

Axel Ziemke

Gehirn und Bewusstsein

Eine Unterrichtsreihe über die Analytische Philosophie des Geistes

Materialsammlung

M 1 Merkmale mentaler Zustände

Es gibt vier Merkmale geistiger Phänomene, deretwegen es unmöglich schien, sie mit unserer „wissenschaftlichen“ Auffassung in Einklang zu bringen, nach der die Welt aus materiellen Dingen besteht. Und genau diese vier Merkmale haben das Körper-Geist-Problem wirklich schwierig gemacht [...]

Das wichtigste dieser Merkmale ist Bewusstsein. Ich, während ich dies schreibe, tue das mit Bewusstsein, und Sie, während Sie es lesen, tun das ebenfalls mit Bewusstsein. Es ist halt eine simple Tatsache, dass die Welt solche bewussten Zustände und Ereignisse des Geistes enthält; es ist allerdings schwer zu sehen, wie es möglich ist, dass rein materielle Systeme Bewusstsein haben. Wie könnte so etwas sein? Wie könnte es beispielsweise sein, dass dieser graue und weiße Glibber in meinem Schädel Bewusstsein hat? [...]

Das zweite dieser Merkmale des Geistes, mit dem schwer zurechtzukommen ist, ist das, was Philosophen und Psychologen „Intentionalität“ nennen: das, womit unsere Geisteszustände auf andere Gegenstände oder Sachverhalte gerichtet sind, von ihnen handeln, sich auf sie beziehen, über sie gehen. „Intentionalität“ bezieht sich, beiläufig gesprochen, nicht bloß auf Intentionen (also Absichten), sondern auch auf Überzeugungen, Wünsche, Hoffnungen, Befürchtungen, Liebe, Hass, Begierde, Ekel, Scham, Stolz, Irritation, Amüsiertheit und alle anderen Geisteszustände (ob bewusst oder unbewusst), die sich auf die außergeistige Welt beziehen oder von ihr handeln. Die Frage nach der Intentionalität ist nun der Frage nach dem Bewusstsein sehr ähnlich. Wie kann dieses Zeug in meinem Kopf von irgendetwas handeln? Wie kann es sich auf irgendetwas beziehen? Letzten Endes besteht der Inhalt meines Schädels ja aus „Atomen in der Leere“, wie auch die gesamte übrige Realität aus Atomen in der Leere besteht. Wie können denn nun, grob gesagt, Atome in der Leere irgendetwas repräsentieren?

Das dritte Merkmal des Geistes, für das sich anscheinend schwer ein Platz in einer wissenschaftlichen Auffassung von der Welt finden lässt, ist die Subjektivität von Geisteszuständen. Diese Subjektivität ist z.B. dadurch gekennzeichnet, dass ich meine Schmerzen fühlen kann, ein anderer aber nicht. Ich sehe die Welt von meinem Standpunkt aus; Sie sehen sie von Ihrem Standpunkt aus. Meiner selbst und meiner inneren Geisteszustände bin ich mir als etwas bewusst, das vollkommen unterschieden ist vom Selbst und den inneren Zuständen anderer Menschen. Seit dem 17. Jahrhundert denken wir uns die Wirklichkeit als etwas, das allen kompetenten Beobachtern gleichermaßen zugänglich sein muss – d.h. wir nehmen an, die Realität müsse objektiv sein. Wie sollen wir nun der Realität subjektiver Geisteszustän-

de in einer wissenschaftlichen Auffassung Rechnung tragen, nach der Realität etwas vollständig Objektives ist?

Schließlich ist da ein viertes Problem, das Problem der geistigen Verursachung. Wir alle nehmen an – und das gehört zu den Annahmen des gesunden Menschenverstandes –, dass es für unser Verhalten wirklich darauf ankommt, welche Gedanken und Gefühle wir haben. Ich entschlief mich beispielsweise dazu, meinen Arm zu heben und – siehe da – mein Arm geht hoch. Doch wenn unsere Gedanken und Gefühle wahrhaft geistig sind, wie können sie sich dann auf irgendetwas Materielles auswirken? Wie könnte sich aus etwas Geistigem eine materielle Veränderung ergeben? Sollen wir annehmen, dass unsere Gedanken und Gefühle irgendwie chemische Wirkungen auf unser Hirn und das restliche Nervensystem hervorbringen? Wie könnte so etwas sein? Sollen wir annehmen, dass Gedanken in der Lage sind, sich um die Axone zu wickeln oder die Dendriten zu schütteln oder in die Zelle zu schlüpfen und auf den Zellkern loszugehen?

Dieser vier Merkmale wegen – Bewusstsein, Intentionalität, Subjektivität und geistige Verursachung – wirkt das Körper-Geist-Problem so schwierig. Und dennoch möchte ich sagen: Es handelt sich bei ihnen allen um wirkliche Merkmale unseres geistigen Lebens. Nicht jeder Geisteszustand hat alle vier. Aber jede befriedigende Theorie des Geistes und der Körper-Geist-Beziehungen muss alle vier berücksichtigen. Wenn Ihre Theorie am Ende irgendeines dieser Merkmale leugnet, dann wissen Sie, dass Sie irgendwo einen Fehler gemacht haben müssen.

(aus: John R. Searle, Geist, Hirn und Wissenschaft, Frankfurt a.M.: Suhrkamp 1986, S. 14-16)

M 2 Kritik des Dualismus

Auf Grund der Probleme eines Cartesianischen Dualismus, besonders seiner Unfähigkeit, einen angemessenen Ansatz für die Beziehung zwischen Geist und Körper zu entwickeln, wird weitestgehend angenommen, dass ein Substanzdualismus, in welcher Form auch immer, nicht mehr zur Diskussion steht. Das heißt nicht, dass keine seriösen Forscher mehr Substanzdualisten wären. Doch meiner Erfahrung nach sind die meisten Substanzdualisten, die ich kenne, Menschen, die diesen Standpunkt aus religiösen Gründen vertreten oder als Teil eines religiösen Glaubens. Es ist eine Konsequenz des Substanzdualismus, dass unsere Seele weiter leben kann, wenn unser Körper zerstört ist; und das macht diese Sichtweise attraktiv für Anhänger von Religionen, die an ein Leben nach dem Tod glauben. Doch unter den meisten Forschern auf diesem Gebiet wird der Substanzdualismus nicht als eine ernsthafte Möglichkeit betrachtet [...]

Alle Formen des Substanzdualismus erben Descartes' Problem, einen kohärenten Ansatz für die kausalen Beziehungen zwischen Seele und Körper zu geben, aber moderne Versionen haben noch ein zusätzliches Problem. Es scheint unmöglich, den Substanzdualismus mit der modernen Physik in Übereinstimmung zu bringen. Die Physik sagt, dass die Menge von Materie bzw. Energie im Universum konstant ist; aber der Substanzdualismus scheint zu implizieren, dass es noch eine andere

Form von Energie gibt, geistige Energie oder spirituelle Energie, die von der Physik nicht erfasst wird. Wenn also der Substanzdualismus wahr wäre, dann müsste eines der fundamentalsten Gesetze der Physik, der Energieerhaltungssatz falsch sein [...]

Es ist wichtig zu verstehen, was für eine extreme Doktrin der Substanzdualismus ist. Dem Substanzdualismus zufolge sind unsere Gehirne und Körper nicht wirklich bewusst. Ihr Körper ist genau so eine unbewusste Maschine wie Ihr Auto oder Ihr Fernsehgerät. Ihr Körper ist im gleichen Sinne lebendig wie Pflanzen lebendig sind, aber es gibt kein Bewusstsein in ihm. Vielmehr wurde Ihre bewusste Seele irgendwie mit Ihrem Körper verbunden und bleibt mit ihm verbunden bis Ihr Körper stirbt und Ihre Seele entweicht. Sie sind identisch mit Ihrer Seele und bewohnen nur zufällig und vorläufig diesen Körper.

Das Problem mit dieser Sichtweise ist, dass es angesichts dessen was wir über die Welt wissen, schwer ist, all das als eine wissenschaftliche Hypothese ernst zu nehmen. Wir wissen, dass Bewusstsein im Menschen überhaupt nicht existieren kann ohne bestimmte physikalische Prozesse, die im Gehirn ablaufen. Wir könnten im Prinzip in der Lage sein, Bewusstsein auch in einer anderen physischen Substanz zu erzeugen, aber derzeitig haben wir noch keine Ahnung, wie man das machen könnte. Doch die Idee, dass Bewusstsein ohne ein physisches Substrat hervorgebracht werden könnte, scheint, wenn sie auch denkbar sein mag, völlig außer Frage als wissenschaftliche Hypothese.

(übersetzt aus: John R. Searle, Mind. A Brief Introduction, New York: Oxford Univ. Press 2004, S. 29-30)

M 3 Mentale Zustände als physikalische Zustände des Gehirns

Jedes geistige Phänomen – ob bewusst oder unbewusst, visuell oder auditiv, Schmerzen, Kitzel, Jucken, Gedanken, ja tatsächlich die Gesamtheit unseres geistigen Lebens – ist von Vorgängen im Gehirn verursacht [...]

