

schiedlichen Grundauffassungen und methodischen Idealen geprägt worden ist. Ein wichtiger Grundzug analytischer Philosophie ist die besondere Beachtung sprachlicher und logischer Aspekte philosophischer Probleme. Diese Wendung zum Sprachlichen, der sogenannte *linguistic turn* (↑Sprachphilosophie, analytische), war durch die Arbeiten des Logikers G. Frege (1848–1924) ermöglicht worden, die insbesondere auf Russell, Wittgenstein, Carnap und die Mitglieder des Wiener Kreises nachhaltigen Einfluß hatten. Frege hatte gezeigt, daß die alte, im wesentlichen auf Aristoteles zurückgehende Logik nur ein winziger Ausschnitt eines riesigen Gebiets ist. Bei einigen von denen, die Freges Resultate zur Kenntnis nahmen, machte sich die Hoffnung breit, es ließen sich mit den neuen, reichhaltigeren logischen Hilfsmitteln viele der traditionellen philosophischen Probleme mit formaler Strenge oder zumindest auf einem neuen Niveau argumentativer Durchsichtigkeit und intersubjektiver Verbindlichkeit behandeln.

Was analytische Philosophie sei, läßt sich nicht bündig charakterisieren. Ohnehin würde man besser sagen: „sogenannte analytische Philosophie“, denn das Etikett „analytisch“ paßt inzwischen nur noch schlecht. Es stammt aus einer Zeit – die zu Anfang dieses Jahrhunderts begann und in den fünfziger Jahren zu Ende ging –, als in gewissen Kreisen die Auffassung herrschte, die eigentliche philosophische Arbeit bestehe darin, die Bedeutung gewisser wichtiger Begriffe durch eine Analyse zu klären und gegebenenfalls zu präzisieren. Eine derartige Konzeption von Philosophie wurde (zu gewissen Zeiten) z.B. von Bertrand Russell, G. E. Moore, Ludwig Wittgenstein, Rudolf Carnap, A. J. Ayer, Gilbert Ryle und vielen logischen Empiristen vertreten; bei genauerer Betrachtung stellt sich allerdings heraus, daß verschiedene Philosophen sehr unterschiedliche Ideen davon hatten, was eine Begriffs- oder Bedeutungsanalyse eigentlich ist. Besonders kraß ist der Gegensatz zwischen den sogenannten „idealsprachlichen“ Konzeptionen, wie z.B. der Carnaps, die auf eine Analyse wissenschaftlicher Begrifflichkeit mit Hilfe formallogischer Konstruktssprachen abzielen, und so-

### **Philosophie des Geistes, analytische** (analytical philosophy of mind)

#### I Analytische Philosophie

Analytische Philosophie ist eine philosophische Richtung, die durch Denker mit sehr unter-

nannten „normalsprachlichen“ Konzeptionen, wie z. B. der Ryles, denen zufolge die in gewöhnlicher Prosa abgefaßte Analyse problemträchtiger Alltagssprachlicher Begriffe im Vordergrund steht (↑ Sprachphilosophie, analytische).

Seit den frühen fünfziger Jahren verlor die bis dato vieldiskutierte Frage nach dem angemessenen Analysetyp mehr und mehr an Interesse. Dies lag im wesentlichen daran, daß Quines große Herausforderung nicht überzeugend beantwortet wurde, die darin bestand, den Unterschied zwischen apriorischer Begriffsanalyse und empirischer Hypothese exakt anzugeben. W. V. Quine hatte 1953 in seiner kleinen Arbeit *Two Dogmas of Empiricism* die These vertreten, daß der Begriff „analytisch“ letztlich leer, ohne greifbaren Sinn sei. Der Unterschied zwischen reiner Begriffsanalyse einerseits, die Philosophen zu apriorischen Einsichten in notwendige Zusammenhänge führen könnte, und andererseits der Erörterung inhaltlicher Fragen, durch die Wissenschaftler zu aposteriorischen Einsichten in die kontingente Beschaffenheit der Welt gelangen, dieser Unterschied ist Quine zufolge selbst eine Fiktion. Die meisten sogenannten analytischen Philosophen scheinen sich dieser Auffassung inzwischen angeschlossen zu haben.

Analytische Philosophie (das „sogenannte“ möge der Leser nun selbst hinzulesen) ist also nicht durch eine bestimmte Methode charakterisiert. Auch gibt es keinen Kanon gemeinsamer philosophischer Grundannahmen. Was heutige analytische Philosophen eint, ist im wesentlichen dies: (i) eine gemeinsame Ahnengalerie, die bei Frege beginnt; (ii) die Neigung, philosophische Fragen und Lösungsvorschläge unter Einsatz logischer Hilfsmittel zu erhellen; und (iii) gemeinsame (wiewohl weitgehend implizite) Standards für die Qualität philosophischer Argumentationen. Kurz, analytische Philosophie ist eher ein gewisser philosophischer Stil mit einer bestimmten Tradition als eine inhaltlich oder methodisch geeinte Strömung.

Bis in die sechziger und siebziger Jahre hinein bildeten Fragen der philosophischen Bedeutungstheorie das Zentrum der Diskussionen. Das waren Fragen wie: Worin besteht sprachliche Be-

deutung? Ist die Bedeutung eines Satzes ein abstrakter Gegenstand? Oder eine komplexe Verhaltensdisposition? Oder etwas ganz anderes? Ist sie wesentlich intersubjektiv? Lassen sich für gewöhnliche Sprachen wie das Deutsche oder das Englische vollständige und präzise ↑ Semantiken entwickeln? Wie sähe eine semantische Theorie einer gewöhnlichen Sprache aus, was wären ihre Grundbegriffe?

## II Philosophie des Geistes

Seit etwa 25 Jahren jedoch werden Fragen der Philosophie des Geistes für nicht weniger wichtig gehalten. Dies hat verschiedene Gründe, u. a. den, daß einflußreiche sprachphilosophische Ansätze den Bedeutungsbegriff in einen sehr innigen Zusammenhang mit Begriffen brachten, die im Rahmen einer philosophischen Theorie des menschlichen Geistes zu untersuchen und zu klären wären. Ein gutes Beispiel liefert der bedeutungstheoretische Ansatz von P. Grice, demzufolge die Bedeutung eines Satzes dadurch bestimmt ist, was mit der Äußerung des Satzes gemeint wird; dies wiederum muß laut Grice mit Rückgriff auf Begriffe wie „etwas beabsichtigen“, „etwas glauben“ und dergleichen expliziert werden. Der Versuch, den ↑ Bedeutungsbegriff zu analysieren, führt zu Begriffen, die in den Arbeitsbereich Philosophie des Geistes fallen. Es gab natürlich auch Gründe ganz anderer Art für den Aufstieg der Philosophie des Geistes zu einem zentralen Themengebiet in der analytischen Philosophie, doch würde es hier zu weit führen, darauf näher einzugehen.

Wovon die Philosophie des Geistes handeln soll, ist durch das Genitivattribut angedeutet: vom **Geist**. „Geist“ bezeichnet hier keine unkörperliche Sache, sondern ist zunächst einmal nichts weiter als eine bequeme Sammelbezeichnung für beliebige Phänomene – Eigenschaften, Zustände, Vorgänge, Ereignisse –, die dem Denken, Wollen, Gestimmtsein und Empfinden eines Menschen, seiner Persönlichkeit oder seinem Charakter zuzurechnen sind. (Wenn ich hier von „Phänomen“ spreche, dann ist das ganz untechnisch gemeint und soll keineswegs beinhalten, daß es sich um etwas handeln muß, das bewußt

ist.) In der analytischen Philosophie gibt es keine unumstrittene Definition dafür, was ein geistiges Phänomen ist; es besteht aber offenbar große Einigkeit darüber, daß zwei Bereiche geistiger Phänomene besondere Aufmerksamkeit verdienen:

(i) intentionale Zustände, Ereignisse, Vorgänge (↑Einstellung, intentionale, ↑Gehalt, intentionaler);

(ii) Zustände, Ereignisse und Vorgänge mit qualitativem Charakter (↑Qualia).

