

Thomas Metzinger

KRITERIEN FÜR EINE THEORIE ZUR LÖSUNG DES LEIB-SEELE-PROBLEMS

Nachdem der analytische Behaviorismus Ryle'scher Prägung als Programm in der Philosophie des Geistes zusammengebrochen war, entstand auf der Suche nach einem neuen philosophischen Paradigma in den letzten drei Jahrzehnten eine besonders lebhafte und fruchtbare Diskussion. Speziell was das Leib-Seele-Problem - also die kausale Relation zwischen mentalen und neurophysiologischen Zuständen - angeht, wurde eine Fülle neuer Positionen und Ansätze entwickelt. Ziel dieses Beitrags ist es, durch eine kurze kritische Analyse¹ der wichtigsten Theorien zu zeigen, daß keine von ihnen derzeit als befriedigend angesehen werden kann. Außerdem werde ich versuchen, von dieser Kritik zur Aufstellung eines Minimalkatalogs von Kriterien zu gelangen, die eine zukünftige Theorie über die Relation zwischen psychischen und physischen Ereignissen erfüllen muß.

Ich werde zu diesem Zweck die fraglichen philosophischen Theorien in fünf Großgruppen einteilen, innerhalb derer dann freilich weiter differenziert werden muß: Identitätstheorie, Emergentistischer Materialismus, Supervenienz-Theorie, dualistischer Interaktionismus und Funktionalismus.

DIE IDENTITÄTSTHEORIE

Für die These der *Identität* mentaler und neurophysiologischer Prozesse können im wesentlichen sechs Argumente als positive Gründe angeführt werden. Die *Korreliertheit* der beiden Klassen von Vorgängen als zentrale Prämisse ist mittlerweile ein hervorragend gesicherter und allseits belegter

¹ Für eine ausführlichere Analyse vgl. Th. Metzinger, *Neuere Beiträge zur Diskussion des Leib-Seele-Problems*, Kap. 6, Frankfurt-Bern-New York 1985.

Sachverhalt, an dem rationalerweise nicht mehr gezweifelt werden kann. Zweitens braucht die Identitätstheorie das Leib-Seele-Problem nicht zu *lösen* (etwa indem sie auf eine dubiose Form psychophysischer Kausalität verweist), weil sie es *auföst*. Ein drittes Argument für die Hypothese der Identität von Gehirn- und Bewußtseinszuständen ist zudem ihre besondere *Kompatibilität -mit dem gegenwärtigen naturwissenschaftlichen Weltbild*. Dies besagt direkt nichts über die Wahrheit weder der Identitätstheorie noch der herrschenden Kosmologie; jedoch ist Kongruenz mit weithin akzeptierten wissenschaftlichen Theoriensystemen aus wissenschaftstheoretischen Gründen ein starkes indirektes Argument für eine Hypothese. Viertens vereinfacht die Identitätsthese die Situation: Sie hat empirisch exakt denselben Gehalt wie die Korrelationsthese, ist aber deutlich einfacher und schafft keine neuen Probleme bzw. verschiebt lediglich die Ausgangsfrage (wie z.B. Poppers Drei-Welten-Lehre)². Damit werden sekundäre Fragen, wie etwa nach der evolutionstheoretisch-genetischen Erklärung des *Faktums* der Korreliertheit, überflüssig. Der fünfte positive Grund liegt in der vollkommenen Sparsamkeit in bezug auf ontologische und strukturelle Annahmen. Weil im Kontext empiristischer Methodologie die Bedeutung des Ockham'schen Rasiermesser-Prinzips kaum überschätzt werden kann, ist der Rekurs auf Ockhams Rasiermesser ein häufiger, ja typischer, zweiter Schritt in aus dem angelsächsischen Raum stammenden Argumentationen für die Identitätsthese.

Macht man von der Möglichkeit, mentale Zustände über ihre kausale Rolle zu definieren, Gebrauch, so umgeht man schließlich die Schwierigkeit der scheinbaren *kausalen Überdeterminiertheit* von Verhaltensmustern, in die die Korrelationsthese gerät: Bezüglich eines bestimmten motorischen Outputs hat sowohl ein bestimmter Gehirnvorgang als auch das ihm zugeordnete mentale Phänomen dieselbe kausale Rolle inne. Kausale Effektivität als Definiens mentaler Zustände bewahrt ein wichtiges Element des logischen Behaviorismus, macht sie jedoch zu *realen* Zuständen und hat gegenüber den kontrafaktischen Relationen des logischen Behaviorismus den Vorteil, nicht von vornherein an der Unfähigkeit zu scheitern, das Episodenhafte, die *Prozessualität* der meisten und den *qualitativen Charakter* mancher mentalen Zustände zu erklären.

Einer der bekanntesten gegen die Identitätstheorie vorgebrachten Einwände stammt von Saul Kripke³. Im ersten Schritt seines Arguments attackiert Kripke vor dem Hintergrund seiner eigenen Semantik-Theorie die behauptete *Kontingenz* der Identität: Wenn Gehirnereignisse und mentale Phänomene

² Vgl. K. Popper, J. Eccles, *Das Ich und sein Gehirn*, München 1982; im Original: *The Self and its Brain*, London 1977.

³ Vgl. S. Kripke, *Identity and Necessity*, [In:] *Identity and Individuation*, Hrsg. M. Munitz, New York 1971 und auch S. Schwartz, *Naming, Necessity and Natural Kinds*, Ithaca 1977.

identisch miteinander sein sollen, dann können sie überhaupt nur *notwendig* miteinander identisch sein. Das gilt zumindest für alle wahren Identitätsaussagen, in denen die beiden Terme *starre Designatoren* sind, die in allen möglichen Welten dieselben Referenten besitzen. Sie funktionieren wie Eigennamen, d.h. sie spezifizieren ihre Referenzobjekte *essentiell*. Wenn zwei durch starre Bezeichnungsausdrücke erfaßte Referenzobjekte - etwa „mein Gehirnvorgang vom Typ A“ und „mein Schmerz B“ - identisch sind, dann sind sie es in allen möglichen Welten. Es ist nicht möglich, daß Gehirnvorgänge eines bestimmten Typs in *einer* kontrafaktischen Situation mit Schmerzen eines gewissen Typs identisch sind und in einer *anderen* nicht.

Das eigentliche Problem liegt in Kripkes zweiter Prämisse, die auf einer Verwechslung von Analytizität und Imaginierbarkeit beruht. Kripke sagt: Wir alle können uns *vorstellen*, daß wir Schmerzen des Typs B haben, ohne daß sich in unserem Gehirn Vorgänge des fraglichen Typs A ereignen. Also handelt es sich hier um eine *kontingente* Relation, folglich, so Kripkes Konklusion, können Schmerzen keine Gehirnvorgänge sein. Man muß sich, um den Fehler in der zweiten Prämisse zu demonstrieren, nicht auf eine Diskussion darüber einlassen, was wir uns vorstellen können und was nicht. (Eine Kritik der *ersten* Prämisse würde, um fair zu sein, ausführlich auf Kripkes Semantik eingehen müssen.) Der Rekurs auf das, was wir uns angeblich alle vorstellen können, dient hier dazu, unser aller cartesianische Intuitionen (geprägt durch die unserer Alltagssprache inhärente dualistische Theorie) auf subtile Weise in ein Argument einfließen zu lassen. Kripke versucht dabei, einer Intuition (der *Kontingenz-Intuition*) den Charakter einer analytischen Wahrheit zu verleihen.