Wenn aber Schmerzen und andere Phänomene von Vorgängen im Gehirn verursacht werden, dann möchte man wissen: Was sind Schmerzen? Was sind sie wirklich? Nun, im Falle von Schmerzen liegt die Antwort auf der Hand: Es sind Empfindungen unangenehmer Art. Diese Antwort genügt uns allerdings nicht, denn sie sagt uns nicht, wie Schmerzen in unsere umfassende Vorstellung von der Welt hineinpassen. Zu unserer ersten Behauptung – dass Schmerzen und andere geistige Phänomene von Hirnvorgängen verursacht sind – müssen wir eine zweite Behauptung hinzunehmen:

Schmerzen und andere geistige Phänomene sind nichts anderes als Eigenschaften des Hirns (und vielleicht des übrigen zentralen Nervensystems).

Eines der Hauptziele dieses Kapitels ist es zu zeigen, wie es sein kann, dass diese beiden Behauptungen zugleich wahr sind. Wie kann es sowohl sein, dass das Hirn den Geist verursacht, als auch trotzdem, dass der Geist bloß eine Reihe von Eigenschaften des Gehirns ist? [...]

Es ist verlockend zu denken, dass immer zwei wohl unterschiedene Ereignisse vorliegen müssen, falls A B verursacht: das eine Ereignis ist als Ursache identifiziert, das andere als die Wirkung. Es ist verlockend anzunehmen, alle Verursachung funktioniere genau so, wie das beim Aufprall einer Billardkugel auf eine andere der Fall ist. Auf Grund dieses groben Modells von Kausalbeziehungen zwischen Hirn und Geist akzeptieren wir leicht eine Form des Dualismus; wir neigen zu der Auffassung, Ereignisse in einem materiellen Bereich (dem „Körperlichen“) verursachten Ereignisse in einem Immateriellen Bereich (dem „Geistigen“). Doch das scheint mir der Fehler zu sein. Und den Fehler beseitigt man, indem man sich einen differenzierteren Begriff der Verursachung besorgt. Zu diesem Zweck werde ich mich für einen Augenblick von den Beziehungen zwischen Geist und Hirn abwenden, um Kausalbeziehungen anderer Art in der Natur zu betrachten.

In der Physik ist die Unterscheidung von Mikro- und Makroeigenschaften von Systemen geläufig – die kleinen und die großen Maßstäbe. Betrachten wir beispielsweise den Schreibtisch, an dem ich gerade sitze, oder das Glas Wasser vor mir. Jedes Objekt ist aus Mikropartikeln zusammengesetzt. Die Mikropartikel haben Eigenschaften auf der Ebene der Moleküle und Atome, und zudem auf der tieferen Ebene der Teilchen. Aber jedes Objekt hat auch gewisse Eigenschaften, die Oberflächeneigenschaften oder globale Eigenschaften des materiellen Systems sind: So ist der Tisch fest, das Wasser flüssig und das Glas durchsichtig. Viele der globalen oder Oberflächeneigenschaften lassen sich durch das Verhalten der Bestandteile auf der Mikroebene kausal erklären [...]

Meiner Ansicht nach liefert das ein völlig normales Modell für die Erklärung der rätselhaften Beziehungen zwischen Geist und Hirn. Im Falle der des Flüssigseins, der Festigkeit und der Durchsichtigkeit haben wir überhaupt keine Schwierigkeiten anzunehmen, dass die Oberflächeneigenschaften durch das Verhalten der Bestandteile auf der Mikroebene verursacht sind, und zugleich akzeptieren wir, dass die Oberflächenphänomene einfach Eigenschaften des fraglichen Systems sind. Am klarsten wird dies m.E. so formuliert: Die Oberflächeneigenschaft wird sowohl vom Verhalten der Mikrobestandteile verursacht, als auch gleichzeitig in dem System realisiert, das aus den Mikrobestandteilen besteht. Es gibt da eine Ursache-Wirkungs-Beziehung, doch zugleich sind die Oberflächeneigenschaften einfach Eigenschaften höherer Ebene desselben Systems, dessen Verhalten auf der Mikroebene diese Eigenschaften verursacht [...]

Genau so – und genau in demselben Sinne von „verursacht“ und „realisiert“ – sind geistige Phänomene von Hirnvorgängen auf der Ebene der Nervenzellen oder neuronalen Moduln verursacht, und zugleich in diesem aus Nervenzellen bestehenden System realisiert. Und wie wir die Mikro-Makro-Unterscheidung für jedes beliebige materielle System brauchen, genau so – und aus genau denselben Gründen brauchen wir diese Unterscheidung auch für das Hirn. Wir können zwar von einem System von Teilchen sagen, seine Temperatur sei 10°C oder es sei fest oder es sei flüssig; wir können aber nicht von einem einzelnen Teilchen sagen, es sei fest, es sei flüssig oder seine Temperatur sei 10°C [...] Genau so verhält es sich mit dem Hirn (soweit wir überhaupt etwas darüber wissen): Wir können zwar von einem

einzelnen Hirn sagen: „Dieses Hirn hat Bewusstsein“ oder „Dieses Hirn erlebt gerade Durst oder Schmerz“, wir können aber nicht von einer einzelnen Nervenzelle im Gehirn sagen: „Diese Nervenzelle hat Schmerzen, diese hat ein Dursterlebnis“. Um es zu wiederholen: Es gibt zwar enorme empirische Geheimnisse über die Einzelheiten der Funktionsweise des Hirns, es gibt aber keine logischen, philosophischen oder metaphysischen Hindernisse, die Beziehung zwischen Geist und Hirn auf eine Weise zu erklären, die uns von der übrigen Natur her völlig vertraut ist. Nichts ist verbreiteter in der Natur als dass Oberflächeneigenschaften eines Phänomens sowohl von einer Mikrostruktur verursacht als auch in ihr realisiert sind. Und genau solche Beziehungen manifestieren sich in der Beziehung vom Geist zum Gehirn.

(aus: John R. Searle, *Geist, Hirn und Wissenschaft*, Frankfurt a.M.: Suhrkamp 1986, S. 17-22)

M 4 Wie es ist, eine Fledermaus zu sein

Es ist das Thema “Bewusstsein, welches das Leib-Seele-Problem wirklich vertrackt macht. [...] Die jüngste Welle reduktionistischer Euphorie hat einige Analysen mentaler Phänomene und mentalistischer Begriffe hervorgebracht, die dazu bestimmt sind, die Möglichkeit irgendeiner Art von Materialismus, psychophysischer Identifikation oder Reduktion zu erklären. Aber die behandelten Probleme sind solche, die dieser und anderen Arten von Reduktion gemein sind. Was aber das Leib-Seele-Problem auszeichnet und von dem Wasser-H₂O-Problem oder von dem Turingmaschinen-IBM-Maschinen-Problem [...] oder von dem Gen-DNS-Problem unterscheidet, wird außer Acht gelassen [...]. Ich werde zu erklären versuchen, warum die gewöhnlichen Beispiele uns nicht helfen, das Verhältnis des Mentalen zum Physischen zu verstehen – warum wir gegenwärtig in der Tat keine Konzeption von etwas haben, das eine Erklärung der physikalischen Natur eines mentalen Phänomens wäre [...]

Bewusste Erfahrung ist ein weit verbreitetes Phänomen. Sie taucht auf vielen Ebenen tierischen Lebens auf, obgleich wir nicht sicher sein können, dass sie in einfacheren Organismen vorkommt, und es sehr schwer ist, im allgemeinen zu sagen, was Indizien für sie liefert. [...] Zweifellos taucht sie in zahllosen Formen auf, die für uns ganz und gar unvorstellbar sind – auf anderen Planeten in anderen Sonnensystemen überall im Universum. Aber ganz gleich wie die Formen voneinander abweichen mögen: Die Tatsache, dass ein Organismus *überhaupt* bewusste Erfahrung hat, heißt im wesentlichen, dass es irgendwie *ist*, dieser Organismus zu *sein*. Es mag weitere Implikationen bezüglich der Form der Erfahrung geben; es mag sogar (obwohl ich es bezweifle) Implikationen bezüglich des Verhaltens des Organismus geben. Grundsätzlich aber hat ein Organismus bewusste mentale Zustände dann und nur dann, wenn es irgendwie *ist*, dieser Organismus zu *sein* – wenn es irgendwie *für* diesen Organismus ist. Wir können dies den subjektiven Charakter von Erfahrung nennen.

[...]

Eine Theorie der physikalischen Grundlage des Mentalen muss vieles erklären. Dies scheint mir aber das Schwerste zu sein. [...] Wenn der Physikalismus verteidigt werden soll, müssen phänomenologische Eigenschaften selbst physikalisch erklärt werden. Wenn wir aber ihren subjektiven Charakter untersuchen, scheint so etwas unmöglich zu sein. Der Grund dafür ist, dass jedes subjektive Phänomen mit einer einzelnen Perspektive verbunden ist; und es scheint unvermeidlich, dass eine objektive physikalische Theorie von dieser Perspektive abstrahieren wird. [...] Um die Verknüpfung zwischen Subjektivität und einer Perspektive zu veranschaulichen und um die Wichtigkeit subjektiver Eigenschaften deutlich zu machen, wird es helfen, den Sachverhalt anhand eines Beispiels zu untersuchen, das die Verschiedenheit der subjektiven und objektiven Betrachtungsweise klar herausstellt.