Intentionale Phänomene sind dadurch charakterisiert, daß sie einen Inhalt haben, von etwas handeln, sich auf etwas beziehen. Wer glaubt, hofft oder wünscht, der glaubt, hofft oder wünscht etwas; läßt sich dieses Etwas – der intentionale Gehalt oder Inhalt – in einem Daß-Satz angeben, dann wird der entsprechende intentionale Zustand auch „propositionaler“ Zustand oder „propositionale Einstellung“ genannt. (Dabei ist der Terminus *propositional* ganz untechnisch zu verstehen: Eine ↑**Proposition** ist einfach das, was ein Daß-Satz beinhaltet; das muß selbst nichts irgendwie Sprachartiges sein, und was es genau ist, darüber gibt es sehr unterschiedliche Auffassungen; keine davon wird hier als richtig vorausgesetzt.) Nicht alle intentionalen Geistesphänomene sind propositionaler Art; wenn einer liebt, sucht, verehrt usw., dann mag es sein, daß sich nicht in einem Daß-Satz angeben läßt, worum es sich bei dem handelt, was da geliebt, gesucht oder verehrt wird.

Unter (ii) zu rechnen sind Phänomene wie: ein Schmerz, eine Farbempfindung, ein Kitzel, ein Jucken, eine Geruchs- oder Geschmackswahrnehmung. Solche geistigen Phänomene haben oftmals keinen intentionalen Gehalt, sie haben aber immer einen qualitativen Charakter: sie treten nicht auf, ohne daß einem dabei irgendwie ist. Intentionale Phänomene hingegen treten häufig ohne so etwas auf. (Fast jeder befindet sich über die meiste Zeit hin in dem intentionalen Zustand: zu wissen, wie er heißt; doch gewöhnlich ist damit kein charakteristisches Gefühl verbunden.)

Viele geistige Phänomene fallen in beide Klassen zugleich. Meine Geruchswahrnehmung in diesem Moment hat einen qualitativen Charakter

und intentionalen Gehalt zugleich: es riecht nach einer Zigarre. Es erscheint nicht immer möglich, eine Aufsplitterung solch eines Phänomens in eine intentionale und eine rein qualitative Komponente vorzunehmen.

Intentionalität und qualitativer Charakter werden in der analytischen Philosophie für die beiden Hauptmerkmale des Geistigen gehalten. Geistige Phänomene, die weder Inhalt noch qualitativen Charakter haben, haben in der Diskussion der letzten 25 Jahre keine Rolle gespielt. Daß Intentionalität und qualitativer Charakter in der Tat sehr grundlegende Aspekte des Geistigen sind, leuchtet ein, wenn man bedenkt, daß zum Thema „intentionale Geistesphänomene“ alles gehört, was mit Denken (mit mentaler Repräsentation) im weitesten Sinn zu tun hat, und zum Thema „Geistesphänomene mit qualitativem Charakter“ alles, was mit Gefühl (mit nichtintentionalem Bewußtsein) im weitesten Sinn zusammenhängt. Der Begriff der Überzeugung steht – gewissermaßen stellvertretend für die gesamte Begrifflichkeit intentionaler Geistesphänomene – im Zentrum der Intentionalitätsdiskussion, der Begriff des Schmerzes entsprechend im Zentrum der Diskussion über qualitativen Charakter.

Was sind nun die Grundprobleme einer philosophischen Theorie des Geistes? Welche Fragen hätte eine philosophische Theorie der Intentionalität bzw. des qualitativen Charakters zu beantworten? Die leitende Frage der meisten Betrachtungen innerhalb der analytischen Philosophie ist: Wie lassen sich diese Phänomene und die Eigenschaften, die wir ihnen charakteristischerweise zuzuschreiben geneigt sind, in ein naturwissenschaftliches Gesamtbild der Wirklichkeit einfügen? Zwei Aspekte dieser Frage lassen sich als Teilfragen unterscheiden. Erstens ein metaphysischer Aspekt: Was ist das Wesen dieser Phänomene (wohlgemerkt: dieser Phänomene, wie sie sich vor der theoretischen Spekulation präsentieren)? Sind sie letztlich materiell? Wenn nicht, in welcher Beziehung stehen sie zur materiellen Welt der körperlichen Gegenstände und ihrer physikalischen Eigenschaften? Zweitens ein erkenntnistheoretischer Aspekt: Ist über-

haupt eine externe, intersubjektive Erkenntnis über solche Phänomene möglich? Wie paßt die besondere Weise, in der einer Person nur ihre eigenen Geistesphänomene gegeben sind, mit einer wissenschaftlichen Vorgehensweise zusammen?

Diese beiden Aspekte hängen natürlich eng miteinander zusammen. Wird beispielsweise die Auffassung vertreten, geistige Phänomene seien letztlich schlicht und einfach gewisse zentralnervöse Zustände und Prozesse, dann fällt es schwer zu sehen, weshalb diese Phänomene besondere Probleme aufwerfen sollten. Oder wenn als Antwort auf die metaphysische Frage gesagt wird, daß es solche Phänomene gar nicht gibt, dann verschwindet jede erkenntnistheoretische Frage. Wird andererseits angenommen, daß über geistige Phänomene prinzipiell nur introspektives Wissen möglich sei, dann ist zum Beispiel die metaphysische Identifizierung von Geistesphänomenen mit Dispositionen zu beobachtbarem Verhalten ausgeschlossen. Erkenntnistheorie und Metaphysik stehen in Wechselwirkung zueinander.

### III Metaphysische Grundpositionen

In der analytischen Diskussion haben die folgenden metaphysischen Grundauffassungen eine gewisse Rolle gespielt und spielen sie zum Teil immer noch.

**1. Dualismen** (↑Leib-Seele-Problem). Als „Dualismus“ wird in diesem Zusammenhang normalerweise eine Lehre wie die von René Descartes (1596–1650) bezeichnet, wonach der Geist des Menschen eine immaterielle, nicht ausgedehnte Substanz ist, die in irgendeinem schwer begreiflichen Zusammenhang mit dem Leib dieses Menschen steht; geistige Phänomene sind Eigenschaften dieser Substanz; der Geist und seine Eigenschaften können unabhängig vom Körper – ja, unabhängig von der gesamten materiellen Welt – existieren. Diese Auffassung wurde in unterschiedlichen Varianten vertreten, z. B. in einer interaktionistischen (d. h., es wird eine kausale Interaktion von Geist und Körper angenommen) und in einer parallelistischen (Geist und Körper sind in sich geschlossene Systeme ohne Kausalkontakt).

Ein derartiger Substanzen-Dualismus ist eigentlich nur noch Gegenstand philosophischer Betrachtungen; in der systematischen Diskussion spielt er vornehmlich die Rolle eines Schreckgespensts. Es gibt allerdings andere Positionen, die manchmal mit dem Terminus „Dualismus“ etikettiert, aber dennoch diskutiert werden, so z. B. der Eigenschaftsdualismus, demzufolge es zwar keine konkreten, nichtmateriellen Gegenstände gibt, wohl aber geistige Eigenschaften, die sich nicht auf physische Eigenschaften zurückführen lassen. Zu nennen ist hier auch der Epiphänomenalismus (↑Epiphänomenalismus). Dies ist eine dualistische Auffassung, nach der geistige Phänomene bloße Epiphänomene der materiellen Welt sind: Sie haben materielle Ursachen, aber keine materiellen Wirkungen. Ein Beispiel: Eine körperliche Verletzung ruft zwar das eigentlich Geistige am Schmerz, seinen qualitativen Charakter, hervor, aber nicht dieser qualitative Charakter – sondern eine Reihe physiologischer Faktoren – bewirkt, daß der Verletzte stöhnt.

**2. Nichtmaterialistische Monismen.** Monistischen Auffassungen zufolge gibt es in der Welt nur eine Art von Substanz. Zwei Varianten des Monismus, die nur noch in Übersichtsdarstellungen erwähnt werden, sind der immaterielle Monismus von G. Berkeley (1685–1753), wonach es in der Welt nur Geist und keine Materie gibt, und der neutrale Monismus, der um die Jahrhundertwende von Denkern wie E. Mach, W. James und B. Russell vertreten wurde. Ein neutraler Monist behauptet, daß Geist und Materie sich aus derselben Substanz zusammensetzen, daß dieser zugrundeliegende Stoff selbst jedoch weder materiell noch geistig sei. Berkeleys Immaterialismus ist schwer zu glauben, der neutrale Monismus ist schwer zu verstehen (wenn man versucht, ihn als eine echte Alternative zum Dualismus und zum Materialismus zu begreifen).