Der erste wirklich schwerwiegende Einwand hängt mit zwei spezifischen Qualitäten zusammen, die Mentalphänomenen innewohnen und bei ontologisch neutralen⁴ und reduktionistischen Analysen verlorengehen. Die solche Qualitäten bezeichnenden Terme der mentalistischen Beschreibungsebene lassen sich nicht in solche niedriger liegender Ebenen (wie denjenigen der Physik oder der Neurophysiologie) übersetzen, die naturwissenschaftliche Sprache hat ein ärmeres Vokabular als unser mentalistisches Alltagsidiom. Die erste von ihnen ist die *phänomenologische Qualität von Sinneswahrnehmungen*. Wenn ich einen roten Kreis sehe, dann kann die Qualität der „Röte“ in einer physikalistischen Beschreibung jenes Wahrnehmungsvorgangs nicht mehr enthalten sein. „Röte“ finden wir weder „da draußen“ (sozusagen als Eigenschaft der Dinge in der Welt), noch als Eigenschaft der Wellen, durch

⁴ Vgl. J. Smart, *Sensations and Brain Processes*, „Philosophical Review“ 1959, No. 68; ders., *Futher Remarks on Sensations and Brain Processes*, „Philosophical Review“ 1959, No. 70; ders., *Brain Processes and In corrigibility*, „Australasian Journal of Philosophy“ 1962, No. 60; ders., *Materialism*, „Journal of Philosophy“ 1963, No. 60; aber auch H. Feigl, *Mind-body, not a Pseudo-problem*, [in:] S. Hook, *Dimensions of Mind*, New York 1960.

deren Modulation die Information über den spezifischen Objektzustand meine Sinnesorgan erreicht (in diesem Fall der Lichtwellen), noch auf meiner Netzhaut, noch in irgendeinem Teil meines Gehirns. Das Wesen naturwissenschaftlicher Beschreibungen besteht darin, von Erscheinungen zu abstrahieren und so in Richtung Objektivität und Verallgemeinerungsfähigkeit fortzuschreiten. Darum kann eine Qualität wie „Röte“ (oder „Süße“, oder „Härte“, oder „Glattheit“) – obwohl sie erlebt wird – in physikalistischen Beschreibungen von sinnlichen Wahrnehmungsvorgängen nicht mehr auftauchen.

Eng verwandt mit dem *Qualia-Problem* ist das *Subjektivitäts-Problem*⁵. Erlebnisse und Erfahrungen besitzen eine spezifische *Perspektivität*: Sie besitzen eine gewisse Innenansicht, es sind meine Erfahrungen und sie zeichnen sich durch einen besonderen subjektiven Charakter aus. Dieses „*wie es ist*“, Erlebnisse zu haben, scheint *essentiell* subjektiv zu sein, d.h. es geht in objektiven Erlebnisbeschreibungen notwendig verloren und erweist sich als gegenüber intentionalen oder funktionalen Erklärungssystemen als analyseresistent. Thomas Nagels Überlegungen zu diesem Problem sind deshalb bedeutend, weil sie fundamentale Schwierigkeiten des wissenschaftlichen Realismus in der Philosophie des Geistes implizieren. Wenn Nagel recht hat, dann gibt es *nicht-verobjektivierbare Elemente des Mentalen*; eine wissenschaftliche Anthropologie könnte also prinzipiell niemals vollständig sein. Nagels „*wie es ist, ein x zu sein*“ ist auch durch Phantasie und Intuition schwer zu erschliessen. Vielleicht kann ich mir vorstellen, wie es für *mich* wäre, ein x zu sein; woran wir aber interessiert sein müssen, ist, herauszufinden, wie es für x ist, ein x zu sein.

Nagels Argument beruht auf der Kritik an einer falschen, von Vertretern der Identitätstheorie immer wieder eingesetzten Analogie, die letztlich auf eine *petitio principii* hinausläuft. Sie besteht in der Gegenüberstellung von Sätzen wie „Blitze sind elektrische Entladungen“ und Sätzen wie „Mentale Ereignisse sind physikalische Ereignisse“. Offensichtlich funktioniert das Identitätszeichen hier auf zwei verschiedene Arten und Weisen. Daß ein Gedanke ein Gehirnvorgang *ist*, heißt in bezug auf das Wörtchen „*ist*“ etwas vollkommen anderes, als daß Wärme die kinetische Aktivität von Molekülen *ist*. Wir verstehen nicht genau, was wir eigentlich meinen, wenn wir sagen: „*Mentale Vorgänge sind mit neurophysiologischen Vorgängen identisch*“. Was wir nicht verstehen ist, wie dieser Satz *wahr* sein kann. In Nagles Worten:

„Es fehlt eine Konzeption davon, wie sich ein mentalistischer und ein physikalistischer Term auf ein und denselben Gegenstand beziehen können, und den gewöhnlichen Analogien mit theoretischen Identifikationen in anderen Gebieten gelingt es nicht, sie zu verschaffen. Wenn wir

⁵ Ich beziehe mich hier und im folgenden auf das Problem in der von Th. Nagel in seinem berühmten Aufsatz von 1974 (*What is Like to Be a Bat?* „*Philosophical Review*“ 1983) formulierten Form; vgl. P. Bieri, *Analytische Philosophie des Geistes*, Königstein 1981.

nämlich die Bezugnahme mentalistischer Terme auf physikalische Ereignisse anhand des gewöhnlichen Modells konstruieren, tauchen entweder wieder separate subjektive Ereignisse auf als die Wirkungen, durch die die mentalistische Bezugnahme auf physikalische Ereignisse gesichert ist, oder wir erhalten eine falsche Erklärung dafür, wie mentalistische Terme bezeichnen (zum Beispiel die eines kausalen Behavioristen“)⁶.

Ein weiteres Argument gegen die Identitätsthese ist erkenntnistheoretischer Natur und wurde von Karl Popper in Anlehnung an Haldane vorgetragen⁷. Poppers Argumentation, die davon ausgeht, daß wenn unsere mentalen Prozesse durch Gesetze der Neurophysiologie determiniert sind, sie nicht zu wahren Erkenntnissen im Sinne der Wissenschaft führen können, trifft moderne Formen des Materialismus nicht wirklich, da sie ihm einen dualistischen Wahrheitsbegriff unterschiebt. Eine der sehr starken Prämissen des Arguments ist die Hypostasierung logischer Regeln und Gültigkeitsstandards in einer von Popper postulierten *Welt 3*; die materialistische Strategie müßte hier eine nominalistische Darstellung von *Welt 3* sein.

Hier taucht nicht nur das Problem der Relation zwischen Personen und Propositionen („*propositional attitudes*“) auf, sondern auch das der *Rationalität des Physikalismus*. Ist der Physikalismus eine rationale Position, da er aus von logischen Kriterien geleiteten Überlegungen hervorgegangen ist – oder sind die Gesetze des Denkens (die für die Entstehung der gesamten modernen materialistischen Literatur verantwortlich sind) ausschließlich Gesetze der *Neurophysiologie*, wie etwa Mario Bunge⁸ meint?