Ich nehme an: wir alle glauben, dass Fledermäuse Erlebnisse haben. Schließlich sind sie Säugetiere, und es gibt keinen größeren Zweifel daran, dass sie Erlebnisse haben als daran, dass Mäuse, Tauben oder Wale Erlebnisse haben. Ich habe Fledermäuse gewählt statt Wespen oder Flundern, weil man das Vertrauen darauf, dass es da Erlebnisse gibt, schrittweise verliert, wenn man den phylogenetischen Baum zu weit nach unten klettert. Obwohl Fledermäuse uns näher verwandt sind als diese anderen Arten, weisen sie einen Sinnesapparat und eine Reihe von Aktivitäten auf, die von den unsrigen so verschieden sind, dass das Problem, das ich stellen möchte, besonders anschaulich ist. [...]

Ich habe gesagt, das Wesentliche an dem Glauben, dass Fledermäuse Erlebnisse haben sei, dass es irgendwie ist, eine Fledermaus zu sein. Heute wissen wir, dass die meisten Fledermäuse [...] die Außenwelt primär durch Radar oder Echoortung wahrnehmen, indem sie das von Objekten in ihrer Reichweite zurückgeworfene Echo ihrer raschen und kunstvoll modulierten Hochfrequenzschreie registrieren. Ihre Gehirne sind dazu bestimmt, die Ausgangsimpulse mit dem darauf folgenden Echo zu korrelieren. Die so erhaltene Information befähigt Fledermäuse, eine genaue Unterscheidung von Abstand, Größe, Gestalt, Bewegung und Struktur vorzunehmen, die derjenigen vergleichbar ist, die wir beim Sehen machen. Obwohl der Fledermaus-Radar klarerweise eine Form von Wahrnehmung ist, ist er in seinem Funktionieren keinem der Sinne ähnlich, die wir besitzen. Auch gibt es keinen Grund zu der Annahme, dass es subjektiv so wie irgendetwas ist, das wir erleben oder das wir uns vorstellen können. Das scheint für den Begriff davon, wie es ist, eine Fledermaus zu sein, Schwierigkeiten zu bereiten. Wir müssen überlegen, ob uns irgendeine Methode erlauben wird, das Innenleben der Fledermaus aus unserem eigenen Fall zu erschließen, und falls nicht, welche alternativen Methoden es geben mag, um sich davon einen Begriff zu machen.

Unsere eigene Erfahrung liefert die grundlegenden Bestandteile für unsere Phantasie, deren Spielraum deswegen beschränkt ist. Es wird nicht helfen, sich vorzustellen, dass man Flughäute an den Armen hätte, die einen befähigten, bei Anbruch der Dunkelheit und im Morgengrauen herumzufliegen, während man mit dem Mund Insekten finge; dass man ein schwaches Sehvermögen hätte und die Umwelt mit einem System reflektierter akustischer Signale aus dem Hochfrequenzbereich wahr-

nähme und dass man den Tag an den Füßen nach unten hängend in einer Dachkammer verbrächte. Insoweit ich mir dies vorstellen kann (was nicht sehr weit ist), sagt es mir nur, wie es für *mich* wäre, mich so zu verhalten, wie sich eine Fledermaus verhält. Das aber ist nicht die Frage. Ich möchte wissen, wie es für eine *Fledermaus* ist, eine Fledermaus zu sein. Wenn ich mir jedoch dies nur vorzustellen versuche, bin ich auf die Ressourcen meines eigenen Bewusstseins angewiesen, und diese Ressourcen sind für das Vorhaben unzulänglich [...]

Die Perspektive, um die es mir geht, ist nicht etwas, das nur einem einzelnen Individuum zugänglich ist. Es handelt sich vielmehr um einen Typus. es ist oft möglich, eine andere als die eigenen Perspektive einzunehmen, so dass das Erfassen von solchen Tatsachen nicht auf den eigenen Fall beschränkt ist. Es gibt einen Sinn, in dem phänomenologische Tatsachen völlig objektiv sind: eine Person kann von einer anderen Person wissen oder sagen, welche Qualität das Erlebnis der anderen hat. Dennoch sind sie in dem Sinne subjektiv, dass diese objektive Zuschreibung von Erlebnissen nur für jemanden möglich ist, der dem Objekt der Zuschreibung ähnlich genug ist, um dessen Perspektive einnehmen zu können – um sozusagen die Zuschreibung in der ersten Person ebenso gut zu verstehen wie die in der dritten. Je verschiedener das andere Wesen von einem selbst ist, desto weniger Erfolg kann man von diesem Versuch erwarten. In unserem eigenen Fall nehmen wir die maßgebliche Perspektive ein; wir werden aber dann, wenn wir uns ihr von einer anderen Perspektive nähern würden, ebenso viele Schwierigkeiten haben wie dann, wenn wir versuchten, die Erlebnisse einer anderen Spezies zu verstehen, ohne deren Perspektive einzunehmen.

Dies ist für das Leib-Seele-Problem unmittelbar relevant. Wenn nämlich Erlebnis-tatsachen – Tatsachen bezüglich dessen, wie es für den Organismus ist, dieser Organismus zu sein – nur einer bestimmten Perspektive zugänglich sind, dann ist es ein Rätsel, wie der wahre Charakter von Erlebnissen in der Funktionsweise dieses Organismus entdeckt werden könnte. Diese Funktionsweise gehört zu einer Perspektive objektiver Tatsachen *par excellence* – einer Art von Tatsachen, die aus verschiedenen Perspektiven und von Individuen mit verschiedenen Wahrnehmungssystemen beobachtet und verstanden werden können. Nichts Vergleichbares steht der Perspektive im Wege, dass menschliche Wissenschaftler Wissen über die Neurophysiologie der Fledermaus erwerben; und intelligente Fledermäuse oder Marsmenschen könnten vielleicht mehr über das Gehirn lernen, als wir jemals lernen werden.

Für sich allein ist das noch kein Argument gegen Reduktion. Ein Wissenschaftler vom Mars ohne Verständnis für visuelle Wahrnehmung könnte den Regenbogen, den Blitz oder die Wolke als physikalische Phänomene verstehen, obwohl er niemals fähig wäre, die menschlichen Begriffe von Regenbogen, Blitz oder Wolke zu verstehen oder den Ort, den diese Dinge in unserer phänomenalen Welt einnehmen. Die objektive Natur der Dinge, die durch diese Begriffe herausgegriffen werden, könnte von ihm verstanden werden, weil diese Dinge nicht, wie die Begriffe von ihnen, an eine besondere Perspektive und an eine besondere visuelle Phänomenologie geknüpft sind: Sie können von einer bestimmten Perspektive aus beobach-

tet werden, sind aber von ihr unabhängig; daher können sie auch von anderen Perspektiven aus erfasst werden, entweder von den gleichen oder von anderen Organismen [...]

Im Falle von Erlebnissen scheint auf der anderen Seite die Verbindung mit einer besonderen Perspektive viel enger zu sein. Es ist schwierig zu verstehen, was mit dem objektiven Charakter eines Erlebnisses gemeint sein könnte – unabhängig von der besonderen Perspektive, von der aus ein Subjekt sie erfasst. Was bliebe letzten Endes von der Weise übrig, wie es ist, eine Fledermaus zu sein, wenn man die Perspektive der Fledermaus entfernte? Wenn aber Erlebnisse nicht zusätzlich zu ihrem subjektiven Charakter eine objektive Natur haben, die von vielen verschiedenen Perspektiven aus erfasst werden kann, wie kann man dann annehmen, dass ein Marsmensch, der mein Gehirn untersuchte, physikalische Prozesse beobachten könnte, die meine mentalen Prozesse wären, nur eben von einer anderen Perspektive aus? Wie schließlich, könnte sie ein menschlicher Physiologe von einer anderen Perspektive aus beobachten?

[...] Welche Moral sollten wir diesen Überlegungen entnehmen, und was sollte als nächstes getan werden? Es wäre ein Fehler zu schließen, dass der Physikalismus falsch sein muss. Nichts ist durch die Unangemessenheit der physikalistischen Hypothesen bewiesen, die eine fehlerhafte objektive Analyse des Mentalen voraussetzen. Es wäre richtiger zu sagen, dass der Physikalismus eine Position ist, die wir nicht verstehen können, weil wir gegenwärtig keine Konzeption davon haben, wie er wahr sein könnte. Vielleicht wird man es für unbegründet halten, eine solche Konzeption als eine Voraussetzung dafür zu verlangen, dass man den Physikalismus verstehen kann. Schließlich mag man sagen, dass die Idee des Physikalismus klar genug ist: Mentale Zustände *sind* Zustände des Körpers, mentale Ereignisse *sind* physikalische Ereignisse. Wir wissen nicht, welche physikalischen Zustände und Ereignisse es sind; das sollte uns jedoch nicht davon abhalten, die Hypothese zu verstehen. Was könnte klarer sein als die Wörter „ist“ und „sind“?