**3. Materialistische Monismen.** Der Materialismus (↑Materialismus) der philosophischen Tradition besagte, daß es nur eine einzige Art von Substanzen in der Welt gibt, und zwar materielle. Analytische Materialisten ziehen es vor, die problematische Substanz/Akzidenz-Terminologie

zu vermeiden, und charakterisieren ihre Position lieber in einer Weise, die explizit auf die Physik Bezug nimmt. Zum Beispiel in folgender Weise: „Alles, was in der Welt geschieht, ist mit den grundlegenden Gesetzen der Physik (oder gar nicht) vollständig erklärbar“; „alles, was in der Welt existiert, setzt sich ganz und gar aus Bestandteilen zusammen, von denen in den grundlegenden Gesetzen der Physik die Rede ist“; oder einfach: „Jede Veränderung in der Welt hat eine physikalische Beschreibung.“ Wegen der wesentlichen Bezugnahme auf die Physik wird der analytische Materialismus häufig auch als „Physikalismus“ bezeichnet. Analytische Materialisten glauben daran, daß die Physik für jedwede empirisch wissenschaftliche Erkenntnis grundlegend ist. Es besteht unter ihnen jedoch keine Einigkeit darüber, in welchem Sinne genau die Physik grundlegend ist. Und es bleibt zumeist unbestimmt, was eigentlich „die Physik“ ist. Was wäre, wenn die meisten theoretischen Physiker plötzlich anfangen, wilde Theorien (über masselose Psychonen und dergleichen) zu entwickeln, die leicht erkennbarerweise nichts weiter wären als dualistische Theorien in neuer Aufmachung? Hierin scheint ein gewisses Dilemma zu liegen. Denn ein Physikalismus, der es den Grillen künftiger Physiker überläßt, was Physik ist, ist vor derartigen Überraschungen offenbar nicht geschützt; wer aber nicht bereit ist, den Physikern zu überlassen, was Physik ist, maßt sich Erstaunliches an – man nannte so jemanden besser einen Metaphysikalisten. Kurz, die Blankoscheck-Haltung des analytischen Materialismus gegenüber der Physik ist nicht unproblematisch.

Im Laufe der letzten dreißig Jahre haben sich eine Reihe unterschiedlicher materialistischer Auffassungen herausgebildet. Es ist instruktiv, drei von ihnen anfangs hervorzuheben: den Begriffsreduktionismus, den Eliminativismus und die These von der Globalsupervenienz. Diese drei begrenzen das Feld möglicher analytischer Materialismen.

(i) *Materialistischer Begriffsreduktionismus.* Diese Auffassung besagt, daß es für jeden Begriff, der ein geistiges Phänomen bezeichnet, einen in rein physikalischer Terminologie formulierten

Begriff gibt, der denselben Inhalt hat. Kurz gesagt: Mentalistische Begrifflichkeit ist vollkommen – ohne Inhaltsverlust – in physikalische Begrifflichkeit übersetzbar.

Der logische ↑Behaviorismus war eine einflußreiche Variante dieser Auffassung. Heutzutage scheint jedoch niemand mehr den Begriffsreduktionismus zu vertreten. Denn inzwischen hat man eingesehen, daß die Forderung der Bedeutungsgleichheit unerfüllbar ist. Außerdem ist die Forderung auch überflüssig stark, denn es sollte einem Materialisten ja damit genug sein, daß geistige Phänomene de facto physisch sind. Daß sie mit begrifflicher Notwendigkeit physisch sind, ist eine unnötige metaphysische Verschärfung.

(ii) *Materialistischer Eliminativismus.* Diese Auffassung besagt, daß jeder Begriff, der anscheinend ein geistiges Phänomen bezeichnet, in Wirklichkeit gar nichts bezeichnet. Diese Auffassung wird oft folgendermaßen begründet: Mentalistische Begriffe gehören einer bestimmten Theorie – nämlich der Alltagspsychologie – an, und diese Theorie ist strenggenommen falsch; wissenschaftlich ist sie jedenfalls nicht akzeptabel und wird über kurz oder lang durch eine neurophysiologische Theorie ersetzt werden, mit der sich bessere Vorhersagen und Erklärungen machen lassen. Die mentalistische Begrifflichkeit wird aus einer derartigen neurophysiologischen Theorie vollständig eliminiert sein. Es wird sich dann herausstellen, daß all dieses Reden von geistigen Phänomenen keine andere sachliche Grundlage hat als unsere Unkenntnis der eigentlich relevanten Faktoren.

(iii) *Die These von der Globalsupervenienz* (↑Supervenienz). Sie besagt, grob skizziert: Das geistige Geschehen hängt vom physischen Zustand der Welt insgesamt ab; damit zwei Welten unterschiedliche geistige Phänomene enthalten, müssen sie auch unterschiedliche physische Phänomene enthalten. Anders ausgedrückt: Es ist ausgeschlossen, daß zwei Welten physikalisch gesehen vollständig gleichartig sind und daß dennoch in der einen Welt ein geistiges Phänomen vorkommt, das in der andern nicht vorkommt. Diese These fängt ein, was man auch als „Minimal-Materialismus“ bezeichnen könnte.

Es ist allerdings umstritten, ob diese These nicht zu schwach ist, um selbst schon als materialistisch zu gelten. Geistige Phänomene hängen dieser These zufolge von der physischen Welt als ganzer ab; sie „supervenieren“ – wie es im Fachjargon heißt – über der Gesamtheit der Welt. Die Pointe liegt dabei in folgendem: Mit der Supervenienzthese geht keine Festlegung auf irgendeinen reduktiven Zusammenhang zwischen Geistigem und Körperlichem einher; es wird nicht behauptet, daß sich irgendein geistiges Phänomen auf eine eingrenzbar Klasse physischer Phänomene zurückführen läßt. Supervenienz-Thesen behaupten keine Reduzibilität oder Identität, sondern nur Kovarianz. Und die These von der Globalsupervenienz besagt nur dies: Die geistigen Phänomene physikalisch gleichartiger Welten kovariieren miteinander. Die damit behauptete Abhängigkeit des Geistigen vom Physischen ist allerdings sehr holistisch und somit unerfreulich schwach; denn die materielle Basis der Welt insgesamt legt im Lichte dieser These auf völlig unspezifizierte Weise die geistigen Phänomene insgesamt fest. Der psychophysische Zusammenhang im einzelnen bleibt völlig ungeklärt. Mit der These von der Globalsupervenienz ist also z. B. die abwegige Idee verträglich, daß in einer Welt, in der nur ein einziges, winziges Detail auf einem fernen Planeten anders ist als in der wirklichen Welt, ganz und gar andere geistige Phänomene vorhanden sind als in der wirklichen Welt.

Alle drei Extrempositionen haben gravierende Schwächen. Der Begriffsreduktionismus ist nachweislich zu stark, die These von der Globalsupervenienz ist unerfreulich schwach, der Eliminativismus ist flagrant kontraintuitiv. Es hat nicht an Versuchen gefehlt, innerhalb des geschilderten Rahmens andere Positionen zu entwickeln.

(iv) *Typ-Identitätstheorie*. Dies ist eine Abmilderung des Begriffsreduktionismus. Es wird nun nur noch behauptet, daß es zu jedem Begriff, unter den eine bestimmte Klasse geistiger Phänomene fällt, einen (sei's auch sehr komplexen) physikalischen Begriff mit derselben Extension gibt. Lax gesagt: Zu jeder natürlichen Art im Bereich des Geistigen gibt es auch eine de facto

deckungsgleiche natürliche Art im Bereich des Physischen; die Forderung, der mentalistische und der physikalische Begriff müßten inhaltsgleich sein, ist weggefallen.

(v) *Token-Identitätstheorie*. Dies ist wiederum eine Abschwächung der Typ-Identitätstheorie. Es wird nun nur noch behauptet, daß es zu jeder mentalistischen Beschreibung eines einzelnen, konkreten geistigen Phänomens (wie z. B. des Gedankens, der Harvey am 22. 2. 1993 um Schlag Mitternacht durchzuckte) auch eine physikalische Beschreibung gibt, die sich auf dasselbe konkrete Phänomen bezieht. Die Forderung, daß es zu jedem Begriff, unter den geistige Phänomene fallen, ein koextensives physikalisches Prädikat gibt, ist weggefallen.

