

IMPORTANT: When citing this article, please refer to the print-version:

Köhler, W. / H. Mutschler (eds.): *Ist der Geist berechenbar?* Darmstadt: 2004, 58-76.

ZUR KRITIK DES FUNKTIONALISMUS

Godehard Brüntrup

1. Grundlegende Unterscheidungen

Der Funktionalismus war in den letzten Jahrzehnten ein einflußreiches Forschungsprogramm in der Philosophie des Geistes und der theoretischen Psychologie. Man sollte unterscheiden zwischen dem Funktionalismus als *methodologischem* und heuristischem Prinzip empirischer Forschung einerseits und dem Funktionalismus als *metaphysischer* These andererseits. Die metaphysische These besagt, daß das Wesen aller mentalen Entitäten allein durch ihre kausal-funktionale Einbettung in der Welt bestimmt wird. Am deutlichsten wird dies beim Computerfunktionalismus, der jeden mentalen Zustand durch seine computationale Rolle definiert, so wie sich auch jeder einzelne Zustand einer Turingmaschine allein durch seine funktionale Einbettung in das Gesamt der Maschinentafel bestimmen läßt. Die Grundintuition ist, daß zwei Entitäten, die dieselbe funktionale Rolle einnehmen, identisch sind. Das ist eine folgenreiche Annahme, die vielfältige Probleme aufwirft, auf die noch einzugehen sein wird. Der methodologische Funktionalismus vermeidet solche metaphysischen Festlegungen und sieht sich nur als eine operational erfolgreiche Strategie, Entitäten herauszugreifen und zu beschreiben. Das Herausgreifen von Entitäten anhand ihrer funktionalen Einbettung in das Netzwerk der Wechselwirkungen ist die Methode der Naturwissenschaft *par excellence*. Daher ist der methodologische Funktionalismus in der Naturwissenschaft unverzichtbar. Die Naturwissenschaft beschreibt ein komplexes Netzwerk relationaler Strukturen, gleichsam eine gewaltige Maschinentafel, eine relational bestimmte Welt. Mit Russell könnte man nun die metaphysische Frage nach dem intrinsischen Charakter der hierbei verbundenen Relata stellen: Was wird eigentlich durch all diese Relationen korreliert? Der methodologische Funktionalist kann diese metaphysische Frage offen lassen und sich auf die Feinmechanik der funktionalen Struktur konzentrieren. Ihn interessiert die spekulative Frage Russells nicht, sondern es zählt nur, was für wissenschaftliche Untersuchungen operationalisierbar (funktional erfaßbar) ist. Die Chance, daß auf diesem Weg explanatorisch nützliche Brückenprinzipien oder eventuell sogar Brückengesetze zwischen den funktionalen Strukturen des Geistes und den funktionalen Strukturen der neurophysiologischen Korrelate gefunden werden, ist durchaus hoch. Warum sollte die in anderen Bereichen der Natur bewährte Strategie beim Mentalen, das Teil dieser Natur ist, versagen?

Ich will eine Kritik am metaphysischen Funktionalismus mit einer positiven Einschätzung des methodologischen Funktionalismus verbinden. Die Problematik des Funktionalismus liegt in der Versuchung, eine erfolgreiche empirische Forschungsstrategie zur Metaphysik zu erheben. Außerdem werde ich zwei Arten des Funktionalismus unterscheiden, die eine unterschiedlich starke Reduktionsthese vertreten. Für beide werde ich zu zeigen versuchen, daß sie in Schwierigkeiten geraten, wenn sie metaphysisch aufgefaßt werden. Ebenso werde ich für beide anhand von Beispielen zeigen, daß sie methodologisch aufgefasst fruchtbare Forschungsprogramme grundlegen.

2. Schwach-reduktiver Funktionalismus

2.1 *Metaphysische Lesart des schwach-reduktiven Funktionalismus*

Der schwach-reduktive Funktionalismus nimmt ein (mindestens) dreigliedriges Schema an: Mental

IMPORTANT: When citing this article, please refer to the print-version:

Köhler, W. / H. Mutschler (eds.): *Ist der Geist berechenbar?* Darmstadt: 2004, 58-76.

- Funktional - Physisch. Während zwischen den mentalen Eigenschaften und den funktionalen Eigenschaften eine Typen-Identität besteht, die Reduktion zuläßt, besteht zwischen der funktionalen Ebene und der physischen Realisierung nur eine Token-Identität, die multiple Realisierung und damit keine Reduktion zuläßt. Die vertretene Reduktionsthese ist also abgeschwächt, sie reicht nicht bis ganz hinunter zur physischen Ebene. Der funktionalen Ebene wird damit eigenes ontologisches Gewicht beigemessen. Dies ist das klassische funktionalistische Programm, das dem Computermodell des Geistes zugrunde liegt. Die funktionale Ebene ist multipel realisierbar (in Gehirn und Rechenmaschine) und hat daher eine gewisse Autonomie.

Die Pointe dieses Funktionalismus liegt darin, daß sich die allgemeinen Gesetze der Psychologie zumindest für alle in unserer Welt naturgesetzlich möglichen physischen Systeme, die mit Mentalität ausgestattet sind, erforschen lassen, ohne auf die konkrete physische Realisierung Rücksicht nehmen zu müssen. Diese ideale Psychologie kann uns bessere Kausalerklärungen bezüglich unseres mentalen Lebens geben als selbst die Physik es vermag. Hilary Putnam behauptete in seiner funktionalistischen Periode, daß die nicht-funktionale (physische) Ebene sowohl ontologisch wie auch explanatorisch irrelevant sei. Natürlich sind ontologische und explanatorische Unabhängigkeit nicht dasselbe. Dieser Unterschied wurde nicht selten verwischt. Ein Textbeleg von Putnam mag dies verdeutlichen:

„... the ultimate constituents don't matter ... only the higher level structure matters ... The same explanation will go in any world (whatever the micro- structure) in which those higher level structural features are present. In that sense this explanation is autonomous Goodness is not a subjective matter ... An explanation is superior if it is more general.“
(Putnam 1975, 296)

Die explanatorische Nützlichkeit der funktionalen Ebene ist relativ unstrittig. Es ist die metaphysische These, die Probleme aufwirft. Die ontologische Unabhängigkeit der funktionalen Ebene kann durch folgenden Gedanken verdeutlicht werden.

- 1) Der funktionale Zustand F kann in der aktuellen (physikalistisch aufgefaßten) Welt w^* durch eine Vielzahl von Elementen aus der Menge aller in w^* möglichen physischen Zustände realisiert werden.
- 2) In einer (dualistisch konzipierten) Nachbarwelt W_1 hätte F durch eine Vielzahl von geordneten Zustandspaaren realisiert werden können (wobei jedes Paar zusammengesetzt wäre aus einem dort möglichen nicht-physischen und einem dort möglichen physischen Zustand).
- 3) Der funktionale Zustand F hätte in einer (nicht-physikalistisch konzipierten) Welt W_2 durch eine Vielzahl von Elementen aus der Menge aller in W_2 möglichen nicht-physischen Zustände realisiert werden können.

Der so verstandene Funktionalismus beantwortet die Frage, was ein bestimmter Zustand ist, unabhängig von der ontologischen Frage, wie er realisiert ist. Er trennt die Frage nach der Natur mentaler Zustände von der Frage ihrer Realisierung in unserer Welt. Aus diesem Grunde konnte Putnam auch behaupten, daß es auf die zugrunde liegende physische Realisierung gar nicht ankäme. Die höhere Ebene der funktionalen Zustände bietet das lohnendere Forschungsobjekt. Diese Sicht ist auch mit einem ontologischen Physikalismus vereinbar, wenn man davon ausgeht, daß es in unserer aktuellen Welt nur physische Realisierungen von funktionalen Zuständen gibt. Der physikalistische Funktionalismus impliziert dann die bereits erwähnte Typen-Identität zwischen mentaler und funktionaler Ebene zusammen mit einer Token-Identität zwischen funktionaler und physischer Ebene. Soweit die kurze Charakterisierung der These. Die metaphysische Lesart dieses schwach-reduktiven Funktionalismus wirft aber eine Reihe von Problemen auf, die in der Literatur seit langem diskutiert werden. Im Folgenden sollen die Kernprobleme herausgearbeitet werden.

