

e-Journal Philosophie der Psychologie	MENTALE EIGENSCHAFTEN UND MENTALE SUBSTANZEN Antworten der Analytischen Philosophie auf das 'Leib-Seele-Problem' von Ansgar Beckermann (Bielefeld)
---	---

1.

Das moderne Menschenbild – das Bild, das wir Menschen von uns selbst haben – ist ganz wesentlich geprägt von den Überlegungen des französischen Philosophen René Descartes. Was sind die Hauptaspekte dieses Cartesischen Menschbildes? Bekanntlich unterscheidet Descartes strikt zwischen zwei Arten von Substanzen – ausgedehnten Substanzen (*res extensae*) und denkenden Substanzen (*res cogitantes*). Nur beim Menschen gehen diese Substanzen eine enge Verbindung ein.

1. Menschen bestehen – anders als alle anderen Wesen auf der Welt, anders auch als alle Tiere – aus zwei Dingen: einem Körper und einer Seele, einer *res extensa* und einer *res cogitans*.

Diese These beinhaltet unter anderem einen radikalen Bruch mit der antiken Auffassung der Seele. Denn für die Antike ist die Seele in erster Linie *Lebensprinzip* – sie ist das, was dafür verantwortlich ist, dass ein Ding lebendig ist. Das Wort 'Psyche' kommt von 'psychein', was soviel heißt wie 'hauchen' oder 'atmen'. Die Seele ist also verantwortlich für all die Fähigkeiten, die Lebewesen von toten Dingen unterscheiden – für die Fähigkeiten, die schon Pflanzen besitzen, sich zu ernähren, zu wachsen und sich fortzupflanzen; für die Fähigkeiten der Tiere, wahrzunehmen, zu begehren und sich zu bewegen, und schließlich auch für die dem Menschen vorbehaltene Fähigkeit zu denken.

Mit Descartes ändert sich diese Auffassung der Seele radikal. Zum einen ist für ihn die Seele ein wenn auch immaterielles Einzelding – eine denkende Substanz, die das eigentliche Selbst des Menschen ausmacht. Zum anderen kappt Descartes den antiken Zusammenhang zwischen Seele und Leben. Die, wenn man so will, elementaren Fähigkeiten von Lebewesen – die vitalen Fähigkeiten von Pflanzen und Tieren – bedürfen für ihn keiner Zurückführung auf eine Seele. Diese Fähigkeiten sind in seinen Augen rein mechanisch erklärbar. Nicht mechanisch erklärbar sind nach Descartes jedoch die im engeren Sinne *mentalen* Fähigkeiten – die Fähigkeiten, zu denken und zu sprechen. Diese sind für ihn ohne Seele undenkbar. Sie bedürfen nicht nur einer Erklärung durch die Seele; sie sind vielmehr in einem bestimmten Sinne Fähigkeiten eben dieser Seele, und nicht etwa Fähigkeiten des mit dieser Seele verbundenen Körpers. Man kann nicht denken oder sprechen, ohne Ideen zu haben. Und das Haben von Ideen ist ein Zustand der Seele, nicht des Körpers. Damit ergeben sich zwei weitere Aspekte.

2. Mentale Eigenschaften – die Eigenschaften zu zweifeln, einzusehen, zu bejahen, zu verneinen, zu wollen, nicht zu wollen, bildlich vorzustellen und zu empfinden – sind Eigenschaften der Seele, nicht des Körpers. (*Meditationen II 8*)

3. Diese Eigenschaften sind die einzigen Eigenschaften des Menschen, die nicht mechanisch erklärt werden können.

Dieser außerordentlich geraffte Rückblick auf den Descartesschen Seelenbegriff lässt sicher viele Nuancen seiner Position unberücksichtigt. Dafür treten in ihm die beiden Grundfragen, die nach wie vor die philosophische Debatte um das Leib-Seele-Problem bestimmen, jedoch besonders prägnant zu Tage:

Frage 1:

Welche *Art von Dingen* haben eigentlich mentale Eigenschaften? Sind die Träger dieser Eigenschaften *immaterielle*, nicht-physische Dinge oder sind es ganz normale Lebewesen aus Fleisch und Blut?

Frage 2:

Was ist die Natur mentaler *Eigenschaften*? Lassen sich mentale Eigenschaften naturwissenschaftlich erklären? Oder handelt es sich um *ontologisch eigenständige* Eigenschaften, die nicht auf physische Eigenschaften zurückgeführt werden können?

Die erste Frage verweist auf das Problem *mentaler Substanzen*, die zweite Frage dagegen auf das Problem *mentaler Eigenschaften*.

Substanz-Dualisten – wie Platon, Descartes oder in neuester Zeit Richard Swinburne – vertreten die Auffassung, dass es neben Molekülen, Steinen, Bäumen und Sternen – also neben den physischen Gegenständen – auch immaterielle, nicht-physische Dinge gibt und dass es diese immateriellen Dinge sind, die ein geistiges Leben haben, die denken, fühlen und sich entscheiden. Für Substanz-Dualisten gilt also: Träger mentaler Eigenschaften sind nicht physische, sondern immaterielle Dinge. *Gegner des Substanz-Dualismus* (*Physikalisten* im weiteren Sinne) sind dagegen der Meinung, dass es in der Welt nur physische Dinge gibt und dass die Träger mentaler Eigenschaften *nicht* immaterielle Seelen, sondern bestimmte physische Dinge – in unserer Welt: bestimmte Lebewesen – sind.

Eigenschafts-Dualisten – wie etwa Karl Popper – müssen nicht zugleich Substanz-Dualisten sein, d.h. sie können durchaus die Auffassung teilen, dass es in der Welt nur physische Dinge gibt und dass bestimmte Lebewesen die Träger mentaler Eigenschaften sind. Für den Eigenschafts-Dualisten ist nur die These charakteristisch, dass mentale *Eigenschaften* in dem Sinne *eigenständig* sind, dass sie sich nicht auf physische Eigenschaften zurückführen lassen, während der *Eigenschafts-Materialist* bzw. *-Physikalist* (der *Physikalist* im engeren Sinne) wiederum die gegenteilige Auffassung vertritt. Seine These lautet, dass die durch mentale Prädikate ausgedrückten Eigenschaften nicht eigenständig sind, dass sie vielmehr – auf die eine oder andere Weise – auf physische Eigenschaften zurückgeführt werden können.

