Kaiserlichen Akademie der Wissenschaften*, Bd. 117. Wien: Tempsky.
- Walde, B. (2008). Die neurowissenschaftliche Wende in der Philosophie des Geistes. In M. Wollschläger (Hrsg.), *Hirn, Herz, Seele, Schmerz: Psychotherapie zwischen Neurowissenschaften und Geisteswissenschaften* (S. 209-224). Tübingen: DGVT-Verlag.
- Walter, S. (2008). Mentale Verursachung: Kausale Exklusion als Argument gegen den nicht-reduktiven Physikalismus. In M. Peschl & A. Batthyany (Hrsg.), *Geist als Ursache: Mentale Verursachung im interdisziplinären Diskurs* (S. 103–119). Würzburg: Königshausen und Neumann.
- Walter, S. (2009). Realisierung und mentale Verursachung. *Deutsche Zeitschrift für Philosophie*, 57, 689-708.
- Walter, S. (2009a). Nicht-reduktiver Physikalismus, mentale Verursachung und Realisierung. In M. Backmann & J. Michel (Hrsg.), *Physikalismus, Willensfreiheit, Künstliche Intelligenz* (S. 51-66). Paderborn: Mentis.
- Weiner, B. (1976). Attributionstheoretische Analyse von Erwartung x Nutzen-Theorien. In H. D. Schmalt & W. U. Meyer (Hrsg.), *Leistungsmotivation und Verhalten* (S. 81-100). Stuttgart: Klett.
- Wimsatt, W. C. (1974). Complexity and Organization. In K. F. Schaffner & R. S. Cohen (eds.), *Boston Studies in the Philosophy of Science* (pp. 67-86). Dordrecht: Reidel.
- Wimsatt, W. C. (1976). Reductionism, levels of organization, and the mind-body problem. In G. Globus, G. Maxwell & I. Savodnik (eds.), *Consciousness and the Brain* (pp. 205-267). New York: Plenum.

- Wimsatt, W. C. (1979). Reduction and reductionism. In P. D. Asquith & H. Kyburg, (eds.), *Current research in philosophy of science* (pp. 352–377). East Lansing: Philosophy of Science Association.
- Wimsatt, W. C. (1980). Reductionistic research strategies and their biases in the units of selection controversy. In T. Nickles (ed.), *Scientific discovery: case studies* (pp. 213-259). Dordrecht: D. Reidel.
- Wimsatt, W. C. (1986). Heuristics and the Study of Human Behavior. In D. W. Fiske & R. Shweder (eds.), *Metatheory in Social Science: Pluralisms and Subjectivities* (pp. 293-314). Chicago: University of Chicago Press.
- Wimsatt, W. C. (2006). Reductionism and Its Heuristics: Making Methodological Reductionism Honest. *Synthese*, 151, 445-475.
- Wimsatt, W. C. (2007). *Re-Engineering Philosophy for Limited Beings*. Cambridge: Harvard University Press.
- Witte, E. H. (2013). Vom Wert und Unwert der Promotion. *Wissenswert*, 01/2013, 33-38.
- Witte, E. H. & Brandt, V. (2010). Social psychological research: the comparison of four journals. *Hamburger Forschungsbericht aus der Sozialpsychologie*, Nr. 89.
- Wittgenstein, L. (1953). *Philosophische Untersuchungen*. Frankfurt am Main: Suhrkamp.
- Wolff, C. (1720/1740). *Der Vernünfftigen Gedancken von Gott, der Welt und der Seele des Menschen, auch allen Dingen überhaupt*. Franckfurt am Mayn: Andreä und Hort.
- Wolfradt, U. (2012). Theoretische Psychologie: Eine historische Betrachtung eines Begriffes und seiner Inhalte. In A Stock, H.-P. Brauns & U. Wolfradt (Hrsg.), *Historische Analysen theoretischer und empirischer Psychologie* (S. 23-36). Frankfurt am Main: Peter Lang.