IMPORTANT: When citing this article, please refer to the print-version:

Köhler, W. / H. Mutschler (eds.): *Ist der Geist berechenbar?* Darmstadt: 2004, 58-76.

2.1.1 Mentale Verursachung?

Ein hartnäckiges Problem liegt darin, daß aufgrund des Fehlens von Reduzierbarkeit von der funktionalen zur physikalischen Ebene die kausale Wirksamkeit der funktionalen Eigenschaften innerhalb des physikalistischen Rahmens nicht verständlich gemacht werden kann. Die Token-Identität zwischen funktionaler und physischer Ebene genügt nicht, um die kausale Wirksamkeit mentaler Zustände oder Eigenschaften zu begründen. Die Token-Identität ist zu schwach, um Reduktion zu erlauben. Unter der Annahme der kausalen Geschlossenheit des physischen Bereichs bleibt der irreduziblen funktionalen Ebene dann keine kausale Wirksamkeit. Gesteht man ihr eine eigenständige kausale Wirksamkeit zu, so erhielte man eine Abwärtsverursachung, welche die kausale Geschlossenheit des Physischen durchbräche. Dieses Problem läßt sich auch über den Gedanken der multiplen Realisierbarkeit der funktionalen Eigenschaften oder Zustände aufzeigen:

- 1) Funktionale Beschreibungen individuieren Zustände oder Eigenschaften, die eine Vielzahl von Realisationsmöglichkeiten abdecken.
- 2) Wenn ein Zustand oder eine Eigenschaft eine Vielzahl von Realisationsmöglichkeiten abdeckt, dann ist es nicht der Zustand oder die Eigenschaft, sondern die einzelnen Realisationen, die kausal wirksam sind.
- 3) Also individuieren funktionale Beschreibungen keine kausal wirksamen Zustände.

Man kann dieses Problem auch anders formulieren: Die funktionale Eigenschaft F ist eine kausale Rolle, die physische Eigenschaft P ist ein Inhaber dieser kausalen Rolle. Die entscheidende Frage lautet: Wie kann F als die kausale Rolle selbst noch kausal wirksam sein? Das ganze Kausalgeschehen läuft auf der Ebene des Inhabers ab. Man kann auch sagen, daß F gegenüber P eine Eigenschaft zweiter Ordnung ist. Hätte diese höhere Ebene eine eigene kausale Wirksamkeit, so wäre die kausale Geschlossenheit der ersten (= physischen) Ebene durchbrochen und der Physikalismus aufgehoben. Die funktionale Ebene wäre dann von der Art einer (emergenten) konfiguralen Kraft mit Abwärtsverursachung. Dafür haben wir aber vermutlich keine empirischen Belege, allenfalls einige Theorien außerhalb des Mainstreams wie Bohms „Quantenpotential“ oder Sheldrakes „morphogenetische Felder“ bieten mögliche Kandidaten an. Der ontologische Status einer eigenständigen funktionalen Ebene oberhalb der Ebene der Realisation bleibt daher äußerst unklar. Es spricht vieles dafür, daß es in der Natur neben den fundamentalen Wechselwirkungen des Physischen keine weiteren eigenständigen Kräfte gibt.

Eine weitere Schwierigkeit liegt in der Lokalisierung des mentalen Gehalts. Wenn die von Putnam und Kripke vorgebrachten Argumente korrekt sind, nach denen Bedeutungen nicht lokal „im Kopf“ sind, sondern eine extern-faktische und eine soziale Komponente haben, dann sind sie nicht bloß funktional charakterisierte Programmmustereigenschaften des Gehirns. Funktionale Gehirnzustände sind eng lokal implementiert. Wenn aber mentaler Gehalt in diesem Sinne „weit“ ist, dann kann er keine kausale Rolle mehr spielen. Selbst wenn die Programmmustereigenschaften über ihre syntaktische Form im Gehirn eine lokale kausale Rolle ausüben könnten, so wäre es nicht mehr der weite mentale Gehalt als solcher, der hier kausal wirksam wäre. Dann läge folglich auch keine mentale Verursachung vor. Sogar wenn die Genese dieser internen Programmeigenschaften nur durch Einbezug der Umwelt erklärt werden kann, so laufen sie doch, einmal entstanden, ohne jeden externen Bezug ab. Gehalt (weit) und Kausalität (lokal, eng) gehen nicht zusammen:

- 1) Der weite Gehalt der mentalen Eigenschaften oder Zustände eines Systems ist nicht supervenient gegenüber seinen lokalen Eigenschaften oder Zuständen.
- 2) Die kausalen Kräfte wirken lokal, sind also supervenient gegenüber den lokalen Eigenschaften oder Zuständen.
- 3) Also: Der weite Gehalt der mentalen Eigenschaften hat keine kausale Kraft.

IMPORTANT: When citing this article, please refer to the print-version:

Köhler, W. / H. Mutschler (eds.): *Ist der Geist berechenbar?* Darmstadt: 2004, 58-76.

2.1.2 Der utopische Charakter des Funktionalismus

Putnam, einer der Begründer des Psychofunktionalismus, hat bei seiner späteren Abkehr vom funktionalistischen Programm ein weiteres Problem herausgearbeitet. In für ihn typischer Manier wendet er die Kritik, die er einst als Funktionalist gegen die Typen-Identitätstheorie vorgebracht hatte, nun gegen den Funktionalismus selbst: Es gibt nicht nur keine Eins-zu-Eins-Relation zwischen der funktionalen Ebene und ihrer Realisierung, sondern auch zwischen der mentalen und der funktionalen Ebene gibt es eine „Viele-zu-Viele-Beziehung“. Dieser Umstand macht sogar das schwache funktionalistische Reduktionsprogramm praktisch undurchführbar, da es immer eine Vielzahl möglicher funktionaler Korrelate eines mentalen Zustands gibt und umgekehrt. Die Interpretationsvielfalt läßt sich nicht auflösen. Der Grundgedanke dieser Kritik besagt, daß wir für eine Reduzierung des Mentalen auf das Funktionale über eine Art „General-Algorithmus“ der ganzen menschlichen Rationalität verfügen müßten. Es riecht nach unreflektiertem Szientismus, an die Durchführbarkeit dieser Idee zu glauben. Betrachten wir einen beliebigen mentalen Zustand: die Überzeugung, daß eine Katze auf einer Matte sitzt. Gemäß der funktionalistischen Theorie müßte sich ein funktionaler Zustand bestimmen lassen, in dem sich alle Wesen mit dieser Überzeugung befinden. Es müßte sich um eine enorm komplexe Beschreibung handeln, welche lückenlos die notwendigen und hinreichenden Bedingungen für diesen Zustand angibt. Es scheint äußerst unwahrscheinlich, daß sich genau eine solche funktionale Beschreibung angeben läßt. Wie verhält es sich mit einem Wesen, das die Überzeugung „a cat sits on a mat“ hat? Befindet es sich im gleichen (oder zumindest einem äquivalenten) funktionalen Zustand? Wenn die verwendeten Ausdrücke synonym sind und sich auf dasselbe beziehen, dann müßte das der Fall sein. Es könnte aber sein, daß sich die Worte „cat“ und „mat“ auf ganz andere Entitäten beziehen. Die komplizierte Frage der Unbestimmtheit der Übersetzung kommt hier ins Spiel. Kann es eine rein funktionalistische Lösung dieses Problems geben, ohne in einen Zirkel zu geraten?