2.

Wenn wir die Geschichte der Philosophie betrachten, hat das Problem des Substanz-Dualismus eine erheblich größere Rolle gespielt als das des Eigenschafts-Dualismus. Auch viele Nichtphilosophen interessiert am Leib-Seele-Problem in erster Linie die eng mit dem Problem mentaler Substanzen zusammenhängende Frage, ob es ein Leben nach dem Tode gibt. In der gegenwärtigen Philosophie spielt der Substanz-Dualismus dagegen kaum noch eine Rolle. Woran liegt das?

Nun, sicher hat das auch mit dem Zeitgeist zu tun. Die meisten Philosophen – zumindest im Bereich der Analytischen Philosophie – sind heute von einer naturalistischen Grundeinstellung geprägt. Zu den Prämissen ihres Denkens gehört die Annahme, dass in der Natur alles mit rechten Dingen zugeht. Aber das ist mehr als nur eine subjektive Meinung. Denn es gibt zwei Hauptgründe, die dafür verantwortlich sind, dass es heute kaum noch Vertreter des Substanz-Dualismus gibt:

1. das nach wie vor ungelöste Problem der kausalen Interaktion von physischen und nicht-physischen Dingen;
2. die große Anzahl theoretischer Probleme, die mit der Annahme der Existenz immaterieller Geister verbunden sind.

Ich will mich hier auf das erste Problem konzentrieren, das deshalb so virulent ist, weil der Substanz-Dualismus – wenn überhaupt – nur in seiner interaktionistischen Variante plausibel ist. Geist und Körper müssen in der Lage sein, kausal aufeinander einzuwirken: Wenn Licht auf die Retina meines Auges trifft und entsprechende neuronale Signale vom Auge zum Gehirn gesandt werden, bewirkt das in meinem Geist einen Wahrnehmungseindruck; und wenn ich meinen Arm heben will, dann bewirkt mein Geist in meinem Gehirn das Feuern bestimmter Motoneuronen, das seinerseits dazu führt, dass sich mein Arm hebt. Aber wie soll es möglich sein, dass mein Geist auf meinen Körper und mein Körper auf meinen Geist kausal aufeinander einwirken?

Da ist zunächst die Frage, an welcher Stelle der Geist kausal auf den Körper einwirkt. Bei willentlichen Handlungen z. B. sind der Interaktionsthese zufolge die Wünsche und Überzeugungen des Handelnden für bestimmte Körperbewegungen kausal mitverantwortlich. Diese Körperbewegungen gehen jedoch auf die Kontraktionen und Relaxationen bestimmter Muskeln zurück, die von der Änderung der Aktivität der motorischen Einheiten der Muskeln hervorgerufen werden. Diese werden durch Impulse aus den motorischen Vorderhornzellen des Rückenmarks aktiviert, die ihrerseits bei ziel- Mentale Eigenschaften und mentale Substanzen gerichteten Willkürbewegungen durch Neuronen der motorischen Hirnrinde aktiviert werden. In den peripherienahen Bereichen scheint es also keine psychophysische Verursachung zu geben, d.h. wenn überhaupt, scheint das Eingreifen des Geistes auf bestimmte zentrale Regionen des Gehirns beschränkt zu sein.

In der Tat hat schon Descartes die Ansicht vertreten, dass der Geist nur an einer einzigen Stelle, nämlich in der *Zirbeldrüse* auf das Gehirn einwirken könne, und zwar nur dadurch, dass er die Richtung der Bewegung der *spiritus animales* (kleiner, sehr leichter Teilchen, die sich in den nach Descartes' Auffassung hohlen Nervenbahnen bewegen) ein wenig verändere. Trotz aller Unterschiede hat diese Auffassung eine große Ähnlichkeit zu entsprechenden neueren Ansätzen – wie etwa dem von John Eccles. Denn Eccles zufolge kann der Geist ebenfalls nur mit einem bestimmten Teil der Hirnrinde (dem Liaisonhirn) direkt interagieren. Dabei soll diese Interaktion so vonstatten gehen, dass der Geist kleine funktionelle Einheiten (Module) in diesem Bereich abtastet und die Aktivität einiger Module, die ihm aufgrund ihrer "Offenheit" zugänglich sind, "leicht modifiziert, wobei die Module dann gemeinsam auf diese geringfügigen Änderungen reagieren und diese gemeinsame Reaktion durch die Assoziations- und Kommissurenfasern weiterleiten" (*Gehirn und Geist*, 173). In neueren Arbeiten vermutet Eccles die Geist/Materie-Interaktion sogar in "*Dendronen*", also in Einheiten, die noch wesentlich kleiner als die Module sind.

Im Hinblick auf diese Auffassung stellen sich jedoch die gleichen Fragen wie im Hinblick auf die Auffassung Descartes'. Aus welchen Gründen eigentlich kann der Geist nur auf bestimmte, sehr zentrale Bereiche des Gehirns einwirken? Wenn er überhaupt kausal in den Bereich des Physischen eingreifen kann, dann sollte dies doch an einer Stelle nicht problematischer sein als an einer anderen. Außerdem: Warum nehmen beide Autoren an, dass der Geist nur sehr geringe Veränderungen in der physischen Welt bewirkt? Reicht seine Kraft nicht weiter? Oder "verbirgt" sich der Geist mit Absicht, indem er physische Gegenstände nur auf empirisch kaum feststellbare Weise beeinflusst? Beide Antworten scheinen wenig plausibel.