- Wundt, W. (1888). Ueber Ziele und Wege der Völkerpsychologie. *Philosophische Studien*, 4, 1-27.
- Wundt, W. (1894). Über psychische Causalität und das Princip des psychophysischen Parallelismus. *Philosophische Studien*, 10, 1–124.
- Wundt, W. (1896/1911). *Grundriß der Psychologie*. Leipzig: Engelmann.
- Wundt, W. (1900-1920). *Völkerpsychologie. Entwicklungsgesetze von Sprache, Mythos und Sitte*, 10 Bde. Leipzig: Engelmann u. Kröner.
- Wundt, W. (1907). Über Ausfrageexperimente und über die Methoden zur Psychologie des Denkens. *Philosophische Studien*, 3, 301-360.
- Wundt, W. (1913). *Die Psychologie im Kampf ums Dasein*. Leipzig: Kröner.
- Wundt, W. (1921). *Kleine Schriften, Band 3*. Leipzig: Engelmann.
- Wünsch, G. (2000). *Einführung in die Philosophie der Chemie. Studienbuch für Chemiker und an Chemie Interessierte*. Würzburg: Königshausen & Neumann.
- Yanchar, S. C. & Slife, B. D. (1997). Pursuing Unity in a Fragmented Psychology: Problems and Prospects. *Review of General Psychology*, 1, 235-255.
- Zeki, S. (1990). A Century of cerebral achromatopsia. *Brain*, 113, 1721-1777.
- Zeki, S. (1993). *A Vision of the Brain. The Visible World and the Cortex*. Oxford: Blackwell.

Abbildungsverzeichnis

<u>Nr.</u>	<u>Titel</u>	<u>Quelle</u>
1	Blind men and an elephant	Martha Adelaide Holton & Charles Madison Curry, <i>Holton-Curry readers</i> , Rand McNally & Co. (Chicago), p. 108 (Unbekannter Illustrator, 1914)
2	Konditionierungsmodell PTBS	Heim & Meinlschmidt, 2003, S. 38
3	Psychophysische Kausalität im Sinne des Interaktionismus	Selbst erstellt
4	Kausale Überdetermination im Interaktionismus	Selbst erstellt
5	Epiphänomenalistischer Dualismus	Selbst erstellt
6	Neckerscher Würfel	Wikipedia, Abruf: 22.03.2013 http://en.wikipedia.org/wiki/File:Necker_cube_and_impossible_cube.svg
7	Konzeptform „3D-Objekt“ und Kipp-bild „Vase vs. Gesichter“	<i>3D-Objekt</i> : Mausfeld, 2006, S. 8 <i>Kippbild</i> : Wikipedia, Abruf: 15.04.2014

http://commons.wikimedia.org/wiki/File:cup_or_faces_paradox.svg

- | | | |
|----|--|-------------------------|
| 8 | Die funktionale Architektur des Wahrnehmungssystems | Mausfeld, 2010a, S. 103 |
| 9 | Explanatorische Lücke der Wahrnehmungsforschung | Mausfeld, 2006, S. 11 |
| 10 | „4D-Objekt“ | Mausfeld, 2006, S. 8 |
| 11 | Hierarchie natürlicher Ordnungsebenen nach Morgan (1923) | Selbst erstellt |
| 12 | Hierarchie von Reduktionsebenen nach Oppenheim und Putnam (1958) | Selbst erstellt |
| 13 | Hierarchie an grammatikalischen Beschreibungsebenen | Selbst erstellt |
| 14 | Abfolge von Informationsverarbeitungsschritten nach Sternberg | Anderson, 2001, S. 13 |

- | | | |
|----|---|--|
| 15 | Ausschnitt aus dem Drei-Schichten-Modell der Intelligenz nach Caroll (1993) | Selbst erstellt |
| 16 | Ebenen des visuellen Prozessierens nach Marr (1982) | Selbst erstellt |
| 17 | Geschichtetes Netzwerk mit versteckten Einheiten (links); rekurrentes Netzwerk (rechts) | Kupfermann, 1996, S. 367 |
| 18 | Computational-konnektionistisches Modell | Selbst erstellt |
| 19 | Ebenenmodell der Emotionen | Selbst erstellt |
| 20 | Organisationsebenen des Nervensystems | Churchland & Sejnowski, 1994/1997, S. 15 |
| 21 | Ebenen als lokale Maxima an Regelmäßigkeit und Vorhersagbarkeit | Wimsatt, 1976, S. 240 |
| 22 | Die formale Struktur mechanistischer Ebenen | Craver, 2007, S. 189 |