Problematisch wird hier die Frage nach der zu rationalen Erkenntnisprozessen gehörenden Freiheit, auch zu *anderen* Entscheidungen bezüglich wissenschaftlicher oder philosophischer Fragestellungen kommen zu könne – und zwar aus *Gründen* und nicht im Rahmen neurophysiologischer Statistik. Ein wissenschaftlicher Materialismus, muß seine eigene Existenz auf befriedigende Weise erklären, und es ist zumindest fraglich, ob er dabei um die These herumkommt, daß *Gründe* qua Mengen analytischer Zustände *Ursachen* sein können.

Eine weitere Schwierigkeit trifft die Identitätstheorie in Gestalt der Armstrong/Lewis-Strategie⁹ der Identifikation von Mentalzuständen mit ihrer kausalen Rolle bezüglich der Verhaltensgenese. Diese frühe Variante der

⁶ Nagel, *What is Like... dt.: Wie ist es eine Fledermaus zu sein?* [In:] Bieri, *Analytische Philosophie...*, S. 270.

⁷ Vgl. Popper, Eccles, *Das Ich und sein Gehirn...*, S. 105 ff.; J. Haldane, *The Inequality of Man*, London 1932, S. 157; Metzinger, *Neuere Beiträge...*, S. 126 ff.

⁸ M. Bunge, *Das Leib-Seele Problem*, Tübingen 1984, S. 203 (im Original: *The Mind-body Problem*, Oxford 1980).

⁹ Vgl. D. Armstrong, *A Materialist Theory of the Mind*, London 1968; D. Lewis, *An Argument for the Identity Theory*, „*Journal of Philosophy*“ 1966, No. 63; Metzinger, *Neuere Beiträge...*, S. 34 ff.

Identitätstheorie erfaßt alle jene Zustände nicht, die kein beobachtbares Verhalten hervorrufen und da sie sich in den anomalen Monismus und funktionalistische Analysen hinein fortentwickelt hat, vererbt sie diesen beiden Ansätzen diese Schwierigkeit. Diese Problem liesse sich aber wahrscheinlich durch die Einführung des Begriffs „Mikro-Verhalten“ (als molekulare Hirnaktivitäten) lösen. Schwerwiegender scheint, daß die Identitätsthese hier grob kontraintuitiv wird, indem sie die Leib-Seele-Frage zum Scheinproblem¹⁰ degradiert. Wir wollen ja gerade wissen, warum das Leib-Seele-Problem auch als solches empfunden wird, woher unser Eindruck rührt, daß das Mentale *qua* Mentales unsere Handlungen verursacht. Zumindest müßte hier eine neue Terminologie entwickelt werden, die frei von allen mentalistischen Begriffen ist, dasselbe deskriptive Potential besitzt und es ermöglicht, durch cartesianische Intuitionen motivierte Einwände wie das Privatheits-Argument¹¹ oder die Inkorrigitätsthese¹² zu entkräften.

Ein fünfter Einwand speist sich aus der *Symmetrie der Identitätsrelation* und wurde erstmals von Paul Feyerabend vorgetragen¹³. Identitätssätze können in zwei Richtungen gelesen werden, und dadurch geht die zentrale materialistische Intuition des Primats des Physikalischen verloren: Es würde nicht nur für jeden mentalen Zustand eine wahre physikalistische Beschreibung existieren, sondern auch für jeden Gehirnzustand eine wahre mentalistische Beschreibung. Diese Einsicht hat zur Entwicklung des Rorty'schen *eliminativen Materialismus*¹⁴ geführt und zur Suche nach einem *asymmetrischen Relativstyp*, wie ihn die *Supervenienz-Theorie*¹⁵ konstruiert.

Wenn wir wissen wollen, was allen mentalen Phänomene eines Typs gemeinsam ist, muß der Physikalismus ein *genereller* sein. Er muß uns

¹⁰ Eine gute Analyse zur Frage des Scheinproblems bietet Feigl, *Mind-body...*

¹¹ Vgl. K. Baier, *Smart on Seansations*, „Australasian Journal of Philosophy“ 1962, No. 60; R. Rorty, *Der Spiegel der Natur*, Frankfurt 1981; ders., *Leib-Seele Identität. Privatheit und Kategorien*, [In:] Bieri, *Analytische Philosophie...* (im Original: *Mind-body Identity. Privacy and Categories*, „Review of Metaphysics“ 1965, No. 19; ders., *Inkorrigierbarkeit als Merkmal des Mentalen*, [In:] Bieri, *Analytische Philosophie...* (im Original: *Incorrigibility as the Mark of the Mental*, „Journal of Philosophy“ 1970, No. 67; Metzinger, *Neuere Beiträge...*, S. 31 ff.; 70.

¹² Rorty, *Inkorrigierbarkeit als Merkmal...*

¹³ Vgl. P. Feyerabend, *Materialism and the Mind-body Problem*, „Review of Metaphysics“ 1963, No. 17; aber auch ders., *Mentale Ereignisse und das Gehirn*, [In:] Bieri, *Analytische Philosophie...* (im Original: P. Feyerabend, *Comment: Mental Events and the Brain*, „Journal of Philosophy“ 1963, No. 60.

¹⁴ Vgl. Rorty, *Der Spiegel der Natur...*; ders., *Leib-Seele Identität...*; ders., *Inkorrigierbarkeit als Merkmal...*; aber auch P. Churchland, *Eliminative Materialism and Propositional Attitudes*, „Journal of Philosophy“ 1981, No. 78; Metzinger, *Neuere Beiträge...*, S. 65 ff., 227 ff.

¹⁵ Vgl. J. Kim, *Causality, Identity and Supervenience in the Mind-Body Problem*, „Midwest Studies“ 1974, No. 4; ders., *Supervenience and Nomological Incommensurables*, „American Philosophical Quarterly“ 1978, No. 15/2; ders., *Psychophysical Supervenience*, „Philosophical Studies“ 1982, No. 41.

Aussagen über mentale Universalien gestatten. Hilary Putnam hat mit ihrer *Multirealisierbarkeit* eine der wichtigsten Eigenschaften mentaler Phänomene erkannt und bereits in den sechziger Jahren darauf hingewiesen, daß es empirisch extrem unplausibel ist, daß ein allgemeiner Typ von mentalem Zustand wie etwa „Hunger“ bei all den verschiedenen Lebewesen, die ihn besitzen können, auf dieselbe Art und Weise realisiert ist¹⁶. Eine *partikuläre* Identitätsthese jedoch, die Identitäten nur für je einzelne Ereignisse behauptet, bleibt gemessen an den ursprünglichen Motiven und Zielen sehr unbefriedigend. Eine solche „*token-identity-thesis*“, wie sie etwa von Davidson¹⁷ in Anschluß an die Lewis'sche Argumentation für die Identitätsthese¹⁸ expliziert wurde, läßt keine generellen Aussagen über das *Wesen* mentaler Zustände mehr zu. So ist der generelle Physikalismus empirisch nicht einleuchtend, der partikuläre Physikalismus als These zu schwach, um den Anforderungen einer Philosophie der Psychologie gerecht zu werden.

Die Argumente 4, 5 und 6 scheinen nahezulegen, daß eine radikale Position, die keine ontologisch neutrale Analyse des mentalistischen Diskurses versucht, sondern schlicht die These aufstellt, daß mentalistische Terme *überhaupt nicht* referieren, eine große Menge von Schwierigkeiten elegant umgeht. Aber auch der Versuch, mentale Zustände wie „altmodische theoretische Entitäten“¹⁹ zu eliminieren, birgt Schwierigkeiten in sich. Wie Rortys Antipoden-Beispiel²⁰ zeigt, enthält sogar die Subjekt-Prädikat-Struktur unserer Grammatik noch Konnotationen einer Substanz-Attribut-Relation bzw. verleiht sie dem logischen Subjekt eine Art metaphysischer Substantialität. Schließlich würde uns auch durch die Elimination des mentalistischen Terms „Ich“ deutlich deskriptives Potential verlorengehen. Man muß deshalb auch den *eliminativen Materialismus* als radikalste Version der Identitätsthese als gescheitert ansehen.