(vi) *Materialistischer Instrumentalismus*. Dies läßt sich als der Versuch auffassen, den Eliminativismus abzumildern. Weder wird behauptet, daß geistige Phänomene physisch real sind, noch wird behauptet, daß es sie überhaupt nicht gibt. Es gibt sie – so sagt der materialistische Instrumentalist gerne – in dem Sinne, in dem es z. B. Breitengrade gibt. Es sei jedenfalls sehr nützlich (und in manchem wichtigen Zusammenhang vielleicht sogar unverzichtbar), mentalistische Begrifflichkeit zu Voraussagen und Erklärungen menschlichen Verhaltens zu verwenden. Der Eliminativismus ist von dieser Warte aus betrachtet insofern verfehlt, als er die Verzichtbarkeit dieser Begrifflichkeit behauptet; die Identitätstheorien sind insofern verfehlt, als sie geistigen Phänomenen dieselbe stramme Realität zuschreiben wie Phänomenen der Physik.

(vii) *Die These von der lokalen Supervenienz*. Ein Versuch, zu einer gehaltvolleren Supervenienz-These zu gelangen als (iii). Die Idee ist hier, daß die geistigen Eigenschaften einer Person (insbesondere ihre intentionalen Zustände) über den physikalischen Eigenschaften dieser Person und ihrer Umgebung supervenieren. Der materielle Bereich, der die Basis der geistigen Phänomene einer Person ausmacht, ist nun nicht mehr die ganze große Welt, sondern ist nun eingegrenzt auf den Leib und seine Umgebung. Eine naheliegende Frage lautet natürlich: Wie weit reicht diese Umgebung räumlich und zeitlich?

(viii) *Die These von der Geist/Hirn-Supervenienz.* Diese Auffassung geht insofern über (vii) hinaus, als nun das Hirn einer Person allein zur materiellen Basis aller ihrer geistigen Phänomene erklärt wird; die räumliche und zeitliche Umgebung spielt keine Rolle. In der Diskussion wurde der Gehalt dieser These oft folgendermaßen veranschaulicht: Physische Doppelgänger – d.h. Personen von exakt gleichartiger physikalischer Beschaffenheit, also auch mit physikalisch ununterscheidbaren Hirnen – haben dieselben geistigen Phänomene. (Zur Erinnerung: Wie bei allen Supervenienz-Thesen wird auch hier nicht behauptet, daß sich die supervenienten Phänomene auf Phänomene der Basis zurückführen lassen, mit ihnen identisch sind.)

#### IV Der Funktionalismus

(↑Funktionalismus). Materialistische Identitätstheorien müssen die Frage beantworten, wodurch diejenigen physischen Phänomene, die mit geistigen Phänomenen identisch sind, kennzeichnend sind, und zwar so, daß in dieser Kennzeichnung keine mentalistische Begrifflichkeit mehr verwendet wird. Ein Beispiel: Harvey hat in diesem Moment einen bestimmten Schmerz. Dieser Schmerz ist gemäß identitätstheoretischer Voraussetzung identisch mit irgendeinem Teil von Harveys momentaner physischer Befindlichkeit. Wodurch ist nun dieser physische Zustand bestimmt, der mit Harveys Schmerz identisch ist? Anders formuliert: Was macht einen physischen Zustand zu einem Schmerz oder zu einer Überzeugung oder zu einem sonstigen geistigen Phänomen?

Die einflußreichste zeitgenössische Antwort auf diese Frage hat der Funktionalismus angeboten. Der Funktionalismus läßt sich am besten verstehen, wenn man berücksichtigt, daß er aus dem logischen Behaviorismus hervorgegangen ist, der bis in die sechziger Jahre in der analytischen Philosophie weit verbreitet war.

**1. Der logische Behaviorismus.** Er besagte, daß geistige Phänomene nichts anderes seien als Dispositionen zu beobachtbarem Verhalten. Schmerzen haben hieße demzufolge: zu bestimmten Verhaltensweisen neigen, z.B. zum

Verzerren des Gesichts, Stöhnen und so weiter. Diese Auffassung verträgt sich trefflich mit dem Materialismus; denn wenn man unterstellt, daß Dispositionen zu beobachtbarem Verhalten im Organismus physisch realisiert sind, dann hat man eine Antwort auf die obige Frage: Ein physischer Zustand ist ein Schmerz, wenn er die physische Realisierung derjenigen Verhaltensdisposition ist, die für Schmerz kennzeichnend ist. (Es muß allerdings erwähnt werden, daß nicht alle logischen Behavioristen Materialisten waren; G. Ryle z. B. hielt den Materialismus für genauso grundverkehrt wie den Dualismus.)

Eine Schwierigkeit mit der behavioristischen Ausgestaltung des analytischen Materialismus lag allerdings in folgendem. Der logische Behaviorismus war ein Begriffsreduktionismus. Der Sinn eines beliebigen mentalistischen Begriffs „G“ sollte sich erschöpfend durch Manifestationsgesetze folgender Art angeben lassen: „Wenn ein beliebiger Organismus x das Phänomen G aufweist, dann gilt: x wird die Reaktion  $R_1$  aufweisen, wenn x in einer Situation vom Typ  $S_1$  ist; x wird  $R_2$  aufweisen, wenn x in  $S_2$  ist; ...; x wird  $R_n$  aufweisen, wenn x in  $S_n$  ist.“ Die für G charakteristische Verhaltensdisposition sollte sich angeben lassen, ohne daß dabei mentalistisches Vokabular verwendet würde; d.h., die Situationen und Reaktionen, von denen in den Manifestationsgesetzen die Rede ist, sollten sich letztlich rein physikalisch beschreiben lassen. Doch bei vielen intentionalen Geistesphänomen erwies sich dies als nicht durchführbar. Wenn jemand eine bestimmte Überzeugung  $\bar{U}$  hat, dann wird er nur dann eine bestimmte Reaktion  $R_i$  aufweisen, wenn er auch entsprechende Wünsche (und gegebenenfalls passende weitere Überzeugungen) hat; wenn jemand einen bestimmten Wunsch  $W$  hat, dann wird er nur dann eine gewisse Reaktion  $R_j$  aufweisen, wenn er auch entsprechende Überzeugungen hat. Es ist in diesen Fällen nicht zu sehen, wie sich die Manifestationsgesetze ohne Rückgriff auf mentalistisches Vokabular formulieren lassen könnten.

**2. Die funktionalistische Lösung.** Diese Schwierigkeit wurde im Funktionalismus dadurch überwunden, daß man nun nicht mehr da-

nach trachtete, für einzelne Geistesphänomene reduktive behavioristische Charakterisierungen zu finden, sondern eine eher holistische Auffassung geistiger Phänomene entwickelte. Geistige Phänomene sind nicht (und schon gar nicht einzeln genommen) durch ihren Zusammenhang mit physischen Reizen und Reaktionen charakterisiert, sondern durch die kausale (oder funktionale) Rolle, die sie im Verbund mit anderen geistigen Phänomenen im Gesamtsystem eines Organismus spielen. Die Frage, was einen physischen Zustand zu einem bestimmten geistigen Phänomen (z. B. einer gewissen Überzeugung) macht, konnte nun so beantwortet werden: Es ist seine charakteristische Rolle in dem gesamten Kausalgefüge des Organismus. Jedes geistige Phänomen ist dadurch gekennzeichnet, welche charakteristischen Ursachen und Wirkungen es hat, und zu diesen Ursachen und Wirkungen gehören auch solche, die selbst wiederum geistige Phänomene sind.

Dieser Auffassung zufolge lassen sich geistige Phänomene auf zweierlei Weise thematisieren: als etwas Geistiges und als etwas Materielles. Wird ein geistiges Phänomen als etwas Geistiges thematisiert, dann geht es um eine abstrakte funktionale Rolle; wird es als Bestandteil der materiellen Welt thematisiert, dann geht es um die konkrete physische Realisierung dieser Rolle in Organismen, die dieses Phänomen aufweisen. Ein attraktiver Zug dieser Auffassung liegt darin, daß nun die Möglichkeit zugelassen werden kann, daß ein und dasselbe geistige Phänomen auf ganz unterschiedliche Weisen physisch realisiert ist. Eine außerirdische Kreatur mit einer ganz andern chemischen Konstitution als der unsern könnte dennoch in derselben Weise einen Geist besitzen wie wir.

