Aus:
Zeitschrift für philosophische Forschung 50 (1996)

Ansgar Beckermann, Bielefeld

Eigenschafts-Physikalismus

1. Das Leib-Seele-Problem – Zwei Aspekte

Menschen führen nicht nur ein biologisches, sondern auch ein geistiges Leben. Sie haben Erfahrungen und Gefühle, Wünsche und Überzeugungen. Sie können Wissen über ihre Umwelt erwerben und dieses Wissen überlegt bei der Planung von Handlungen einsetzen. Kurz: Sie besitzen die Fähigkeit zu denken, also einen Verstand. Und sie haben Bewußtsein. Beides wird häufig in der Formel zusammengefaßt, daß sich Menschen dadurch von (den meisten) anderen Wesen unterscheiden, daß sie einen *Geist* bzw. eine *Seele* besitzen.

Wenn man die Tatsache, daß Menschen mentale Fähigkeiten und Eigenschaften besitzen, in dieser Weise mit Hilfe von Substantiven wie 'Geist' oder 'Seele' ausdrückt, entsteht jedoch die Gefahr, daß das Leib-Seele-Problem auf einen Teilaspekt reduziert wird. Denn dann scheint dieses Problem nur noch in der Frage zu bestehen: Gibt es neben den uns geläufigen materiellen Dingen auch noch ganz andere immaterielle, nicht-physische Dinge (Geister oder Seelen)? Neben dieser gibt es aber auch noch andere und vielleicht sogar interessantere Fragen – etwa die Frage, wie sich die mentalen Zustände und Eigenschaften von Menschen zu ihren physischen Zuständen und Eigenschaften verhalten, ob jene vielleicht mit diesen identifiziert oder auf sie zurückgeführt werden können. Grundsätzlich muß man also mindestens zwei Aspekte des Leib-Seele-Problems unterscheiden – zwei Aspekte, die man z.B. so erläutern kann.

Im Hinblick auf offene Sätzen wie

(1) *x* fühlt starke Schmerzen

und

(2) *x* glaubt, daß es morgen regnen wird,

also offene Sätze, die mentale Eigenschaften ausdrücken, kann man mindestens zwei Fragen stellen.[1]

---

[1] Diese Unterscheidung orientiert sich an der Unterscheidung von drei Fragen in Schiffer (1986, 127f.): 1. Auf welche Art von Dingen treffen offene Sätze wie (1) und (2) zu? 2. Welchen ontologischen Status haben die durch offene Sätze wie (1) und (2) ausgedrückten Eigenschaften? 3. Wie verhalten sich einzelne *Token* der durch offene Sätze wie (1) und (2) ausgedrückten Eigenschaften oder Zustände zu einzelnen Token physischer Eigenschaften oder Zustände? Aus Gründen, die im Laufe der Argumentation deutlich werden sollen, habe ich die dritte Frage hier nicht berücksichtigt.

Frage 1: Auf welche Art von *Dingen* treffen diese offenen Sätze zu? Oder anders ausgedrückt: Welche Art von Dingen sind die Träger mentaler Eigenschaften?

Substanz-Dualisten wie Platon, Descartes oder – in neuester Zeit – Sir John Eccles behaupten, daß es neben Steinen, Bäumen und Planeten – also neben den physischen Gegenständen – auch noch immaterielle, nicht-physische Dinge gibt und daß es diese immateriellen Dinge sind, die ein geistiges Leben haben, die denken, fühlen und sich entscheiden. Für Substanz-Dualisten gilt also: Offene Sätze wie (1) und (2) können nicht auf physische Dinge zutreffen; es gibt aber Dinge, auf die sie zutreffen, nämlich immaterielle Dinge (*res cogitantes*). Der *Substanz-Materialist* behauptet dagegen, daß es in der Welt nur physische Dinge gibt und daß, wenn es überhaupt so etwas wie ein geistiges Leben gibt, bestimmte Organismen, also physische Dinge, Träger dieses Lebens sind. Für ihn treffen offene Sätze wie (1) und (2), wenn überhaupt, auf physische Gegenstände zu.

Frage 2: Welche Art von *Eigenschaften* werden durch die offenen Sätze (1) und (2) ausgedrückt? Beziehungsweise: Welchen ontologischen Status haben die durch offene Sätze wie (1) und (2) ausgedrückten mentalen Eigenschaften?

Hier vertreten *Eigenschafts-Dualisten* die Auffassung, daß mentale Eigenschaften in dem Sinne *eigenständig* sind, daß sie sich *nicht* auf physische Eigenschaften zurückführen lassen, während *Eigenschafts-Materialisten* oder *Physikalisten* wiederum die gegenteilige Auffassung vertreten: Die durch offene Sätze wie (1) und (2) ausgedrückten mentalen Eigenschaften sind mit physischen Eigenschaften identisch bzw. können auf solche Eigenschaften zurückgeführt werden und sind insofern letzten Endes selbst physische Eigenschaften.[2]

Da der Substanz-Dualismus in der gegenwärtigen Diskussion kaum noch eine Rolle spielt, werde ich mich in diesem Aufsatz nur mit dem Problem des Eigenschafts-Dualismus beschäftigen[3] und dabei (abgesehen vom letzten Abschnitt) auch nur mit einem Teilaspekt, nämlich mit der Frage: Was genau ist der Inhalt der Behauptung eines *Eigenschafts-Physikalisten*? D.h., was bedeutet es, zu behaupten, daß mentale Eigenschaften auf physische Eigenschaften zurückgeführt werden können?

[2] Die beiden Probleme des Substanz- und des Eigenschafts-Dualismus machen zusammen allerdings auch noch nicht das ganze Leib-Seele-Problem aus. Insgesamt lassen sich mindestens vier Teilprobleme unterscheiden (vgl. hierzu besonders Churchland (1988)): 1. Das *ontologische* Teilproblem ('Was ist die Natur des Mentalen?'); 2. das *epistemische* Teilproblem ('Wie können wir etwas über das mentale Leben anderer Wesen wissen? Hat das Wissen um die eigenen mentalen Zustände einen privilegierten Status?'); 3. das *semantische* Teilproblem ('Wodurch erhalten mentale Ausdrücke ihre Bedeutung?'); und schließlich das *methodologische* Teilproblem ('Welches ist die beste Methode, um geistige Phänomene zu untersuchen?'). Die Probleme des Substanz- und des Eigenschafts-Dualismus erschöpfen nur das erste, das ontologische Teilproblem.

[3] Im folgenden werde ich daher statt der Ausdrücke 'Eigenschafts-Dualismus' und 'Eigenschafts-Physikalismus' häufig nur die Ausdrücke 'Dualismus' bzw. 'Physikalismus' verwenden.

2. Die klassischen Antworten – Semantischer Physikalismus und Identitätstheorie

In der Analytischen Philosophie des Geistes waren die beiden klassischen Antworten auf diese Frage der *Semantische Physikalismus* (*Logische Behaviorismus*)[4] und die *Identitätstheorie*. Die Position des *Semantischen Physikalismus* ist mit aller Klarheit zuerst von R. Carnap formuliert worden:[5]

(SemPh) Zu jedem Satz *S* einer beliebigen Wissenschaft – und damit auch zu jedem Satz der Psychologie – gibt es einen Satz *S′* der physikalischen Sprache, der dieselbe Bedeutung hat wie *S*.[6]

Diese Formulierung impliziert zwar nicht, daß es für jeden mentalen Ausdruck einen bedeutungsgleichen Ausdruck der physikalischen Sprache gibt; aber sie ist doch von vielen Autoren so verstanden worden. Statt durch (SemPh) kann man diese Position daher auch so charakterisieren:

(SemPh′) Zu jedem mentalen Prädikat *M* gibt es einen offenen Satz $A(x)$ der physikalischen Sprache, so daß der Satz

Für alle x: x ist M genau dann, wenn $A(x)$

analytisch wahr ist.[7]

Die Problematik dieser These zeigt sich allerdings sofort, wenn man versucht, sie an einem konkreten Fall durchzuführen. Wie z.B. könnte der offene Satz $A(x)$ für das Prädikat 'x möchte ein Bier trinken' aussehen? Logische Behavioristen haben etwa folgendermaßen versucht, dieses Prädikat mit Hilfe einer Reihe von Verhaltensdispositionen zu analysieren:

(3) x möchte genau dann ein Bier trinken, wenn gilt:
wenn x zu Hause ist, holt er sich ein Bier aus dem Kühlschrank, und

[4] Der Ausdruck 'Semantischer Physikalismus' scheint mir für die hier gemeinte Position treffender als der verbreitetere Ausdruck 'Logischer Behaviorismus'. Denn dieser Ausdruck beinhaltet eine zu starke Einschränkung auf Verhaltensaspekte, die ja nicht den ganzen Bereich dessen ausmachen, was in physikalischer Sprache ausgedrückt werden kann.

[5] Vgl. bes. Carnap (1932a; 1932b).