Ich schließe nun den Abschnitt über die Identitätstheorie. Aus dem bis jetzt Gesagten geht hervor, daß sie beim heutigen Stand der Untersuchungen nicht mehr als akzeptabel gelten kann, da sie zumindest von den Argumenten 2 und 6 voll getroffen wird. Dieses Scheitern des Programms der ontologischen Neutralisierung bzw. der Elimination mentalistischer Ausdrücke hat zu dem Versuch geführt, *nicht-reduktive Formen des Materialismus* zu entwickeln, Positionen also, die den Primat des Physikalischen und das volle deskriptive Potential der psychologischen Beschreibungsebene zu wahren versuchen. In den nächsten beiden Abschnitten geht es um zwei solcher Ansätze.

¹⁶ Vgl. H. Putnam, *Mind, Language and Reality*, Vol. 2, Cambridge/MA 1975.

¹⁷ D. Davidson, *Mentale Ereignisse*, [In:] Bieri, *Analytische Philosophie ...* (im Original: *Mental Events*, [In:] D. Davidson, *Essays on Actions and Events*, New York 1980).

¹⁸ D. Lewis, *An Argument for the Identity Theory*, „*Journal of Philosophy*“ 1966, No. 63.

¹⁹ Dieser Ausdruck kommt von Bieri, *Analytische Philosophie...*, S. 45.

²⁰ Vgl. Rorty, *Der Spiegel der Natur...*

EMERGENTISTISCHER MATERIALISMUS

Ein Vorteil der These, mentale Zustände wären emergente Eigenschaften, die in Biosystemen ab einer gewissen Komplexitätsstufe auftreten, ist ihre heuristische Potenz. Indem sie mentale Zustände zu Zuständen konkreter Dinge macht, öffnet sie den weiten Bereich mentaler Phänomene endlich durch Bereitstellung des notwendigen Begriffsrahmens für die Naturwissenschaften. Der emergentistische Materialismus ist, so sein Begründer Mario Bunge, empirisch plausibel und steht in Einklang mit so wichtigen Disziplinen wie Entwicklungspsychologie, Neurophysiologie und Evolutionsbiologie (d.h. er kann die *graduelle* Entstehung von Mentalität erklären). Vor allem jedoch erhebt er den Anspruch, die qualitative Vielfalt unseres Mentallebens nicht einfach „einzuplanieren“²¹.

Trotzdem erscheinen Bunes Beiträge zur neueren philosophischen Diskussion des Leib-Seele-Problems wenig überzeugend; die Inkohärenz seiner These läßt sich an zwei Einwänden zeigen.

Erstens führt Bunge den für ihn zentralen Begriff der „Emergenz“ nur kurz und in völlig unpräziser Weise²² ein. Emergenz ist jedoch immer Emergenz *in bezug* auf eine bestimmte Menge von Eigenschaften oder nomologischen Verallgemeinerungen. Ist die emergente Eigenschaftsfamilie in Begriffen der primären Eigenschaftsmenge *definierbar*, dann ist sie auch *reduzierbar*. Und Bunge tut im Großteil seiner Arbeit nichts weiter, als akribisch einen psychologischen Begriff nach dem anderen biologistisch *umzudefinieren*. Er verwendet größte Mühe auf etwas, das seine erklärten Intentionen zunichte macht. Wie Feyerabend bereits vor zwei Jahrzehnten sagte²³, kann man der dualistisch „infizierten“ Alltagssprache ihre Inkompatibilität mit einem reduktiven Materialismus allein dadurch nehmen, daß man sie systematisch umdefiniert. Genau dies tut Bunge und macht so sein Projekt eines nicht-reduktiven Materialismus zunichte.

Mit der in aller Deutlichkeit vertretenen These, die Gesetze des Denkens seien neurophysiologische und keine logischen²⁴, setzt sich Bunge erkenntnistheoretischen Gegenargumenten wie denen Poppers aus. Wenn auch die Produkte des Denkens von der neurophysiologischen Beschreibungsebene *allein* erfaßt werden sollen, dann sind auch wissenschaftliche Theorien nicht mehr wahr oder falsch, denn „Wahrheit“ ist kein neurophysiologischer

²¹ Es ist dieses Vorgehen, daß Bunge im englischen Originaltext auch als 'bulldozing of qualitative variety' bezeichnet. Vgl. Bunge, *The Mind-body Problem...*

²² Vgl. Metzinger, *Neuere Beiträge...*, S. 237 ff.

²³ Feyerabend, *Materialism...*

²⁴ Vgl. Bunge, *Das Leib-Seele Problem...*, S. 203.

Begriff. Wenn es stimmt, daß Gründe keine Ursachen sein können, dann ist auch der emergentistische Materialismus *ausschließlich* Produkt eines hochkomplexen neurophysiologischen Prozesses und kann keinen theoretischen Wahrheitsanspruch stellen.

DIE SUPERVENIENZ-THEORIE

Der Begriff der „Supervenienz“ stammt ursprünglich aus der Ethik²⁵ und wurde in Zusammenhang mit dem Leib-Seele-Problem das erste Mal von Donald Davidson²⁶ eingesetzt. Im allgemeinen bezeichnet er eine asymmetrische Relation zwischen zwei Familien von Eigenschaften, wobei in einem recht starken Sinne die Determiniertheit der einen Eigenschaftsmenge durch die andere angenommen wird. In Bezug auf das psychophysische Problem heißt Supervenienz:

a) sobald die physischen Eigenschaften eines Systems feststehen, sind dadurch auch alle seine psychologischen Eigenschaften fixiert;

b) es ist unmöglich, daß sich zwei Systeme in ihren psychischen Eigenschaften unterscheiden, ohne sich gleichzeitig in ihren physischen Eigenschaften zu unterscheiden.

Es gibt verschiedene Varianten dieser These²⁷, die ich an dieser Stelle aber nicht näher benennen werde. Von den mentalen Zuständen eines Organismus wird nun ausgesagt, daß sie in obigem Sinne *supervenient* bezüglich der Menge gegenwärtiger physischer Eigenschaften sind. Die Basismenge (physischer oder neurophysiologischer Eigenschaften) determiniert so die superveniente Menge (mentaler Eigenschaften) ohne bikonditionale Gesetze, die die „Kluft“ zwischen beiden Mengen überbrücken und dabei strenge, korrelative Zuordnungen einzelner Elemente etablieren. Die Supervenienz-Theorie ist philosophische Explikation zweier Intuitionen. Angenommen, wir könnten einen perfekten Doppelgänger eines Menschen schaffen, indem wir seinen Körper Molekül für

²⁵ Hier ist es die These, daß sobald die deskriptive Eigenschaften einer Handlung feststehen, auch ihre moralischen Eigenschaften fixiert sind; und daß außerdem sich zwei Handlungen nicht in ihren moralischen Eigenschaften unterscheiden können, ohne sich gleichzeitig in Hinblick auf ihre deskriptive Eigenschaften zu unterscheiden. Vgl. R. Hare, *The Language of Morals*, London 1952; G. Moore, *Philosophical Studies*, London 1922. In der Ästhetik finden sich verwandte Relationstypen: Es gibt keine hinreichenden nicht-ästhetischen Bedingungen für die Zuschreibung ästhetischer Eigenschaften. Vgl. F. Sibley, *Aesthetic Concepts*, „Philosophical Review“ 1959, No. 68.