Das zweite Problem liegt in der Tatsache, dass das kausale Einwirken des Geistes auch dann, wenn seine Wirkungen nur sehr gering sind, den Erhaltungssätzen der Physik widerspricht. Schon Descartes war sich dieses Problems bewusst. Da er den genauen Gehalt der Erhaltungsgesetze noch nicht kannte, glaubte er allerdings, dieses Problem durch die Annahme lösen zu können, dass der Geist nur die Richtung, aber nicht die Größe der Geschwindigkeit der *spiritus animales*

verändere. Spätestens seit Leibniz ist jedoch klar, dass diese Annahme das Problem nicht löst, da auch eine Richtungsänderung den Impuls der betroffenen Teilchen verändert und daher den Satz der Impulserhaltung verletzt. Heute können wir dies sogar noch allgemeiner fassen. Selbst wenn man annimmt, dass die Wirkung des Geistes nicht in der Veränderung des Impulses eines oder mehrerer Teilchen besteht, würde das kausale Eingreifen des Geistes in ein physikalisches System auf jeden Fall eine Änderung des Energiezustandes dieses Systems implizieren und daher mit dem Energieerhaltungssatz in Konflikt kommen. Es scheint so, als hätte die Annahme, dass die kausalen Wirkungen des Geistes nur sehr gering und daher kaum merklich sind, auch den Sinn, diesem Problem auszuweichen. Aber natürlich kann man, falls das so ist, bestenfalls von einer Scheinlösung sprechen.

Außerdem: Es geht nicht nur um die Erhaltungssätze der Physik, sondern viel grundsätzlicher um die *allgemeine* Geltung der Naturgesetze. Jedes Eingreifen des Geistes in die Welt physischer Körper führt zu einer Veränderung in der physischen Welt, für die es keine physischen, sondern nur außer-physische Ursachen gibt. Wenn der Geist kausal in die Körperwelt eingreift, passiert es also zumindest manchmal, dass z.B. Neurotransmitter ausgeschüttet oder Ionenkanäle geöffnet werden, ohne dass es dafür eine physikalische oder chemische Ursache gäbe. Wenn der Geist kausal in die Körperwelt eingreift, müssen daher selbst die grundlegenden Gesetze der Naturwissenschaften in dem Sinne unvollständig sein, als zumindest manchmal Dinge geschehen, die von diesen Gesetzen nicht erfasst werden. Es gibt aber keinerlei empirischen Grund zu glauben, dass es solche Gesetzeslücken tatsächlich gibt. Vielmehr sprechen alle empirischen Befunde für die kausale Geschlossenheit der Körperwelt: Jedes physische Ereignis hat, wenn es überhaupt eine Ursache hat, eine physische Ursache. Und es ist klar, dass diese kausale Geschlossenheit des Physischen mit einem Eingreifen von Seiten des Geistes absolut unvereinbar wäre.

Das schwerwiegendste mit der Annahme des Einwirkens des Geistes auf den Körper verbundene Problem liegt jedoch in der Frage, wie ein solcher kausaler Zusammenhang *überhaupt* gedacht werden kann. Auch hier war Descartes der erste, der die Tragweite dieses Problems erkannt hat. Für ihn stellte sich dieses Problem sogar mit besonderer Schärfe, da sein Hauptanliegen darin bestand, den Gegensatz von Körper und Geist besonders klar herauszustellen. Wenn Körper und Geist, wie Descartes annahm, zwei grundsätzlich verschiedene Substanzen sind, die keinerlei Eigenschaften gemeinsam haben, dann wird die Möglichkeit einer kausalen Interaktion völlig unverständlich. Für Descartes sogar in besonderer Weise, da seiner Auffassung nach jede Ursache ihre Wirkung in gewisser Weise in sich enthalten muss. Für Descartes blieb daher nur der Ausweg zu sagen: Die Erfahrung zeigt uns zwar, *dass* es eine Interaktion von Körper und Geist gibt, aber der Verstand ist außerstande, zu begreifen, *wie* diese Interaktion möglich ist (Briefe an Elisabeth von Böhmen vom 21. Mai und 28. Juni 1643).

Auch wenn man Descartes' Theorie der Kausalität nicht teilt, ergibt sich aus der angenommenen Verschiedenheit von Körper und Geist und insbesondere aus der angenommenen Unräumlichkeit des Geistes für die Möglichkeit der kausalen Interaktion eine Reihe schwerwiegender Fragen (vgl. McGinn 1982, 24f.). Wie kommt es, dass mein Geist auf mein Gehirn, aber auf kein anderes Gehirn einwirken kann? Warum kann der Geist nur auf das Gehirn und nicht direkt auf die Muskeln oder sogar auf die Welt einwirken? Warum werden die Fähigkeiten des Geistes durch bestimmte Gehirnverletzungen beeinträchtigt? Warum ist das Wirken des Geistes überhaupt an das Vorhandensein eines hinreichend komplexen Gehirns gebunden? Sicher, man kann die Unbeantwortbarkeit dieser Fragen ähnlich wie Descartes darauf zurückführen, dass die Interaktion

von Körper und Geist eben ein Geheimnis ist. Aber ich denke, dass sie eher zeigt, dass es sich hier um eine inkohärente Idee handelt.

Auf die theoretischen Probleme, die mit der Annahme mentaler Substanzen einhergehen, kann ich hier nur ganz kurz eingehen. Zunächst einmal geht es um die besonders von Peter Strawson aufgeworfenen Fragen: Welche Gründe könnten dafür sprechen, dass es (außer mir) noch andere Geister gibt? Wie lassen sich Geister voneinander unterscheiden? Und mit Hilfe welcher Kriterien lässt sich die Identität von Geistern feststellen?