Ich glaube aber, dass es genau diese scheinbare Klarheit des Wortes „ist“ ist, die trägt. Wenn wir gewöhnlich gesagt bekommen „X ist Y“, wissen wir, wie so etwas wahr sein könnte. Es hängt dann aber von einem begrifflichen oder theoretischen Hintergrund ab und wird nicht durch „ist“ alleine vermittelt [...] Zum Beispiel bekommt man heute in jungen Jahren gesagt, dass alle Materie in Wirklichkeit Energie ist. Obwohl man weiß, was „ist“ bedeutet, können sich die meisten niemals einen Begriff davon machen, was diese Behauptung wahr macht, weil ihnen der theoretische Hintergrund fehlt. Die Situation des Physikalismus ist gegenwärtig ähnlich wie die der Hypothese, dass Masse Energie ist, wenn sie von einem vorsokratischen Philosophen geäußert worden wäre. Wir haben nicht einmal ansatzweise eine Konzeption davon, wie er wahr sein könnte. Um die Hypothese zu verstehen, dass ein mentales Ereignis ein physikalisches Ereignis ist, müssen wir mehr als nur das Wort „ist“ verstehen. Es fehlt eine Konzeption davon, wie sich ein mentalistischer und ein physikalischer Term auf ein und denselben Gegenstand beziehen können und den gewöhnlichen Analogien mit theoretischen Identifikationen in anderen Gebieten gelingt es nicht, sie zu verschaffen.

(Übersetzt aus: Thomas Nagel, *What is it like to be a bat? Philosophical Review* 83, 435-450 (1974))

M 5 Mary

Mary ist eine brillante Wissenschaftlerin, die, warum auch immer, gezwungen ist, die Welt aus einem schwarz-weißen Raum über einen Schwarz-Weiß-Fernseher kennen zu lernen. Sie spezialisiert sich auf die Neurophysiologie des Sehens und erwirbt, stellen wir uns vor, alle physikalische Information, die man darüber gewinnen kann, was sich abspielt, wenn wir reife Tomaten oder den Himmel sehen und Worte wie „rot“ oder „blau“ verwenden. Sie entdeckt beispielsweise, welche Wellenlängenkombinationen des Himmels die Retina stimulieren und wie dies über das zentrale Nervensystem die Kontraktion der Stimmbänder und das Ausstoßen von Luft aus den Lungen bewirkt, das die Aussprache des Satzes „Der Himmel ist blau“ zur Folge hat [...].

Was wird geschehen, wenn Mary aus ihrem schwarz-weißen Raum befreit wird oder ihr ein Farbfernseher gegeben wird? Wird sie etwas *lernen* oder nicht? Es scheint offensichtlich, dass sie etwas über die Welt und unser visuelles Erleben dieser Welt lernen wird. Dann ist es aber unumgänglich, dass ihr vorangehendes Wissen unvollständig war. Sie verfügte aber über alle physikalische Information. Also gibt es mehr als das und der Physikalismus ist falsch.

Klarerweise könnte dieselbe Form von Argumenten des unvollständigen Wissens für das Schmecken, das Hören, die Körperempfindung und allgemein all die verschiedenen mentalen Zustände entwickelt werden, die phänomenale Merkmale oder Qualia aufweisen. Die Schlussfolgerung ist in jedem dieser Fälle, dass die Qualia von der physikalistischen Betrachtung ausgelassen werden. Und die polemische Stärke des Arguments des unvollständigen Wissens besteht darin, dass es schwer ist, seine zentrale Behauptung zu leugnen, man könne über alle physikalische Information verfügen ohne alle Information zu haben, die man haben muss.

(Übersetzt aus: Frank Jackson, *Epiphenomenal Qualia, Philosophical Quarterly*, 32 (1982), 127-36)

M 6 Subjektive Erfahrung

Keine noch so umfangreiche Beschreibung, so ausführlich sie sich auch gestalten mag, wird je imstande sein, subjektive Erfahrung vollständig darzustellen. Eine Menge Philosophen haben die Farbwahrnehmung als Beispiel herangezogen, um diesen Standpunkt klar zu machen. Keine wissenschaftliche Beschreibung neuraler Mechanismen zur Farbunterscheidung – so vollkommen sie auch gelingen mag – wird Sie je verstehen lassen, wie es sich für jemand anderen anfühlt, eine bestimmte Farbe wahrzunehmen [...]

Befindet sich eine befriedigende wissenschaftliche Behandlung von Bewusstsein damit für immer außer Reichweite? [...] Die Antwort hierauf hat unserer Ansicht nach mit einer realistischen Einschätzung dessen zu tun, was wissenschaftliche Erklärungen ganz allgemein zu leisten vermögen und was nicht. Wissenschaftliche Erklärungen [...] können die Eigenschaften des Phänomens erklären und sogar be-

gründen, warum das Phänomen unter eben diesen Bedingungen stattfindet. Keine wissenschaftliche Beschreibung oder Erklärung aber kann das real vorhandene Ding selbst ersetzen. Jeder von uns trägt dieser Tatsache Rechnung, wenn er etwas beschreibt, beispielsweise einen Wirbelsturm: Um welche Art von physikalischem Prozess es sich handelt, warum er die Eigenschaften hat, die er hat, und unter welchen Bedingungen er sich bildet. Niemand aber würde erwarten, dass eine wissenschaftliche Beschreibung eines Hurrikans selbst ein Hurrikan ist oder einen solchen verursacht.

Warum sollten wir nicht genau dieselben Maßstäbe an das Bewusstsein anlegen? Wir sollten eine adäquate Beschreibung dessen liefern, um was für einen physikalischen Vorgang es sich dabei handelt, warum es die Eigenschaften hat, die es hat, und unter welchen Umständen es entsteht. Wie wir sehen werden, ist am Bewusstsein nichts, was einer angemessenen wissenschaftlichen Beschreibung jener speziellen Art von neuronalen Abläufen, durch die es zustande kommt, im Wege stünde.

(Gerald Edelman, Giulio Tononi, aus: Gehirn und Geist. Wie aus Materie Bewusstsein entsteht, München: Deutscher Taschenbuchverlag 2004)

M 7 Der eliminative Materialismus

In dem Sinne, in dem „Reduktion“ für das Mind-Body-Problem relevant ist, ist Reduktion in erster Linie eine Beziehung zwischen Theorien. Im einfachsten Falle steht eine Theorie, die reduzierte Theorie T_R , in einer bestimmten Beziehung zu einer grundlegenderen Theorie T_G . Behauptungen, dass sich ein Phänomen P_R auf ein anderes Phänomen P_G reduzieren lässt, sind ableitbar aus der grundlegenderen Annahme, dass die Theorie, die das erstgenannte Phänomen kennzeichnet, reduziert werden kann auf die Theorie, die das als Zweites genannte Phänomen zum Gegenstand hat [...].

Wenn wir sagen, dass Reduktion eine Beziehung zwischen Theorien ist, dann müssen wir diese Beziehung näher kennzeichnen. Unter welchen Bedingungen kann man von einer Theorie sagen, dass sie auf eine andere reduziert wurde? Die entscheidende Annahme der Antwort der logischen Empiristen war, dass die Theorie T_R , um auf die Theorie T_G reduziert zu werden, aus der Theorie T_G *logisch abgeleitet* werden muss – allerdings unter Einbeziehung einiger Zusatzannahmen. Diese Zusatzannahmen beinhalteten bereichsspezifische Randbedingungen, begrenzende Annahmen, Annäherungen usw. Ebenso entscheidend waren aber die Überbrückungsprinzipien. Nach Überzeugung des logischen Empirismus war es die Funktion der Überbrückungsprinzipien, Eigenschaften, die durch die Gesetze von T_R beschrieben werden, mit Eigenschaften zu verbinden, die durch die Gesetze von T_G beschrieben werden. Im einfachsten Fall würden diese Überbrückungsprinzipien die Eigenschaften der reduzierten Theorie mit denen der grundlegenderen Theorie *identifizieren*.

[...]

In der Tat gibt es aber nicht ein einziges Beispiel in der Wissenschaftsgeschichte, das exakt diesem logisch-empiristischen Muster einer Reduktion folgt, und einige

herausragende Fälle entsprechen ihm überhaupt nicht [...] Das Problem ist, dass Theorien immer wieder korrigiert und modifiziert werden mussten, um aus der grundlegenden Theorie abgeleitet werden zu können. Angesichts der vorgenommenen Modifikationen war es natürlich nicht die alte Theorie selbst, die abgeleitet wurde, sondern bestenfalls eine *korrigierte Version* der alten Theorie. Theorien bewegen sich also auf einem Spektrum von Möglichkeiten, wie viel Korrektur und Umarbeitung sie benötigen, um eine ableitbare Form zu erhalten. Manche benötigen relativ geringfügige Korrekturen, um reduziert werden zu können (z.B. die Optik nach ihrer Reduktion auf die Theorie elektromagnetischer Wellen), bei anderen ist hingegen so viel Korrektur notwendig, dass fast nichts außer einigen belanglosen Verallgemeinerungen erhalten bleibt. Zwei solche Beispiele sind die Phlogiston-Theorie der Verbrennung oder die Theorie, dass Nervenranke von Dämonen besessen seien. [...] Das Spektrum weist entsprechend an einem Ende Theorien auf, die nach ihrer Reduktion weitgehend erhalten geblieben sind, am anderen Ende Theorien, die größtenteils ersetzt wurden, und die verschiedensten Fälle dazwischen [...]