[6] 'Bedeutung' wird hier im Sinne des Fregeschen Terminus 'Sinn' verwendet; für den Fregeschen Ausdruck 'Bedeutung' werde ich im Folgenden 'Bezug' benutzen.

[7] Entsprechendes gilt natürlich auch für mehr als einstellige Prädikate.

wenn *x* in der Kneipe ist, bestellt er sich ein Bier, und
wenn man *x* ein Bier anbietet, nimmt er es sofort an,
usw.

Der erste kritische Punkt dieses Analyseversuchs ist offenbar das 'usw.'. Es ist zwar richtig, daß der Wunsch, ein Bier zu trinken, mit einer ganzen Reihe von Verhaltensdispositionen einhergeht; aber welche dieser Dispositionen sind für die Bedeutung des zu analysierenden Prädikats relevant? Wie es scheint, gibt es keine Antwort auf diese Frage, die nicht auch ein Stück Willkür beinhaltet.

Entscheidend ist jedoch ein zweiter Punkt. Wenn man genau hinsieht, wird nämlich sofort klar, daß es nicht einmal *faktisch* wahr ist, daß jeder, der ein Bier trinken möchte, ein Bier aus dem Kühlschrank holt, wenn er zu Hause ist, ein Bier bestellt, wenn er in der Kneipe ist, und ein angebotenes Bier sofort annimmt. Selbst wenn ihm der Sinn eindeutig nach einem Bier steht, wird er das erste nämlich *nicht* tun, wenn er glaubt, daß im Kühlschrank gar kein Bier ist; das zweite *nicht*, wenn ihm etwas anderes wichtiger ist, wenn er etwa sofort nach Hause will; und das dritte *nicht*, wenn er einen Grund hat, das Bier nicht anzunehmen, etwa weil er gerade erzählt hat, daß er ein Jahr lang keinen Tropfen Alkohol mehr anrühren wird. Wenn man versucht, alle diese Nebenbedingungen mit zu berücksichtigen, kommt man jedoch zu einer Analyse wie

(3′) *x* möchte genau dann ein Bier trinken, wenn gilt:
wenn *x* zu Hause ist, holt er sich ein Bier aus dem Kühlschrank, *falls er glaubt*, daß sich im Kühlschrank ein Bier befindet, und
wenn *x* in der Kneipe ist, bestellt er sich ein Bier, *falls er keinen wichtigeren Wunsch hat*, der damit unvereinbar ist, und
wenn man *x* ein Bier anbietet, nimmt er es sofort an, *falls er keinen Grund hat*, das Bier abzulehnen,
usw.

Aber diese Analyse enthält im Analysans eine ganze Batterie von mentalen Ausdrücken, auf die ganz offensichtlich nicht verzichtet werden kann. D.h., es ist unmöglich, in *rein physikalischer Sprache* einen offenen Satz zu formulieren, der dieselbe Bedeutung hat wie das Prädikat '*x* möchte ein Bier trinken'.

Die *Identitätstheorie*, die Ende der 50er und Anfang der 60er Jahre besonders von U.T. Place und J.J.C. Smart entwickelt wurde,[8] war eine explizite Reaktion auf das Scheitern des Semantischen Physikalismus. Ihre Ausgangsfrage lautete: Wie ist Physikalismus möglich, wenn man nicht davon ausgehen kann, daß es zu jedem mentalen Ausdruck einen bedeutungsgleichen Ausdruck der physikalischen Sprache gibt? Bei der Beantwortung dieser Frage stützten sich Place und

8 Die wichtigsten Arbeiten zur Identitätstheorie finden sich in dem Sammelband Borst (1970). Zu sehr ähnlichen Überlegungen vgl. auch Feigl (1958).

Smart auf eine Grundeinsicht Freges: Zwei sprachliche Ausdrücke können, auch wenn sie nicht bedeutungsgleich sind, durchaus denselben Bezug haben.[9] Entsprechend formulierten sie die Grundthese der Identitätstheorie:

(IdTh) Mentale Zustände sind in dem Sinne identisch mit Gehirnzuständen, daß es für jedes mentale Prädikat *M* einen neurobiologischen Ausdruck *N* gibt, so daß *M de facto* dieselbe neuronale Eigenschaft bzw. denselben Gehirnzustand bezeichnet wie *N*, obwohl *M* und *N* nicht dieselbe Bedeutung haben.

Zur Erläuterung dieser These konnten Place und Smart auf eine Reihe von gut etablierten theoretischen Identitätsbehauptungen aus der Physik verweisen:

– Wasser ist identisch mit H_2O,
– Blitze sind identisch mit elektrischen Entladungen,
– die Temperatur eines Gases ist identisch mit der mittleren kinetischen Energie seiner Moleküle.[10]

Auch die Identitätstheorie wurde jedoch von Anfang an recht kritisch beurteilt – diesmal allerdings weniger aus theoretischen als vielmehr aus empirischen Gründen. Eine Identitätsaussage der Form '*M* bezeichnet dieselbe Eigenschaft wie *N*' kann nämlich nur dann wahr sein, wenn das allgemeine Bikonditional 'Für alle *x*: *x* hat *M* genau dann, wenn *x N* hat' *zumindest empirisch* wahr ist. Und eben dies ist für mentale Zustände sehr fraglich.

Wir wissen heute, daß ein bestimmter mentaler Zustand bei verschiedenen Personen mit ganz unterschiedlichen neuronalen Zuständen korreliert sein kann. Ja, sogar bei derselben Person kann sich diese Korrelation im Laufe der Zeit dramatisch verändern. Nach Gehirnverletzungen z.B. können andere Teile des Gehirns die Funktionen des geschädigten Gewebes übernehmen. Außerdem wissen wir, daß sich die Neurophysiologie der meisten Tiere mehr oder weniger stark von der unsrigen unterscheidet. Soll allein daraus schon folgen, daß diese Tiere nicht dieselben mentalen Zustände haben wie wir? Und wie steht es schließlich mit Marsmenschen und Robotern? Sollen diese Wesen schon deshalb kein dem unsren vergleichbares mentales Leben haben, weil ihr 'Gehirn' nicht aus Nervenzellen, sondern z.B. aus Silizium-Chips besteht?

Alle diese Überlegungen machen es ziemlich unwahrscheinlich, daß jedem mentalen Zustand genau ein neurophysiologischer Zustand entspricht. Und eben deshalb steht die Identitätstheorie auf einer äußerst schwachen empirischen Grundlage.

9 Vgl. bes. Frege (1892).

10 Zur Frage, ob in dem letzten Fall wirklich eine Identität vorliegt, siehe unten Abschnitt 6.

3. Die Suche nach Alternativen – Token-Identität und Supervenienz

Nach dem Scheitern des Semantischen Physikalismus und der Identitätstheorie setzte sich in der Analytischen Philosophie des Geistes als erfolgversprechende Alternative zunächst der *Funktionalismus* durch, der die positiven Aspekte des Logischen Behaviorismus aufnahm, ohne seine Schwächen zu teilen. Seine These lautet:[11]

(Fkt) Mentale Zustände sind durch ihre *kausale Rolle* charakterisiert, d.h. durch ihre kausalen Beziehungen zu Wahrnehmungen (*inputs*), anderen mentalen Zuständen und Handlungen (*outputs*).

Bei dieser These, die zunächst ja (wie die These des Logischen Behaviorismus) nur eine These über die Bedeutung mentaler Ausdrücke ist, blieb allerdings seltsam im Dunkeln, inwiefern der *Funktionalismus* überhaupt eine physikalistische Position darstellt bzw. welche Antwort er auf die Frage gibt, was es denn heißen soll, daß mentale Zustände oder Eigenschaften auf physische Zustände oder Eigenschaften zurückführbar sind. Für viele schien das Stichwort '*Token-Identität*' den Ausweg zu markieren. Doch was genau ist mit diesem Stichwort gemeint? Und liefert die mit ihm assoziierte These wirklich eine tragfähige Antwort auf die Frage, was Eigenschafts-Physikalismus ist?

Diese Fragen sind schwerer zu beantworten, als die geläufige Verwendung des Ausdrucks 'Token-Identität' vermuten läßt. Und das auch deshalb, weil sich unter dem Deckmantel dieses Wortes zwei ganz verschiedene Positionen verstecken. Die erste, und wenn man so will, authentische Erläuterung, was mit Token-Identität gemeint sein soll, stammt von D. Davidson:[12]

(TIT_1) Jedes einzelne *mentale Ereignis* ist mit einem einzelnen *physischen Ereignis* numerisch identisch.

Aber diese These ist wegen der mit ihr verbundenen nicht leicht verständlichen Ereignistheorie nicht nur selbst schwer zu verstehen, sie gibt auch nicht die geringste Antwort auf die Frage, wie sich mentale zu physischen *Eigenschaften* verhalten. Dies ist allerdings auch kein Wunder. Denn Davidson macht gar keinen Hehl daraus, wie sehr ihm schon diese Frage zuwider ist. Für ihn als dezidierten Nominalisten sind Fragen nach der Natur von Eigenschaften (ebenso wie die Annahme, Eigenschaften könnten identisch sein!) von vornherein verfehlt. Aber diese Einstellung hilft weniger nominalistisch gesonnen Philosophen natürlich nicht weiter. Für sie ist Davidsons Version der Token-Identitäts-These einfach eine *ignoratio elenchi*; sie liefert keine Antwort auf die gestellte Frage.