²⁶ Vgl. Davidson, *Mentale Ereignisse...*

²⁷ Z.B. Haugelands Konzeption einer 'schwachen' oder 'globalen' Supervenienz. Vgl. J. Haugeland, *Weak Supervenience*, „American Philosophical Quarterly“ 1982, No. 19/1.

Molekül rekonstruieren – etwa so wie es uns Captain Kirk und seine Mannen in der beliebten Fernsehserie „Raumschiff Enterprise“ vorführen, wenn sie ihren Körper an einen weit entfernten Ort *beamen*. Wenn die Supervenienz-These wahr ist, würden Original und Doppelgänger auch in allen psychologischen Hinsichten identisch miteinander sein.

Ein zweite Intuition ist die „Mikro-Makro-Intuition“: Betrachten wir physikalische Objekte mittlerer Größenordnung (z.B. einen Tisch), so nehmen wir an ihnen gewisse makroskopische Eigenschaften wie Solidität und Farbigkeit wahr. Diese Eigenschaften sind nicht durch nomologische Korrelationen von universeller Gültigkeit mit den *mikroskopischen* Eigenschaften, die ihnen zugrundeliegen, verknüpft. Da es keine bikonditionalen Verknüpfungen gibt, die *universelle* Gültigkeit besitzen (die Festigkeit und Farbe eines Tisches können aus verschiedenen Materialien resultieren), sind Makro-Eigenschaften nicht auf Mikro-Eigenschaften reduzierbar oder in Begriffe der mikroskopischen Ebene übersetzbar: Makro-Eigenschaften erfordern eine eigene Beschreibungsebene. Trotzdem glauben wir nicht, daß Tische in irgendeinem metaphysischen Sinne unabhängige Entitäten sind, daß sie ontologisch auf einer anderen Ebene liegen als Moleküle oder daß Farbe und Festigkeit eines Tisches noch nicht determiniert sind, sobald seine Mikro-Eigenschaften feststehen.

Wenn psychologische und neurophysiologische Prädikate so Bestandteile zweier inkompatibler Deskriptionsebenen sind, dann kann – ähnlich wie die Gestalt einer Welle nicht mehr *kausal* durch die Konfiguration der Wassermoleküle, aus denen sie besteht, erklärt werden kann – die Determination mentaler Eigenschaften auch nicht mehr als *Verursachung von unten* erklärt werden, da weder homonome noch heteronome Verallgemeinerungen („bridge-laws“) die beiden Beschreibungsebenen verbinden können. Hier bietet sich ein neuer Typ von Erklärungsstrategie an: die „Instantiierungs-Erklärung“. Wenn wir die Eigenschaft E eines Systems S erklären wollen, stellen wir nun die Frage: Was bedeutet es für S, E zu *instantiieren*? Diese Frage beantworten wir, indem wir die Eigenschaften der Komponenten von S analysieren, sowie *die Art und Weise ihrer strukturellen Organisation* angeben. Neu ist nun, daß es kein Ursache-Wirkung-Paar mehr gibt, wie wir es aus kausalen Erklärungsmodellen kennen („Was waren die Bedingungen für die *Zustandsänderung* von S?“). Neu ist auch die Frage, ob diese Strategie wirklich mehr *erklärt*, oder ob sie nur eine semantische Reduktion vornimmt.

Instantiierungs-Erklärungen müssen sich aus Gesetzen herleiten, die die Eigenschaften der Komponenten genau kennzeichnen. Müssen diese Eigenschaften selbst wieder durch eine darunter liegende nomologische Ebene gekennzeichnet werden, entsteht eine *Supervenienz-Hierarchie* bei der die *Basis-Ebene* schließlich aus einer Eigenschaftsmenge besteht, welche selbst für eine Instantiierungs-Erklärung nicht mehr geeignet ist oder sie nicht mehr

benötigt (z.B. Elektromagnetismus, Gravitation etc.). Supervenienz als Relation bietet sich also auch für differenzierte Schichtentheorien der Wirklichkeit, gerade auch der psychologischen, an. Dies ist einer ihrer Vorteile; weiterhin ist sie auf den ersten Blick *nicht-reduktiv* und kann auch die Multirealisierbarkeit mentaler Zustände durch verschiedene Basismengen plausibel machen.

Jedoch darf, wenn eine Instantiierungs-Theorie zweier Eigenschaftsmengen wirklich *informativ* sein soll, in der Analyse der Basismenge nicht auf Eigenschaften der supervenienten Menge rekuriert werden. Soll eine Gesamterklärung eines Systems von der Physik bis hinauf zur Psychologie geleistet werden, ist die Basismenge also von besonderer Bedeutung. Für den Übergang von der Neurophysiologie zu Psychologie allein gilt dies jedoch nicht.

Schließlich erbt die Supervenienz-Theorie als „Epiphänomenalismus ohne Kausalität“ auch erkenntnistheoretische Schwierigkeiten ihrer Vorgänger. Was manche Systeme zu *rationalen Agenten* macht, müßte nun aufgrund mit neurophysiologischen Daten präzisiert Instantiierungs-Erklärungen entschieden werden. Die kausale Determiniertheit von Ereignissen der untersten Basismenge einmal gesetzt, folgt daraus auch die Determiniertheit der Geschichte der Rationalität und der Geschichte der Wissenschaft.

Neben der Gefahr einer historizistischen Selbstaufhebung gibt es auch noch das Problem der semantischen Umwandlung des mentalistischen Idioms bzw. der extremen Konstraintuivität der Supervenienz-These. Unser ganzes Selbstverständnis beruht auf der Vorstellung, daß ein mentales Ereignis, wie etwa eine Entscheidung, durch ein anderes mentales Ereignis, wie etwa eine Einsicht, qua mentales Ereignis verursacht wird. Die Supervenienz-Theorie aber behauptet, daß unsere Entscheidung supervenient in bezug auf unseren Hirnzustand ist, vollkommen durch ihn determiniert – sie *könnte* in diesem gegenwärtigen Augenblick keine andere sein. Sie muß zumindest noch eine Zusatzerklärung liefern, die befriedigend erklärt, wie und warum jene Illusion des Mentalen als durch Mentales q u a Mentales hervorgerufen in uns entsteht.

INTERAKTIONISTISCHER DUALISMUS

Prominenteste Vertreter der Theorie einer Wechselwirkung zwischen zwei ontologisch distinkten Bereichen im Menschen sind Karl R. Popper und Sir John Eccles²⁸. Für Popper ist der Mensch Bürger zweier Welten mit Einsicht in eine dritte. *Welt 1* ist die Welt physikalischer zu beschreibender Prozesse, in sie gehören Bäume und Steine, aber auch Hochhäuser und menschliche Gehirne. *Welt 1* ist *kausal nicht abgeschlossen*, sie steht Einflüssen aus *Welt 2*, der Welt

²⁸ Vgl. Popper, Eccles, *Das Ich und sein Gehirn...*

psychischer Prozesse und Ereignisse offen. *Welt 2* ihrerseits partizipiert an einer überindividuellen nicht-physikalischen *Welt 3*. *Welt 3* ist die Welt der Erzeugnisse des menschlichen Geistes; zu ihr gehören z.B. Philosophien, Mythen, Märchen oder unentdeckte mathematische Probleme.