Wie sollte man also die Bedingungen für eine intertheoretische Reduktion spezifizieren, wenn das logisch-empiristische Verständnis von Reduktion nicht flexibel genug und falsch ist? Im Prinzip so: Konstruiere im Rahmen der neuen reduzierenden Theorie T_G ein Analogon T_R^* der Gesetze der zu reduzierenden Theorie T_R . Das Analogon T_R^* , nicht aber T_R selbst kann dann logisch aus der reduzierenden Theorie T_G abgeleitet werden [...] Allgemein wird diese analoge Theorie in Hinblick auf die Zuordnung von Begriffen der alten Theorie zu Begriffen der neuen Theorie, Gesetzen der alten Theorie zu Sätzen (nicht notwendigerweise Gesetzen) der neuen Theorie konstruiert. Wenn die Reduktion erfolgreich ist, wird die neue Theorie die alte erklären. Sie wird erklären, warum die alte Theorie so gut funktionierte, wie sie es tat, und wo sie scheitern musste [...]

In dem Maße, in dem die neue Theorie die alte grundlegend ersetzt, verändert sich auch die Auffassung darüber, was erklärt werden muss. Ein vorangehender Konsens über das zu Erklärende löst sich auf und neue Explananda, die sich aus der neuen Theorie ergeben, beginnen die Forschung zu bestimmen. In der Ptolemäischen Theorie beispielsweise war die Erde das Zentrum des Universums, um das sich die himmlischen Sphären drehten [...] Tatsächlich meinte man beobachten zu können, wie sich die Sphären täglich um die unbewegte Erde drehten – eine Beobachtung die völlig klar und unproblematisch schien. Ptolemäische Theoretiker hatten also die Frage zu beantworten, was die Himmels sphären zum Drehen bringt [...] Als die Newtonsche Theorie aber schließlich ohne die Annahme himmlischer Sphären auskam, musste auch der Mechanismus ihrer Bewegungen nicht mehr erklärt werden und die Behauptungen, die Sphären und ihre Bewegung beobachtet zu haben, erwiesen sich als missverständlich und falsch [...]

Die Relevanz all dessen für eine künftige umfassende Theorie des Geist/Gehirns ist diese: Was wir im Rahmen einer intuitiven Psychologie als offensichtlich oder beobachtbar annehmen, garantiert keineswegs, dass es auch wahr ist [...]. Praktisch alle Argumente gegen den Reduktionismus, wenn sie nicht nur Unklarheiten über

den Anspruch intertheoretischer Reduktionen sind, beruhen darauf, Aspekte unseres alltäglichen Verständnisrahmens mentaler Zustände und Prozesse als korrekt und nicht reduzierbar anzunehmen [...] Der Kürze halber ersetze ich diese lang gewundene Bezeichnung mit einem kurzen Schlagwort, nämlich „Alltagspsychologie“. Mit Alltagspsychologie meine ich jenes grobe Netz von Begriffen, Verallgemeinerungen und Daumenregeln, die wir gemeinhin nutzen, um menschliches Verhalten zu erklären und vorherzusagen [...]

Die Angriffe gegen das reduktionistische Programm einer umfassenden Theorie des Geist/Gehirns, in der die psychologische Theorie letztendlich durch die neurobiologische Theorie erklärt wird, leiten sich aus zwei Quellen ab: (1) der Überzeugung, dass ein Teil der Verallgemeinerungen der wissenschaftlichen Psychologie solcherart sind, dass sie einer Reduktion widerstehen und (2) der Überzeugung, dass die Alltagspsychologie substanziell korrekt und nicht reduzierbar auf die Neurobiologie ist. Die Verallgemeinerungen der wissenschaftlichen Psychologie, die dabei als problematisch erachtet werden, stammen vor allem aus der Sozialpsychologie und der Kognitionspsychologie [...] beruhen auf einer Theorie mentaler Repräsentationen, die selbst wiederum der Alltagspsychologie entstammt.

Obwohl uns die Alltagspsychologie grundlegend vertraut und offensichtlich erscheint, und obwohl wir die Kategorien der Alltagspsychologie auf unsere Beobachtungen beziehen können, bleibt es dennoch zutreffend, dass die Alltagspsychologie ein theoretischer Rahmen ist und daher ein Rahmen, dessen Richtigkeit in Frage gestellt und überprüft werden muss. Die Angemessenheit der Alltagspsychologie ist in keiner Weise dadurch gesichert, dass sie überwältigend offensichtlich, auf Beobachtungen oder Introspektionen anzuwenden oder sogar angeboren sei, wenn letzteres der Fall sein sollte. Wenn wir verstehen, dass die Alltagspsychologie kein erkenntnistheoretisches Privileg, keine Immunität gegen Überarbeitung und Korrektur verdient, dann beginnen wir zu sehen, dass ihre Verallgemeinerungen und Kategorien korrigiert und verbessert werden können.

Wenn wir erst einmal erkannt haben, dass die Alltagspsychologie nicht immun gegenüber wissenschaftlichem Fortschritt ist, eröffnet das die Möglichkeit, dass sich die Verallgemeinerungen der wissenschaftlichen Psychologie, die letztlich auf die Neurowissenschaft reduziert werden können, weit über die „Wahrheiten“ der Alltagspsychologie hinaus entwickelt haben könnten [...]. Sollte dies die Richtung der Koevolution von Psychologie und Neurowissenschaft sein, werden künftige Wissenschaftshistoriker feststellen, dass die Alltagspsychologie durch die Neurobiologie eher größtenteils ersetzt als erhaltend reduziert sein wird [...].

Mit „eliminativer Materialismus“ meine ich die Position, die behauptet:

1. dass die Alltagspsychologie eine Theorie ist;
2. dass sie eine Theorie ist, die solche Probleme aufwirft, dass sie unter Umständen grundlegend überdacht oder sogar völlig ersetzt werden muss (daher „eliminativ“) und
3. dass die Theorie, die die Alltagspsychologie letztlich ersetzen wird, eine reife Neurowissenschaft sein wird (daher „Materialismus“)

(Übersetzt aus: Patricia Churchland, *Neurophilosophy. Toward a Unified Science of the Mind/Brain*, Cambridge: MIT Press 1986, S. 278ff)

M 8 Der nicht-reduktionistische Physikalismus

Wir bezeichnen unsere Position als „nicht-reduktionistischen Physikalismus“, weil wir zwar einerseits der Auffassung sind, dass auch mentale Phänomene letztendlich physikalische Phänomene sind, die im Gehirn ablaufen, dass aber die Untersuchung dieser mentalen Phänomene mit psychologischen Methoden nicht weniger sinnvoll ist als die Untersuchung der physikalischen Prozesse im Gehirn mit neurobiologischen Methoden. Es ist eine nicht zu leugnende empirische Erkenntnis, dass alle mentalen Phänomene eindeutig Prozessen im Gehirn zugeordnet werden können: Jedes mentale Phänomen geht mit messbaren neuronalen Prozessen einher. Reizungen des Gehirns bewirken sehr oft auch mentale Phänomene. Krankhafte Veränderungen oder Verletzungen des Gehirns haben auch seelische Veränderungen zur Folge. Mentales und Neuronales sind also zwei Perspektiven desselben Geschehens. Es gibt keinen Grund in ihnen etwas substantiell Verschiedenes zu sehen. Es gibt aber auch keinen Grund, einer der beiden Perspektiven die Berechtigung abzuspochen. Schaut man sich ein Fußballspiel im Fernsehen an und verfolgt parallel dazu eine Übertragung im Radio, so wird man weder behaupten, dass es ein visuelles und ein akustisches Fußballspiel gäbe, noch dass eine Form der Übertragung weniger real als die andere sei [...].