Dies gilt in dieser Form jedoch nicht für die zweite Version der Token-Identitäts-These, die sich zum Beispiel bei Churchland (1988, 37) findet:

11 Vgl. bes. Putnam (1960; 1967a; 1967b) und Fodor (1968).

12 Vgl. bes. Davidson (1970).

(TIT_2) Jede *Instantiierung* einer mentalen Eigenschaft oder eines mentalen Zustands ist numerisch identisch mit der *Instantiierung* einer physischen Eigenschaft oder eines physischen Zustands.

Allerdings: Auch diese Version gibt Anlaß zu ontologischen Bauchschmerzen. Zum einen deshalb, weil der ontologische Status von *Eigenschaftsinstantiierungen* – manche sprechen hier auch von *Tropen* – nicht besonders klar ist; zum anderen aber einfach deshalb, weil einige Autoren ihre ontologische Sorgfaltspflicht nicht ernst genug genommen und statt von Eigenschaftsinstantiierungen einfach von Eigenschafts- oder Zustands-Token gesprochen haben, obwohl der Begriff eines Token eindeutig für Gegenstände (einzelne Exemplare eines bestimmten Typs von Gegenständen) reserviert ist.[13]

Angesichts dieser Situation lag es nahe, die Frage nach der Beziehung zwischen mentalen und physischen Eigenschaften erneut aufzunehmen und nach einer physikalistischen Antwort jenseits von Semantischem Physikalismus und Identitätstheorie zu suchen. Da diese beiden Positionen als *reduktionistisch* eingestuft wurden,[14] hieß das Ziel also: Wie kann ein *nicht-reduktionistischer* Physikalismus aussehen?

Das Stichwort, daß von jetzt an die Diskussion für viele Jahre bestimmte – der Begriff der *Supervenienz* – taucht ebenfalls schon in den Überlegungen Davidsons (1970, 214) auf.

> Although the position I describe denies there are psychophysical laws, it is consistent with the view that mental characteristics are in some sense dependent, or supervenient, on physical characteristics. Such supervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respect, or that an object cannot alter in some mental respect without altering in some physical respect.

Eigentlich stammt der Supervenienz-Gedanke jedoch aus der Praktischen Philosophie. Denn dort ist man sich zwar weitgehend einig, daß normative und wertende Prädikate nicht in physikalischer Sprache definiert werden können; auf der anderen Seite scheint es aber doch eine klare Abhängigkeit des Normativen vom Physischen zu geben. Denn mit unseren moralischen Intuitionen ist es nicht vereinbar, anzunehmen, daß von zwei physisch völlig gleichen Gegenständen (z.B. von zwei gleich großen, gleich roten, gleich schmeckenden Äpfeln mit der gleichen chemischen Zusammensetzung und der gleichen physischen

13 Zur Kritik dieser Nachlässigkeit siehe Kemmerling (1996).

14 Dabei wird offenbar angenommen, daß alle reduktionistischen Positionen implizieren, daß jeder mentalen Eigenschaft bzw. jedem mentalen Zustand *genau eine* physische Eigenschaft bzw. *genau ein* physischer Zustand entspricht. Jede Position, die von der Multirealisierbarkeit mentaler Zustände ausgeht, ist diesem Begriff zufolge also *a fortiori* nicht-reduktionistisch. Diese Art, den Begriff des Reduktionismus einzuengen, ist im Grunde theoretisch verfehlt.

Geschichte) der eine gut und der andere schlecht oder daß, wenn zwei Personen in der gleichen Situation das gleiche tun, das Tun des einen moralisch gerechtfertigt, das Tun des anderen dagegen moralisch verwerflich sein könnte. Grundsätzlich gilt also offenbar: Zwei Dinge oder Ereignisse können sich nur dann in ihren moralischen oder Werteigenschaften unterscheiden, wenn es zwischen ihnen *auch* einen physischen Unterschied gibt.

Auf der Grundlage dieser Idee hat besonders Jaegwon Kim[15] versucht, seine Version eines *nicht-reduktionistischen* Physikalismus zu formulieren. Der Grundgedanke dieser Version – der Supervenienz-Theorie – lautet daher:

(ST) Das Mentale ist insofern vom Physischen abhängig, als es keinen mentalen Unterschied ohne einen physischen Unterschied geben kann.

Diese (vermeintliche) Version des Physikalismus ist in den 80er Jahren ausführlich diskutiert worden, wobei sich gezeigt hat, daß es zumindest drei verschiedene Arten gibt, die Grundthese dieser Theorie zu verstehen: im Sinne *schwacher*, im Sinne *starker* und im Sinne *globaler* Supervenienz:

(SchST) Das Mentale ist insofern vom Physischen abhängig, als es *in dieser Welt* keinen mentalen Unterschied ohne einen physischen Unterschied gibt.

(StST) Das Mentale ist insofern vom Physischen abhängig, als es *in allen möglichen Welten* keinen mentalen Unterschied ohne einen physischen Unterschied gibt.

(GlST) Es gibt keine zwei Welten, die sich in allen physischen Eigenschaften gleichen, aber in ihren mentalen Eigenschaften unterscheiden.

Kim hat darüber hinaus immer argumentiert, daß man unter einer bestimmten Voraussetzung[16] die ersten beiden Thesen auch so formulieren kann:

(SchST′) Für jede mentale Eigenschaft M gilt: wenn ein Gegenstand x M hat, dann gibt es eine physische Eigenschaft P mit: x hat P und für alle y: wenn y P hat, dann hat y auch M.
(Formal: $\forall M \forall x(Mx \rightarrow \exists P(Px \wedge \forall y(Py \rightarrow My))))$

(StST′) Für jede mentale Eigenschaft M gilt: wenn ein Gegenstand x M hat, dann gibt es eine physische Eigenschaft P mit: x hat P und *notwendigerweise* für alle y: wenn y P hat, dann hat y auch M.
(Formal: $\forall M \forall x(Mx \rightarrow \exists P(Px \wedge \Box\forall y(Py \rightarrow My))))$

15 Die wichtigsten Aufsätze Kims finden sich inzwischen in dem Band Kim (1993b).

16 Die Voraussetzung der Abgeschlossenheit von Eigenschaften unter Booleschen Operationen, wobei auch unendliche Konjunktionen und Disjunktionen erlaubt sein sollen.

Unabhängig davon, ob die genannte Voraussetzung zutrifft, scheinen mir diese Formulierungen insofern angemessener, als erst an ihnen klar wird, warum Kim der Auffassung sein konnte, daß die Supervenienz-Theorie möglicherweise eine Antwort auf die Frage liefert, wie ein *nicht-reduktionistischer* Physikalismus aussehen könnte. Inhaltlich gesehen laufen die beiden Formulierungen (SchST′) und (StST′) nämlich darauf hinaus, daß alle mentalen Eigenschaften bzw. Zustände eine physische Basis haben, so daß man den Kerngedanken der Supervenienz-Theorie auch so fassen kann:

(ST′) Jeder mentale Zustand M hat eine *physische Basis*. D.h., für jeden mentalen Zustand M gibt es eine Menge $P = \{p_1, \ldots, p_n, \ldots\}$ von physischen Zuständen, für die (in dieser bzw. in allen von dieser Welt zugänglichen Welten) gilt:
(a) M tritt nur auf, wenn einer der Zustände $p_i \in P$ vorliegt;
(b) für alle Zustände $p_i \in P$ gilt: Wenn p_i auftritt, dann tritt auch M auf.

Offenbar ist das Erfülltsein von (ST′) eine *notwendige* Bedingung für jede physikalistische Position; aber reicht diese Bedingung allein schon aus für eine adäquate Analyse eines nicht-reduktionistischen Physikalismus?

In den 80ern wurde zunächst einhellig festgestellt, daß die Relationen der schwachen und der globalen Supervenienz zu schwach sind, um eine physikalistische Position zu begründen, die diesen Namen verdient. Ende der 80er und Anfang der 90er wurde dann jedoch aufgrund eines Aufsatzes von Grimes und mehrerer sehr kritischer Aufsätze von Kim selbst der ganze Ansatz der Supervenienz-Theorie einer grundlegenden Kritik unterzogen.[17]

Erstens, so wurde dabei festgestellt, kann eine Position nur dann als physikalistisch bezeichnet werden, wenn sie eine *asymmetrische* Abhängigkeit des Mentalen vom Physischen beinhaltet; selbst wenn alle mentalen Zustände eine physische Basis haben, schließt das aber nicht aus, daß auch das Umgekehrte gilt. Zweitens kann es sein, daß die Bedingung (ST′) erfüllt ist, weil sowohl das Mentale wie das Physische von etwas Drittem in der Weise abhängen, in der dem Physikalismus zufolge das Mentale vom Physischen abhängen sollte. Und drittens schließlich wurde nun auch klar, das das alleinige Erfülltsein von (ST′) schon deshalb keine physikalistische Position begründen kann, weil es mit eindeutig dualistischen Positionen wie dem Parallelismus oder dem Epiphänomenalismus ohne weiteres vereinbar ist. Die entscheidende theoretische Schwäche der Supervenienz-Theorie wird jedoch erst deutlich, wenn man diese Theorie mit den Broadschen Überlegungen zum Emergenz-Begriff konfrontiert.