Doch diese Liste lässt sich noch erheblich erweitern: Können Geister wahrnehmen und, wenn ja, was nehmen sie wahr? Wenn ein Mensch einen Baum sieht, dann ist der visuelle Eindruck, den er dabei hat, von seinen *physischen* Eigenschaften abhängig – davon, aus welcher Richtung er auf diesen Baum schaut und wie weit der Baum von ihm entfernt ist, und ebenso davon, wie seine Augen und sein Gehirn das einfallende Licht verarbeiten. Geister haben aber keine physischen Eigenschaften. Was bestimmt also ihre Wahrnehmungseindrücke, wenn sie überhaupt welche haben? Können Geister ultraviolettes Licht sehen oder Töne von mehr als 16000 Hz hören? Gibt es Eigenschaften (Magnetismus, elektrische Ladung), die Geister wahrnehmen können, obwohl wir das nicht können? Ähneln das Wahrnehmungsvermögen von Geistern vielleicht eher dem von Fledermäusen als dem unseren? Können Geister vielleicht sogar alles wahrnehmen? (Und was würde das heißen?)

Ähnliche Fragen stellen sich im Hinblick auf die Verständigungsfähigkeit von Geistern. Wie bringen sie es fertig, miteinander zu kommunizieren? Sicher können sie dazu kein physisches Medium verwenden. Könnte es also sein, dass sie sich ihre Gedanken direkt, ohne die Unterstützung durch ein Medium mitteilen? Aber wie könnte das geschehen? Vielleicht hat ein Geist einfach den Eindruck, dass ihm ein anderer Geist etwas mitteilen will. Aber wie ist es, wenn ihm verschiedene Geister gleichzeitig etwas mitteilen möchten? Empfängt er ihre Botschaften gleichzeitig oder nacheinander? Und wie kann er unterscheiden, wer ihm etwas mitteilt, und wie kann er den Fall, dass ihm etwas mitgeteilt wird, von dem unterscheiden, dass er nur den Eindruck hat, dass ihm etwas mitgeteilt wird?

All dies sind Fragen, auf die es einfach keine Antwort zu geben scheint und die damit deutlich machen, dass irgend etwas an der Konzeption absolut immaterieller Geister grundlegend verkehrt ist. Und genau *das* ist der Grund dafür, dass es heute kaum noch Vertreter des Substanz-Dualismus gibt. Der Substanz-Dualist behauptet, dass es neben den physischen Dingen auch immaterielle Dinge gibt, die die Träger mentaler Eigenschaften sind. Aber nicht nur die Annahme, es könnte Geister im Sinne immaterieller Dinge geben, auch die Annahme, Geister könnten wahrnehmen, sich erinnern, denken und ihre Gedanken einander mitteilen, scheint einfach inkohärent.

3.

Doch damit ist das Leib-Seele-Problem noch nicht vollständig gelöst. Es bleibt die Frage nach der Natur mentaler Eigenschaften. Lassen sich mentale *Eigenschaften* doch – auf die eine oder andere Weise – auf physische Eigenschaften zurückführen? Sind mentale Eigenschaften mit physischen Eigenschaften identisch oder auf solche Eigenschaften reduzierbar? Oder handelt es sich um *ontologisch eigenständige* Eigenschaften, die in keiner Weise auf physische Eigenschaften zurückgeführt werden können?

Auf den ersten Blick scheint hier viel für den Eigenschafts-Dualismus zu sprechen. Denn wie um alles in der Welt soll die Eigenschaft, dass ich jetzt Schmerzen empfinde, oder die Eigenschaft, dass

ich eine bestimmte Überzeugung oder einen bestimmten Wunsch habe, auf physische Eigenschaften reduzierbar bzw. mit physischen Eigenschaften identisch sein? Wenn man hier eine Lösung finden will, muss man allerdings zuerst versuchen, genau herauszufinden, welche Arten von mentalen Eigenschaften es gibt und welche Merkmale dieser verschiedenen Eigenschaften für unsere dualistischen Intuitionen verantwortlich sind.

C. McGinn hat hervorgehoben, dass es eine Eigenheit mentaler Eigenschaften ist, dass wir sie unter zwei auf den ersten Blick ganz verschiedenen Arten von Bedingungen zuschreiben. Von uns selbst sagen wir auf der Grundlage unserer eigenen, wenn man so will "inneren" Erfahrung, dass wir z. B. Schmerzen fühlen, einen bestimmten Farbeindruck haben oder Freude empfinden. Anderen dagegen schreiben wir solche mentalen Eigenschaften aufgrund ihres Verhaltens und ihrer verbalen Äußerungen zu. (Diese beiden Arten der Zuschreibung werden häufig als Zuschreibungen aus der Perspektive der ersten Person bzw. aus der Perspektive der dritten Person bezeichnet.) Aus dieser Tatsache ergibt sich in gewisser Weise ein Problem. Denn offenbar verwenden wir Begriffe wie "Schmerz" bei Zuschreibungen aus der Perspektive der ersten Person in derselben Bedeutung wie bei Zuschreibungen aus der Perspektive der dritten Person; wenn das nicht so wäre, wäre ein Dialog wie "Sie scheinen starke Schmerzen zu haben?" "Ja, seit heute morgen habe ich fürchterliche Kopfschmerzen" sinnlos. Andererseits ist es aber plausibel, anzunehmen, dass die Bedeutung von Begriffen mit den charakteristischen Bedingungen ihrer Zuschreibung eng verbunden ist. Also scheint die Bedeutung mentaler Begriffe nicht einheitlich zu sein: sie scheinen immer eine subjektive Komponente zu enthalten, die die Bedingungen der Zuschreibung aus der Perspektive der ersten Person reflektiert, und eine objektive Komponente äußerer Kriterien, in der die Bedingungen der Zuschreibung aus der Perspektive der dritten Person zum Ausdruck kommen. Allerdings ist es sicher so, dass bei manchen mentalen Begriffen die eine Komponente mehr im Vordergrund stehen wird als die andere. Und vielleicht ist es sogar so, dass generell eine der beiden Komponenten mehr Gewicht hat.

