

Geist und Welt

Herausgegeben von Renate Breuninger und Achim Stephan

Interdisziplinäre Schriftenreihe des Humboldt-Studienzentrums
Universität Ulm

Beiträge zur Philosophie
Band 18

Humboldt-Studienzentrum
Universität Ulm
2001

Ist Bewußtsein reduktiv erklärbar?

Ansgar Beckermann

Ist Bewußtsein reduktiv erklärbar?

1.

Was ist das überhaupt – Bewußtsein? Beim Gebrauch des *Substantivs* „Bewußtsein“ sollte man eine gewisse Vorsicht walten lassen. Die deutsche Sprache kennt dieses Substantiv nur innerhalb weniger Redewendungen – „Dann verlor er das Bewußtsein“, „Nach vielen Stunden kam er wieder zu Bewußtsein“, „Im Bewußtsein seiner schweren Verantwortung“ usw. Außerhalb dieser Redewendungen kommt das Wort dagegen praktisch nicht vor; und deshalb sollte man skeptisch werden, wenn Philosophen Sätze formulieren wie „Menschen haben ein Bewußtsein“. „Bei Bewußtsein sein“ ist gebräuchliches Deutsch, „ein Bewußtsein haben“ nicht.¹ Diese Bemerkungen mahnen zwar zu besonderer Vorsicht, helfen in der Sache jedoch nicht viel weiter. Wovon reden wir also, wenn wir fragen, ob Bewußtsein reduktiv erklärbar ist? Fragen wir vielleicht so: Was fehlt dem, der das Bewußtsein verloren hat?

¹ Noch verheerender ist die philosophische Substantivierung des an sich harmlosen Wörtchens „selbst“. Völlig in Ordnung sind Sätze wie „Der Bundeskanzler kam selbst“ oder „Das kann doch jeder selbst erledigen“; aber was ist wohl mit einem Satz gemeint wie „Jeder Mensch hat ein Selbst“?

Bewußtlose Personen sind nicht ansprechbar; sie hören offenbar nicht, was man ihnen sagt – jedenfalls nicht so, wie wir normalerweise hören; sie sehen und fühlen auch nicht so, wie wir normalerweise sehen und fühlen; und ihnen fehlt das normale Schmerzempfinden. Bewußtlose sind also in ihrer Wahrnehmungsfähigkeit stark beeinträchtigt; ihnen fehlt, was man das normale Wahrnehmungsvermögen nennen könnte. Und: Bewußtlose liegen in der Regel im Bett, auf einer Bahre oder sogar auf der Straße. Sie können sich also nicht einmal aufrecht halten. Um so mehr fehlt ihnen die Fähigkeit, auf ihre Umwelt angemessen zu reagieren – z.B. sich selbst zu ernähren, einer Gefahr auszuweichen oder auf Fragen zu antworten. Kurz gesagt: Bewußtlosen fehlen die *normalen Fähigkeiten, wahrzunehmen und sich situationsangemessen zu verhalten*.

Ein zweiter Aspekt dessen, was wir mit Bewußtsein verbinden, hat mit dem Wort selbst zu tun. Denn in diesem Wort steckt ganz offensichtlich der Bestandteil „Wissen“. Um welche spezielle Art von Wissen handelt es sich hier? Wenn eine Person aus der Bewußtlosigkeit erwacht, wird man sie vielleicht fragen: „Können Sie mich hören?“ Oder: „Wie viele Finger sind das?“ Aber man wird sie auch fragen: „Spüren Sie das?“, wobei man sie leicht mit einer Nadel sticht. Und mit dieser Frage will man nicht klären, ob die Person ihre Umwelt wieder normal wahrnehmen kann, sondern ob sie auch wieder über ihre eigenen mentalen Zustände informiert ist. Bewußtsein hat also offenbar etwas mit *Selbstwissen* oder *Selbstkenntnis* zu tun. Bei Bewußtsein zu sein, impliziert, über die eigenen mentalen Zustände informiert zu sein und über sie Auskunft geben zu können.