Wenn die ganze intentionale Begrifflichkeit auf eine funktionale Begrifflichkeit reduziert werden soll, dann müßten auch Begriffe wie „Synonymie“, oder „Koextensionalität“ funktional bestimmt werden. Es ist gewöhnlich ein Akt gutwilliger Interpretation (principle of charity), mit dem wir dem englischsprachigen Sprecher hier eine Überzeugung über Katzen und Matten zusprechen. Im Rahmen einer funktionalistischen Theorie kann man sich aber nicht allein auf die Interpretationspraxis berufen. Wir müßten vermutlich zunächst alle nur erdenklichen Ein- und Ausgaben in beiden Systemen (dem deutschen und dem englischen Sprecher) testen, um herauszufinden, ob sie sich wirklich im selben Zustand befinden. Da der Gehalt ihrer Überzeugungen auch durch die Umwelt bestimmt wird, genügt es nicht, die beiden Systeme isoliert von ihrer Umwelt zu betrachten. Zwei isoliert betrachtete identische Systeme können in unterschiedlicher Umwelt verschiedene mentale Gehalte haben. Sprachliche Bedeutungen sind eben nicht bloß im Kopf, sondern auch abhängig von der Umgebung des Sprechers. Was ein Wort bedeutet, hängt ganz wesentlich davon ab, worauf es sich in der Welt bezieht. Es kann uns auch nicht genügen, daß sich die Systeme rein äußerlich so verhalten, als ob sie im selben Zustand wären. Dann böte der Funktionalismus keinen Fortschritt gegenüber dem Behaviorismus. Es muß daher gezeigt werden, daß auch ihre interne Struktur der Informationsverarbeitung, die Feinmechanik des Geistes, äquivalent wäre. Wegen des holistischen Zusammenhangs mentaler Gehalte müßten wir das gesamte interne Netz ihrer Überzeugungen überprüfen. Eine Überzeugung gewinnt ihren bestimmten Gehalt erst im Ensemble mit vielen anderen Überzeugungen. Man kann propositionale Gehalte nicht in Isolation voneinander bestimmen. Um das ganze Netz der Überzeugungen zu überprüfen, müßte die funktionalistische Theorie aber reich genug sein, um jede noch so seltene Überzeugung individuieren und beschreiben zu können, inklusive schwer entdeckbarer abweichender Überzeugungen, die ja bei einer „radical translation“ rein pragmatisch

IMPORTANT: When citing this article, please refer to the print-version:

Köhler, W. / H. Mutschler (eds.): *Ist der Geist berechenbar?* Darmstadt: 2004, 58-76.

nicht auszuschließen sind. Selbst wenn wir unsere funktionalistische Theorie des Geistes auf menschliche Wesen beschränken wollten, benötigen wir eine funktionalistische Theorie der menschlichen Rationalität als Ganzer. Insgesamt erweist sich daher die funktionalistische Reduktion also als ein extrem utopisches Projekt. Die Interpretationspraxis läßt sich nicht in einer funktionalistischen Theorie einfangen. Wir können unsere eigenen interpretativen Fähigkeiten niemals vollständig formalisieren, weil wir jeweils innerhalb einer interpretativen Praxis formalisieren. Die Zuschreibung von Überzeugungen ist stets eine relative und damit für neue kulturelle Entwicklungen offene Tätigkeit. Eine reduktionistische Theorie, mit der sich auf einer sub-mentalen und sub-intentionalen Ebene objektiv und vorgängig zur Relativität einer interpretativen Praxis angeben ließe, in genau welchem mentalen Zustand ein Wesen sich befindet, liegt außerhalb unseres epistemischen Horizonts. Man muß also konstatieren, daß wir mentale Gehalte nicht eindeutig identifizieren und auch ihre funktionale Realisierung auf verschiedene Weise konstruieren können. Es liegt auch hier (wie zwischen funktionaler und physischer Ebene) eine „many-to-many“ Relation vor. (Putnam 1987, 251-77)

Eine Ausarbeitung der Typen-Identitätstheorie zwischen mentalen und funktionalen Eigenschaften oder Zuständen bleibt deshalb ein utopisches Programm. Der Funktionalist könnte einwenden, daß die Anerkennung unserer epistemischen Grenzen noch nicht die Falschheit des Funktionalismus impliziert. Liegt nicht doch hinter all unseren Interpretationen die eine und einzige funktionale Struktur der Wirklichkeit? Um diesem Einwand entgegenzutreten, läßt sich Putnams Gedanke noch weiter radikalieren: Funktionale Strukturen sind immer Interpretationen. Die Rede von einer „funktionalen Struktur an sich“ jenseits aller Interpretationen ist problematisch. Bereits Searle hatte gegen den Computerfunktionalismus argumentiert:

„The really deep objection to computational theories of mind can be stated quite clearly. ... Nothing is intrinsically computational. Computation exists only relative to some agent or observer who imposes a computational interpretation on some phenomenon.“
(Searle 1993, 16)

2.2 Methodologische Lesart des schwach-reduktiven Funktionalismus

Obwohl der schwach-reduktive Funktionalismus als metaphysische Theorie also massive Begründungsprobleme aufwirft, scheint er mir rein methodologisch als Forschungsheuristik fruchtbar zu sein. Die Naturwissenschaften bestimmen jeden Gegenstand über seine kausal-funktionale Rolle. Hat man ihn einmal so definiert, läßt sich auf einer tieferen Ebene nach einer kleinteiligeren Struktur suchen, welche dieselbe kausal-funktionale Rolle einnimmt. In diesem allgemeinen Sinne ist die funktionalistische Betrachtungsweise ein unabdingbares Fundament der Naturwissenschaft, und das Geistige wird sich nur in dem Maße erforschen lassen, als es sich mit dieser Methode fassen läßt. Es ist unstrittig, daß sich die meisten mentalen Zustände auch funktional charakterisieren lassen. Der Funktionalismus - in einem sehr weiten Sinne - ist daher die alternativlose Methodologie jeder Naturwissenschaft des Geistes.

Wie verhält es sich aber mit dem enger definierten Programm des schwach-reduktiven Funktionalismus? Auch er kann ein erfolgversprechendes Forschungsprogramm sein. Als Beispiel kann man die Konnektionismus-Klassizismus-Debatte nennen, also die Frage, ob die Informationsverarbeitung im menschlichen Gehirn nach dem Modell eines klassischen Computers oder eines massiv parallel-verarbeitenden neuronalen Netzes abläuft. Jerry Fodor und Zenon Pylyshyn haben starke Argumente dafür vorgetragen, daß gewisse mentale Fähigkeiten von Menschen besser verständlich gemacht werden können, wenn man zumindest die Implementation einer klassischen funktionalen Computerarchitektur in unserem neuronalen Netzwerk annimmt. (Fodor/Pylyshyn 1998) Sie begründen das unter anderem mit der so genannten „Systematizität“

IMPORTANT: When citing this article, please refer to the print-version:

Köhler, W. / H. Mutschler (eds.): *Ist der Geist berechenbar?* Darmstadt: 2004, 58-76.