Was haben wir mit dem Aufweis einer strikten Parallelität zwischen mentalen Prozessen und Gehirnprozessen gewonnen? Wir können mit Hilfe dieser empirischen Einsichten eine Reihe von Annahmen widerlegen und einige der bekannten philosophischen Standpunkte als sehr unplausibel erscheinen lassen. Zu den ersteren gehört die weit verbreitete Anschauung, Mentales sei *absolut privat*, d.h. nur mir selbst zugänglich und sonst niemandem. Wegen dieser grundlegenden Unsicherheit des „Fremdpsychischen“ sei die Übereinstimmung zwischen mentalen Ereignissen und Hirnprozessen *grundsätzlich* nicht festzustellen. Diese Auffassung ist nicht gerechtfertigt. In der Experimentalpsychologie ist der Bericht über innere Erlebnisse, die *Introspektion*, ebenso eine Quelle von Erkenntnis wie die Reaktionszeitmessung und die Registrierung evozierter Potentiale. Wir können beispielsweise anhand der Registrierung evozierter Potentiale und dieser Introspektion feststellen, ob bestimmte Hirnregionen (z.B. Area V4) immer dann aktiv sind, wenn eine Versuchsperson eine Farbempfindung hat. Man kann diese Experimente mit mehreren Personen durchführen und sehen, ob Berichte und Registrierungen voneinander abweichen. Eine derartige Vorgehensweise unterscheidet sich nicht wesentlich von den in den Natur- und Biowissenschaften üblichen Verfahren der Korrelation von Phänomenen. Werden auf diese Weise psychologische und neurobiologische Methoden miteinander verbunden, wird sich das sogenannte „Körper-Geist-Problem“ in nicht allzu ferner Zukunft nicht mehr stellen.

(aus: Gerhard Roth und Helmut Schwegler: Das Geist-Gehirn-Problem aus der Sicht der Hirnforschung und eines nicht-reduktionistischen Physikalismus, Ethik und Sozialwissenschaften 6 (1995)1, S 72)

M 9 Der Funktionalismus

Funktionalismus ist die Ansicht, dass das, was einen Gedanken, einen Wunsch, einen Schmerz (oder irgendeinen anderen mentalen Zustand) ausmacht, nicht auf seiner inneren Beschaffenheit beruht, sondern auf seiner Funktion oder Rolle innerhalb des kognitiven Systems, zu dem er gehört. Präzise formuliert bestimmen funktionalistische Theorien die Identität eines mentalen Zustands durch seine kausalen Beziehungen zu sensorischen Stimulationen, anderen mentalen Zuständen und Verhalten.

So könnte, um ein hoffnungslos vereinfachtes Beispiel zu nennen, eine funktionalistische Theorie Schmerz charakterisieren als einen Zustand, der von einer Verletzung des Körpers verursacht wird, der den Glauben hervorbringt, dass etwas mit dem Körper nicht in Ordnung ist, und den Wunsch, diesen Zustand los zu sein, und der, in Abwesenheit stärkerer andersartiger Bedürfnisse, Zucken und Stöhnen verursacht. Entsprechend dieser Theorie wären alle und nur diejenigen Wesen fähig, Schmerz zu empfinden, deren innere Zustände diese Bedingungen erfüllen oder diese Rolle spielen können.

Nehmen wir an, dass es bei Menschen eine bestimmte Art neuronaler Aktivität (z.B. C-Faser-Stimulation) gibt, die diese Bedingungen erfüllt. Wenn dem so ist, dann können Menschen einfach durch Stimulation dieser Fasern Schmerzen empfinden. Jedoch erlaubt es diese Theorie auch, bei Wesen mit einer völlig anderen physischen Konstitution mentale Zustände anzunehmen: Wenn es silikon-basierte Zustände hypothetischer Marsmenschen oder anorganische Zustände hypothetischer Androiden gibt, die eben diese Bedingungen erfüllen, dann können auch diese Wesen Schmerzen empfinden. Funktionalisten sagen oft, dass Schmerz in verschiedenen Systemen in Form verschiedener physischer Zustände realisiert oder „multipel realisiert“ sein kann. Da diese Beschreibungen sich nur auf die kausalen Beziehungen zu Stimulationen, Verhalten und aufeinander beziehen, wäre es auch möglich, dass nicht-physische Zustände in irgendwelchen Systemen diese relevanten Rollen spielen. Daher ist der Funktionalismus auch mit dualistischen Ansätzen verträglich, die davon ausgehen, dass mentale Zustände physikalische Zustände verursachen und von ihnen verursacht werden.

Wenn der Funktionalismus auch streng genommen neutral gegenüber dem Materialismus und Dualismus ist, wird er insbesondere von Materialisten als sehr attraktiv empfunden, da viele Materialisten es für überwältigend wahrscheinlich halten, dass die Zustände, die die betreffenden Rollen spielen, physische Zustände sind. Wenn dem so ist, dann kann der Funktionalismus eine materialistische Alternative zur psycho-physischen Identitätstheorie sein, der Position also, dass jeder mentale

Zustand identisch mit einem bestimmten neuronalen Zustand ist. Diese bislang dominante materialistische Theorie des Geistes zieht die Annahme nach sich, dass keine anderen Wesen unsere Empfindungen, Annahmen und Wünsche teilen können, wie ähnlich ihr Verhalten und ihre innere Organisation der unsrigen auch sein möge. Dies ist eine Konsequenz, die viele Materialisten nicht für plausibel halten. Auf diese Weise wird der Funktionalismus mit seiner Annahme der multiplen Realisierbarkeit mentaler Zustände als eine umfassendere, weniger „spezies-chauvinistische“ und somit plausiblere Theorie betrachtet, die mit dem Materialismus kompatibel ist.

(übersetzt aus: Janet Levin, *Functionalism*, <http://plato.stanford.edu/entries/functionalism>)

M 10 Das chinesische Zimmer

Der Geist hat mehr als eine Syntax, er hat eine Semantik. Kein Computerprogramm kann jemals ein Geist sein, und zwar aus dem einfachen Grunde, weil ein Computerprogramm bloß syntaktisch ist und der Geist mehr als bloß syntaktisch ist. Der Geist ist semantisch – semantisch in dem Sinne, dass er mehr hat als eine formale Struktur: Er hat einen Gehalt.

Um dies zu veranschaulichen, habe ich ein Gedankenexperiment entwickelt. Man stelle sich vor, ein paar Computerprogrammierer hätten ein Programm geschrieben, das einen Computer in die Lage versetzt, das Verständnis des Chinesischen zu simulieren. Wenn dem Computer also beispielsweise eine Frage auf Chinesisch vorgelegt wird, wird er die Frage mit seinem Gedächtnis (seiner Datenbasis) konfrontieren und passende Antworten auf Chinesisch produzieren. Nehmen wir der Argumentation halber einmal an, dass die Antworten des Computers genau so gut sind wie die eines chinesischen Muttersprachlers. Versteht der Computer dann allein deshalb schon Chinesisch? Versteht er wirklich Chinesisch – in dem wörtlichen Sinne, in dem Chinesen Chinesisch verstehen? Nun, stellen Sie sich vor, Sie wären in ein Zimmer eingesperrt, in dem mehrere Körbe mit chinesischen Symbolen stehen. Und stellen Sie sich vor, dass Sie (wie ich) kein Wort Chinesisch verstehen, dass Ihnen aber ein auf Deutsch abgefasstes Regelwerk für die Handhabung dieser chinesischen Symbole gegeben worden wäre. Die Regeln geben rein formal – nur mit Rückgriff auf die Syntax und nicht die Semantik der Symbole – an, was mit den Symbolen gemacht werden soll. Eine solche Regel mag lauten: „Nimm ein Kritzel-Kratzel-Zeichen aus Korb 1 und lege es neben ein Schnörkel-Schnarkel-Zeichen aus Korb 2.“ Nehmen wir nun an, dass irgendwelche andern chinesischen Symbole in das Zimmer gereicht werden, und dass Ihnen noch zusätzliche Regeln dafür gegeben werden, welche chinesischen Symbole jeweils aus dem Zimmer herauszureichen sind. Die hereingereichten Symbole werden von den Leuten draußen „Fragen“ genannt, und die Symbole, die Sie dann aus dem Zimmer herausreichen, „Antworten“ – aber dies geschieht ohne Ihr Wissen. Nehmen wir außerdem an, dass die Programme so trefflich und ihre Ausführung so brav sind, dass Ihre Ant-

worten sich schon bald nicht mehr von denen eines chinesischen Muttersprachlers unterscheiden lassen. Da sind Sie nun also in Ihrem Zimmer eingesperrt und stellen Ihre chinesischen Symbole zusammen; Ihnen werden chinesische Symbole hereingereicht und darauf hin reichen Sie chinesische Symbole heraus. In so einer Lage, wie ich sie gerade beschrieben habe, könnten Sie einfach dadurch, was Sie mit den formalen Symbolen anstellen, kein bisschen Chinesisch lernen.

Der Witz der Geschichte ist nun schlicht folgender: Weil Sie ein formales Computerprogramm ausführen, verhalten Sie sich aus der Sicht eines Außenstehenden genau so, als verstünden Sie Chinesisch – und dennoch verstehen Sie nicht ein Wort Chinesisch. Wenn aber die Ausführung eines passenden Computerprogramms *in Ihrem Fall* nicht ausreicht, um Chinesisch zu verstehen, dann reicht das auch bei *keinem digitalen Computer* aus. Auch hierfür lässt sich der Grund ganz einfach angeben. Wenn Sie kein Chinesisch verstehen, dann könnte auch kein Computer Chinesisch verstehen; denn kraft seiner Ausführung eines Programms hat kein digitaler Computer irgendetwas, das Sie nicht haben. Der Computer hat – genau wie Sie – nichts außer einem formalen Programm für die Handhabung uninterpretierter chinesischer Symbole. Um es zu wiederholen: Ein Computer hat eine Syntax, aber keine Semantik. Der ganze Witz der Parabel mit dem Chinesischen Zimmer besteht darin, uns an etwas zu erinnern, das wir die ganze Zeit über gewusst haben. Eine Sprache verstehen oder überhaupt sich in gewissen Geisteszuständen befinden, dazu gehört mehr als dass man bloß ein paar formale Symbole hat. Dazu gehört, dass diese Symbole eine Interpretation (oder eine Bedeutung) haben. Und ein digitaler Computer kann laut Definition einfach nur formale Symbole haben, denn die Funktionsweise des Computers ist, wie bereits gesagt, dadurch definiert, dass er gewisse Programme ausführen kann. Und diese Programme lassen sich rein formal angeben – d.h. sie haben keinen semantischen Gehalt.