17 Grimes (1988) und Kim (1989; 1990). Vgl auch Kutschera (1992).

4. Emergenz, Supervenienz und Realisierung

Broad entwickelte diese Überlegungen als Beitrag zur Debatte um das Vitalismus-Problem, das um die Jahrhundertwende die Gemüter vieler Wissenschaftler genauso bewegte wie heute das Leib-Seele-Problem. Zwei Positionen standen sich gegenüber: der *Mechanismus*, der annahm, daß auch die für Lebewesen charakteristischen Eigenschaften (Ernährung, Wachstum, Wahrnehmung, zielgerichtetes Verhalten, Fortpflanzung, Morphogenese) mechanisch, d.h. letzten Endes physikalisch erklärt werden können, und der *Vitalismus*, der im Gegenteil annahm, daß dies nicht möglich sei und daß man zur Erklärung des Lebens eine besondere Substanz oder einen besonderen Stoff postulieren müsse – eine Entelechie oder einen *élan vital*. Broad versuchte, mit seiner Theorie der emergenten Eigenschaften zwischen diesen beiden Polen Platz für eine dritte, vermittelnde Position zu schaffen.

Sein erster Schritt war dabei, darauf hinzuweisen, daß das Vitalismus-Problem nur der Spezialfall eines sehr allgemeinen Problems ist – des Problems, wie sich die *Makroeigenschaften* eines komplexen Systems zu seiner *Mikrostruktur* verhalten, d.h. zu den *Eigenschaften* der Teile, aus denen das System besteht, und der *Anordnung* dieser Teile. Was diese Relation betrifft, gibt es natürlich die Möglichkeit (die die Position des Vitalismus ausmacht), daß sich eine Makroeigenschaft F eines Systems nicht durch seine Mikrostruktur, sondern nur durch die Annahme einer weiteren Substanz erklären läßt. Aber wenn wir diese Möglichkeit einmal außer acht lassen, gibt es nach Broad eben nicht nur *eine* weitere Möglichkeit (die des Mechanismus), sondern *zwei*: die Eigenschaft F kann *mechanisch erklärbar*, sie kann aber auch *emergent* sein.

> Put in abstract terms the emergent theory asserts that there are certain wholes, composed (say) of constituents A, B, and C in a relation R to each other; that all wholes composed of constituents of the same kind as A, B, and C in relations of the same kind as R have certain characteristic properties; that A, B, and C are capable of occurring in other kinds of complex where the relation is not of the same kind as R; and that the characteristic properties of the whole R(A,B,C) cannot, even in theory, be deduced from the most complete knowledge of the properties of A, B, and C in isolation or in other wholes which are not of the form R(A,B,C). The mechanistic theory rejects the last clause of this assertion. (Broad 1925, 61)

Broads Begriffe der mechanischen Erklärbarkeit und der Emergenz kann man also folgendermaßen auf den Punkt bringen:

(ME) Eine Makroeigenschaft F eines komplexen Systems S mit der Mikrostruktur $[C_1, \ldots, C_n; R]$ (S besteht aus den Teilen $C_1, \ldots, C_n$ in der Anordnung R) ist genau dann *mechanisch erklärbar*, wenn gilt:
- (a) Der Satz ‘Für alle x: wenn x die Mikrostruktur $[C_1, \ldots, C_n; R]$ hat, dann hat x die Makroeigenschaft F’ ist ein wahres Naturgesetz, und
- (b) F kann (wenigstens im Prinzip) aus der vollständigen Kenntnis all der Eigenschaften deduziert werden, die die Komponenten $C_1, \ldots, C_n$ isoliert oder in anderen Anordnungen haben.

(E) Eine Makroeigenschaft F eines komplexen Systems S mit der Mikrostruktur $[C_1, \ldots, C_n; R]$ ist genau dann *emergent*, wenn gilt:
- (a) Der Satz ‘Für alle x: wenn x die Mikrostruktur $[C_1, \ldots, C_n; R]$ hat, dann hat x die Makroeigenschaft F’ ist ein wahres Naturgesetz, aber
- (b) F kann nicht (nicht einmal im Prinzip) aus der vollständigen Kenntnis all der Eigenschaften deduziert werden, die die Komponenten $C_1, \ldots, C_n$ isoliert oder in anderen Anordnungen haben.

In Beckermann (1992b) habe ich untersucht, warum Broad für die Bedingung (b) die angegebene komplizierte Formulierung gewählt hat. Hier kann ich jedoch nur das Ergebnis dieser Überlegungen wiedergeben: Broad möchte sicherstellen, daß bei der Deduktion der Eigenschaft F nicht auf bestimmte *ad hoc*-Eigenschaften der Systemkomponenten oder auf bestimmte *ad hoc*-Gesetze, sondern nur auf die fundamentalen Eigenschaften der Komponenten von S und auf die für Gegenstände mit diesen Eigenschaften allgemein geltenden Naturgesetze zurückgegriffen wird.

Auch wenn dieser Punkt geklärt ist, bleibt allerdings noch die Frage, was Broad wohl meint, wenn er davon spricht, daß F aus bestimmten Eigenschaften der Komponenten von S mit Hilfe bestimmter Gesetze *deduziert* werden kann. Auch hier möchte ich ohne große Begründung einen Interpretationsvorschlag machen. Meiner Meinung nach meint Broad damit, daß man auf der Grundlage der fundamentalen Eigenschaften der Komponenten von S und der für Gegenstände mit diesen Eigenschaften allgemein geltenden Naturgesetze zeigen kann, daß S (bzw. $[C_1, \ldots, C_n; R]$ selbst[18]) alle für die Eigenschaft F charakteristischen Merkmale besitzt.[19]

Insgesamt denke ich also, daß man die beiden Definitionen (ME) und (E) auf die folgende Weise noch klarer formulieren kann:

(ME′) Eine Makroeigenschaft F eines komplexen Systems S mit der Mikrostruktur $[C_1, \ldots, C_n; R]$ ist genau dann *mechanisch erklärbar*, wenn
- (a) der Satz ‘Für alle x: wenn x die Mikrostruktur $[C_1, \ldots, C_n; R]$ hat, dann hat x die Makroeigenschaft F’ ein wahres Naturgesetz ist und wenn es
- (b) wenigstens im Prinzip möglich ist, auf der Grundlage der fundamentalen Eigenschaften der Komponenten von S und der für Gegenstände mit diesen Eigenschaften allgemein geltenden Natur-

18 Dieser Zusatz ist nötig für den Fall, daß es sich bei F um eine Eigenschaft zweiter Stufe im Sinne vom Kim (1992b und 1996) handelt.

19 Auf die Frage, was damit im Einzelfall gemeint sein kann, komme ich noch zurück.

gesetze zu zeigen, daß S (bzw. $[C_1, \ldots, C_n; R]$ selbst) alle für die Makroeigenschaft F charakteristischen Merkmale besitzt.

(E′) Eine Makroeigenschaft F eines komplexen Systems S mit der Mikrostruktur $[C_1, \ldots, C_n; R]$ ist genau dann *emergent*, wenn
(a) auf der einen Seite der Satz 'Für alle x: wenn x die Mikrostruktur $[C_1, \ldots, C_n; R]$ hat, dann hat x die Makroeigenschaft F' ein wahres Naturgesetz ist, wenn es auf der anderen Seite aber
(b) nicht einmal im Prinzip möglich ist, auf der Grundlage der fundamentalen Eigenschaften der Komponenten von S und der für Gegenstände mit diesen Eigenschaften allgemein geltenden Naturgesetze zu zeigen, daß S (bzw. $[C_1, \ldots, C_n; R]$ selbst) alle für die Makroeigenschaft F charakteristischen Merkmale besitzt.

Diese Definitionen sind für mich aus zwei Gründen außerordentlich interessant. Erstens weil sie die bemerkenswerte Pointe haben, daß nach Broad sowohl mechanisch erklärbare als auch emergente Makroeigenschaften über den Mikrostrukturen der Systeme supervenieren, deren Eigenschaften sie sind. Broad zufolge haben also sowohl mechanisch erklärbare als auch emergente Makroeigenschaften eine Basis in diesen Mikrostrukturen. Der Unterschied zwischen mechanisch erklärbaren und emergenten Eigenschaften läßt sich also *nicht* mit Hilfe des Supervenienzbegriffs erläutern. Denn dieser Unterschied besteht nur darin, daß – obwohl *beide* Arten von Eigenschaften über den Mikrostrukturen der Systeme supervenieren, in denen sie auftreten – mechanisch erklärbare Eigenschaften *außerdem* aus diesen Mikrostrukturen deduziert werden können, emergente Eigenschaften dagegen nicht.