**3. Die Computeranalogie** (↑ Computtheorie des Geistes). Dem Funktionalismus zufolge ist der Geist, als Geist, wesentlich ein funktionales System. Solch ein funktionales System überführt in Abhängigkeit vom inneren Ausgangszustand, in dem es sich befindet, Inputs in Outputs. Das spezielle System Geist überführt aufgrund seiner inneren Organisation (zu der Faktoren wie Stimmungen, Ziele, Vorwissen, Erinnerungen usw.

gehören) Sinneswahrnehmungsinputs in Verhaltensoutputs. Diese funktionale Betrachtungsweise legt den Vergleich des Geistes mit einem Computer, der mit einem bestimmten Programm geladen ist, nahe. Der programmierte Computer läßt ja ebenfalls die beiden relevanten Betrachtungsweisen zu: erstens die funktionale, wonach er wesentlich durch seine Input/Output-Beziehungen und die abstrakte algorithmische Organisation charakterisiert ist; zweitens die physikalische Betrachtungsweise, wonach er ein Gerät mit bestimmten materiellen Eigenschaften (z. B. Maße, Gewicht, verarbeitete Materialien usw.) ist.

Der Funktionalismus legt also eine möglicherweise interessante Strukturanalogie zwischen Geist und Computer nahe, und ein Thema in der analytischen Philosophie des Geistes der letzten zwanzig Jahre war es, wie weit diese Analogie eigentlich reicht. Zwei besonders extreme Thesen verdienen in diesem Zusammenhang Erwähnung. Erstens die starke KI-These, wonach ein Informationsverarbeitungsgerät, dessen funktionale Organisation der eines Menschen hinreichend ähnlich ist, einen Geist hat. Zweitens die These der Computer-Psychologie (*computational psychology*), daß wesentliche geistige – insbesondere kognitive – Prozesse tatsächlich Symbolverarbeitungsabläufe (in der Manier eines klassischen Von-Neumann-Rechners) sind, in deren Verlauf neuronal realisierte Symbole gemäß gewissen Algorithmen transformiert werden. Heftig umstritten ist im Lager der Computer-Psychologen, ob es möglich oder überhaupt nötig ist, die Symbole, von denen man annimmt, daß sie im Gehirn verarbeitet werden, als Träger eines begrifflichen Inhalts zu betrachten. Insbesondere J. Fodor und Z. Pylyshyn haben die Auffassung vertreten, dies sei möglich und für die Zwecke einer wissenschaftlichen Computer-Psychologie auch erforderlich; diese Position ist die „repräsentationalistische“ Variante der These von der Existenz einer Sprache des Geistes. Vertreter einer rein syntaktischen Variante bestreiten Möglichkeit und Nutzen einer inhaltlichen Deutung der hypothetisch unterstellten Symbolverarbeitungsprozesse.



Die Diskussion um die Frage, ob und inwiefern der menschliche Geist einem Computer gleicht, hat durch das Aufkommen konnektionistischer Modelle für geistige Phänomene neue Aspekte hinzugewonnen (↑ Repräsentation im Konnektionismus). Insbesondere scheint es nun, zumindest für gewisse kognitive Leistungen (wie Mustererkennung), möglich zu sein, Computersimulationen zu entwickeln, die nach Auffassung einiger Autoren ganz ohne Repräsentation oder mit ganz anderen Repräsentationsformen operieren als Simulationen mit Hilfe klassischer Computer. Konnektionistische Systeme verheißen gegenüber klassischen Computern eine Reihe von Vorteilen: Insbesondere stellt sich ihnen – weil sie Information nicht „explizit“ repräsentieren – das sogenannte *frame*-Problem nicht; und sie haben insofern vielleicht einen gewissen Anflug von biologischer Glaubwürdigkeit, als man in ihren Abläufen eine strukturelle Ähnlichkeit zu neuronalen Prozessen erblicken kann. Der zweite Punkt wird einen echten Funktionalisten nicht beeindrucken; denn der in seinen Augen zentrale Unterschied zwischen funktionaler Rolle einerseits und ihrer physischen Realisierung andererseits wird dadurch eher verwischt oder in den Hintergrund gedrängt.

In der Diskussion um konnektionistische und klassische Computer-Simulationen kognitiver Fähigkeiten und anderer Geistesphänomene werden oft unterschiedliche Fragestellungen nicht voneinander differenziert. Einerseits gibt es die begriffliche, philosophische Frage: Was sind geistige Phänomene in einer rein materiellen Welt; wie ist so etwas möglich? Und die funktionalistische Antwort darauf besagt: Es sind materielle Phänomene, insofern sie als Träger funktionaler Rollen betrachtet werden. Wird der Geist im Hinblick auf diese Frage und diese Antwort mit einem Computer verglichen, dann zielt der Vergleich zunächst einmal nur darauf ab, daß es im programmierten Computer ebenfalls materielle Phänomene gibt, die wir als Träger funktionaler Rollen (z. B. als die der Ziffer 5 beim Rechnen oder die des Königs beim Schach) betrachten. Dabei kommt es aber auf die Art der physischen Realisierung insofern gar nicht an, als

physikalisch andersartige materielle Phänomene dieselbe funktionale Rolle realisieren können. Andererseits gibt es die empirische, psychologische Frage: Wie spielen sich die geistigen Phänomene X, Y und Z bei einem normalen Menschen ab; wie läßt sich so etwas präzise modellieren? Und eine gerne herangezogene methodische Empfehlung besagt: „Man schreibe ein Programm für einen Computer, so daß damit auf diesem Computer simuliert wird, wie sich das ausnimmt, wenn eine normale Person X, Y und Z hat“. Hier wird der Geist gar nicht mit einem Computer verglichen, sondern der Computer wird benutzt, um eine psychologische Spekulation über Geistesphänomene möglichst effektiv zu gestalten. Eine dritte Frage ist neurophysiologischer Art: Was geschieht in einem normalen menschlichen Hirn, das einer psychisch normalen Person gehört, die gerade das geistige Phänomen X aufweist? Auch darüber könnte man mit Computer-Simulationen spekulieren, dabei würde man aber wiederum nicht den Geist mit einem Computer vergleichen. Die Vermischung von philosophischen, psychologischen und neurophysiologischen Erkenntniszielen hat viel dazu beigetragen, die Analogie von Geist und Computer eher unklar als erhellend zu machen. Was am materialistischen Funktionalismus als einer philosophischen Lehre über das Wesen geistiger Phänomene wichtig und richtig sein mag, ist auf Geist/Computer-Analogien nicht angewiesen.

**4. Schwierigkeiten des Funktionalismus.** Als eine Theorie von Geistesphänomenen mit qualitativem Gehalt scheint der Funktionalismus vielen von vornherein abwegig. Das Wesen des Schmerzes ist keine funktionale Rolle, die z. B. Hautverletzungen mit Stöhnen in Kausalzusammenhang bringt; das Wesen des Schmerzes besteht darin, daß er wehtut. Und das Wehtun scheint von Hautverletzungen und Stöhnen begrifflich unabhängig zu sein; es ist ja nicht begriffswidrig, sich vorzustellen, daß jemand Hautverletzungen hat, die bewirken, daß er stöhnt, obwohl ihm nichts wehtut. Dies hieße, so scheint es: Hier hat einer Schmerzen (die funktionale Rolle ist realisiert), dem nichts wehtut (kein pas-

sender qualitativer Charakter ist da). Doch so etwas darf der Funktionalist nicht einräumen, weil es begrifflich unmöglich ist. Denn daß einer Schmerzen hat, heißt eben auch: Ihm tut etwas weh. Der qualitative Charakter ist dem Schmerz wesentlich; und er scheint funktionalistisch unerreichtbar. Darauf gibt es für den Funktionalisten offenbar nur zwei mögliche Repliken. Entweder muß er behaupten, daß das gerade eben Ausgemalte sich in Wirklichkeit gar nicht ausmalen läßt; er müßte zeigen, daß die Vorstellung, ein Zustand könne zwar die funktionale Rolle von Schmerz haben, ohne wehzutun, eine versteckte Inkohärenz enthält. Oder er müßte behaupten: Der qualitative Charakter von Geisteszuständen, soweit er funktionalistisch nicht einzuholen ist, ist selbst nichts Geistiges.