Die Drei-Welten-Doktrin ist in hohem Masse kompatibel mit unseren alltäglichen Sprachgewohnheiten und entspricht weitgehend unseren auch durch diese Gewohnheiten konditionierten vorphilosophischen Intuitionen. Außerdem ermöglicht sie als Hierarchie von Erklärungsmodellen eine gute Beschreibung evolutionärer Zusammenhänge. Es ist jedoch fraglich, ob dieser Vorteil mit der Ontologisierung der verschiedenen deskriptiven Ebenen zusammenhängt. Es sind fünf Gründe, die den dualistischen Interaktionismus in der Prägung von Popper und Eccles zu einer schwer zu haltenden Position machen.

Erstens wird die weitgehende Autonomie des Mentalen als zentrale Prämisse immer fraglicher, der Trend der empirischen Forschung ist gegenläufig. Die Determiniertheit mentaler durch neurophysiologische Eigenschaften wird immer augenfälliger. Der traditionelle *egologische Ansatz* in der Bewußtseinsphilosophie, der von der Unteilbarkeit des Ich ausgeht und durch den sich auch Popper leiten läßt, ist angesichts der Untersuchungen an kommissurotomierten Patienten obsolet geworden²⁹. Die Phänomene die auftreten, wenn ein informationsverarbeitendes System wie das menschliche Gehirn durch Maßnahmen auf der „Hardware-Ebene“ in zwei Subsysteme gespalten wird, lassen sich nur systemtheoretisch und nicht wieder egologisch – etwa in Form einer „Doppel-Ich-Hypothese“ – interpretieren.

Die zentrale Aufgabe jedes Interaktionismus ist es, eine über Metaphern hinausgehende detaillierte Erklärung des Modus der Interaktion anzubieten. Eine Antwort auf die Frage, was transphysikalische oder abwärtsgerichtete Kausalität genau heißen könnte, gibt Popper nicht und auch Eccles begnügt sich hier mit mechanistischer Metaphorik. Außerdem wird Descartes' Problem noch ausgeweitet: Erklärt werden muß nun die Interaktion zwischen Objekten in *Welt 1* und *2*, zwischen Objekten in *Welt 2* und *3* („Erkenntnis“) und die Interaktion von *Welt 1* mit *Welt 3* via *Welt 2*. So ergeben sich eine ganze Reihe neuer Fragen, nicht nur bezüglich der drei Kausalitätstypen, sondern auch bezüglich der Räumlichkeit und Zeitlichkeit der drei hypothetischen Welten.

²⁹ Vgl. J. Bogen, *The Other Side of the Brain I. Dysgraph and Dyscopia Following Cerebral Commissurotomy*, „Bulletin of the Los Angeles Neurological Societies“ 1969, No. 34 und ders., *The Other Side of the Brain II. An Appositional Mind*, „Bulletin of the Los Angeles Neurological Societies“ 1969, No. 34; Metzinger, *Neuere Beiträge...*, Kap. 4; T. Nagel, *Brain-bisection and the Unity of Consciousness*, [In:] J. Glover, *Philosophy of Mind*, Oxford 1976; Popper, Eccles, *Das Ich und sein Gehirn...*, S. 379 ff.; R. Sperry, *Lateral Specialisation in the Surgically Separated Hemispheres*, [In:] *The Neurosciences: Third Study Program*, Hrsg. F.O. Schmitt, F.G. Warden, MIT Press, Cambridge/MA 1973.

Drittens kann so auch die immer wieder behauptete kausale Effektivität des Mentalen nicht mehr begründet werden. Viertens ist der Interaktionismus (der immerhin die Energieerhaltungssätze der Physik verletzt) keine testbare wissenschaftliche Hypothese im Sinne des kritischen Rationalismus: Er kann seine Falsifikationsbedingungen nicht angeben. Die fünfte Schwierigkeit ist eher technischer Natur. Wie Daniel Dennett bemerkt hat³⁰, steht der Beitrag von Popper und Eccles fast völlig isoliert in der neueren Diskussion, er rezipiert kaum eine der relevanten Theorien. Das führt dazu, daß Popper häufig uninteressante Formen des Monismus attackiert und philosophische Strohmänner angreift – Positionen, die so heute längst nicht mehr vertreten werden. Dies ist bedauerlich, denn in einer Situation, in der der Dualismus auf philosophischer Ebene allseits so stark unter Beschuß steht, würde man sich allein aus heuristischen Gründen – besonders von zwei so prominenten Autoren – vehemente Gegenangriffe von Seiten des Dualismus wünschen.

FUNKTIONALISMUS

Während alle bis jetzt genannten Positionen nur Versuche darstellen, für klassische Leib-Seele-Theorien neue Argumentationsstrategien zu entwickeln, speist sich der Funktionalismus aus einer neuen Intuition: Mentales und Physikalisches verhalten sich zueinander wie *software* und *hardware*. Die heuristische Potenz der Computer-Metapher hat sich in den letzten zwei Jahrzehnten in der Philosophie des Geistes und der Kognitionspsychologie als enorm erwiesen. Die These, mentale Zustände seien die Input und Output eines Systems mit einander verknüpfenden *funktionalen* Zustände desselben, welche in Form physikalischer Strukturen *realisiert* sind, entgeht dem Vorwurf von Kategorienfehlern, allen Einwänden aus dem Begriff der Identität sowie essentialistischen Argumenten. Als *nicht-ontologischer Dualismus* kommt er unserer cartesianischen Konditionierung ein Stück weit entgegen, obwohl er durch die Vereinbarkeit mit dem partikularen Physikalismus in der Lage ist, die Identität zwischen einzelnen mentalen/funktionalen *tokens* und einzelnen neurophysiologischen *tokens* beim Menschen zu postulieren und so zum *funktionalen Materialismus* zu werden. Als Theorie ist er empirisch nicht unplausibel, erklärt die starke Interdependenz und die extraorganismische Relationalität mentaler Zustände und birgt schließlich auch die Hoffnung auf eine *universelle* (d.h. spezies- und „hardware-unabhängige“) *Psychologie* in sich.

³⁰ D. Dennett, *Review: The Self and its Brain – An Argument for Interactionism*, „Journal of Philosophy“ 1979, No. 76/2 einen guten Überblick über Dennetts ausgezeichnete Beiträge zu der hier skizzierten Debatte gibt D. Dennett, *The Intentional Stance*, Cambridge/MA 1987.

Unter der schwer übersehbaren Vielzahl der in letzter Zeit entstandenen Varianten lassen sich starke und schwache Funktionalismus-Thesen unterscheiden. Letztere sehen in der funktionalen Analyse mentaler Phänomene ein Instrument zur Konstruktion und Präzisierung *empirischer* Hypothesen. Der *semantische Funktionalismus* dagegen enthält eine stärkere These über die *Bedeutung* mentalistischer Terme: Er nimmt Funktionalität und funktionale Analysierbarkeit als Wesensmerkmal *aller* mentalen (besonders kognitiver) Zustände an. Ihmzufolge wäre es möglich, Computer so zu programmieren, daß sie Probleme auf dieselbe Art und Weise verstehen und lösen wie Menschen. Solche Maschinen würden Problemlösungsverhalten nicht nur simulieren, sie besäßen *selbst* psychologische Eigenschaften und vor allem würden sie diese auch *erklären*: Die funktionale Analyse ihrer internen Zustände würde die richtige *psychologische Theorie* der betreffenden kognitiven Leistungen für alle möglichen Realisierungen liefern.