Abgesehen von den schon genannten Problemen läßt sich auf der Grundlage des Supervenienzbegriffs also besonders deshalb keine Version des Physikalismus formulieren, weil dieser Begriff gar keine spezifische Position beinhaltet, sondern eine Reihe deutlich verschiedener Positionen umfaßt. Die Annahme, daß mentale Eigenschaften emergent sind, impliziert ebenso, daß sie über physischen Eigenschaften supervenieren, wie die Annahme, daß sie mechanisch erklärbar bzw. physisch realisiert sind. Letzten Endes gilt dies sogar für den Semantischen Physikalismus und für die Identitätstheorie. Auch diesen beiden Positionen zufolge supervenieren mentale über physischen Eigenschaften.

Das Erfülltsein der Bedingung (ST′) ist also eine notwendige Bedingung für alle Versionen des Eigenschafts-Physikalismus, aber auch nicht mehr.[20] Oder anders ausgedrückt: Der Begriff der Supervenienz bildet eine Art von Dach, unter dem sich alle physikalistischen Positionen (und vielleicht noch einige andere) aufhalten; er selbst stellt aber keine eigene Position dar.

20 Diese Auffassung wird inzwischen auch von Kim geteilt; vgl. Kim (1996). Eine Argumentation für die Auffassung, daß Supervenienz allein nicht ausreicht, findet sich auch in Horgan (1993).

Der zweite interessante Aspekt der Broadschen Definitionen liegt für mich darin, daß in der Bedingung (b) der Definition (ME) ein überzeugender und sehr allgemeiner *Identitäts- und Realisierungsbegriff* enthalten ist. Denn was anderes soll es heißen, daß F mit G identisch oder durch G realisiert ist, als daß der Besitz von G (immer oder zumindest in einem bestimmten System S) dazu führt, daß S alle für die Eigenschaft F charakteristischen Merkmale besitzt, bzw. daß G selbst (immer oder zumindest in einem bestimmten System S) alle für die Eigenschaft F charakteristischen Merkmale besitzt? Und deshalb scheint es mir durchaus sinnvoll, im Sinne Broads die Begriffe der (Eigenschafts-)Identität und der Realisierung so zu definieren:

(I) Eine Eigenschaft F ist genau dann mit der Eigenschaft G *identisch*, wenn ein Gegenstand x F dann und nur dann hat, wenn er G hat, und wenn sich aus den für Gegenstände mit der Eigenschaft G allgemein geltenden Naturgesetzen ergibt, daß Gegenstände, die die Eigenschaft G besitzen, alle für F charakteristischen Merkmale besitzen bzw. daß G selbst alle für F charakteristischen Merkmale besitzt.

(R) Eine Eigenschaft F ist genau dann im System S durch die Eigenschaft G *realisiert*, wenn S die Eigenschaft G hat und wenn sich aus den für Gegenstände mit der Eigenschaft G allgemein geltenden Naturgesetzen ergibt, daß S, wenn es die Eigenschaft G hat, alle für F charakteristischen Merkmale besitzt bzw. daß G im System S alle für F charakteristischen Merkmale besitzt.[21]

5. Was ist Eigenschafts-Physikalismus?

Wenn wir die bisherigen Überlegungen zusammenfassen, ergeben sich also nicht drei, sondern vier mögliche Relationen, in denen mentale Eigenschaften oder Zustände zu physischen Eigenschaften oder Zuständen stehen können.

(1) *Semantischer Physikalismus*
Alle mentalen Prädikate sind in physikalischer Sprache definierbar.

21 Offensichtlich muß auch noch die Möglichkeit berücksichtigt werden, daß F in einunddemselben System zu *verschiedenen* Zeiten durch *verschiedene* Eigenschaften realisiert wird. Eigentlich müßte die Definition (R) also in der folgenden Weise ergänzt werden:
(R′) Eine Eigenschaft F ist genau dann zur Zeit t im System S durch die Eigenschaft G *realisiert*, wenn S zu t die Eigenschaft G hat und wenn sich aus den für Gegenstände mit der Eigenschaft G allgemein geltenden Naturgesetzen ergibt, daß S, wenn es zu t die Eigenschaft G hat, zu t alle für F charakteristischen Merkmale besitzt bzw. daß G zur Zeit t im System S alle für F charakteristischen Merkmale besitzt.

(2) *Identitätstheorie*
Jede mentale Eigenschaft ist identisch mit einer physischen Eigenschaft.

(3) *Realisierungstheorie*
Jede mentale Eigenschaft hat eine physische Basis, durch die sie realisiert ist.

(4) *Emergenztheorie*
Jede mentale Eigenschaft hat eine physische Basis; sie ist aber nicht mit dieser Basis identisch und auch nicht durch sie realisiert.

Um die Position des Vollblutdualismus nicht unberücksichtigt zu lassen, sollten wir jedoch auch noch die fünfte Möglichkeit hinzufügen:

(5) *Vollblutdualismus*
Mentale Eigenschaften haben nicht einmal eine physische Basis.

Welche dieser fünf Positionen können als physikalistisch gelten? Welche nicht? Bei einigen ist die Antwort leicht. Denn die Positionen (1) und (2) sind ebenso eindeutig physikalistisch wie die Position (5) nicht-physikalistisch bzw. dualistisch ist. Aber wie steht es mit den Positionen (3) und (4)?

Meiner Meinung nach ist auch die Position (3) – die Realisierungstheorie – eindeutig als physikalistisch zu klassifizieren. Dies zeigt sich schon daran, daß die Identitätstheorie eigentlich nur eine Variante dieser Theorie ist – eine Variante, die dann zutrifft, wenn (zufälligerweise) eine mentale Eigenschaft nur durch eine einzige physische Mikrostruktur realisiert wird. So unsinnig es wäre, anzunehmen, daß es neben der mittleren kinetischen Energie der Moleküle eines Gases *zusätzlich* auch noch die Eigenschaft der Temperatur dieses Gases gibt, so unsinnig wäre es – falls die Realisierungstheorie zutrifft – anzunehmen, daß es neben ihren physischen Realisierungen auch noch mentale Eigenschaften in einem robusten Sinne gibt.[22] Die Annahme, es gäbe solche Eigenschaften, ginge völlig ins Leere, da alle Aufgaben, die diese Eigenschaften übernehmen könnten, schon von ihren Realisierungen erledigt werden. Wenn F eine eigenständige, robuste Eigenschaft ist, dann bedeutet das, daß F einen eigenen Beitrag zur Welt leistet, daß es in der Welt einen Unterschied macht, ob ein Gegenstand F hat oder nicht hat. Wenn nun aber F durch eine Mikrostruktur $[C_1, \ldots, C_n; R]$ realisiert ist, welchen neuen Beitrag könnte F dann noch leisten? Nehmen wir einmal (das Unmögliche) an, ein Gegenstand a hätte die Mikrostruktur $[C_1, \ldots, C_n; R]$, aber nicht die Makroeigenschaft F. Was würde diesen Fall von dem unterscheiden, daß a sowohl die Mikrostruktur $[C_1, \ldots, C_n; R]$ als auch die Eigenschaft F hat? Offenbar nichts. Denn wenn a die Mikrostruktur $[C_1, \ldots, C_n; R]$ hat, dann hat a voraussetzungsgemäß alle Merkmale, die für F charakteristisch sind, und dann läuft in der Welt daher alles so ab, als hätte a F.[23]

Dies bedeutet allerdings nicht, daß im Falle der Wahrheit der Realisierungstheorie auch auf mentales *Vokabular* völlig verzichtet werden könnte (oder gar müßte). Ganz im Gegenteil, wenn mentale Eigenschaften *vielfältig* realisiert sind, dann benutzen wir mentale Prädikate, um alle die Dinge zusammenzufassen, in denen dieselben mentalen Eigenschaften realisiert sind und die sich aus diesem Grunde z.B. auf die gleiche Weise verhalten. Multirealisierbarkeit impliziert geradezu, daß es Gesetze gibt, die sich *nicht* in physikalischer Sprache formulieren lassen; denn sie impliziert, daß sich Dinge auf die gleiche Weise verhalten, die im Hinblick auf ihre physische Struktur keinerlei Gemeinsamkeiten aufweisen. Wenn die Realisierungstheorie zutrifft, bedeutet das also nicht, daß mentale Prädikate überflüssig sind, sondern nur, daß wir die Annahme aufgeben müssen, daß die Dinge, auf die ein mentales Prädikat zutrifft, eine robuste Eigenschaft gemeinsam haben. Vielleicht ist es diese Position, die vernünftiger Weise mit dem Ausdruck ‘nicht-reduktionistischer Physikalismus’ bezeichnet werden sollte.[24]

Allerdings gibt es eine Reihe von Autoren, die meinen, daß eigentlich die Emergenztheorie diesen Titel verdiene.[25] Doch diese Auffassung scheint mir völlig verfehlt; meiner Meinung nach ist die Emergenztheorie überhaupt keine physikalistische Position. Physikalismus beinhaltet nicht nur die These, daß der Bereich des Physischen in dem Sinne fundamental ist, daß er alles andere determiniert, sondern auch die Auffassung, daß es in einem robusten Sinne nur Physisches gibt – nur physische Gegenstände und nur physische Eigenschaften; oder anders ausgedrückt, daß es neben den physischen keine weiteren eigenständigen Eigenschaften gibt. Es ist jedoch gerade eine Pointe der Emergenztheorie, daß es neben den (fundamentalen) physischen Eigenschaften auch noch andere Eigenschaften gibt, Eigenschaften, die sich nicht auf physische Eigenschaften zurückführen lassen.