Wenn wir auf die Frage nach den verschiedenen Arten mentaler Eigenschaften oder Zustände zurückkommen, ist es sinnvoll, zwei große Gruppen zu unterscheiden: *Empfindungen* auf der einen und *intentionale Zustände* (propositionale Einstellungen) auf der anderen Seite. Zu den Empfindungen gehören körperliche Empfindungen wie Schmerzen, Kitzel und Übelkeit ebenso wie Wahrnehmungseindrücke – wie der Eindruck einer bestimmten Farbe, das Klangerlebnis des Brummens einer Hummel oder das Geschmackserlebnis beim Essen einer süßen Birne. Zwischen diesen beiden Gruppen gibt es zwar eine Reihe von Unterschieden; trotzdem ist es sinnvoll, sie zusammenzufassen. Denn alle Empfindungen scheinen im Wesentlichen durch ihre phänomenalen Eigenschaften definiert zu sein. Für Empfindungen ist ihr qualitativer Gehalt charakteristisch – das, *was* man fühlt, wenn man eine solche Empfindung hat; die Art, *wie* es ist, eine solche Empfindung zu haben. (In der neueren Literatur spricht man hier häufig von den mit Empfindungen verbundenen *Qualia*).

Die zweite große Gruppe mentaler Zustände ist die Gruppe der Zustände, die einen intentionalen Inhalt haben und bei deren Zuschreibung wir deshalb "dass"-Sätze verwenden – etwa, wenn wir Hans eine bestimmte Überzeugung zuschreiben, indem wir sagen "Hans glaubt, *dass* die Erde rund ist". Auch innerhalb der Gruppe der intentionalen Zustände gibt es erhebliche Unterschiede, z.B. zwischen kognitiven Einstellungen wie Überzeugungen und Einstellungen wie Wünschen, Absichten und Befürchtungen, die auch eine konative oder affektive Komponente haben. Allen intentionalen Zuständen ist aber gemeinsam, dass sie durch jeweils zwei Aspekte gekennzeichnet sind: durch die

Art des Zustandes – glauben, wünschen, hoffen, etc. – und durch ihren intentionalen Gehalt – das, was geglaubt, was gewünscht oder was gehofft wird.

Offenbar sind intentionale Zustände anders als Empfindungen nicht durch spezifische phänomenale Qualitäten charakterisiert. Denn bei der Zuschreibung von intentionalen Zuständen spielen äußere Kriterien und insbesondere das Verhalten von Personen eine entscheidende Rolle. Man kann nicht die Absicht haben, ein Auto zu kaufen, ohne zumindest die Tendenz zu haben, das zu tun, was zum Erwerb eines Autos notwendig ist. Und wenn man feststellen will, ob jemand die Überzeugung hat, dass es draußen kalt ist, fährt man in der Regel ebenfalls am besten, wenn man darauf achtet, ob er sich warm anzieht, wenn er nach draußen geht.

Empfindungen und intentionale Zustände unterscheiden sich auch im Hinblick auf den Aspekt der Bewusstheit. Ein Schmerz, den man nicht mehr fühlt, hört auf zu existieren. Empfindungen sind insofern an Bewusstsein gebunden. Überzeugungen und Wünsche dagegen kann man auch dann haben, wenn man sich ihrer nicht bewusst ist. Unbewusste Wünsche zählen seit Freud sogar zu den Hauptfaktoren bei der Erklärung bestimmter psychischer Phänomene. Aber auch unabhängig von der Rolle, die un- oder vorbewusste intentionale Zustände in der Psychoanalyse spielen, würden wir sicher sagen, dass jemand auch dann weiß, wie alt er ist, wenn er darüber im Augenblick nicht nachdenkt, während ein Jucken, das er nicht aktuell verspürt, einfach nicht vorhanden ist.

Bei Empfindungen scheinen es nun gerade ihre Bewusstheit und ihre Subjektivität (d. h. die Tatsache, dass sie im wesentlichen durch ihre phänomenalen Qualitäten charakterisiert sind) zu sein, die es schwierig erscheinen lassen, diese mentalen Zustände mit physikalischen Zuständen zu identifizieren oder auf physikalische Zustände zurückzuführen. Lassen wir Empfindungen deshalb zunächst beiseite und wenden uns den intentionalen Zuständen zu. Für intentionale Zustände ist Bewusstheit keine notwendige Vorbedingung und sie sind auch nicht durch bestimmte phänomenale Qualitäten, sondern eher durch die Art charakterisiert, in der sie unser Verhalten beeinflussen. Intentionale Zustände sind daher vielleicht besser auf physische Zustände reduzierbar oder mit physischen Zuständen identifizierbar als Empfindungen.

Auch intentionale Zustände haben aber einen Aspekt, der in diesem Zusammenhang zu Problemen führt. Denn bei der Zuschreibung spielen Prinzipien der Rationalität eine Rolle, für die ebenfalls im Bereich des Physischen kein Platz zu sein scheint. Wenn wir einer Person bestimmte Wünsche und Überzeugungen zuschreiben, dann geht es uns auch darum, das Verhalten dieser Person als *sinnvoll* zu verstehen – als ein Verhalten, das relativ zu den Wünschen und Überzeugungen dieser Person *rational* ist.¹ Aus diesem Grunde können wir Personen nur intentionale Zustände zuschreiben, die auch im Verhältnis zueinander rational sind, die also insgesamt ein kohärentes und konsistentes Muster von Einstellungen bilden. Nur wenn wir diese Standards der Rationalität anlegen, kann es uns gelingen, eine Einstellung oder ein Verhalten als sinnvoll zu erweisen. Rationalitätserwägungen dieser Art scheinen aber ebenfalls auf den Bereich des Psychischen beschränkt zu sein. Bei der Charakterisierung der Zustände unseres ZNS oder unseres Körpers im allgemeinen spielen solche Erwägungen normalerweise jedenfalls keine Rolle. Wie bei den Empfindungen ihr phänomenaler Charakter scheint es also bei den intentionalen Zuständen das mit ihnen untrennbar verbundene Element der Rationalität unmöglich zu machen, diese Zustände auf physische Zustände zu reduzieren oder mit physischen Zuständen zu identifizieren. Lässt sich dieses Problem lösen?