Ein anderer Einwand richtet sich gegen die zu reichhaltige, gewissermaßen zu leicht herbeizuführende Realisierbarkeit funktionaler Rollen. Einmal angenommen, es wäre bekannt, wie Zahnschmerz bei einem normalen Menschen physisch realisiert ist; im Kortex trete ein charakteristisches neuronales Muster vom Typ M auf. Der Funktionalismus scheint zu lehren: Jede beliebige Realisierung des abstrakten Musters M ist Zahnschmerz. Nun läßt sich M, als ein abstraktes Muster, aber auch mit vom Herbstwind durchwehten Tannenzapfen in einem hinreichend großen Wald realisieren; es wäre aber doch absurd, dem Tannenwald Zahnschmerzen zuzuschreiben. Dieser berühmte Einwand (er stammt von N. Block) richtet sich nur gegen eine sehr extreme Variante des Funktionalismus, die besagt, daß es bei der funktionalen Rolle und ihrer Realisierung nur darauf ankommt, was sozusagen auf der Innenseite des jeweiligen Systems geschieht. Doch nicht jede Form des Funktionalismus ist darauf festgelegt zu behaupten: „Wenn die funktionale Rolle R von einem bestimmten System S realisiert wird, dann wird R auch von jedem beliebigen System S\* realisiert, das zu S in einer rein innenbezüglichen Isomorphie steht.“ Ein Funktionalist könnte also auf diesen zweiten Einwand – mit Recht, soweit ich sehe – erwidern: „Natürlich hat ein Tannenwald keine Zahnschmerzen. Aber ein vorstellbarer Organismus, der Zähne

hätte, könnte Zahnschmerzen haben, auch wenn er Tannenzapfen statt Neuronen im Schädel hätte. Es kommt nicht darauf an, ob die Schmerzfunktion – unter der selbstverständlichen Voraussetzung, daß sie da ist – mit Tannenzapfen oder mit Nervenzellen realisiert ist.“ Der Funktionalismus ist ja keine Antwort auf die erkenntnistheoretische Frage, woran Zahnschmerz (sei es „von innen“, sei es „von außen“) zu erkennen ist. Er liefert eine Antwort auf die metaphysische Frage, was Zahnschmerz ist. Er läßt zwar zu, daß bei gewissen Systemen das Gewoge von Tannenzapfen Zahnschmerz ist; er ist aber nicht auf die Behauptung festgelegt, daß jeder sich in jener charakteristischen Manier M wiegende Tannenwald Zahnschmerzen hat.

Ein weiterer berühmter Einwand gegen den Funktionalismus – John Searles Argument mit dem **Chinesisch-Zimmer** (↑ Chinesisch-Zimmer-Szenario) – hat es mit intentionalen Geisteszuständen zu tun; er ist aber nur eine konfusere Variante des eben diskutierten Mißverständnisses. (Searle stellt sich vor, in einem Zimmer, in das ihm fremdartige Schriftzeichen hineingereicht werden, die er nicht versteht, all die syntaktischen Operationen durchzuführen, die ihm von einem Buch mit den entsprechenden Anweisungen vorgegeben sind. Am Ende des Vorgangs reicht er jeweils wieder einen Zettel mit Schriftzeichen, die er ebenfalls nicht versteht, aus dem Zimmer. Was er nicht weiß, ist dies: Auf den hineingereichtenzetteln stehen Fragen auf chinesisches; auf den von ihm hinausgereichtenzetteln stehen Antworten auf chinesisches; die Folge von Operationen, die er bei der Ausführung der Anweisungen im Buch durchläuft, entspricht Schritt für Schritt den Operationen, die ein Computer durchlaufen würde, der mit einem Programm zur Simulation chinesisches Sprachbeherrschung geladen ist. Trotzdem versteht Searle im Zimmer kein Chinesisches.) Dieses unterbeschriebene Gedankenexperiment läuft darauf hinaus, daß jemand allein deswegen noch lange kein Chinesisches verstünde, weil er bloß die Innenseite der funktionalen Rolle des Chinesisch-Verstehens (im Rahmen der klassischen Computeranalogie) impersonifizierte. Das ist richtig, aber irrelevant. Wird bloß

das innerseitige Muster irgendeines Systems kopiert, das Chinesisch versteht (oder irgendeine anderweitige Leistung erzeugt), dann wird damit allein natürlich noch nicht diese Leistung erbracht. Das Innere der schwarzen Schachtel konstituiert die Leistung des Systems nicht, sondern führt sie nur aus. Der Funktionalismus behauptet anderes nur da, wo er schlecht formuliert ist oder als Pappkamerad herhalten muß.

#### V Die Anti-Individualismusdebatte

(↑Anti-Individualismus). In der analytischen Diskussion um die Natur intentionaler Geistesphänomene gibt es eine Frage, die sich in den vergangenen fünfzehn Jahren als grundlegend erwiesen hat. Wie sind intentionale Geistesphänomene individuiert: individualistisch oder nicht? Diese sehr abstrakt wirkende Frage ist so außerordentlich wichtig, weil es in ihr um die metaphysische Autonomie des Geistes geht. Die philosophische Tradition der Neuzeit ist weitgehend dem Individualismus verpflichtet, und auch den meisten zeitgenössischen Philosophen und Psychologen kommt es wie eine unbestreitbare Selbstverständlichkeit vor, daß geistige Phänomene individualistisch individuiert sind.

Eine Eigenschaft, die einer Person zukommt, heißt „individualistisch individuiert“, wenn die Beschaffenheit der Person allein entscheidend dafür ist, ob sie diese Eigenschaft hat oder nicht; nichts, was sich außerhalb der Person abspielt, hat Einfluß darauf, ob sie diese Eigenschaft besitzt oder nicht. (Beispiele für individualistisch individuierte Eigenschaften sind etwa: einen leeren Magen zu haben, eine Frau zu sein, einen Ruhepuls von 60 Schlägen zu haben; Beispiele für nicht-individualistisch individuierte Eigenschaften: verheiratet zu sein, englischer Staatsangehöriger zu sein, einen durchschnittlichen Ruhepuls zu haben.)

L. Wittgenstein, H. Putnam und T. Burge haben darauf aufmerksam gemacht, daß intentionale Geisteszustände wie Überzeugungen, Erwartungen und dergleichen oftmals nicht-individualistisch individuiert sind. Insbesondere Burge hat dies mit sehr sorgfältig entwickelten Gedankenexperimenten unterschiedlicher Art darge-

legt. Stellen wir uns z. B. vor, daß Otto glaubt, Aluminium sei ein Leichtmetall; Otto hat einen Doppelgänger, der sich von Otto physikalisch gesehen nicht unterscheidet. Der Doppelgänger lebt in einer Umgebung, in der es kein Aluminium gibt, sondern nur Zwalum, ein anderes Metall, das so viel Ähnlichkeit mit Aluminium hat, daß nur ein Fachmann den Unterschied feststellen könnte. In der Sprachgemeinschaft des Doppelgängers wird das Wort „Aluminium“ benutzt, um über Zwalum zu reden. Wenn Ottos Doppelgänger nun sagt „Aluminium ist ein Leichtmetall“, dann bekundet er damit seine Überzeugung, daß Zwalum ein Leichtmetall ist, und nicht (wie Otto es mit der phonetisch gleichartigen Äußerung tut), daß Aluminium ein Leichtmetall ist. Diese beiden Personen hätten unterschiedliche Überzeugungen. Es hängt also offenbar nicht von der Person allein ab, welche Überzeugung sie hat, auch die Beschaffenheit ihrer Umgebung hat Einfluß darauf. Mithin sind zumindest einige intentionale Zustände nicht-individualistisch individuiert.

Von den meisten Philosophen wird inzwischen eingeräumt, daß die gewöhnliche Verwendung von Ausdrücken, mit denen wir intentionale Zustände bezeichnen (Ausdrücke wie z. B. „glauben, daß“, „wollen, daß“ usw.), auf eine anti-individualistische Individuierung dieser Zustände hinweisen. Heftig umstritten ist jedoch, welche philosophischen Lehren daraus zu ziehen sind. Der Anti-Individualismus (oder Externalismus) besagt, daß diese Beobachtung philosophisch ernstzunehmen sei, weil sie darauf hinweise, daß es eben zum Wesen einiger intentionaler Zustände gehöre, daß sie nicht-individualistisch individuiert sind. Der Individualismus hingegen insistiert darauf, daß eine den Zwecken der Psychologie angemessene Individuierung intentionaler Geisteszustände individualistisch sein müsse; die Psychologie benötige also ein Vokabular, das sich von dem der gewöhnlichen Sprache in dieser wichtigen Hinsicht unterscheidet. Innerhalb des individualistischen Lagers gibt es grundsätzlich unterschiedliche Auffassungen darüber, ob und in welcher Weise eine wissenschaftliche Psychologie von intentionalen Gei-

stesphänomenen zu handeln hat. Der Intentionalitätsrealismus ist die Lehre, daß es intentionale Phänomene wirklich gibt und daß in psychologischen Erklärungen auf sie Bezug genommen werden muß; eine individualistische Psychologie muß also eine neue Begrifflichkeit für diese Phänomene entwickeln. Dem Eliminativismus zufolge ist dies vermutlich unmöglich und jedenfalls überflüssig, denn eine wissenschaftliche Psychologie brauche überhaupt keine Begriffe, die intentionale Phänomene bezeichnen.