Dies wird besonders deutlich, wenn man solche emergenten Makroeigenschaften betrachtet, zu deren charakteristischen Merkmalen auch gehört, daß sie bestimmte (physische) Wirkungen haben. Nehmen wir als Beispiel die Eigenschaft, magnetisch zu sein. Zu den charakteristischen Merkmalen dieser Eigen-

22 Zur Unterscheidung von robusten und nicht-robusten (pleonastischen) Eigenschaften vgl. z.B. Schiffer (1987).

23 Hier zeigt sich ganz deutlich, daß die Annahme, a hätte die Mikrostruktur $[C_1, \ldots, C_n; R]$, aber nicht die Makroeigenschaft F, in sich widersprüchlich ist. Denn wenn a alle Merkmale besitzt, die für F charakteristisch sind, dann besitzt a auch F.

24 Vgl. zu dieser Argumentation Beckermann (1992a, 19f.). Allerdings gibt es auch Autoren, die die Position (3) lieber als ‘schwachen oder lokalen Reduktionismus’ bezeichnen würden, und für diese Bezeichnung gibt es ebenfalls gute Gründe, z.B. die große theoretische Nähe zur Identitätstheorie. Wenn man diese Bezeichnung wählt, muß man jedoch die oben in der Anm. 14 angesprochene Auffassung aufgeben, der Reduktionismus impliziere, daß jeder mentalen Eigenschaft *genau eine* physische Eigenschaft entspreche.

25 Siehe z.B. Kim (1992; 1993a).

schaft gehört, daß magnetische Gegenstände Eisenfeilspäne in ihrer Nähe anziehen und daß sie einen Strom in Kreisleitern induzieren, durch die sie geführt werden. Was würde es also bedeuten, wenn Magnetismus eine emergente Eigenschaft wäre? Einerseits, daß nur Gegenstände mit bestimmten Mikrostrukturen magnetisch sind; andererseits aber auch, daß nicht diese Mikrostrukturen für die genannten Wirkungen verantwortlich sein können. Wenn Magnetismus eine emergente Eigenschaft wäre, würde das also bedeuten, daß immer dann, wenn ein magnetischer Gegenstand Eisenfeilspäne anzieht oder Strom in einem Kreisleiter induziert, nur die Eigenschaft, magnetisch zu sein, zur Erklärung dieser Wirkungen herangezogen werden könnte. Magnetismus müßte in diesem Fall also eine eigenständige robuste Eigenschaft sein, da bestimmte Wirkungen, die magnetische Gegenstände auf ihre Umgebung haben, eben nur auf diese Eigenschaft und nicht auf die Mikrostrukturen zurückgeführt werden können, die die Basen dieser Eigenschaft bilden. Die Eigenschaft, magnetisch zu sein, würde in diesem Fall also offensichtlich einen eigenen Beitrag zur Welt leisten.

Aber auch, wenn es nicht zu den charakteristischen Merkmalen einer emergenten Eigenschaft gehört, bestimmte (physische) Wirkungen zu haben, kann man auf ähnliche Weise argumentieren. Denn auch in diesem Fall gibt es mindestens ein charakteristisches Merkmal, daß sich nicht auf die zugrundeliegenden Mikrostrukturen zurückführen läßt. Auch in diesem Fall gibt es also etwas Neues, das über diese Mikrostrukturen hinausgeht. Auch in diesem Fall muß die emergente Eigenschaft also eine von ihren physischen Basen unabhängige Realität besitzen.

Meiner Meinung nach ist also die Position (3) noch physikalistisch, während die Position (4) nicht mehr als physikalistisch bezeichnet werden kann. Und diese Einschätzung scheint mir auch ganz im Einklang mit der Tradition zu stehen. Emergenz-Theoretiker von Broad zu Popper haben sich immer als Gegner eines Physikalismus verstanden, der davon ausgeht, daß alle Phänomene in der Welt auf physikalische Phänomene zurückgeführt werden können.

Für die Charakterisierung des Eigenschafts-Physikalismus ergibt sich damit das folgende Fazit:

(EigPh) Eine Theorie bzgl. der Natur mentaler Eigenschaften ist genau dann physikalistisch, wenn sie impliziert:
- (a) mentale Eigenschaften haben eine physische Basis, und
- (b) mentale Eigenschaften sind nicht emergent, d.h. sie sind zumindest physisch realisiert.

6. Probleme des Physikalismus

Wenn man die Probleme verstehen will, mit denen der Physikalismus heute konfrontiert ist, muß man noch einmal auf den Begriff der Realisierung zurückkommen. Wie ist dieser Begriff genau zu verstehen? D.h., wie kann man im

Einzelfall zeigen, daß S, falls es die Mikrostruktur $[C_1, \ldots, C_n; R]$ besitzt, alle für die Makroeigenschaft F charakteristischen Merkmale hat?

Nehmen wir noch einmal das Beispiel der Temperatur. Wir wissen heute, daß Temperatur *nicht identisch* ist mit der mittleren kinetischen Energie der Moleküle eines Gases; vielmehr gilt nur, daß Temperatur *in Gasen* durch die mittlere kinetische Energie ihrer Moleküle *realisiert* ist. In *festen Körpern* dagegen ist sie durch die Vibrationen von Molekülen realisiert; im *Plasma* auf andere Weise, da Plasma weder Moleküle noch vollständige Atome enthält; und im Vakuum schließlich durch die Verteilung elektromagnetischer Strahlung. Aber wie kann man das zeigen? Wie kann man z.B. zeigen, daß Temperatur in Gasen durch die mittlere kinetische Energie ihrer Moleküle realisiert ist?

Entscheidend ist hier *erstens*, daß der Begriff der Temperatur *allein* durch die *kausale Rolle* dieser Eigenschaft bestimmt ist. Wenn jemand fragen würde, was wir mit dem Ausdruck 'Temperatur' meinen, dann würden wir antworten: Temperatur ist die Eigenschaft von Körpern, die in uns bestimmte Wärme- bzw. Kälteempfindungen hervorruft; die dazu führt, daß die Quecksilbersäule in Thermometern, die mit diesen Körpern in Berührung kommen, steigt oder fällt; deren Steigen oder Fallen in Gasen mit festem Volumen ein entsprechendes Steigen oder Fallen des Druckes bewirkt; die bestimmte chemische Reaktionen auslöst, usw.

Und entscheidend ist *zweitens*, daß sich die kausalen Beziehungen, die für die Eigenschaft der Temperatur charakteristisch sind, in den für diese Eigenschaft geltenden *Gesetzen* widerspiegeln. Dies gilt in erster Linie natürlich für grundlegende Gesetze wie das Gesetz von Boyle und Charles

(4) $p \cdot V = N \cdot k \cdot T$.[26]

Aber darüber hinaus auch für eine ganze Reihe anderer, z.T. ganz harmlos aussehender Gesetze wie etwa:

(5) Wenn die Temperatur eines Körpers T Kelvin beträgt und man diesen Körper mit einem Quecksilberthermometer in Berührung bringt, dann steigt die Quecksilbersäule des Thermometers auf die Marke T.

Wenn man zeigen will, daß Temperatur in Gasen durch die mittlere kinetische Energie ihrer Moleküle realisiert ist, d.h. daß ein Gas, dessen Moleküle eine bestimmte mittlere kinetische Energie besitzen, alle Merkmale aufweist, die für eine bestimmte Temperatur charakteristisch sind, muß man daher nachweisen, daß die angeführten Gesetze in idealen Gasen auch gelten, wenn man in ihnen den Ausdruck 'T' durch einen Ausdruck für die mittlere kinetische Energie er-

26 In dieser Formel steht p für den Druck und N für die Zahl der Moleküle des Gases, V für sein Volumen, k für die Boltzmann-Konstante und T für seine absolute Temperatur.

setzt. D.h., dann muß man zeigen, daß z.B. für ideale Gase außer den Gesetzen (4) und (5) auch die folgenden Gesetze wahre Naturgesetze sind:

(4′) $p \cdot V = N \cdot k \cdot 2/3k \cdot E$.[27]

(5′) Wenn die mittlere kinetische Energie eines Gases $2/3k \cdot E$ Joule beträgt und man dieses Gas mit einem Quecksilberthermometer in Berührung bringt, steigt die Quecksilbersäule des Thermometers auf die Marke *T*.