kognitiver Fähigkeiten. Beispielsweise gibt es Gruppen semantisch verbundener mentaler Zustände, von denen gilt, daß ein Organismus, wenn er einen davon realisieren kann, sich auch in den anderen befinden kann. Man findet beispielsweise keine Lebewesen, die denken können „Peter liebt das Mädchen“ ohne auch denken zu können „Das Mädchen liebt Peter“. Fodor und Pylyshyn nehmen an, daß es eine funktionale Struktur gibt, aufgrund derer viele kognitive Fähigkeiten von Lebewesen in diesem Sinne systematisch und damit offen für immer neue Kombinationen sind. Sie versuchen aufzuzeigen, daß eine klassische funktionale Struktur die beste Erklärung für dieses Phänomen anbietet. Es geht also darum, welche Beschreibungen explanatorisch nützlich sind und welche besser sind als mögliche Alternativen. Die ontologische Frage nach der vom Erklärungskontext unabhängigen „Natur“ der beschriebenen Zustände ist demgegenüber sekundär. Auf jeden Fall ist nicht gefordert, daß eine explanatorisch erfolgreiche Beschreibung eine vollständige Abbildung der geistunabhängigen Welt sein muß. Die methodologische Lesart des schwach-reduktiven Funktionalismus ist offen für die Möglichkeit, daß der funktionalistisch verstehbare Aspekt des Geistes nur einen Teil des Gesamtphänomens darstellt. Sie liefert keine Metaphysik des Geistes, sondern eine Forschungsstrategie, die es erlaubt geistige Phänomene zu operationalisieren und damit empirischer Überprüfbarkeit zugänglich zu machen.

Es fragt sich nun, ob die Probleme des metaphysisch verstandenen Funktionalismus überwunden werden können, wenn man einfach auf eine ontologisch eigenständige funktionale Ebene verzichtet.

3. Stark-reduktiver Funktionalismus

3.1. Metaphysische Lesart des stark-reduktiven Funktionalismus

Der stark-reduktive Funktionalismus, wie er beispielsweise von David Lewis vertreten wurde, nimmt keine ontologisch unabhängige funktionale Ebene an. Er nimmt damit die oben entwickelte Einsicht auf, daß funktionale Strukturen zunächst auf der Ebene der Beschreibung anzusiedeln sind. Versucht wird eine Reduktion mentaler Zustände auf physische Zustände durch Zuhilfenahme einer funktionalen Beschreibung. Wenn die funktionale Beschreibung präzise genug ist und auf eine bestimmte natürliche Art eingeschränkt ist (z.B. Schmerz beim Menschen), kann sogar eine Typen-Identität zwischen der funktionalen Ebene und physischen Ebene angenommen werden. Der Kerngedanke ist, daß wir anhand der funktionalen oder kausalen Rolle einen Zustand herausgreifen, dessen Essenz dann durch empirische Forschung bestimmt wird. Wenn man die kausal-funktionale Rolle von Genen bestimmt hat, kann man auf der molekularbiologischen Ebene nach etwas suchen, das exakt diese Rolle einnimmt. Wenn man es gefunden hat (die DNA), dann läßt sich die klassische Genetik auf eine molekularbiologische Ebene reduzieren. Im Gegensatz zum schwach-reduktiven Funktionalismus geht es hier also nicht um die Wahrung der Autonomie einer funktionalen Ebene unabhängig von ihrer Realisierung, sondern um die mittels der Identität von funktionalen oder kausalen Rollen erreichte Reduktion auf eine möglichst basale physische Ebene. Die Grundintuition läßt sich schematisch verdeutlichen:

- . Der mentale Zustand M = Inhaber der kausal-funktionalen Rolle KR
- . Der neuronale Zustand N = Inhaber der kausal-funktionalen Rolle KR
- . Also: M=N

Wir greifen einen mentalen Zustand über seine kausale Rolle heraus. Dazu genügt unsere Alltagstheorie. Durch Forschung entdecken wir einen physischen Zustand, der dieselbe kausale Rolle einnimmt (wissenschaftliche Theorie). Aufgrund der Transitivität von Identität kann man schließen, daß die beiden Zustände identisch sind. Die funktionalistische Grundeinsicht der multiplen Realisierbarkeit funktionaler Zustände kann diese Theorie immer noch aufrechterhalten, wenn sie die Reduzierbarkeit nur innerhalb der Grenzen einer natürlichen Art zuläßt.

IMPORTANT: When citing this article, please refer to the print-version:

Köhler, W. / H. Mutschler (eds.): *Ist der Geist berechenbar?* Darmstadt: 2004, 58-76.

Es drängt sich nun sofort eine Frage auf: Ist die unterste Reduktionsebene, z.B. der neurophysiologische Zustand N, selbst rein funktional bestimmt? In diesem Fall wäre es unmöglich, daß ein Wesen in einer anderen Welt N hat, N aber in dieser anderen Welt eine andere kausale Rolle inne hätte. Positiv möglich wäre es aber, daß in einer anderen möglichen Welt dieselbe kausale Rolle anders realisiert wäre. Damit wäre Reduktion im strikten Sinne unmöglich, und wir wären wieder beim schwach-reduktiven Funktionalismus angelangt.

Auf der unteren Ebene sollte also ein Zustand anhand seiner kausalen Rolle herausgegriffen werden, dessen Wesen nicht allein durch seine kausale Rolle definiert ist. Die kausale Rolle ist hier Identifikationskriterium, nicht Identitätskriterium.

Die interessante Konsequenz dieser Annahme ist die Möglichkeit, daß zum Beispiel ein Wesen, das sich in demselben kausal-funktional bestimmten Zustand N befindet, andere nicht kausal-relational bestimmte Eigenschaften haben könnte. Man denke hier an invertierte Spektren oder metaphysische Zombies. Damit ist die größte Herausforderung des Funktionalismus angesprochen: das phänomenale Erleben.

3.1.1 Qualitatives Erleben: Von Mary und Zombies

Der schwerwiegendste Einwand gegen jede Art des Funktionalismus besteht darin, daß er den Aspekt des qualitativen Erlebens (Qualia) nicht zu erfassen scheint. Die mentalen Entitäten werden beim stark-reduktiven Funktionalismus zwar via Reduktion gerettet und damit das Problem der mentalen Verursachung entschärft. Aber um welchen Preis? Jaegwon Kim sprach in diesem Zusammenhang von der „Vietnam Metapher“: „Saving a village by destroying it.“ Man vernichtete ein Dorf, um es vor den Vietcong zu „retten.“ (Kim 1993, 367)

Die Intuition, daß bei der funktionalistischen Reduktion etwas verloren gegangen ist, läßt sich so formulieren:

- 1) Alle physikalischen Begriffe sind funktionale Begriffe.
- 2) Wenn phänomenale Wahrheiten a priori in funktionalen Wahrheiten enthalten sein sollen, dann muß es eine Analyse der phänomenalen Konzepte in funktionalen Ausdrücken geben.
- 3) Es gibt keine Analyse phänomenaler Begriffe in funktionalen Ausdrücken.
- 4) Also: Phänomenale Wahrheiten sind nicht a priori in physikalischen Wahrheiten enthalten.

Zur Begründung dieser Analyse wurden Argumente wie das sogenannte „Argument vom unvollständigen Wissen“ und das sogenannte „Argument der Möglichkeit metaphysischer Zombies“ entwickelt, auf die im folgenden eingegangen werden soll.

Ersteres Gedankenexperiment führt eine mit komplettem naturwissenschaftlichen Wissen über die Physiologie der Farbwahrnehmung ausgestattete Neurophysiologin namens „Mary“ ein. Sie hat aber vor einem bestimmten Zeitpunkt t noch niemals eine Farbe gesehen. Das Argument des unvollständigen Wissens läßt sich folgendermaßen rekonstruieren:

- 1) Mary kennt vor t alle physikalischen Tatsachen.
- 2) Mary lernt zu t etwas.
- 3) Also erwirbt Mary zu t satzhaftes (propositionales) Wissen.
- 4) Also lernt Mary zu t eine neue Tatsache kennen.
- 5) Also gibt es Tatsachen, die keine physikalischen Tatsachen sind.
- 6) Wenn der Physikalismus wahr ist, dann gibt es keine nicht-physikalischen Tatsachen.
- 7) Also ist der Physikalismus falsch.