Im Zentrum dieser Debatte steht das Problem des Inhalts geistiger Zustände. Intentionale Geisteszustände wie die Überzeugung, daß Aluminium ein Leichtmetall ist, sind ja offenbar deshalb anti-individualistisch individuiert, weil es von externen Faktoren abhängt, welches ihr Inhalt (der im Daß-Satz angegeben wird) ist. Man kann die zentrale Überlegung kurz so zusammenfassen: Erstens sind solche Geisteszustände durch ihre Inhalte individuiert; zweitens sind die Inhalte abhängig davon, wie die außenweltliche Umgebung beschaffen ist; also hat die außenweltliche Umgebung Einfluß auf die Individuation dieser Geisteszustände, sie sind, anders gesagt, nicht-individualistisch individuiert. Anti-Individualismus und Eliminativismus stimmen in dieser Überlegung überein; der Eliminativist betrachtet die Nicht-Individualität intentionaler Geisteszustände allerdings als einen Hinweis darauf, daß mit ihnen wissenschaftlich gesehen etwas im argen ist, während der Anti-Individualist sie als einen Hinweis darauf betrachtet, daß die wissenschaftliche Psychologie sich selbst und ihren eigentlichen Gegenstand mißversteht, wenn sie sich einer individualistischen Methodologie verschreibt. Der individualistische Intentionalitätsrealismus à la J. Fodor versucht, am zweiten Schritt der Überlegung anzusetzen und eine Konzeption des Inhalts von Geisteszuständen zu entwickeln, so daß solch ein Inhalt – in der Diskussion zumeist als „enger Inhalt“ bezeichnet – nicht von der Beschaffenheit der außenweltlichen Umgebung abhängig ist. Der oben erwähnte Otto und sein Doppelgänger z. B. hätten Geisteszustände, deren „weiter“ Inhalt zwar verschieden, deren „enger“ Inhalt hingegen

identisch wäre. Der enge Inhalt soll dasjenige am Inhalt eines intentionalen Geisteszustands sein, was übrigbleibt, wenn die kontingente Beschaffenheit der außenweltlichen Umgebung gewissermaßen ausgeblendet wird. (Ein Vorschlag besteht darin, den engen Inhalt als eine Funktion aufzufassen, die für die verschiedenen Kontexte dem Geisteszustand den jeweils passenden weiten Inhalt zuordnet; solch eine Funktion würde z. B. den Geisteszuständen von Otto und seinem Doppelgänger für jeden Kontext denselben weiten Inhalt zuordnen; eine Verschiedenheit der weiten Inhalte gibt es nur, wenn Otto und sein Doppelgänger sich in verschiedenen Kontexten befinden.) Solcherlei Inhalt wäre es, wonach sich, dem Vorschlag Fodors zufolge, die Individuierung intentionaler Geisteszustände im Rahmen einer wissenschaftlichen Psychologie zu richten hätte. Dieser Ansatz zu einer Rettung des Individualismus bringt eine Reihe schwerwiegender Probleme mit sich; insbesondere ist fraglich, ob die Idee eines engen Inhalts überhaupt in der gewünschten Weise theoretisch entwickelt werden kann.

Die anti-individualistische Position steht ebenfalls vor erheblichen Schwierigkeiten. Zum einen stellt sich die Frage: Wie kann es sein, daß wir ohne weiteres wissen, was wir denken, wenn der Inhalt unserer Gedanken auch von äußeren Faktoren abhängt, die uns oft nicht ohne weiteres zugänglich sind? Denn wenn es – wie der Anti-Individualismus besagt – von äußeren Faktoren  $F_1 \dots F_n$  abhängt, ob eine Person P den Gedanken hat, daß Aluminium ein Leichtmetall ist, oder den Gedanken, daß Zwalum ein Leichtmetall ist, dann scheint es so, als könne P nicht ohne weiteres wissen, was er denkt: den einen Gedanken oder den andern. Kurz, die sogenannte Erste-Person-Autorität, die eine Person laut herrschender Meinung bezüglich ihrer Gedanken hat, scheint im Lichte der anti-individualistischen Auffassungen unerklärlich. Neben dieser erkenntnistheoretischen Schwierigkeit gibt es ein grundlegendes Problem, das die Metaphysik des Geistes und die Wissenschaftstheorie der Psychologie betrifft. Insbesondere J. Fodor hat es gegen den Anti-Individualismus geltend ge-

macht. Folgendes Prinzip wird vielerseits für unbestreitbar gehalten: „Die einzig natürliche und für eine wissenschaftliche Psychologie angemessene Individuierung von Geisteszuständen ist eine, die sich nach den Kausalkräften der Geisteszustände richtet.“ Wenn zwei Geisteszustände Wirkungen genau der gleichen Art haben, dann sollten sie als Geisteszustände desselben Typs klassifiziert werden. Eine vernünftige Psychologie sollte also Geisteszustände ihren Kausalkräften entsprechend individuieren, gleichgültig was für eine Individuierung durch die umgangssprachliche Begrifflichkeit nahegelegt wird. Für genauso unbestreitbar wird von vielen folgende Supervenienz-These gehalten: „Die Kausalkräfte der Geisteszustände einer Person supervenieren über dem neuronalen Zustand dieser Person.“ Aus dieser These folgt, daß zwei neurophysiologisch ununterscheidbare Personen (wie Otto und sein Doppelgänger) Geisteszustände mit identischen Kausalkräften haben; und mit dem Prinzip der wissenschaftlichen Individuierung ergibt sich daraus, daß sie sich in denselben Geisteszuständen befinden. Die Diskussion um diese Fragen ist erst in den letzten Jahren in Gang gekommen und wird ein wichtiger Streitpunkt in der Philosophie des Geistes bleiben.

## VI Die Rückkehr des Repräsentationalismus.

Im Rationalismus und Empirismus des 17. und 18. Jahrhunderts herrschte die Auffassung vor, der menschliche Geist enthalte – als unmittelbare Objekte seiner Betätigungen – Ideen, und geistiger Bezug auf die außergeistige Welt sei durch den repräsentationalen Gehalt dieser Ideen vermittelt. Einige Philosophen dieser Zeit dachten sich Ideen als eine Art geistige Bildchen, die dem Geist zeigen, wovon sie handeln. Daß jemand an einen Hasen denkt, heißt im Lichte solcher Lehren etwa folgendes: Sein Geist perzipiert eine Idee, die ihm einen Hasen zeigt; daß jemand einen Hasen fürchtet, heißt entsprechend: Sein Geist perzipiert eine Idee, die ihm einen Hasen zeigt, und diese Perzeption ist von einer Empfindung der Furchtsamkeit begleitet. Kurz, geistige

Phänomene bestehen darin, daß man Ideen in einem bestimmten psychischen Modus hat.

Solch ein repräsentationalistischer Ansatz muß natürlich die Frage beantworten, was denn eine Idee I gerade zu einer Idee des Gegenstands x macht. Darauf gab es im wesentlichen drei Antworten: (i) Zwischen x und I besteht eine geeignete Kausalverknüpfung; (ii) zwischen x und I besteht eine geeignete Ähnlichkeitsbeziehung; (iii) Gott hat I zu einer Idee von x gemacht. Die dritte Antwort ist fromm, aber nicht sehr erhellend. Die beiden ersten Antworten werfen eine Reihe von Problemen auf. Wie kann es Ideen von Sachen geben, die gar nicht existieren; wie kann es geschehen, daß etwas fehlrepräsentiert wird; wie kann es sein, daß wir Ideen von Abstrakta haben? Eine als noch gravierender empfundene Schwierigkeit der traditionellen repräsentationalistischen Ansätze war der ihnen implizit (und manchmal auch explizit) innewohnende Dualismus. Denn letztlich ist ja gar nicht verständlich, wie materielle Substanzen auf geistige Phänomene einwirken oder ihnen ähnlich sein könnten.