Wenn man Gesetze wie (4′) und (5′) *Bildgesetze* der ursprünglichen Gesetze (4) und (5) nennt, gilt also: Wenn man von einer Makroeigenschaft *F*, *die allein durch ihre kausale Rolle* – d.h. durch eine Menge *L* von für *F* charakteristischen Gesetzen – *definiert ist*, zeigen will, daß sie durch eine Mikrostruktur $[C_1, \ldots, C_n; R]$ realisiert ist, dann muß man zeigen, daß auch alle Bilder der Gesetze von *L* – d.h. alle Gesetze, die aus den Elementen von *L* entstehen, wenn man in ihnen den Ausdruck '*F*' durch den Ausdruck 'hat die Mikrostruktur $[C_1, \ldots, C_n; R]$' ersetzt – wahre Naturgesetze sind. Bzw., genauer: Wenn man von einer Makroeigenschaft *F*, die allein durch ihre kausale Rolle – d.h. durch eine Menge *L* von für *F* charakteristischen Gesetzen – definiert ist, zeigen will, daß sie *in einem System S* durch eine Mikrostruktur $[C_1, \ldots, C_n; R]$ realisiert ist, dann muß man zeigen, daß für jedes Gesetz von *L* der Form 'Für alle *x*: … *x* hat *F* ---' das *systemrelative* Bild '… *S* hat die Mikrostruktur $[C_1, \ldots, C_n; R]$ ---' eine wahre gesetzesartige Aussage darstellt.

Genau an dieser Stelle ist es sinnvoll, noch einmal auf den Funktionalismus zurückzukommen. Denn das, was gerade über Realisierung gesagt wurde, hilft auch zu klären, worauf diese Position eigentlich hinausläuft. Ich hatte schon darauf hingewiesen, daß der Funktionalismus zunächst nur eine These über die Bedeutung mentaler Ausdrücke ist: Mentale Zustände und Eigenschaften sind allein durch ihre kausale Rolle definiert – durch ihre kausalen Beziehungen zu Wahrnehmungen (*inputs*), anderen mentalen Zuständen und Handlungen (*outputs*). Zu einer physikalistischen Position wird diese Theorie erst, wenn man als zweite These hinzufügt: Alle mentalen Zustände und Eigenschaften sind physikalisch realisiert. Tatsächlich ist der Funktionalismus von den meisten Autoren in diesem Sinne verstanden worden. Zugleich mit dem Funktionalismus wurde daher die Idee der Realisierung geboren – jedoch ohne daß dies auch explizit ausgesprochen worden wäre, d.h. ohne daß die Vertreter dieser Theorie erkannt hätten, daß ihre Position genau genommen aus zwei recht verschiedenen Thesen bestand, und ohne daß sie versucht hätten, die Idee der Realisierung allgemein und präzise zu formulieren.

In den wenigen Versuchen, den Inhalt des Funktionalismus systematisch zu formulieren, zeigt sich allerdings genau das Muster, das ich gerade erläutert

27 In dieser Formel steht jetzt *E* für die mittlere kinetische Energie der Moleküle des Gases.

habe:[28] Auch für den Funktionalismus ist die Ableitung systemrelativer Bilder der für mentale Zustände charakteristischen Gesetze entscheidend. Denn zu zeigen, daß mentale Zustände in einem System *S* physisch realisiert sind, heißt für den Funktionalisten zu zeigen, daß es in *S* physische Zustände gibt, die zueinander und zu den möglichen *inputs* und *outputs* genau in den kausalen Beziehungen stehen, die für die entsprechenden mentalen Zustände charakteristisch sind. Denn die erste These des Funktionalismus besagt ja gerade, daß mentale Zustände durch ihre kausale Rolle charakterisiert sind und durch sonst nichts.

Doch dieses Muster liefert eben nur eine Teilantwort, nämlich lediglich eine Antwort auf die Frage, wie man von Makroeigenschaften, *die allein durch ihre kausale Rolle definiert sind*, zeigen kann, daß sie durch bestimmte Mikrostrukturen realisiert sind. Immer noch offen ist die Frage, wie man dies für andere Makroeigenschaften zeigen kann – Eigenschaften, deren Merkmale sich nicht in ihren kausalen Rollen erschöpfen. Und genau in dieser Frage steckt das Hauptproblem für den gegenwärtigen Physikalismus; denn zumindest bei einer bestimmten Gruppe von Merkmalen – den sogenannten Qualia – gibt es bisher nicht einmal den Ansatz einer überzeugenden Antwort.

Worum geht es dabei? Seit Anfang der 60er Jahre der Funktionalismus populär wurde, hat es immer wieder Kritiker gegeben, die vehement für die Auffassung gestritten haben, daß es zumindest eine Gruppe von mentalen Zuständen gibt, die nicht (allein) durch ihre kausale Rolle definiert sind.[29] Empfindungen (*sensations*) wie ein bestimmter Schmerz, der Wahrnehmungseindruck einer roten Tomate, einer schrillen Trompete oder einer süßen Birne haben zwar auch eine kausale Rolle, aber diese ist für sie nicht entscheidend. Entscheidend ist für diese Zustände ein ganz anderes charakteristisches Merkmal: ihr qualitativer Gehalt; die Art und Weise, wie es sich anfühlt, in diesen Zuständen zu sein.[30] Alles, was sich anfühlt wie ein Schmerz, ist ein Schmerz, und alles, was sich anfühlt wie der Eindruck des Klangs einer Klarinette, ist der Eindruck des Klangs einer Klarinette. Aber, und darin liegt das Problem für den Physikalisten, offenbar gibt es keine Möglichkeit zu zeigen, daß der Zustand, in dem sich ein System befindet, wenn es eine bestimmte Mikrostruktur hat, einen bestimmten qualitativen Gehalt besitzt, ja daß mit diesem Zustand überhaupt ein qualitativer Gehalt verbunden ist. Und was noch schlimmer ist, wir können dies nicht nur nicht zeigen, wir haben nicht einmal die geringste Ahnung, was es heißen würde, dies zu zeigen.

J. Levine hat dieses Problem sehr schön veranschaulicht. Wenn man, so argumentiert er, die beiden Identitätsbehauptungen

28 Vgl. bes. Lewis (1972).

29 Z.B. Block, Nagel und Levine, um nur einige zu nennen.

30 Diese Charakterisierung geht hauptsächlich auf Nagel (1974) zurück.

(6) Temperatur ist die mittlere kinetische Energie der Moleküle eines Gases

und

(7) Schmerz ist das Feuern von C-Fasern

vergleicht, dann zeigt sich ein prinzipieller Unterschied.[31] Im demselben Sinne nämlich, indem es auf der einen Seite *undenkbar*[32] ist, daß in einem Gas die mittlere kinetische Energie der Moleküle $6.21 * 10^{-21}$ Joule beträgt, das Gas aber nicht die Temperatur von 300 K hat, scheint es auf der anderen Seite sehr wohl denkbar, daß in meinem Körper die C-Fasern feuern, ich aber keinen Schmerz empfinde. Nach Levine liegt dies daran, daß die Aussage (6) *vollständig explanatorisch* ist, die Aussage (7) dagegen nicht. Was ist damit gemeint?

Zunächst ist auch für Levine die Eigenschaft der Temperatur allein durch ihre *kausale Rolle* charakterisiert. Der Grund für den explanatorischen Charakter von (6) ist daher

> … that our knowledge of chemistry and physics makes intelligible how it is that something like the motion of molecules could play the causal role we associate with heat. Furthermore, antecedent to our discovery of the essential nature of heat, its causal role … *exhausts our notion of it*. Once we understand how this causal role is carried out there is nothing more we need to understand. (Levine 1983, 357 – Hervorh. vom Verf., A.B.)

Mit der Aussage (7) ist es auf den ersten Blick ähnlich. Denn auch mit dem Ausdruck 'Schmerzen' assoziieren wir eine bestimmte kausale Rolle. Schmerzen werden durch die Verletzung von Gewebe verursacht, sie führen dazu, daß wir schreien oder 'Aua' sagen, sie bewirken, daß wir in Zukunft die schmerzauslösende Situation vermeiden, usw. Auch die Aussage (7) erklärt also das Phänomen Schmerz, indem sie uns sagt, welcher neuronale Zustand Träger der entsprechenden kausalen Rolle ist. Der entscheidende Unterschied ist jedoch, daß sich die Bedeutung von 'Schmerz' nicht in einer kausalen Rolle erschöpft.

> However, there is more to our concept of pain than its causal role, there is its qualitative character, how it feels; and what is left unexplained by the discovery of C-fiber firing is *why pain should feel the way it does*! For there seems to be nothing about C-fiber firing which makes it naturally 'fit' the phenomenal properties of pain, any more than it would fit some other set of phenomenal properties. Unlike its functional role, the identification of the qualitative side of pain with C-fiber firing […] leaves the connection between it and what we identify it with completely mysterious. One might say, it makes the way pain feels into merely a brute fact. (ebd.)