¹ Zur Rationalität intentionaler Zustände siehe Davidson (1980), Dennett (1978, 1987) und Cherniak (1985).

4.

Nun, in diesem Punkt scheinen sich Lösungsansätze zumindest abzuzeichnen. Insbesondere die Entwicklung der Computerwissenschaften in den letzten Jahrzehnten hat gezeigt, wie rein physische Systeme aussehen könnten, die den mit intentionalen Zuständen verbundenen Rationalitätsanforderungen genügen. Wenn man versucht, dies am Beispiel von Überzeugungen plausibel zu machen, ist es notwendig, zunächst noch einmal genauer zu klären, was es eigentlich heißen soll, dass Überzeugungen zumindest bis zu einem gewissen Grade rational sein müssen.

Sicher wird man nicht verlangen können, dass jemand, der von etwas, sagen wir p , überzeugt ist, auch von *allem* überzeugt sein muss, was logisch aus p folgt. Wenn das so wäre, wären alle Beweise in der Mathematik sinnlos. Aber man wird doch verlangen, dass wer p glaubt, auch alle *offensichtlichen* Folgerungen aus p glaubt. Wer glaubt, dass alle Katzen eigenwillig sind, und glaubt, dass Mietzi eine Katze ist, der sollte auch glauben, dass Mietzi eigenwillig ist. Und wer glaubt, dass Mietzi intelligent und eigenwillig ist, der sollte auch glauben, dass Mietzi intelligent ist. Wenn die Überzeugungen eines Menschen diesen Bedingungen nicht genügen, würde man sicher daran zweifeln, dass er überhaupt Überzeugungen hat.

Zweitens besagt die Rationalitätsanforderung, dass Überzeugungen grundsätzlich konsistent sein müssen; niemand kann *offensichtlich* Widersprüchliches glauben. Auch hier ist es natürlich möglich, dass jemand etwas glaubt, woraus sich in nichttrivialer Weise ein Widerspruch ableiten lässt. Frege zum Beispiel war von der Wahrheit der Axiome seiner Arithmetik völlig überzeugt, bis Russell nachweisen konnte, dass aus diesen Axiomen ein Widerspruch ableitbar ist. Auf der anderen Seite würde jedoch die Annahme, dass jemand zugleich p und nicht- p glaubt, wieder Zweifel daran aufkommen lassen, dass er überhaupt eine entsprechende Überzeugung hat.

Drittens schließlich besagt die Rationalitätsanforderung für Überzeugungen, dass nicht alle (vielleicht nicht einmal die meisten) Überzeugungen einer Person falsch sein dürfen. Denn in einem solchen Fall könnte man nicht mehr sagen, was diese Person eigentlich glaubt.

Inwiefern haben nun die Computerwissenschaften gezeigt, dass es rein physische Systeme geben kann, deren Zustände diesen drei Rationalitätskriterien genügen? Der entscheidende Punkt ist hier der Nachweis, dass sich der Prozess des logischen Schließens mechanisieren lässt. Genau genommen beruht dieses Resultat der Mechanisierbarkeit logischen Schließens auf zwei einander ergänzenden Ergebnissen der mathematischen Beweistheorie und der Computerwissenschaften. Die Beweistheorie hat gezeigt, dass der Begriff der logischen Folgerung auch rein syntaktisch durch die formale Ableitung von strukturierten Zeichenketten aus strukturierten Zeichenketten charakterisiert werden kann. Und die Computerwissenschaften haben gezeigt, dass solche rein syntaktischen Ableitungen mechanisiert werden können. Zusammen bedeuten diese beiden Ergebnisse, dass es möglich ist, in einem rein physischen System mit Hilfe rein physischer Prozesse (in aller Regel durch Abarbeiten von Inferenzprogrammen) aus einer Menge von Zeichenketten genau die Zeichenketten zu erzeugen, die den logischen Folgerungen aus dieser Menge entsprechen.

Wenn man etwa an ein Computersystem denkt, dessen Kernüberzeugungen in geeigneten Datenstrukturen (z.B. Listen) abgelegt sind, dann ist leicht zu sehen, wie ein Inferenzprogramm dafür sorgen kann, dass die Überzeugungen dieses Systems der ersten Rationalitätsanforderung genügen. Dafür ist nur nötig, dass dieses Programm alle offensichtlichen Folgerungen aus den im Datenspeicher vorhandenen Listen berechnet und ebenfalls im Datenspeicher ablegt oder dass das Programm diese Folgerungen zumindest dann berechnet, wenn dies in einem bestimmten Zusammenhang sinnvoll ist. In ähnlicher Weise kann ein Inferenzprogramm auch für die

Konsistenz der Überzeugungen des Systems sorgen, indem es für jede Liste prüft, ob die Negation dieser Liste aus den übrigen Listen abgeleitet werden kann, und, falls das der Fall ist, die geprüfte Liste oder eine der Listen, die zur Ableitung der Negation notwendig waren, aus dem Datenspeicher entfernt. (An dieser Stelle ist, wie man sofort sieht, allerdings das keineswegs einfache Entscheidungsproblem zu lösen, welche Liste entfernt werden soll.)

Inferenzprogramme der gerade geschilderten Art reichen allerdings nicht mehr aus, wenn es um die Erfüllung der dritten Rationalitätsanforderung geht. Denn damit die Überzeugungen eines Systems in der Regel wahr sind, d. h. z. B. mit seiner Umwelt übereinstimmen, ist es offensichtlich notwendig, dass das System über Wahrnehmungskomponenten verfügt, die es ihm ermöglichen, Informationen über diese Umwelt aufzunehmen. Dies ist, wie sich gezeigt hat, ein sehr schwieriges Problem; aber auch hier ist zumindest in den Grundzügen klar, wie dieses Problem durch bloße Symbolmanipulation in einem Computer gelöst werden kann. Aus den Arbeiten von D. Marr etwa geht zumindest im Prinzip hervor, was notwendig ist, um z. B. aus der Grauwertverteilung eines Fernsehbildes die Umweltszene zu rekonstruieren, die zu dieser Grauwertverteilung geführt hat. Insgesamt kann man daher zu Recht sagen, dass die Computerwissenschaften und die Künstliche Intelligenz- Forschung in den letzten Jahren gezeigt haben, dass es rein physische Systeme geben kann, die den für intentionale mentale Zustände charakteristischen Rationalitätsanforderungen genügen.