Vor allem die Schritte 2), 3) und 4) sind einer intensiven Debatte ausgesetzt, die hier jetzt nicht im Detail dargestellt werden kann. Statt dessen soll gleich noch das sogenannte Zombie-Argument herangezogen werden, um beide Argumentsstränge zu verbinden. Es läßt sich in einem ersten Zugang folgendermaßen rekonstruieren:

IMPORTANT: When citing this article, please refer to the print-version:

Köhler, W. / H. Mutschler (eds.): *Ist der Geist berechenbar?* Darmstadt: 2004, 58-76.

- 1) Wenn der Materialismus wahr ist, dann erzwingen die entsprechenden physisch-funktionalen Eigenschaften das Auftreten der phänomenal-intrinsischen Eigenschaften (das bewußte Erleben) mit Notwendigkeit.
- 2) Wenn 1) der Fall ist, dann ist es nicht möglich, daß in einem System die entsprechenden physisch-funktionalen Eigenschaften realisiert sind, ohne daß darin auch die phänomenal-intrinsischen Eigenschaften realisiert sind.
- 3) Es ist möglich, daß in einem System die entsprechenden physisch-funktionalen Eigenschaften realisiert sind, ohne daß darin auch die phänomenal-intrinsischen Eigenschaften realisiert sind (der Fall des perfekten Roboters, des metaphysischen Zombies).
- 4) Also erzwingen die entsprechenden physisch-funktionalen Eigenschaften nicht mit Notwendigkeit die phänomenal-intrinsischen Eigenschaften. Der Materialismus ist falsch.

Es geht in beiden Gedankenexperimenten um die metaphysisch kontingente Beziehung zwischen den funktional bestimmten physischen Fakten und den phänomenalen Fakten, die über ihren intrinsischen qualitativen Gehalt bestimmt werden. Aus der Konjunktion aller physischen Fakten, nennen wir sie „P“, folgt nichts über ein beliebiges phänomenales Faktum, nennen wir es „Q“. Es ist also möglich, daß P & non-Q. Geschlossen wird in beiden Gedankenexperimenten von der Vorstellbarkeit auf die Möglichkeit, allerdings auf unterschiedliche Weise. Um den Unterschied herauszuarbeiten, muß man positive von negativer Vorstellbarkeit unterscheiden. Positive Vorstellbarkeit liegt vor, wenn man sich begrifflich oder imaginativ eine klare Konzeption oder ein klares Bild von etwas machen kann. Negative Vorstellbarkeit liegt vor, wenn man etwas nicht a priori ausschließen kann. Das Zombieargument behauptet, daß Zombies positiv vorstellbar sind, und daraus folgt direkt die kontingente Beziehung zwischen physischen und phänomenalen Fakten. Das Argument des unvollständigen Wissens behauptet, daß die phänomenalen Fakten nicht a priori aus den physischen Fakten abgeleitet werden können. Daraus folgt, daß P & non-Q nicht ausgeschlossen, also negativ vorstellbar sind.

Das klassische Gegenargument behauptet nun, daß die Explikation der Begriffe „notwendig“ und „möglich“ in einem solchen Argument nicht auf A-priori-Analysen aufbauen darf. Es ist nämlich gar nicht klar, daß aus der A-priori-Vorstellbarkeit auch reale Möglichkeit folgt. Mit einer Variante eines durch Hilary Putnam bekannt gewordenen philosophischen Gedankenexperiments läßt sich dieses Problem verdeutlichen. Nehmen wir an, Forscher von der Erde landen auf einem fremden, der Erde aber zum Verwechseln ähnlichen Planeten (eine Zwillingserde). Dort beobachten sie eine Flüssigkeit, die äußerlich in jeder Beziehung Wasser gleicht. Sie vermuten daher, daß es sich auch um Wasser handelt. Als ausgebildete Chemiker wissen sie aber, daß man sich vom ersten Anschein nicht täuschen lassen darf. Es könnte sich herausstellen, daß die molekulare Struktur der Flüssigkeit eventuell von der des Wassers (H₂O) unterschieden ist. Könnten sie jetzt folgendermaßen schlußfolgern?

- 1) Es ist widerspruchsfrei denkbar, daß diese Flüssigkeit kein Wasser ist.
- 2) Wenn diese Flüssigkeit kein Wasser sein könnte, dann kommt ihr die molekulare Struktur H₂O nicht notwendig zu.
- 3) Wasser kommt die molekulare Struktur H₂O notwendig zu.
- 4) Also ist diese Flüssigkeit nicht mit Wasser identisch.

Ganz offensichtlich handelt es sich bei diesem Schluss von Vorstellbarkeit auf Möglichkeit um einen Fehlschluss. Wenn es sich bei der fraglichen Flüssigkeit um Wasser handelt, dann kommt ihr die molekulare Struktur H₂O notwendig zu. Was diese Flüssigkeit wirklich ist, kann man aber nur a posteriori erkennen, eine A-priori-Analyse reicht nicht aus. Diese Einsicht kann man auf das Leib-Seele-Problem anwenden. Auch hier reicht dann der Nachweis nicht aus, daß die physischen

IMPORTANT: When citing this article, please refer to the print-version:

Köhler, W. / H. Mutschler (eds.): *Ist der Geist berechenbar?* Darmstadt: 2004, 58-76.

Wahrheiten nicht a priori bestimmte mentale Wahrheiten enthalten. Man muß auch noch zeigen, daß es zwischen beiden keine a posteriori notwendige Beziehung gibt. Wenn man durch empirische Forschung entdecken würde, daß phänomenales Erleben eine bestimmte neurophysiologische Struktur ist, dann hätte man, wie im dargestellten Fall des Wassers, a posteriori eine Notwendigkeit entdeckt. Dieses Gegenargument ist, jedenfalls in dieser einfachen Form, unzureichend, da es eine zentrale bedeutungstheoretische Unterscheidung übersieht. Die philosophische Tradition, insbesondere seit Frege, hat nämlich zwei Arten von Bedeutungen unterschieden. Diese Unterscheidung läßt sich leicht an unserem Beispielausdruck „Wasser“ aufzeigen. Nehmen wir an, auf der Zwillingserde handele es sich bei der durchsichtigen, trinkbaren Flüssigkeit tatsächlich nicht um H₂O, sondern um eine komplizierte chemische Verbindung XYZ. Die folgenden Sätze sind beide wahr:

1) Das Wasser auf der Zwillingserde hat die chemische Struktur XYZ.

2) Auf der Zwillingserde gibt es kein Wasser.

Die Sätze sind nicht widersprüchlich, da das Wort „Wasser“ in ihnen in zwei verschiedenen Bedeutungen gebraucht wird. Die Bedeutung von „Wasser“ in 1) ist eine Beschreibung von Wasser, die wir im Kopf haben: die klare, trinkbare, geschmacksneutrale etc. Flüssigkeit. Die Bedeutung von „Wasser“ in 2) ist die chemische Substanz in der Welt: dasjenige in unserer Welt, das wir „Wasser“ getauft haben. Ich will im folgenden mit David Chalmers „Bedeutung“ im ersten Sinn „primäre Intension“ und „Bedeutung“ im zweiten Sinn „sekundäre Intension“ nennen. (Chalmers 2002) Mittels dieser Unterscheidung wird klar, daß a posteriori notwendige Sätze wie „Wasser ist H₂O“ mit der sekundären Intension von „Wasser“ arbeiten. Wenn wir „Wasser“ in seiner primären Intension verstehen, dann ist „Wasser ist H₂O“ gerade nicht notwendig, sondern kontingent.