- | | | |
|----|---|---|
| 23 | Mechanistische Ebenen
des räumlichen Gedächtnisses | Craver, 2007, S. 166 |
| 24 | Das kausale Ausschlussprinzip | Selbst erstellt (in Anlehnung an Kim, 2003, S. 159) |
| 25 | Identitätsbasierte Reduktion | Selbst erstellt (in Anlehnung an Kim, 2003, S. 166) |
| 26 | Hierarchie natürlicher
Ordnungsebenen nach
Morgan (1923) | Selbst erstellt |
| 27 | Farbkreis | Goldstein, 2008, S. 158 |
| 28 | Mechanistische Ebenen
des räumlichen Gedächtnisses | Craver, 2007, S. 166 |
| 29 | Blumengarten | Selbst aufgenommen |
| 30 | Neuronales Geflecht
(links) versus Hierarchische Anordnung in
Form einer Pyramide
(rechts) | Neuronales Geflecht: Barondes, 1995, S. 82
Hierarchie: Selbst erstellt |
| 31 | Psychophysische Kausalität im Sinne des Interaktionismus | Selbst erstellt |

- 32 Verallgemeinertes (links) versus speziesspezifisches Modell multipler Realisierungen (rechts) Polger, 2004, S. 27
- 33 Modell der multiplen Realisierbarkeit Polger, 2004, S. 31
- 34 Absorptionskurven des Lichts der verschiedenen Zapfentypen Wikipedia (Abruf: 15.07.2013) <http://commons.wikimedia.org/wiki/File:Cone-response.svg>
- 35 Kontinuumsmodell der approximativen Theoriereduktion McCauley, 2007, S. 113
- 36 Kontextabhängiges Modell der Theoriereduktion McCauley, 2007, S. 133
- 37 Kontexteffekt der Wahrnehmung Selfridge, 1955 bzw. erweitert durch Anderson, 2001, S. 63
- 38 Kontextabhängiges Modell der Theoriereduktion McCauley, 2007, S. 133



University
of Bamberg
Press

Das zentrale Ziel dieser Arbeit ist die Bereitstellung eines begrifflich-konzeptuellen Fundaments der Psychologie in Form der Formulierung eines allgemeinen Modells einer Architektur des Psychischen. Es ist die Klärung einer längst überfälligen Frage, die Frage nach einer eindeutigen Bestimmung des Gegenstandes der Psychologie und seiner Einordnung in das (Gesamt-)System wissenschaftlicher Größen. Auf der Basis der Teil-Ganzes-Beziehung in Form von Mechanismen unterschiedlicher Hierarchieebenen wird ein allgemeines Modell zur Einpassung von Daten unterschiedlichster Komplexitätsstufen etabliert. Dabei wird sich zeigen, dass das Psychische eine höherstufige Ebene in der natürlichen Organisation eines (menschlichen) Organismus bildet, zugleich aber dennoch mit seinen physischen Grundlagen eine voneinander unablösbare Einheit darstellt. Der Geist ist gewissermaßen das strukturbildende oder formende Prinzip auf allen Ebenen der natürlichen Organisation des Gehirns. Nicht nur trotz, sondern gerade aufgrund der Koextensivität bzw. Identität natürlicher Organisationsebenen handelt es sich beim Psychischen sowohl um ein reales als auch ein kausal wirksames Phänomen. Obgleich jede mentale Eigenschaft neben der psychischen Struktur, wie sie von der Psychologie beschrieben wird, eine Vielzahl (mikro-)struktureller Konfigurationen aufweist (u.a. auf der neuronalen Ebene) und viele Aspekte psychischer Phänomene daher mit Rückgriff auf diese Mikroebenen von z.B. der Neurobiologie erklärt werden können, gibt es weder einen ontologischen noch einen explanatorischen Vorrang der Mikroebenen. Ohne das begrifflich-konzeptuelle Gerüst der Psychologie bleibt auch die methodisch ausgefeilteste Neurowissenschaft der Zukunft bei der Erforschung mentaler Phänomene gänzlich blind.

eISBN 978-3-86309-305-1



9 783863 093051

www.uni-bamberg.de/ubp/