5.

Bleibt die Frage nach der Naturalisierbarkeit von Bewusstsein und phänomenalen Zuständen. Hier gibt es bisher keine in jeder Hinsicht befriedigenden Lösungen. Aber ich will wenigstens die Richtung andeuten, in die manche Überlegungen heute gehen. Bewusstsein scheint vielen der eigentliche Prüfstein zu sein, an dem der Eigenschafts-Physikalismus in der einen oder anderen Weise scheitern muss. Manchmal wird dabei die spezifische Einheit des Bewusstseins besonders betont, manchmal aber auch der eigenartig durchsichtige Charakter des Bewusstseins, der in der Geschichte immer wieder dazu geführt hat, das Bewusstsein mit einer hell erleuchteten Bühne zu vergleichen. Wenn man jedoch von den Wörtern "Bewusstsein" und "Selbstbewusstsein" ausgeht, wird schnell klar, dass Bewusstsein in erster Linie etwas mit Wissen zu tun hat – mit Selbst-Wissen oder Selbst-Kennntnis. Bewusstsein zu haben, heißt, in direkter Weise über die eigenen (mentalen) Zustände informiert zu sein, Wissen über die eigenen (mentalen) Zustände zu besitzen. Aus diesem Grunde spielt das Stichwort "*Metarepräsentationen*" in diesem Zusammenhang eine große Rolle. Wenn Repräsentationen interne Strukturen (z. B. die zuvor erwähnten Listen) sind, in denen Wissen über die Umwelt gespeichert ist, dann sind Metarepräsentationen Strukturen, in denen Wissen über die eigenen Zustände eines Systems und insbesondere über seine Repräsentationen erster Stufe gespeichert ist. Wenn es aber möglich ist, rein physische Systeme mit internen Repräsentationen zu konzipieren, die Informationen über die Umwelt der Systeme enthalten, dann scheint der Weg zu Systemen mit internen Metarepräsentationen und damit zu Systemen, die Informationen über sich selbst besitzen und daher in diesem Sinne Bewusstsein haben, zumindest nicht sehr weit zu sein.

Bleibt zum Schluss das Problem der *phänomenalen Qualitäten*. Hier hat die Diskussion bisher am wenigsten Klarheit gebracht. Man kann sogar sagen, dass hier die meisten Philosophinnen und Philosophen inzwischen von der Richtigkeit des Eigenschafts-Dualismus überzeugt sind. Das liegt insbesondere an zwei zentralen Argumenten – dem *Argument der Erklärungslücke* von Joseph Levine und dem *Argument des unvollständigen Wissens* von Frank Jackson. Ich will hier nur auf das

erste Argument etwas ausführlicher eingehen. Ausgangspunkt der Überlegungen Levines sind die beiden Aussagen

- (1) Schmerz ist identisch mit dem Feuern von C-Fasern.
- (2) Die Temperatur eines idealen Gases ist identisch mit der mittleren kinetischen Energie seiner Moleküle.

Levine zufolge gibt es einen grundsätzlichen Unterschied zwischen diesen beiden Sätzen: Der zweite ist 'vollständig explanatorisch', der erste nicht. Auf der einen Seite ist es nämlich in einem bestimmten epistemischen Sinn *undenkbar*, dass in einem Gas die mittlere kinetische Energie der Moleküle sagen wir $6.21 \cdot 10^{-21}$ Joule beträgt, dass dieses Gas aber nicht die entsprechende Temperatur von 300 K besitzt. Auf der anderen Seite scheint es aber *sehr wohl denkbar*, dass ich keine Schmerzen fühle, obwohl meine C-Fasern feuern. Worauf beruht dieser Unterschied?

Levines Antwort lautet: Wenn man uns fragen würde, was wir mit dem Ausdruck 'Temperatur' meinen, dann würden wir antworten:

- (2') Temperatur ist die Eigenschaft von Körpern, die in uns bestimmte Wärme- bzw. Kälteempfindungen hervorruft, die dazu führt, dass die Quecksilbersäule in Thermometern, die mit diesen Körpern in Berührung kommen, steigt oder fällt, die bestimmte chemische Reaktionen auslöst, und so weiter.

Mit anderen Worten: Wir würden Temperatur allein durch eine *kausale Rolle* charakterisieren. Dies würde als Antwort auf die gestellte Frage allerdings nicht ausreichen, wenn nicht noch ein zweiter Punkt hinzukäme:

... unsere Kenntnis der Physik und Chemie [macht] es verständlich ..., wie es dazu kommt, dass etwas wie die Bewegung von Molekülen die kausale Rolle spielen kann, die wir mit Temperatur verbinden. (Levine 1983, 357)

Für den explanatorischen Charakter der Aussage (2) gibt es also zwei Gründe:

1. Unser Begriff von Temperatur erschöpft sich vollständig in einer kausalen Rolle.
2. Die Physik kann verständlich machen, dass die mittlere kinetische Energie der Moleküle eines Gases genau diese kausale Rolle spielt.

Warum ist dann aber die Aussage (1) nicht vollständig explanatorisch? Mit dem Ausdruck 'Schmerzen' assoziieren wir doch ebenfalls eine kausale Rolle: Schmerzen werden durch die Verletzung von Gewebe verursacht, sie führen dazu, dass wir schreien oder wimmern, und sie bewirken in uns den Wunsch, den Schmerz so schnell wie möglich loszuwerden. Dies bestreitet auch Levine nicht. Und er bestreitet auch nicht, dass die Identifikation von Schmerzen mit dem Feuern von C-Fasern den Mechanismus erklärt, auf dem diese kausale Rolle beruht. Dennoch gibt es seiner Meinung nach einen entscheidenden Unterschied.