Aus dieser Analyse ergibt sich somit, daß jeder notwendige A-posteriori-Satz eine kontingente primäre Intension hat. Genau auf diese primäre Intension beziehen sich aber die Vorstellbarkeitsargumente. Der materialistische Einwand besagt also, daß die Vorstellbarkeitsargumente fälschlicherweise von der primären Intension ausgegangen waren. Wären sie korrekt von der sekundären Intension ausgegangen, dann wäre sofort klar gewesen, daß die angenommenen Möglichkeiten gar nicht existieren. Anders ausgedrückt: die phänomenalen Begriffe einerseits und die physisch-funktionalen Begriffe andererseits greifen trotz der A-priori-Unterschiedlichkeit a posteriori exakt diesselben Eigenschaften heraus.

Zwei Argumente lassen sich gegen diesen Einwand vorbringen: Erstens geht es in den Vorstellbarkeitsargumenten um logische Möglichkeit. Den Bereich des logisch Möglichen erreichen wir aber über die primären Intensionen. Keine Einsicht über A-posteriori-Notwendigkeiten der beschriebenen Art schließt irgendeine logisch mögliche Welt aus. Auch wenn es natürlich keine mögliche Welt gibt, in der H₂O nicht H₂O ist, so gibt es doch mögliche Welten, in denen die durchsichtige geschmacksneutrale Flüssigkeit in den Flüssen und Seen nicht H₂O ist. Wenn man diese Unterscheidung trifft, dann greift der Einwand über A-posteriori-Notwendigkeiten gegen ein Vorstellbarkeitsargument wie das der Möglichkeit einer Zombiewelt nicht. Der Funktionalist wird also versuchen, eine Bedeutungstheorie zu entwickeln, die ohne primäre Intensionen auskommt. Auf dieses weite Feld kann hier aber jetzt nicht eingegangen werden. Es soll nur noch ein zweites Gegenargument dargestellt werden, das die Irrelevanz sekundärer Intensionen in diesem Kontext aufzeigen will. Im speziellen Falle des bewußten Erlebens greift die Einführung sekundärer Intensionen überhaupt nicht, da in diesem besonderen Falle primäre und sekundäre Intension zusammenfallen. Ein Zustand in der aktuellen Welt ist genau dann ein bewußter Zustand, wenn er sich irgendwie anfühlt. Der phänomenale Gehalt macht das Wesen dieses Zustands aus.

Auch in einer anderen, kontrafaktisch angenommenen Welt wird ein bewußter Zustand ebenso durch seinen phänomenalen Gehalt bestimmt. Die Unterscheidung zwischen primärer und

IMPORTANT: When citing this article, please refer to the print-version:

Köhler, W. / H. Mutschler (eds.): *Ist der Geist berechenbar?* Darmstadt: 2004, 58-76.

sekundärer Intension bei Begriffen wie „Wasser“ wird durch die Einsicht begründet, daß es in einer anderen, kontrafaktisch angenommenen Welt etwas geben könnte, das zwar wie Wasser aussieht und sich wie Wasser anfühlt, in Wirklichkeit aber gar kein Wasser ist. Diese Unterscheidung macht im Falle bewußten Erlebens gar keinen Sinn. Etwas, was sich genauso anfühlt wie ein bewusstes Erlebnis, ist ein bewusstes Erlebnis, auch in einer kontrafaktisch angenommenen anderen Welt. Es macht einfach keinen Sinn anzunehmen, daß das Schmerzerlebnis eines Wesen in einer anderen möglichen Welt, das sich genauso anfühlt wie mein momentanes Schmerzerlebnis, in Wirklichkeit gar kein Schmerzerlebnis sei. Aus den beiden eben entwickelten Gründen ist die Einführung von a posteriori notwendigen Sätzen keine geeignete Strategie, um den Übergang von der epistemischen auf die modale Ebene zu blockieren.

Das materialismuskritische Argument läßt sich jetzt wesentlich präziser fassen. Ich verwende diejenige Formulierung, die David Chalmers in den letzten Jahren wiederholt in die Diskussion eingebracht hat. (Chalmers [LINK](#)) Es sollen folgende Definitionen gelten:

- . P = die physikalisch-funktionalen Wahrheiten
- . Q = eine phänomenal-intrinsische Wahrheit
- . P', Q' = die primären Intensionen von P, Q

Drei Prämissen sollen eingeführt werden: Die erste formuliert das Ergebnis der Analyse auf der epistemischen Ebene, wie z.B. durch das Argument des unvollständigen Wissens. Die zweite formuliert das Ergebnis der semantischen Analyse, die auf der Unterscheidung von primärer und sekundärer Intension aufbaut. Die dritte liefert eine Explikation des Materialismus.

- 1) Die physikalischen Wahrheiten enthalten nicht a priori die phänomenalen Wahrheiten.
(Prämisse)
- 2) Wenn ein Satz S a posteriori ist, dann hat S eine kontingente primäre Intension.
(Prämisse)
- 3) Wenn der Materialismus wahr ist, dann erzwingen die physikalischen Tatsachen alle wahren Propositionen.
(Prämisse)
- 4) $P \rightarrow Q$ ist a posteriori. [aus 1]
- 5) $P \rightarrow Q$ hat eine kontingente primäre Intension. [aus 2 & 4]
- 6) Es gibt eine mögliche Welt W, in der gilt: P' ist wahr in W, Q' ist falsch in W. [aus 5]
- 7) Es gibt eine mögliche Welt W, in der gilt: P ist wahr in W, Q' ist falsch in W. [aus 6?]
- 8) Folglich ist der Materialismus falsch. [aus 3 & 7]

Bei Schritt 7) wurde ein Fragezeichen angefügt, denn er folgt nur unter der Zusatzannahme: $P' = P$. Diese Zusatzannahme sollte der Physikalist aber einführen. Wenn der Materialist sie nämlich bestritte, so würde behauptet, daß die Sätze, welche die physisch-funktionalen Wahrheiten gemäß der primären Intension beschreiben, das Wesen des Physischen nicht vollständig ausdrücken. Die Zombie-Welt scheint dann physisch identisch mit der unseren zu sein, in Wirklichkeit ist sie aber verschieden. Das würde bedeuten, daß es essentielle Eigenschaften des Physischen geben muß, die durch die physisch-funktionale Beschreibung nicht erfaßt werden. Wenn das der Fall wäre, würden dem Zombie genau diese verborgenen Eigenschaften des Physischen fehlen, und er hätte genau deshalb keine Erlebnisse. Diese These untergräbt aber den Materialismus, da sie einen Aspekt des Physischen annimmt, welcher der funktionalen Begrifflichkeit der auf der Physik aufbauenden Naturwissenschaften entgeht. Genau dieser Aspekt ist es, der für das phänomenale Erleben zuständig ist. Diese Position sieht mehr aus wie eine Zwei-Aspekte-Theorie und hat dementsprechend eine Tendenz zum Panpsychismus (oder wenigstens zum Panprotopsyichismus). Es ist kein Materialismus im engeren Sinn. An dieser Stelle wird jetzt deutlich, wieviel davon abhängt, ob die letzte, tiefste Ebene der Reduktion noch immer rein funktional bestimmt ist. Ist dies

IMPORTANT: When citing this article, please refer to the print-version:

Köhler, W. / H. Mutschler (eds.): *Ist der Geist berechenbar?* Darmstadt: 2004, 58-76.

der Fall, dann geraten wir in die Schwierigkeiten des schwach-reduktiven Physikalismus. Ist dies nicht der Fall, so sind alternativen (nicht-materialistischen) Interpretationen wie dem Panprotopsyichismus Tür und Tor geöffnet.