(aus: John R. Searle, *Geist, Hirn und Wissenschaft*, Frankfurt a.M.: Suhrkamp 1986, S.30-32)

M 11 Absent-Qualia- und Inverted-Qualia-Argumente

Auch würde man die einfachen Vorstellungen nicht deshalb für falsch halten können, wenn durch einen verschiedenen Bau der Organe derselbe Gegenstand in den Seelen verschiedener Menschen gleichzeitig verschiedene Vorstellungen hervorbrächte; wenn z.B. die bei dem einen durch ein Veilchen hervorgebrachte Vorstellung der bei einem andern durch eine gelbe Ringelblume hervorgebrachten gleich wäre, und umgekehrt. Denn dies kann niemals festgestellt werden; da die Seele des Einen nicht in den Leib des Andern eintreten kann, um zu sehen, welche Vorstellungen durch dessen Organe hervorgebracht werden, und deshalb würden selbst in solchem Falle weder die Vorstellungen noch die Namen verwechselt werden, oder eines von beiden falsch werden. Denn alle Gegenstände mit dem Gewebe eines Veilchens brächten dann die Vorstellung hervor, die er Blau nennt, und alle mit dem Gewebe einer Ringelblume die, welche er regelmäßig Gelb nennt. Welcher

Art nun auch diese Vorstellungen in seiner Seele wären, so wäre er doch im Stande, dadurch die Dinge für seine Zwecke ebenso regelmäßig zu unterscheiden, diese Unterschiede zu erkennen und durch die Zeichen erkennbar zu machen, welche die Worte Blau und Gelb gewähren, als wenn seine von diesen beiden Blumen empfangenen Vorstellungen genau dieselben mit denen des andern wären. Indes möchte ich doch annehmen, dass die bei mehreren Menschen durch denselben Gegenstand hervorgebrachten Sinneswahrnehmungen wohl immer beinahe gleich und nicht zu unterscheiden seien. Dafür ließen sich viele Gründe anführen, indes gehört dies nicht zu meiner jetzigen Aufgabe, und ich will daher den Leser nicht damit belästigen, sondern ihn nur erinnern, dass die entgegen gesetzte Annahme, selbst wenn man sie beweisen könnte, für die Vermehrung der Kenntnisse oder die Bequemlichkeiten des Lebens ohne Nutzen sein würde; deshalb braucht man sich damit nicht zu bemühen.

(Übersetzt aus: Locke, J., 1689, Essay Concerning Human Understanding, Oxford: Oxford University Press 1975, II, 32, 15)

[Ein Einwand gegen den Funktionalismus] ist das "inverted qualia argument", das von einer Denkmöglichkeit ausgeht, die ursprünglich von John Locke entwickelt wurde. Nehmen wir an, dass zwei Menschen, Norma und Anorma, jeweils umgekehrte Farbqualitäten empfinden würden. Beide nennen Sie Stoppzeichen, Coladosen und Elmo „rot“, Zuckererbsen, Heinekenflaschen und Kermit den Frosch hingegen „grün“. Jedoch ist Abnormas phänomenale Erfahrung, wenn sie Coladosen sieht, so wie Normas Erfahrung, wenn sie Heinekenflaschen sieht. Wenn Norma eine Coladose sieht, hat sie eine Rotempfindung; wenn Anorma eine Coladose sieht, hat sie eine Grünempfindung. Entsprechend hat Norma, wenn sie Heinekenflaschen sieht, eine Grünempfindung, während Anorma, wenn sie Heinekenflaschen sieht, eine Rotempfindung hat. Qualitativ empfinden die beiden also invertiert zueinander. Obgleich die meisten Menschen dieses Szenario als begrifflich kohärent empfinden, kann ein Funktionalist diese Inversion nicht zum Ausdruck bringen. Beide nennen sie eine Coladose „rot“. Beide bringen sie zum Ausdruck, dass sie die gleiche Farbe wie Stoppzeichen und reife Tomaten hat. Funktional gesprochen besteht keine Möglichkeit, die mentalen Zustände voneinander zu unterscheiden, in denen sich Anorma und Norma befinden, wenn sie eine Coladose sehen. Somit sieht es so aus, als ob funktionale Definitionen die qualitativen Aspekte mentaler Zustände nicht erfassen könnten.

(Übersetzt aus: Qualia, The Internet Encyclopedia of Philosophy, <http://www.iep.utm.edu/q/qualia.htm#H2>)

M 12 Roboter werden uns überholen

Ich behaupte insbesondere, dass Computer eines Tages auf dieselbe Weise werden wahrnehmen, erkennen und denken können wie ein Mensch. Daraus folgt, dass ein hinreichend hoch entwickeltes künstliches System – etwa ein elektronisches – im

Prinzip zu denselben Leistungen fähig ist wie das menschliche Nervensystem einschließlich des Gehirns [...]. Die Auseinandersetzung läuft letztlich auf die Frage hinaus, ob das Verhalten eines Menschen berechenbar und insbesondere mit einem Computer modellierbar ist. Meiner Auffassung nach gibt es kein gutes wissenschaftliches Argument gegen diese These. Im Gegenteil: Alles deutet darauf hin, dass sie zutrifft [...]. Wahrscheinlich werden bereits die nächsten Jahrzehnte eine Entscheidung über meine [...] These bringen. Entweder werden dank ausreichender Rechenkapazität die Roboter Tiere und schließlich Menschen in ihren Fähigkeiten einholen und übertreffen; oder wir entdecken eines Tages spezifische tierische oder menschliche Fertigkeiten, die aus prinzipiellen Gründen auch der rechenstärkste Roboter nicht simulieren kann. Das wäre eine wissenschaftliche Sensation; doch gibt es bislang keinen einzigen Anhaltspunkt für einen solchen prinzipiellen Unterschied. [...]

Zu den psychologischen Leistungen gehören – für Roboter wie für Menschen – Ziele, Überzeugungen, Gefühle und Vorlieben. [...] Simulationsprogramme zu entwickeln, die zu solchen Leistungen fähig sind, ist ein gewaltiges Unternehmen. Tausende von Programmierern werden daran arbeiten, und ebenso viele Roboter werden Erfahrungen sammeln müssen. Das Simulationsprogramm wird mit verschiedenen Sensoren, die den Sinnen des Menschen ähnlich sind, äußere Ereignisse registrieren und inneren Modelle entwickeln, um stets ein möglichst getreues Bild der Umgebung zu haben. Mit ihm kann ein Roboter die verschiedensten Fähigkeiten durch bloße Nachahmung erlernen, ja sogar eine gewisse Art von Bewusstsein erlangen. Auf die Frage seines Besitzers, warum Kerzen auf dem Tisch stehen, würde er beispielsweise sein Simulationsprogramm für das Haus, seinen Herrn und sich selbst aufrufen und darauf antworten, er selbst habe die Kerzen dorthin gestellt, weil sein Besitzer ein Dinner bei Kerzenschein schätze und er – der Roboter – ihm gerne einen Gefallen tun möchte. Auf weiteres Nachfragen würde sich ein – einfaches, aber immerhin – Gemütsleben des Roboters auftun, das konkrete Situationen und die Menschen in seinem Arbeitsbereich umfasst.

(aus: Hans Moravec, Die Roboter werden uns überholen, in: Spektrum der Wissenschaft, Spezial: Forschung im 21. Jahrhundert)

M 13 Zusatzmaterial: Pseudonormales Sehen

A.M.: Die Netzhaut unseres Auges besteht im Wesentlichen aus drei Schichten: der Sinneszellschicht, die merkwürdigerweise die am tiefsten liegende ist, die darüber gelagerte Schicht von Bipolar-, Amakrin- und Horizontalzellen und die Ganglienzellschicht. Die Sinneszellschicht enthält zwei Klassen von Sinneszellen: die sehr lichtempfindlichen, aber auf keine bestimmte Wellenlänge spezialisierten Stäbchen und die weniger lichtempfindlichen, aber auf bestimmte Wellenlängenbereiche spezialisierten Zapfen.

L.Z.: Gut, das kenne ich noch aus der Schule.

A.M.: Klasse. Also weiter: Diese Sinneszellen haben nun eine etwas merkwürdige Funktionsweise. Sie enthalten ein lichtempfindliches Pigment, einen Sehfärbstoff.