Ein dritter Aspekt von Bewußtsein schließlich hat mit dem zu tun, was neuerdings *phänomenales* Bewußtsein genannt wird. Bei Bewußtsein sein heißt auch, Erlebnisse zu haben. Bewußtseinsfähige Wesen zeichnen sich dadurch aus, daß sie mentale Zustände nicht nur haben, sondern daß sie sie *erleben*. Wir haben Schmerzen nicht nur in dem Sinne, in dem wir Übergewicht oder Wünsche und Überzeugungen haben, wir erleben unsere Schmerzen – es fühlt sich auf eine bestimmte Weise an, Schmerzen zu haben. Und es fühlt sich auf andere

Weise an, wenn uns schwindlig ist oder wenn wir Hunger haben. Genauso wie sich Roterlebnisse anders anfühlen als Grünerlebnisse und Süßerlebnisse anders als Wärmeerlebnisse. Alle diese Zustände haben einen je eigenen *qualitativen Charakter*.

Selbst wenn wir es bei dieser kurzen Skizze belassen, haben wir schon drei Aspekte, die alle mit dem verbunden sind, was wir Bewußtsein nennen:²

1. Die normalen Fähigkeiten, wahrzunehmen und sich situationsangemessen zu verhalten.
2. Selbstkenntnis, d.h. Wissen über die eigenen mentalen Zustände.
3. Die Tatsache, daß wir Wahrnehmungseindrücke und körperliche Empfindungen nicht nur haben, sondern erleben; daß es sich auf eine bestimmte qualitative Weise anfühlt, diese Eindrücke und Empfindungen zu haben.

Ein Problem, das an dieser Stelle zu Verwirrung führen kann, besteht darin, daß diese drei Aspekte in *unserem* Leben normalerweise nicht getrennt sind. Wenn jemand wach und bei Bewußtsein ist, dann verfügt er in aller Regel über die normale Wahrnehmungsfähigkeit und die Fähigkeit zum situationsgerechten Handeln ebenso wie über die für Bewußtsein notwendige Selbstkenntnis; und außerdem haben wache Personen auch Erlebnisse, es fühlt sich für sie auf eine bestimmte Weise an, Hunger zu haben oder etwas Rotes zu sehen. Es scheint einfach nicht vorzukommen, daß jemand normal wahrnehmen und handeln kann, ihm dabei aber die notwendige Selbstkenntnis fehlt. Ebenso wie es einfach nicht vorzukommen scheint, daß jemand über die normalen Fähigkeiten wahrzunehmen und zu handeln sowie über die notwendige Selbstkenntnis ver-

² Vgl. hierzu auch Chalmers (1995, sec. II) und Chalmers (1996).

fügt, daß er aber nichts *erlebt*, daß seine Wahrnehmungseindrücke und Empfindungen nicht mit spezifischen Erlebnisqualitäten verbunden sind.³

Begrifflich lassen sich die genannten drei Aspekte von Bewußtsein aber sehr wohl unterscheiden: Begrifflich impliziert die Tatsache, daß jemand normal wahrnehmen und auf seine Umwelt reagieren kann, nicht, daß er auch über die eigenen mentalen Zustände informiert ist. Und begrifflich impliziert die Tatsache, daß jemand über die normalen Fähigkeiten wahrzunehmen und zu handeln sowie über die notwendige Selbstkenntnis verfügt, nicht, daß seine Wahrnehmungseindrücke und Empfindungen mit spezifischen Erlebnisqualitäten verbunden sind. Die genannten drei Aspekte von Bewußtsein müssen daher getrennt erklärt werden. Statt mit der einen Frage „Ist Bewußtsein reduktiv erklärbar?“ haben wir es genau genommen mit drei Fragen zu tun:

1. Können die normalen Fähigkeiten, wahrzunehmen und situationsangemessen zu handeln, reduktiv erklärt werden?
2. Kann Selbstkenntnis reduktiv erklärt werden?
3. Können Erlebnisse – mentale Zustände mit phänomenalen Qualitäten – reduktiv erklärt werden?

Im folgenden will ich mich auf eine Diskussion der dritten Frage beschränken, und zwar aus einem einfachen Grund. In der Philosophie des Geistes hat sich die Auffassung weitgehend durchgesetzt, daß diese Frage den „harten“ Kern⁴ des „Bewußtseinsproblems“ bildet. Die reduktive Erklärbarkeit unserer

³ Dieser letzte Fall wäre der Fall eines philosophischen Zombies – eines Wesens, das sich in seinem Verhalten und in der Art und Weise, wie es mit uns und anderen redet, nicht von einem normalen wachen Menschen unterscheidet, das aber trotzdem keine Empfindungen hat oder genauer: dessen Empfindungszustände keinen qualitativen Charakter haben.