Der materialistische Funktionalist kann nun auf eine doppelte Weise vorgehen. Er kann entweder die epistemische Analyse ablehnen und behaupten, Zombies seien gar nicht vorstellbar, oder er kann die epistemische Analyse akzeptieren und dann den Übergang auf die modale und metaphysische Ebene blockieren.

Der erste Weg, der sogenannte analytische Funktionalismus, behauptet, daß es eine a priori notwendige Beziehung zwischen funktionaler und phänomenaler Begrifflichkeit gibt. Allerdings hat noch niemand diese gesuchten A-priori-Beziehungen aufzeigen können. Einige Philosophen unterscheiden daher unsere faktischen begrifflichen Möglichkeiten von idealen begrifflichen Möglichkeiten. Es könnte durch die mangelhafte Struktur unseres Erkenntnisvermögens begründet sein, daß wir die a priori notwendige Beziehung zwischen phänomenalen Fakten und physischen Fakten nicht einsehen können. Thomas Nagel hat den Verdacht geäußert, daß die phänomenale und die physikalische Begrifflichkeit die Natur der Phänomene, auf die sie referieren, noch nicht hinreichend erfassen. (Nagel 1998, 337-352) Zugleich äußerte er die Hoffnung, daß ein revolutionär neuer begrifflicher Rahmen die notwendige Verbindung zwischen dem Mentalen und dem Physischen einsichtig machen könnte. Colin McGinn hat pessimistischer vermutet, daß aufgrund der unüberwindbar dichotomen Struktur unseres Erkenntnisapparates, der - kantisch gesprochen - in einen äußeren raum-zeitlichen und einen inneren nur zeitlichen Sinn geteilt ist, der gesuchte neue begriffliche Rahmen für Wesen wie uns prinzipiell nicht erreichbar ist. Das hartnäckige Leib-Seele-Problem ist nur ein Spiegel unseres dichotomen Erkenntnisapparates, ein allwissendes Wesen könnte das Leib-Seele-Problem dagegen a priori lösen. (McGinn 1989, 349-266) Das sind massive metaphysische Thesen, die nicht leicht zu begründen sein werden. Wir haben hier die fragwürdige Form von Metaphysik vor uns, die über das spricht, was ihren eigenen Annahmen zufolge epistemisch unzugänglich ist.

Die meisten Philosophinnen und Philosophen haben daher den zweiten Weg beschritten und anerkennen die epistemische Analyse, d.h. die Vorstellbarkeit der Gedankenexperimente. Sie argumentieren, daß nicht alles, was gemäß der primären Intension negativ oder positiv vorstellbar ist, auch metaphysisch möglich ist. Man könnte behaupten, daß metaphysische Möglichkeit an sekundäre Intensionen und damit sekundäre Vorstellbarkeit gebunden ist. Dahinter steht oft eine Theorie der Modalitäten in der Tradition von Saul Kripke oder sogar in der Tradition des Aristoteles. Nach dieser Auffassung ist metaphysische Möglichkeit immer an die aktuelle Welt gebunden. Es ist das Wesen einer Entität in dieser aktuellen Welt, das bestimmt, welchen Veränderungen diese Entität unterliegen kann. Man spricht daher auch von „wissenschaftlichem Essentialismus“. Wenn wir herausfinden wollen, ob etwas metaphysisch möglich ist, dann müssen wir gleichsam den Finger auf eine in dieser Welt durch Ostension herausgegriffene Entität legen und sie durch andere mögliche Welten schieben. In diesem Sinne ist es metaphysisch unmöglich, daß Wasser nicht H₂O ist. Wir können uns fragen, ob etwas mit dem Wesen, das es in dieser Welt hat, in einer anderen möglichen Welt immer noch existieren könne. In diesem Sinne wäre es ebenfalls unmöglich, daß ein in der aktuellen Welt herausgegriffener Mensch in einer anderen Welt ein Zombie wäre. Bewusstsein zu haben ist eine wesentliche Eigenschaft des aktuell existierenden Menschen. Wenn wir Kripkes semantische Analyse mit einer solchen Metaphysik untermauern, dann ist die Frage, was logisch möglich ist, tatsächlich metaphysisch irrelevant. Rein logisch mögliche Possibilia, wie z.B. Zombies oder andere Fabelwesen, sind unsere begrifflichen Konstrukte. Ihre Individualität wird allein durch unsere meist sehr unvollständige Beschreibung

IMPORTANT: When citing this article, please refer to the print-version:

Köhler, W. / H. Mutschler (eds.): *Ist der Geist berechenbar?* Darmstadt: 2004, 58-76.

konstituiert. Solche vagen Ideen, so der Einwand, sagen wenig über das, was wirklich ist. Sie haben kein metaphysisches Gewicht.

Ob dieser Ausweg gegen die Vorstellbarkeitsargumente greift, ist eine schwierige Frage, die über den Gegenstand dieses Aufsatzes weit hinausreicht. Im vorliegenden Kontext interessiert die Frage, ob der Funktionalismus durch diese Argumentation gestützt werden könnte. Durch die starke Trennung der epistemischen Ebene - wie wir etwas beschreiben - und der metaphysischen Ebene - was etwas ist - läßt die Argumentation wiederum das nun schon mehrfach beschriebene Schlupfloch zu. Es könnte sein, daß die komplette physische Beschreibung die Essenz eines Lebewesens gar nicht voll erfaßt. Wir greifen es ja durch Ostension heraus. Dies da, das wir durch Ostension herausgegriffen und physikalisch-funktional beschrieben haben, könnte durchaus auch noch phänomenal-intrinsische Eigenschaften haben, die durch diese Beschreibung nicht erfaßt wurden. Wenn es in allen metaphysisch möglichen Welten Bewusstsein hat, also niemals ein Zombie sein kann, liegt das daran, daß es über genau diese intrinsisch-phänomenalen Eigenschaften verfügt. Der Rekurs auf die Trennung von logischer und metaphysischer Möglichkeit vermag diesem Einwand also nicht zu entgehen. Ein anderes Problem einer solchen Konzeption liegt darin, daß man zu einem modalen Dualismus von logischer und metaphysischer Möglichkeit und Notwendigkeit gezwungen ist. Der gesamte Bereich logischer Möglichkeit und Notwendigkeit wird durch ein Netzwerk von Begriffen bestimmt, die man „rationale Begriffe“ nennen könnte: logische Konsistenz, logische Folgerung, positive Vorstellbarkeit, negative Vorstellbarkeit etc. Sie bestimmen ganz wesentlich das, was wir unter „Rationalität“ verstehen. Wenn man metaphysische Möglichkeit und Notwendigkeit davon abkoppelt, trennt man sie auch von der Rationalität ab. Was möglich und unmöglich ist, wird der rationalen Analyse entzogen, und damit werden natürlich auch unsere modalen Intuitionen unzuverlässig. Auf der anderen Seite ist der Bereich der Rationalität mit seinen logischen Modalbegriffen aus leicht einsehbaren Gründen unverzichtbar. Es bleibt dann nichts anderes übrig, als zwei getrennte Bereiche der Modalität anzunehmen. Ein modaler Dualismus von logischer und metaphysischer Notwendigkeit ist die Folge. Damit wird wiederum der Graben zwischen epistemisch-rationaler und metaphysischer Ebene aufgerissen, der das Erkennungszeichen schlechter Metaphysik ist.