Spätestens seit dem Aufstieg des Behaviorismus in der Psychologie und Philosophie galt der Repräsentationalismus als ein endgültig widerlegter Irrtum. Doch in den siebziger Jahren erhielt er im Rahmen der Computer-Analogie, insbesondere durch die Arbeiten von J. Fodor, neuen Auftrieb. Fodor verbindet seine repräsentationalistische Theorie mit der These, daß es eine Sprache des Geistes gibt, deren Sätze eine syntaktische Konstituentenstruktur besitzen, die dem Satzinhalt angemessen ist. Nur durch diese These sei erklärlich, wie kognitive Leistungen produktiv und systematisch sein können.

Dies ist im wesentlichen nichts anderes als traditioneller Repräsentationalismus ohne Dualismus und dafür mit Computer-Analogie. Daß jemand einen Hasen fürchtet, wird nun so aufgefaßt: Er hat eine logisch strukturierte mentale Repräsentation, die von einem Hasen handelt, und er steht zu dieser mentalen ↑ Repräsentation im psychischen Modus des Fürchtens. Dieser psychische Modus besteht darin, daß die mentale Repräsentation eine gewisse funktionale Rolle spielt, die für das Fürchten charakteristisch ist;

diese funktionale Rolle ist (laut Computer-Analogie) eine algorithmische Beziehung, in der diese mentale Repräsentation zu anderen mentalen Repräsentationen steht. Mentale Repräsentationen sind neuronal realisiert; die funktionale Rolle einer konkreten mentalen Repräsentation *r* (die also nichts anderes ist als eine bestimmte neuronale Konstellation) ist eine Eigenschaft dieses Hirnzustands, aus der sich ergibt, daß dieser Hirnzustand unter den-und-den Umständen zustande kommt und im Verbund mit diesen-und-jenen anderen Hirnzuständen die-und-die Wirkungen hat. Soweit der funktionalistisch/materialistisch geläuterte Repräsentationalismus à la Fodor. Einer seiner entscheidenden Vorzüge gegenüber den Vorläufern aus dem 17. und 18. Jahrhundert liegt darin, daß die neuen mentalen Repräsentationen (im Gegensatz zu den alten Ideen) materielle Entitäten sind und daß eine kausale Theorie des Inhalts somit nicht von vornherein metaphysisch ungereimt ist.

In der Diskussion um diese Lehre steht die Frage im Vordergrund, ob es möglich ist, eine naturwissenschaftlich akzeptable Theorie des repräsentationalen Gehalts mentaler Repräsentationen zu entwickeln. Solch eine Theorie sollte diejenigen Faktoren spezifizieren, an denen es liegt, daß eine mentale Repräsentation gerade den Inhalt hat, den sie hat. Eine Reihe kausaltheoretischer Vorschläge sind in den vergangenen Jahren diskutiert worden; keiner unter ihnen kann als weithin akzeptiert gelten. Als sehr hartnäckig hat sich das sog. Disjunktionsproblem erwiesen: Angenommen, die mentale Repräsentation *r* wird fast immer von *x*-Dingen, manchmal aber auch von *y*-Dingen hervorgerufen; wie kann eine Kausaltheorie des Inhalts mentaler Repräsentationen dann die beiden folgenden Fälle auseinanderhalten: (i) *r* hat denselben Inhalt wie „ist ein *x*-Ding“, wird aber manchmal auch von *y*-Dingen hervorgerufen und ist dann eine Fehlrepräsentation; (ii) *r* hat denselben Inhalt wie „ist ein *x*-Ding oder ein *y*-Ding“.

Zugunsten des Repräsentationalismus wird von manchen Autoren ins Feld geführt, er stelle eine Art wissenschaftlicher Rettung unserer vorwissenschaftlichen Theorie intentionaler Geistes-

phänomene dar. Denn gemäß unserer Alltagstheorie über diese Dinge seien intentionale Zustände im wesentlichen durch zwei Merkmale ausgezeichnet: daß sie einen Inhalt haben und daß sie kausal wirksam sind. Im Computer-Repräsentationalismus werde diesen Merkmalen aufs schönste Rechnung getragen, weil eine mentale Repräsentation ja ein konkreter Hirnzustand ist, der dank seiner syntaktischen Struktur zugleich passenden Inhalt haben und eine entsprechende kausale Rolle spielen kann. Zumindest die Behauptung, intentionale Zustände seien im Lichte unserer psychologischen Alltagstheorie echte Ursachen, wirkt überzogen und ist wohl eher eine philosophische Mißdeutung unserer Erklärungspraktiken („Sie ging ins Kino, weil sie glaubte, der Film sei gut“) als ein ernstzunehmendes Argument für den Repräsentationalismus.

## VII Psychologie

Ein Thema, das für die analytische Philosophie des Geistes von einiger Bedeutung ist, liegt in der Frage nach dem Verhältnis von Alltagspsychologie und wissenschaftlicher Psychologie. Viele Autoren haben die Voraussetzung akzeptiert, es gebe so etwas wie eine vorwissenschaftliche Theorie des Geistes, die der gewöhnliche Mensch benutzt, wenn er mit umgangssprachlichen Begriffen (wie „Schmerz“, „Überzeugung“, „Angst“ usw.) das Tun und Lassen gewöhnlicher Menschen erklärt oder vorhersagt; diese Theorie sei ihren Benutzern zwar nur implizit bekannt, aber es sei dennoch eine richtige Theorie mit Gesetzen und allem, was dazugehört. Eine wichtige Teiltheorie sei z. B. die Glaubens-und-Wollens-theorie menschlichen Handelns. Eliminativisten begründen ihre Extremposition gerne mit Hinweis auf die Unhaltbarkeit dieser psychologischen Alltagstheorie; viele ihrer Gegner beteuern deren Unverzichtbarkeit. Es ist jedoch nicht einmal klar, ob es eine derartige Theorie überhaupt gibt und ob sie – falls es so etwas in irgendeinem interessanten Sinne tatsächlich gibt – denn wirklich eine Ansammlung empirischer Hypothesen ist, die sinnvoll als in Konkurrenz zu den Hypothesen einer wissenschaftlichen Psychologie befindlich betrachtet werden können.

Eine andere Frage in diesem Zusammenhang ist, wieviel thematische und begriffliche Kontinuität zur Alltagspsychologie von einer wissenschaftlichen Psychologie verlangt werden darf bzw. muß. Eine psychologische Theorie eines Geistesphänomens – wie visuelle Wahrnehmung, Erinnerung oder Handlungsplanung – zeichnet sich nicht nur durch ihr Explanandum, sondern auch durch ihre Begrifflichkeit aus. (Würde ein Geistesphänomen rein neurophysiologisch erklärt, dann handelte es sich dabei ja nicht um eine psychologische Theorie; solch eine Erklärung könnte eine psychologische Theorie zum selben Thema allerdings überflüssig machen.) Was sind eigentlich die Charakteristika genuin psychologischer Begrifflichkeit? Dies ist auch im Hinblick auf manche Aspekte kognitionswissenschaftlicher Begriffs- und Theoriebildung eine interessante Frage, zumindest dann, wenn mit Computermodellierungen kognitiver Fähigkeiten und Leistungen der Anspruch erhoben wird, einen Beitrag zur Psychologie dieser Phänomene zu leisten. Diese Frage ist u. a. deshalb wichtig, weil sie so eng mit der Frage zusammenhängt, was eigentlich eine wissenschaftliche Psychologie sein soll. Angesichts der verwickelten Entwicklungen innerhalb des Faches Psychologie der vergangenen 100 Jahre und angesichts der unübersichtlichen Vielfalt derzeitiger Theorie-Ansätze in diesem Fach ist auch dies ein Thema, dem sich eine analytische Philosophie des Geistes zu widmen hat.

A. Kemmerling

Bieri, P. (1981) (Hg.). *Analytische Philosophie des Geistes*. Königstein: Hain.

Block, N. (1980/81) (Hg.). *Readings in the philosophy of psychology*. 2 Bände. Cambridge, Mass.: MIT Press.

Lycan, W. (1990) (Hg.). *Mind and cognition*. Oxford: Basil Blackwell.

Macdonald, C. & Macdonald, G. (1995) (Hg.). *Philosophy of psychology*. Oxford: Basil Blackwell.