Trotz der andauernden Popularität der funktionalistischen Strategie muß sie wegen einer Reihe von Schwierigkeiten bereits als gescheitertes Forschungsprogramm gelten. Da heutige Computer in einer Sprache programmiert werden, können all jene Zustände, die auf linguistischer Ebene nicht adäquat repräsentierbar und in keiner Sprache formalisierbar sind prinzipiell von ihnen nicht realisiert werden. Putnam hat bereits sehr früh darauf hingewiesen³¹, daß die mentalen oder logischen Zustände – in deren Begrifflichkeit eine Turing-Maschine ohne jeden Rekurs auf ihre physikalischen Struktureigenschaften als funktional organisierter, endlicher Automat vollständig beschrieben werden kann – eng mit *Verbalisierung* verknüpft sind. Daß es auf einer Turing-Maschinentafel nicht möglich ist, den *qualitativen* und den *semantischen*³² Gehalt vieler mentaler Zustände adäquat zu repräsentieren, macht den starken Funktionalismus empirisch wenig plausibel. So wird auch die Auskunft des semantischen Funktionalismus über mentale Universalien falsch: Schon allein für eine Philosophie der Humanpsychologie kann *Funktionalität* nicht das dringend gesuchte Wesensmerkmal sein, das allen Mentalphänomenen zukommt, sie essentiell kennzeichnet und sie zu mentalen Phänomenen *macht*.

Weiterhin gibt es eine Reihe von Schwierigkeiten bei der Präzisierung von Zuschreibungskriterien für mentale Zustände. Es scheint, daß der semantische Funktionalismus in manchen Fällen zu *liberal* bei der Zuschreibung psycho-

³¹ Vgl. H. Putnam, *Minds and Machines*, [In:] Hook, *Dimensions of Mind...*

³² Zur Erläuterung der Bedeutung von Repräsentationen gibt es heute wesentlich zwei Strategien: über funktionale Rollen und über kausale Entstehung. Vgl. z.B. Dennett, *The Intentional Stance...*, bes. Kap. 8 u. 9; J. Fodor, *Psychosemantics*, Cambridge/MA 1987; J. Searle, *Geist, Hirn und Wissenschaft*, Frankfurt a. M. 1986 (im Original: *Minds, Brain and Science*, Cambridge/MA 1985); R. Millikan, *Language, Thought and Other Biological Categories*, Cambridge/MA 1984.

logischer Eigenschaften ist. Das *absent-qualia-argument*³³ besagt, daß es nomologisch möglich ist, daß zwei psychologische Zustände funktional identisch sind, einer von ihnen jedoch keinerlei qualitativen Gehalt besitzt: Etwas könnte eine funktional isomorphe Simulation von uns selbst sein, ohne die phänomenalen Erlebnisqualitäten zu kennen, die wir wahrnehmen. Ebenfalls problematisch sind Systeme, deren einzelne Bausteine selbst wieder einer funktionalistische Beschreibung besitzen³⁴. William G. Lycan hat durch die Entwicklung teleologischer Zusatzkriterien eine neue Form des Funktionalismus konstruiert, die er *homunkularen Funktionalismus* oder *Homunktionalismus* nennt³⁵. Diese Variante ist als ausdrückliche Reaktion auf das oben angesprochene Problem zu sehen³⁶.

Der empirische Funktionalismus als *humanpsychologische* Theorie dagegen könnte leicht zu *chauvinistisch* sein. Es könnte funktional zu uns äquivalente Marsmenschen mit wahren psychologischen Beschreibungen geben, deren Verhaltensstrukturen sich stark von uns unterscheiden. Die Kennzeichnung mentaler Zustände durch physikalische Beschreibungen ihres jeweiligen Inputs und Outputs führt im besten Fall zu endlos langen Ketten von Disjunktionen; sie sagt uns nichts über den gemeinsamen Nenner aller so gekennzeichneten realen Zustände – eines der philosophischen Kernprobleme des Behaviorismus.

Ned Block und Jerry Fodor haben außerdem darauf hingewiesen³⁷, sich manche mentalen Phänomene wie Dispositionen oder propositionale Einstellungen nur einer *abstrakten Eigenschaft einer Maschinentafel* korrespondieren und nicht, wie aktuell auftretende Mentalzustände, einem bestimmten *Maschinentafelzustand*³⁸. Weiterhin können am Modell der Turing-Maschine orientierte psychologische Beschreibungen nur in Form von *Listen* gegeben werden, im menschlichen Bewußtsein laufen aber eine große Menge von

³³ N. Block, *Troubles with Funktionalism* (nachgedruckt), [In:] ders., *Readings in Philosophy of Psychology*, Vol. 1, Cambridge/MA 1980.

³⁴ Vgl. die 'homunculi-headed-robots' und das 'Chinese-nation' – Beispiel in N. Block, *Troubles with Funktionalism*. Auch die ursprüngliche Fassung des Searle'schen Chinese-room-Arguments bezieht sich auf ein solches System. Vgl. Searle, *Minds, Brains and Programs*, „The Behavioral and Brain Sciences“ 1980, No. 3.

³⁵ Vgl. W. Lycan, *Consciousness*, Kap. 3,4,5, Cambridge/MA–London 1987.

³⁶ Lycan betont, daß es eine Kontinuität hierarchisch geordneter Ebenen in der Natur gibt und daß man den Funktionalismus nicht zu einem Zwei-Ebenen-Funktionalismus verkürzen darf. So werden in einigen neueren Beiträgen mehrere gestufte Ebenen der Informationsverarbeitung innerhalb des Mentalen angenommen. Um mit Ray Jackendoff zu sprechen: Es gibt eine Entwicklung vom 'Leib-Seele-Problem' hin zum 'Seele-Seele-Problem'. Vgl. R. Jackendoff, *Consciousness and the Computational Mind*, Cambridge/MA–London 1987.

³⁷ N. Block, J. Fodor, *What Psychological States Are Not*, [In:] Block, *Readings in Philosophy...*

³⁸ *Ebd.*, S. 247 ff.

Prozessen *parallel* ab³⁹. Das kann bedeuten, daß die beste funktionalistische Beschreibung eines Organismus von einer *Menge* von miteinander vernetzten probabilistischen Automaten ausgehen muß. Philosophisch stellt diese Erweiterung kein wesentliches Problem dar, in Hinblick auf empirische Fragestellungen etwa der Hirnforschung erscheint sie sogar als besonders interessant. Block und Fodor haben schließlich darauf hingewiesen, daß sich strukturelle Relationen zwischen mentalen Zuständen (wie etwa die partielle Konstitution einer propositionalen Einstellung durch ein andere) in Form von Listen nicht ausgedrückt werden können. Um die internen Relationen zwischen funktionalen Äquivalenten von „glauben, daß *X* wahr ist“ und „glauben, daß *X* und *Y* wahr ist“ zu erfassen, reichen Listen allein nicht aus⁴⁰.

Während diese Einwände nur die „Maschinenversion“ des Funktionalismus treffen, muß auch der allgemeine Funktionalismus als generelle Theorie des Geistes als gescheitert gelten, da eine befriedigende funktionale Analyse von Qualia derzeit nicht in Sicht ist.