31 Daß Levine hier von Identität statt von Realisierung spricht, spielt für die Struktur des Arguments keine Rolle.

32 Vgl. oben Anm. 23.

Der wesentliche Grund dafür, daß die Aussage (7) in den Augen Levines nicht vollständig explanatorisch ist, ist also, daß (a) die Bedeutung des Ausdrucks 'Schmerz' außer einer kausalen Rolle auch einen qualitativen Aspekt umfaßt und daß man (b) allein mit den Mitteln der Physik, Chemie und Neurobiologie *unmöglich zeigen kann*, daß es sich für einen Organismus in der für Schmerzen charakteristischen Weise *anfühlt*, wenn seine C-Fasern feuern. Ja, es scheint außerhalb der Reichweite jeder möglichen naturwissenschaftlichen Erklärung, zu zeigen, daß sich das Feuern von C-Fasern auf irgendeine Weise anfühlt.[33]

Damit ist der zentrale Punkt klar benannt: Empfindungen stellen ein bisher unlösbares Problem für den Physikalismus dar, da sie mit ihrem qualitativen Charakter über ein Merkmal verfügen, von dem, wie es scheint, mit naturwissenschaftlichen Mitteln nicht gezeigt werden kann, daß physische Zustände dieses Merkmal besitzen, ja von dem wir nicht einmal wissen, wie ein solcher Nachweis überhaupt aussehen könnte.[34] Seit der Kritik am Funktionalismus hat es zwar immer wieder Versuche gegeben, dieses Problem auf indirektem Wege sozusagen zu umgehen, und ich selbst habe durchaus Sympathien für diese Vorgehensweise;[35] doch in den Augen vieler Autoren haben alle diese Versuche der Frage, wie qualitative mentale Zustände physikalisiert werden können, ihren theoretischen Stachel bisher nicht nehmen können.

Literatur

Beckermann, A. (1992a) Introduction – Reductive and Nonreductive Physicalism. In: Beckermann/Flohr/Kim, 1-21.

Beckermann, A. (1992b) Supervenience, Emergence, and Reduction. In: Beckermann/Flohr/Kim, 94-118.

Beckermann, A., Flohr, H. & Kim, J. (eds.) (1992) *Emergence or Reduction? – Essays on the Prospects of Nonreductive Physicalism*. Berlin/New York: Walter de Gruyter.

Bieri, P. (1992) Was macht Bewußtsein zu einem Rätsel? *Spektrum der Wissenschaft*, Oktober 1992, 48-56.

Block, N. (1978) Troubles with Functionalism. In: C.W. Savage (ed.) *Perception and Cognition. Minnesota Studies in the Philosophy of Science. Vol. 9*. Minneapolis: University of Minnesota Press, 261-325. Wiederabgedr. in: N. Block (ed.) *Readings in Philosophy of Psychology. Vol. 1*. Cambridge, Mass.: Harvard University Press 1980, 268-305.

33 Diese Argumentation deckt sich fast völlig mit C.D. Broads Argumentation für den absolut emergenten Charakter von Sinnesempfindungen in (1925). Vgl. hierzu Stephan (1993).

34 Dies scheint mir auch der Hauptpunkt der Argumentation in Bieri (1992).

35 Noch in den 70ern z.B. von S. Shoemaker, in neuester Zeit bes. von C. Hardin.

Borst, C.V. (ed.) (1970) *The Mind-Brain Identity Theory*. London: Macmillan.
Broad, C.D. (1925) *The Mind and Its Place In Nature*. London: Routledge and Kegan Paul.
Carnap, R. (1932a) Die physikalische Sprache als Universalsprache der Wissenschaft. *Erkenntnis* 2, 432-465.
Carnap, R. (1932b) Psychologie in physikalischer Sprache. *Erkenntnis* 3, 107-142.
Churchland, P.M. (1988) *Matter and Consciousness*. Rev. ed. Cambridge, Mass.: MIT-Press.
Davidson, D. (1970) Mental Events. In: L. Foster and J.W. Swanson (eds.) *Experience and Theory*. Amherst: Unversity of Massachusetts Press, 79-101. Wiederabgedr. in: D. Davidson, *Essays on Actions and Events*. Oxford: Clarendon Press 1980, 207-225.
Feigl, H. (1958) The 'Mental' and the 'Physical'. In: H. Feigl, M. Scriven & G. Maxwell (eds.) *Concepts, Theories, and the Mind-Body Problem*. *Minnesota Studies in the Philosophy of Science. Vol. 2*. Minneapolis: University of Minnesota Press, 320-492.
Fodor, J. (1968) *Psychological Explanation*. New York: Random House.
Frege,G. (1892) Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* NF 100, 25-50. Wiederabgedr. in: G. Frege, *Funktion, Begriff, Bedeutung. Fünf logische Studien*. Hrsg. und eingel. von G. Patzig. 6. Aufl. Göttingen: Vandenhoek und Ruprecht 1986, 40-65.
Grimes, Th.R. (1988) The Myth of Supervenience. *Pacific Philosophical Quarterly* 69, 152-160.
Hardin, C.L. (1987) Qualia and Materialism: Closing the Explanatory Gap. *Philosophy and Phenomenological Research* 48, 281-298.
Hardin, C.L. (1992) Physiology, Phenomenology, and Spinoza's True Colors. In: Beckermann/Flohr/Kim, 201-219.
Horgan, T. (1993) From Supervenience to Superdupervenience: Meeting the Demands of a Material World. *Mind* 102, 555-586.
Kemmerling, A. (1996) Überzeugungen für Naturalisten. In: G. Meggle (Hg.) *Analyomen 2*. Berlin/New York: Walter de Gruyter.
Kim, J. (1989) The Myth of Nonreductive Materialism. *Proceedings and Addresses of the American Philosophical Association* 63, 31-47. Wiederabgedr. in: Kim (1993b), 265-284.
Kim, J. (1990) Supervenience as a Philosophical Concept. *Metaphilosophy* 21, 1-27. Wiederabgedr. in: Kim (1993b), 131-160.
Kim, J. (1992a) 'Downward Causation' in Emergentism and Nonreductive Materialism. In: Beckermann/Kim/Flohr, 119-138.
Kim, J. (1992b) Multiple Realization and the Metaphysics of Reduction. *Philosophy and Phenomenological Research* 52, 1-26. Wiederabgedr. in: Kim (1993b), 309-335.

Kim, J. (1993a) The Nonreductivist's Troubles with Mental Causation. In: J. Heil and A. Mele (eds.) *Mental Causation*. Oxford: Clarendon Press, 189-210. Wiederabgedr. in: Kim (1993b), 336-357.
Kim, J. (1993b) *Supervenience and Mind*. Cambridge: Cambridge University Press.
Kim, J. (1996) Supervenience, Emergence, and Realization in the Philosophy of Mind. In: M. Carrier and P. Machamer (eds.) *Philosophy and the Sciences of the Mind*.
Kutschera, F.v. (1992) Supervenience and Reductionism. *Erkenntnis* 36, 333-343.
Levine, J. (1983) Materialism and Qualia: The Explanatory Gap. *Pacific Philosophical Quarterly* 64, 354-361.
Lewis, D. (1972) Psychophysical and Theoretical Identifications. *Australasian Journal of Philosophy* 50, 249-258.
Nagel, Th. (1974) What is it like to be a bat? *Philosophical Review* 83, 435-450.
Putnam, H. (1960) Minds and Machines. In: S. Hook (ed.) *Dimensions of Mind*. New York, 138-164. Wiederabgedr. in: Putnam (1975), 362-385.
Putnam, H. (1967a) The Mental Life of Some Machines. In: H. Castañeda (ed.) *Intentionality, Mind, and Perception*. Detroit. Wiederabgedr. in: Putnam (1975), 408-428.
Putnam, H. (1967b). Psychological Predicates. In: W.H. Capitan & D.D. Merrill (eds.) *Art, Mind, and Religion*. Pittsburgh. Unter dem Titel 'The Nature of Mental States' wiederabgedr. in: Putnam (1975), 429-440.
Putnam, H. (1975) *Mind, Language, and Reality. Philosophical Papers. Vol. 2*. Cambridge: Cambridge University Press.
Schiffer, St. (1986) Functionalism and Belief. In: M. Brand & R. Harnish (eds.) *The Representation of Knowledge and Belief*. Tuscon, 127-159.
Schiffer, St. (1987) *Remnants of Meaning*. Cambridge, Mass.: MIT Press.
Shoemaker, S. (1975) Functionalism and Qualia. *Philosophical Studies* 27, 291-315. Wiederabgedr. in: S. Shoemaker, *Identity, Cause, and Mind*. Cambridge: Cambridge University Press 1984, 184-205.
Stephan, A. (1993) C.D. Broads a priori-Argument für die Emergenz phänomenaler Qualitäten. In: H. Lenk und H. Poser (Hrsg.) *Neue Realitäten – Herausforderung der Philosophie. Sektionsbeiträge I zum XVI. Deutschen Kongreß für Philosophie*. Berlin, 176-183