Unser Begriff von Schmerzen umfasst ... mehr als eine kausale Rolle; es gibt auch den *qualitativen Charakter* von Schmerzen, wie es sich anfühlt, Schmerzen zu haben. (Levine 1983, 357)

Levines erster Grund für die These, dass die Aussage (1) nicht vollständig explanatorisch ist, ist also:

1. Unser Begriff von Schmerzen erschöpft sich nicht in einer kausalen Rolle; er umfasst auch einen qualitativen Aspekt – die Art, wie es sich anfühlt, Schmerzen zu haben.

Dies allein ist aber nicht entscheidend. Denn die Aussage (1) könnte immer noch vollständig explanatorisch sein, wenn die Neurobiologie nur verständlich machen könnte, dass sich das Feuern von C-Fasern schmerzhaft anfühlt. Levines zweiter Grund ist daher, dass genau dies nicht der Fall ist.

2. Aus den allgemeinen Gesetzen der Neurobiologie folgt nicht, dass sich das Feuern von C-Fasern auf die für Schmerzen charakteristische Weise – nämlich schmerzhaft – anfühlt.

Offensichtlich gehört zu den zentralen Prämissen dieser Argumentation, dass Schmerzen durch ihren spezifischen qualitativen Aspekt zumindest mitcharakterisiert sind und dass dieser qualitative Aspekt weder durch eine bestimmte kausale Rolle noch durch andere spezifische Verhaltensweisen definiert werden kann. Doch gerade hier können Zweifel aufkommen. Denn diese Prämisse scheint zumindest die Möglichkeit zu implizieren, dass es Wesen gibt, die sich in allen Situationen exakt genau so verhalten wie ich, die genau dasselbe sagen wie ich und die trotzdem nicht in demselben Sinne Schmerzen haben wie ich. Es geht also um das, was neuerdings als die Möglichkeit *philosophischer Zombies* bezeichnet wird.

Aber sind solche Wesen wirklich möglich? Ist es wirklich sinnvoll, z. B. von zwei Wesen, die alle Farbwörter in genau derselben Weise gebrauchen, anzunehmen, dass ihre Farbempfindungen systematisch vertauscht sind? Oder von einem Wesen, das sich genau so verhält wie wir, wenn wir Schmerzen haben, anzunehmen, dass es nichts empfindet, was unseren Schmerzen auch nur entfernt ähnelt? Nun, hier ist die Diskussionslage ziemlich unklar. Viele sind heute – im Anschluss besonders an Thomas Nagels Aufsatz "What is it like to be a bat?" – tatsächlich davon überzeugt, dass alle diese Annahmen sinnvoll sind. Es gibt aber auch andere, die – im Geiste des späten Wittgenstein – meinen, dass diese Annahmen mit unüberwindlichen sprachphilosophischen und erkenntnistheoretischen Problemen konfrontiert sind. Diese Philosophinnen und Philosophen sind der Meinung, dass auch für Empfindungen die äußeren Verhaltenskriterien (die Perspektive der dritten Person) die entscheidende Rolle spielen. Wenn man (wie ich) diese Auffassung teilt, ist auch das Problem phänomenaler Zustände nicht mehr unüberwindlich. Andererseits besteht im Hinblick auf diese Auffassung aber, wie gesagt, alles andere als Einigkeit. Offenbar ist es schwer, der Perspektive der ersten Person so wenig Raum zu lassen. Wenn man an seine eigenen Empfindungen denkt: Es ist doch nicht das Verhalten, das den Schmerz ausmacht. Oder vielleicht doch?

Literatur

Beckermann A: Der Computer – ein Modell des Geistes? In: Krämer S (Hg.), Geist – Gehirn – Künstliche Intelligenz. Zeitgenössische Modelle des Denkens. Berlin/New York 1994

Beckermann A: Analytische Einführung in die Philosophie des Geistes. 2., überarbeitete Aufl. Berlin/New York 2001

Cherniak C: Minimal Rationality. Cambridge, Mass. 1985

Davidson D: Essays on Actions and Events. Oxford 1980 (Dt.: Handlung und Ereignis. Frankfurt am Main 1985)

Dennett D: Brainstorms. Montgomery, Verm. 1978.

Dennett D: The Intentional Stance. Cambridge MA: MIT Press 1987

Descartes R: Discours de la méthode, fr/dt. Hrsgg. und übers. von L. Gäbe. Hamburg 1960

Descartes R: Meditationes de prima philosophia, lt/dt. Auf Grund der Ausgabe von A. Buchenau neu hrsgg. von L. Gäbe, durchgesehen von H.G. Zekl. Hamburg 1977

- Descartes R: Die Leidenschaften der Seele, fr/dt. Hrsgg. und übers. von K. Hammacher. Hamburg 1984
- Descartes R: Über den Menschen und Beschreibung des menschlichen Körpers, übers. von K. E. Rothschuh. Heidelberg 1969
- Descartes R: Briefe. Hrsgg. von M. Bense, übers. von F. Baumgart. Köln 1949
- Eccles J C und Zeier H: Gehirn und Geist. München 1980
- Fodor J A: Representations. Cambridge, Mass. 1981
- Jackson F: Epiphenomenal Qualia. Philosophical Quarterly 32 1982
- Levine J: Materialism and Qualia: The Explanatory Gap. Pacific Philosophical Quarterly 64 (1983)
- Marr D: Vision. New York 1982
- McGinn C: The Character of Mind. Oxford 1982
- Nagel T: What is it like to be a bat? Philosophical Review 83 (1974)
- Strawson P: Individuals. London 1959 (Dt.: Einzelding und logisches Subjekt. Stuttgart 1972)

*

Erstpublikation in: Ulrich Lorenz (Hg.): Philosophische Psychologie. Freiburg/München: Verlag Karl Alber 2003. S.203-221. Nachdruck mit freundlicher Genehmigung des Verlages.