SCHLUSS

Es gibt zur Zeit keine Vorschläge zur Lösung des Leib-Seele-Problems, die akzeptiert werden können. Jede der im Verlauf der neueren Diskussion entwickelten Positionen hat fundamentale Schwierigkeiten, keine von ihnen kann „im Sinne eines historisch relativen Optimums, als die ‘vorläufige Lösung’ betrachtet“⁴¹ werden. Aus dieser allgemeinen Situation den Schluß abzuleiten, die philosophische Diskussion des Leib-Seele-Problems seit dem Zweiten Weltkrieg hätte nichts erbracht und sei insgesamt sinnlos gewesen, wäre jedoch verfehlt.

Sie hat zu wesentlichen Klärungen begrifflicher Natur geführt und sowohl Fragestellungen wie Lösungsversuche stark präzisiert. Die Konturen der Leib-Seele-Frage sind deutlich geworden. Verbindungen zu anderen Forschungsbereichen, wie etwa zur kognitiven Psychologie, zur KI-Forschung, zur Hirn- und Neurophysiologie, zur Wissenschaftstheorie der Psychologie, zur Linguistik usw. sind entdeckt und für eine wechselseitige Befruchtung nutzbar gemacht worden. All das hat dazu geführt, daß das Niveau der Analysen ständig gestiegen ist. Der Fortschritt gegenüber früheren Stationen der Problemgeschichte ist deutlich erkennbar.

³⁹ *Ebd.*

⁴⁰ *Ebd.*

⁴¹ D. Birnbacher, *Gibt es für das Leib-Seele-Problem eine Lösung?* Akten des Neunten Internationalen Wittgenstein-Symposium 1984, Wien 1985, S. 215.

Die zwei wichtigsten Entdeckungen sind die *nomologische Inkommensurabilität*⁴² des Mentalen und Physikalischen sowie die Frage nach der *Theorienneutralität innerer Gegebenheiten*. Wenn es richtig ist, daß wir prinzipiell keine detaillierten Kausalketten zwischen mentalen und neurophysiologischen Prozessen beschreiben können (da wir von unterschiedlichen logischen Subjekten prädizieren), dann müssen wir nach *neuen Relationstypen* suchen, mit denen eine theoretische Erklärung der ein- oder gegenseitigen Determination von Prozessen geleistet werden kann. Die zweite Hauptfrage scheint zu sein, ob es eine Erkenntnis unserer eigenen mentalen Eigenschaften *außerhalb von Sätzen gibt: Was ist der epistemische Gehalt reiner Introspektion? Ist die von Nagel so deutlich hervorgehobene Subjektivität* unserer Erlebnisse bereits „theorieinfiziert“ oder ist sie Resultat prädeskriptiver, vorsprachlicher Erkenntnis einer Eigenschaft, die unseren mentalen Zuständen *essentiell* zukommt? Diesen Problemkomplex so deutlich herausgestellt zu haben ist ein Verdienst Rortys. Wie überall, so scheint auch bei der Frage, ob Inkorrigerbarkeit *Infällibilität im Sinne eines Wahrheitsanspruchs* impliziert, das Hauptproblem in der Nichteinbringung von Intuitionen in Argumente zu liegen.

Die moderne Diskussion hat nicht nur gezeigt, welche Ansätze das Problem *nicht* lösen, sie erlaubt auch die Aufstellung eines Minimalkatalogs von acht materialen Kriterien, die jeder zukünftige Beitrag zur Lösung des Leib-Seele-Problems erfüllen muß.

1. *Empirische Plausibilität*. Eine Theorie zur Relation von Psychischem und Physischem darf nicht mit Fakten kollidieren, sie muß im Gegenteil so viele Fakten wie möglich erklärend integrieren. Ihre Hypothesen müssen prinzipiell testbar sein.

2. Sie muß die empirisch plausible These der *multiplen Realisationsmöglichkeiten* mentaler Zustände berücksichtigen. Physikalische Eigenschaften informationsverarbeitender Systeme können keine Mentalitätskriterien sein.

3. Angesichts der *nomologischen Inkommensurabilität* psychologischer und neurophysiologischer Beschreibungen muß sie (insofern ihre These keine *parallelistische* ist) einen *neuen Relationstyp* anbieten, der an die Stelle kausaler Determination tritt.

4. Sie muß ein Antwort auf die Frage nach *theorienneutralen inneren Gegebenheiten* geben. Die erkenntnistheoretische Status introspektiver Intuitionen muß analysiert und bewertet werden.

5. Die vorgeschlagene Theorie muß das Auftreten *phänomenaler Qualitäten* sowie die *Subjektivität des Mentalen* deskriptiv und explanatorisch erfassen können.

⁴² Vgl. Davidson, *Mentale Ereignisse...*

6. Sie muß Auskunft über mentale Universalien geben: Was sie zu solchen macht und was allen mentalen Phänomene einer Klasse gemeinsam ist. Zusammen mit (2) führt dies zur Beantwortung der Frage nach einer universellen Psychologie.

7. Eine wissenschaftliche Theorie der Relation zwischen Psychischem und Physischem ist immer auch eine Theorie über sich selbst: Sie muß verständlich machen, was es bedeutet, daß manche Systeme mit wahren mentalistischen Beschreibungen rationale Agenten sind und so Wissenschaft überhaupt erst ermöglichen.

8. Wie Daniel Dennett gesagt hat, wird jede Theorie, die einen Fortschritt macht, *anfänglich* kontraintuitiv sein⁴³. Ein plausibler Lösungsvorschlag zum Leib-Seele-Problem sollte jedoch nicht unnötig kontraintuitiv sein und auch erklären, warum wir das Problem als solches *empfinden*. Ist er gezwungen, die fundamentale Falschheit introspektiver Erkenntnisse anzunehmen, so sollte er das Zustandekommen illusionärer Intuitionen so detailliert wie möglich erklären. Es gibt keinen Grund zu der Annahme, daß eine Theorie des Geistes, die obige Kriterien erfüllt, nicht entwickelt werden könnte.

Justus Liebig Universität
Gießen

Thomas Metzinger

KRYTERIA, JAKIE SPELNIAC MUSI TEORIA MAJĄCA ROZWIĄZAĆ PROBLEM RELACJI POMIĘDZY DUSZĄ A CIAŁEM

Po załamaniu się analitycznego behawioryzmu proveniencji Ryle'owskiej rozgorzały dyskusje, w których poszukiwano nowego paradygmatu, który pozwoliłby rozwiązać problem relacji pomiędzy duszą a ciałem, czyli relacji pomiędzy wymiarem mentalnym a wymiarem neurofizjologicznym. W artykule krytycznej analizie poddane zostają najważniejsze współczesne teorie zajmujące się tym problemem: teoria tożsamości, emergentystyczny materializm, teoria superweniencji, dualistyczny interakcjonizm oraz funkcjonalizm. Przeprowadzona analiza prowadzi autora do wniosku, że żadna z teorii nie wyjaśnia rozważanego problemu w sposób zadowalający. Można jednak na podstawie analizowanych teorii sformułować katalog podstawowych kryteriów, jakie spełniać będzie musiała przyszła teoria mająca wyjaśnić relacje pomiędzy zjawiskami fizycznymi a psychicznymi.

⁴³ Vgl. Dennett, *The Intentional Stance...*