3.2 Methodologische Lesart des stark-reduktiven Funktionalismus

Kann der stark-reduktive Funktionalismus trotz dieser Einwände ein vielversprechendes Forschungsprogramm sein? Ohne Zweifel liegen die metaphysischen Intuitionen des stark-reduktiven Funktionalismus vielen erfolgreichen Reduktionsprogrammen in der Geschichte der Naturwissenschaft zugrunde. Die intuitiv einleuchtende These besagt, daß man die Grundstrukturen des Universums nicht auf einer abstrakten Ebene erkennen kann, die unabhängig von der materiellen Realisierung ist. Je mehr man in die Feinmechanik, den Mechanismus der materiellen Strukturen eindringt, desto näher ist man am realen Geschehen. Ich hatte aber versucht aufzuzeigen, daß diese Strategie bei mentalen Zuständen an ihre Grenzen stößt. Wenn man aber auf eine metaphysische Gesamterklärung des Mentalen verzichtet, dann bietet das funktionalistische Programm vielversprechende Perspektiven. Abschließend soll wieder an einem kleinen Beispiel aufgezeigt werden, daß eine nicht-metaphysische Lesart des stark-reduktiven Funktionalismus die aufgezeigten Schwierigkeiten umgeht: Was könnte eine geeignete Strategie sein, um die sogenannten „NCCs“, die neuronalen Korrelate des Bewußtseins zu finden?

Die funktionalistische Strategie läßt sich hier vielversprechend anwenden. Wir sollten zunächst irgendeine funktionale Beschreibung von Bewußtsein formulieren, die wir noch nicht aus der empirischen Forschung, sondern durch begriffliche oder phänomenale Analyse gewinnen. Man braucht eine erste (vorwissenschaftliche) funktionale Operationalisierung, um ein empirisches

IMPORTANT: When citing this article, please refer to the print-version:

Köhler, W. / H. Mutschler (eds.): *Ist der Geist berechenbar?* Darmstadt: 2004, 58-76.

Forschungsprojekt überhaupt durchführen zu können. Wir können zum Beispiel sagen, daß ein Ereignis dann bewußt ist, wenn es direkt und umfassend verfügbar ist für die Kontrolle des Verhaltens. Die These, daß etwas in einen „*global workspace*“ eintreten muss, um dann umfassend für die Verhaltenssteuerung verfügbar zu sein, ist eine der stärksten Intuitionen über die funktionale Rolle des Bewußtseins. Wenn man diese Intuition als eine Art heuristisches Prinzip oder „Brückenprinzip“ annimmt, dann kann man in einem zweiten Schritt empirisch nach dem neuronalen Korrelat globaler Verfügbarkeit suchen. In einem dritten Schritt kann man nun Bewußtsein mit diesem neurophysiologischen Geschehen korrelieren. (Vgl. dazu auch Chalmers 1998, 219-230)

- 1) Bewußtsein ↔ globale Verfügbarkeit (Annahme aus begrifflicher o. phänomenaler Analyse)
- 2) Globale Verfügbarkeit ↔ neuronaler Prozeß (empirische Forschung)
- 3) Bewußtsein ↔ neuronaler Prozeß

Was wir mit „↔“ genau meinen, können wir getrost den Metaphysikern überlassen. Man wird für die Forschungspraxis auf jeden Fall mit einer Beziehung auskommen können, die schwächer ist als Identität, und so die Probleme des metaphysischen Funktionalismus vermeiden. Da wir mehrere Brückenprinzipien formulieren können und wir vermutlich auch mehr als ein neuronales Korrelat finden werden, sind auch hier Eins-zu-Eins-Zuordnungen nicht zu erwarten. Das stark-reduktive metaphysische Programm wird hier also nicht vorausgesetzt. Vor allem kann man durch diese Strategie keine vollständige Erklärung des Bewußtsein erwarten. Das Brückenprinzip wurde ja nur als heuristisches Werkzeug eingeführt. Die Frage, warum globale Verfügbarkeit mit Bewußtsein verbunden, bzw. die Frage nach der Möglichkeit von Zombies mit globaler Verfügbarkeit wurde gar nicht gestellt. Es geht hier also nicht um das metaphysische Ansinnen, A-posteriori-Notwendigkeiten zu entdecken oder das harte Problem des Bewußtseins zu lösen. Es geht darum, Korrelationen zu entdecken, die hoffentlich irgendwann einmal genügend abgesichert sind, um als Naturgesetze formuliert werden zu können. Mehr will ein methodologischer Funktionalismus nicht erreichen. In einem sehr schwachen und heute seltener gebrauchten Sinn von „Reduktion“ könnten die so entdeckten Brückengesetze auch als hinreichend für eine reduktive Erklärung betrachtet werden. Aber wir sollten genau beachten, was wir hier erklären. Wir erklären das Bewußtsein, das wir für unsere Zwecke handlich in eine operationale Definition gepackt hatten, die wir vorgängig zur empirischen Forschung durch begriffliche oder phänomenale Analyse gewonnen hatten. Wir setzen also das Faktum des Bewußtseins voraus. Der Blick auf die entdeckten funktionalen Strukturen im Gehirn würde aus sich heraus niemals zum Begriff des phänomenalen Erlebens führen. Die explanatorische Kluft wird nicht völlig überwunden. Es bleibt ein Rest von „Undurchsichtigkeit von unten nach oben“. Die entdeckten Brückenprinzipien oder Brückengesetze gelten nicht in allen logisch möglichen Welten. Der psycho-physische Nexus in unserer Welt bleibt ein *factum brutum*, das man wie die frühen Emergenztheoretiker mit einer Grundhaltung der „natural piety“ hinzunehmen hat. Die metaphysischen Probleme sind also nicht gelöst. Aber darauf kommt es dem methodologischen Funktionalismus ja gar nicht an. Für die empirische Wissenschaft bleibt er alternativlos und vielversprechend. Die funktionale Komplexität des Gehirns garantiert ihr einen schier unerschöpflichen Forschungsgegenstand.

IMPORTANT: When citing this article, please refer to the print-version:

Köhler, W. / H. Mutschler (eds.): *Ist der Geist berechenbar?* Darmstadt: 2004, 58-76.

Literatur

- Chalmers, David (1998): On the Search for the Neural Correlate of Consciousness, in: ders. *Toward a Science of Consciousness, The Second Tucson Discussions and Debates*, Cambridge.
- Chalmers, David (2002): The Components of Content, in: ders., *Philosophy of Mind: Classical and Contemporary Readings*, Oxford, 608-633
- Chalmers, David: Mind and Modality, (nur im Internet veröffentlicht, 1. September 2010) in: <http://consc.net/papers/mm.html>
- Cramer, K. / Fulda, H.F. / Horstmann, R.P. / Pothast, U. (1987) (Hg.): *Theorie der Subjektivität*, Frankfurt.
- Fodor, J. / Pylyshyn Z. (1998): Connectionism and Cognitive Architecture: A Critical Analysis, in: *Cognition* 28, 3-71
- Kim, J. (1993): The Nonreductivist's Trouble with Mental Causation, in: ders., *Supervenience and Mind. Selected Philosophical Essays*, Cambridge, 336-367
- McGinn, Colin (1989): Can We Solve the Mind-Body Problem?, in: *Mind* 98, 266-349
- Nagel, Thomas (1998): Conceiving the Impossible and the Mind-Body Problem, in: *Philosophy (Royal Institute of Philosophy)* 73, 337-352
- Putnam, Hilary (1975): Philosophy and Our Mental Life, in: ders., *Mind, Language and Reality, Philosophical Papers Vol. 2*, Cambridge MA, 291-303
- Putnam, Hilary (1987): Meaning Holism and Epistemic Holism, in: *Cramer 1987*, 251-277
- Searle, John (1993): The Problem of Consciousness, in: *Social Research* 60, 3-16