Solange Licht auf die Zapfen oder Stäbchen fällt, haben sie ein im Vergleich zu Nervenzellen stark positives Ruhepotenzial, bei dem sie Neurotransmitter ausschütten. Wenn der Farbstoff Licht absorbiert, dann zerfällt er und setzt eine Signalkette in Gang, die nicht zu einer Umkehr, sondern zu einer Hyperpolarisation des Potenzials führt, bei dem kein Neurotransmitter mehr ausgeschüttet wird.

L.Z.: Hä?

A.M.: Das heißt, diese Sinneszellen sind eigentlich im Dunkeln aktiv und werden durch Licht gehemmt. Eigentlich sind unsere Zapfen und Stäbchen also „Dunkel-sinneszellen“! Ich weiß: Das ist ziemlich verwirrend. Ich werde dieses Problem umgehen, indem ich mich immer auf den Zerfall des Sehfärbstoffes beziehe, der ja noch ganz nachvollziehbar im Licht geschieht – auch wenn meine Ausdrucksweise dann etwas umständlich wird.

L.Z.: Ich bitte dich darum!

A.M.: Der Neurotransmitter, den die Zapfen und Stäbchen nun ausschütten oder eben nicht, bewirkt nun eine Hemmung oder Erregung nachgeschalteter Zellen, bei der auch eine erste Verarbeitung erfolgt. Das „Ergebnis“ dieser Vorverarbeitung wird von den Ganglienzellen über den Sehnerv zum Gehirn geleitet. Werden die Sinneszellen nicht mehr belichtet, dann wird der Sehfärbstoff regeneriert. Dies geschieht in der Dunkelheit, aber auch schon dann, wenn wir im Hellen dunkle Oberflächen betrachten. Für die Farbwahrnehmung können wir uns jetzt auf die Zapfen beschränken. In unserer Netzhaut gibt es drei Typen davon, die verschiedene Sehfärbstoffe enthalten: Erythrolabe, Chlorolabe und Cyanolabe. Erythrolabe hat ein Empfindlichkeitsmaximum bei langwelligem Licht, also bei Licht, das ein normaler Mensch unter normalen Bedingungen rot wahrnimmt. Chlorolabe ist am empfindlichsten für Licht mittlerer Wellenlänge, also für Licht, das ein normaler Mensch unter normalen Bedingungen grün wahrnimmt. Cyanolabe hat ein Empfindlichkeitsmaximum bei kurzwelligem Licht, also bei Licht, das ein normaler Mensch unter normalen Bedingungen als blau(violett) empfindet. Man nennt diese drei Zapfentypen nach den entsprechenden Wellenlängen L-, M- und S-Zapfen – für long, medium und short. Zur besseren Orientierung kann man sie aber auch nach den unter normalen Umständen auftretenden Farbbeimpfindungen R-, G- und B-Zapfen nennen - für rot, grün und blau. Wie diese Zapfen mit den folgenden Zellen der Netzhaut verschaltet sind, ist zur Zeit noch nicht klar. Auch die vorhandenen theoretischen Modelle erklären nicht alle Farbphänomene, die man beobachten kann. Ich gehe daher jetzt mal von dem einfachsten denkbaren Modell aus. Diesem Modell zufolge gibt es zwei „Farbkanäle“: einen Rot-Grün-Kanal und einen Blau-Gelb-Kanal. Über den Rot-Grün-Kanal wird die Aktivität der R- und G-Zapfen miteinander verglichen. Dabei entsteht eine Rotempfindung, wenn in den R-Zapfen mehr Sehfärbstoff zerfällt als in den G-Zapfen. Es entsteht eine Grünempfindung, wenn mehr Pigment in den G-Zapfen zerfällt. Komplizierter aufgebaut ist der Blau-Gelb-Kanal. Hier wird die Aktivität der R- und G-Zapfen mit der Aktivität der Blauzapfen verglichen. Es entsteht eine Blauempfindung, wenn mehr Sehfärbstoff in den Blauzapfen zerfällt und eine Gelbempfindung wenn mehr Farbstoff in den R- und G-Zapfen zerfällt – oder eben eine Grün- oder Rotempfindung wenn R-

und G-Zapfen unterschiedlich aktiv sind. Keine Farbempfindung tritt jeweils bei gleicher Erregung aller Zapfen auf. Wir nehmen dann, je nach Lichtintensität, weißes oder graues Licht wahr. Die Ganglienzellen mischen also sozusagen die drei Grundfarben rot, grün und blauviolett zu den anderen uns bekannten Farben.

L.Z.: Halt mal. Die Grundfarben sind doch Gelb, Rot und Blau. Grün kann ich doch aus Blau und Gelb mischen, Gelb hingegen aus keinen anderen Farben.

A.M.: Nein, diese Farbmischungen ergeben sich, wenn du Farbpigmente in einer Palette oder auf dem Papier mischst. Man nennt das dann auch subtraktive Farbmischung. Unserer Farbempfindung liegen die grundlegenden additiven Farbmischungen zugrunde, die sich ergeben, wenn man Licht direkt mischt, also etwa, wenn man zwei bis drei Diaprojektoren auf eine weiße Wand richtet und mit Folien verschiedener Farben versieht. Man kann sich die gleichen Effekte aber auch an einer Bühnenbeleuchtung deutlich machen. Mischt man grünes Licht, das vom Farbstoff der G-Zapfen absorbiert wird, mit rotem Licht, für das der Farbstoff der R-Zapfen empfindlich ist, dann ergibt sich gelbes Licht. Mischt man rotes mit blauem Licht ergibt sich Purpur, mischt man grünes mit blauem Licht, erhält man türkis. Mischt man rotes, blaues und grünes Licht so, dass alle drei Zapfentypen gleichmäßig erregt oder eben gehemmt werden, erhält man weißes Licht.[...]

A.M.: Schauen wir uns zunächst mal an, was bei einem Menschen nicht funktioniert, der, anders als deine Neurobiologin Mary, tatsächlich farbenblind ist. Hierfür kann es viele Ursachen geben. Bei manchen Farbenblinden sind durch einen genetischen Defekt auf dem X-Chromosom die R-Zapfen nicht mit Erythrolabe, sondern „fälschlicherweise“ - wie die G-Zapfen - mit Chlorolabe gefüllt. Die Konsequenz dieser Störung besteht darin, dass G- und R-Zapfen immer gleichermaßen aktiv sind. Da sie geschlechtschromosomal vererbt wird, tritt dieser Defekt fast nur bei Männern auf. Diesen Männern ist eine Rot-Grün-Unterscheidung nicht mehr möglich. Die Farbempfindung beim Anblick roter oder grüner Gegenstände müsste bei ihnen also, je nach Erregung der G- und R-Zapfen im Vergleich zu der Erregung der B-Zapfen im Bereich zwischen Gelb und Blau liegen. Es könnte aber auch die umgekehrte Störung geben, bei der die G-Zapfen fälschlicherweise - wie die R-Zapfen - mit Erythrolabe gefüllt sind. Diese Störung hätte natürlich ebenfalls eine Rot-Grün-Blindheit zur Folge, weil wieder beide Zapfentypen unabhängig von der Wellenlänge des Lichtes die gleiche Erregung aufweisen würden. Wenn beide genetischen Defekte unabhängig voneinander vererbt werden, müssten sie mit dem Produkt ihrer Wahrscheinlichkeiten gleichzeitig auftreten: Die R-Zapfen wären dann mit dem - normalerweise in den G-Zapfen enthaltenen - Chlorolabe gefüllt, die G-Zapfen hingegen mit dem - normalerweise in den R-Zapfen enthaltenen - Erythrolabe. Glaubt man der Theorie, sollten solche Menschen tatsächlich ein umgekehrtes Rot-Grün-Spektrum wahrnehmen: Rot wäre dann die Empfindungsqualität, die dann auftritt, wenn in den R-Zapfen mehr Pigment zerfällt als in den G-Zapfen. Bei diesen Menschen wäre das aber dann der Fall, wenn Licht mittlerer Wellenlänge vorherrscht, das ein Normaler als grün empfinden würde. Grün wäre die Empfindungsqualität, die dann auftritt, wenn in den G-Zapfen mehr Farbstoff

zerfällt als in den R-Zapfen. Das wäre aber dann der Fall, wenn Licht langer Wellenlänge vorherrscht, das ein Normaler als rot empfinden würde. Ein solcher Mensch würde also dort rot sehen, wo der Normalsichtige grün sieht und dort grün, wo der Normalsichtige rot sieht. Da ihr Verhalten aus den von dir schon dargestellten Gründen völlig unauffällig wäre, nennt man diese hypothetisch angenommene Farbwahrnehmungsstörung „pseudonormal“. Wenn es wirklich so sein sollte, dass diese Menschen ein umgekehrtes Farbspektrum sehen würden, läge darin auch der Grund, dass man noch nie einen Menschen gefunden hat, der diese Störung aufweist.

(aus: Axel Ziemke, Im Netzwerk der Unsterblichkeit. Ein Biochemiker und ein Philosoph im Selbstgespräch über Gehirn, Bewusstsein und geistige Welten, Frankfurt a. M. 2007)