⁴ Chalmers (1995, 10).

normalen Wahrnehmungs- und Handlungsfähigkeiten ist nicht wirklich umstritten. Und auch die zweite Frage, ob Selbstkenntnis reduktiv erklärbar ist, würden viele mit einem – vielleicht vorsichtigen – „Ja“ beantworten. Aber der qualitative Charakter von Empfindungen scheint eine unüberwindbare Hürde zu sein, wenn es um die reduktive Erklärbarkeit von Erlebnissen geht. Die Gründe für diese Auffassung möchte ich gerne ausführlicher erläutern.

2.

Keine der drei genannten Fragen kann allerdings zufriedenstellend beantwortet werden, wenn nicht zuvor geklärt ist, was mit dem Ausdruck „reduktiv erklärbar“ gemeint sein soll. Stellen wir uns folgende Situation vor: Bei einem Spaziergang begegnen wir einer bekannten Persönlichkeit – sagen wir Boris Becker; und plötzlich – sozusagen aus dem Nichts – entsteht neben dieser Persönlichkeit eine molekülidentische Kopie – ein Wesen, das aus Molekülen derselben Art aufgebaut ist, die auf genau dieselbe Art angeordnet sind wie beim Original. Wir lassen uns dadurch jedoch nicht irritieren. Vielmehr geht uns ebenso plötzlich die Frage durch den Kopf, welche Eigenschaften Boris Becker und sein molekülidentischer Doppelgänger wohl gemeinsam haben. Offenbar sehen die beiden gleich aus, und beide haben auch dasselbe Gewicht; aber können sie auch gleich gut Tennis spielen und haben sie z.B. dieselben Krankheiten? Das scheint zumindest plausibel; denn zwei molekülidentische Wesen werden wohl über dieselben physischen Fähigkeiten verfügen; und wenn Boris Becker kurzsichtig sein sollte, wird auch sein molekülidentischer Doppelgänger kurzsichtig sein. Schließlich sind seine Augen genau so aufgebaut wie die des Originals. Auf der anderen Seite gibt es aber auch eine Reihe von Eigenschaften, die Boris Becker und seine Kopie sicher nicht gemeinsam haben – z.B. alle historischen und relationalen Eigenschaften. Die Kopie hat niemals Wimbledon

gewonnen; damals hat sie noch gar nicht existiert. Und die Kopie hat auch nicht dieselben Eltern wie Boris Becker; wie es aussieht, hat sie gar keine Eltern. Außerdem ist die Kopie nicht in dieselbe Schule gegangen, und sie war nicht mit derselben Frau verheiratet. Und wie steht es mit den mentalen Eigenschaften? Wenn Boris Becker im Augenblick Zahnschmerzen hat, hat dann auch sein Doppelgänger Zahnschmerzen?

Wir wissen heute schon eine Menge über die neuronalen Grundlagen unserer Schmerzempfindungen. Nehmen wir also an, wir haben herausgefunden, daß für uns Menschen gilt:

(1) Wenn die XYZ-Neuronen in unserem Gehirn feuern, fühlen wir Schmerzen.

Wenn das wirklich so ist, d.h. wenn der Satz (1) wahr ist, dann haben wir offenbar eine Antwort auf unsere Frage. Denn wenn im Gehirn von Boris Becker die XYZ-Neuronen feuern, dann ist das offenbar auch bei seiner Kopie so. Schließlich ist diese Kopie ein molekülidentischer Doppelgänger. Und wenn die XYZ-Neuronen im Gehirn der Kopie feuern, dann fühlt sie aufgrund der Wahrheit von (1) auch Schmerzen.

Allerdings müssen wir uns klar machen, daß es hier einen wichtigen Unterschied gibt, der bisher noch gar nicht in den Blick gekommen ist. Gesetze wie der Satz (1) oder allgemeiner Sätze wie

(2) Wenn im Gehirn einer Person bestimmte Neuronen feuern, dann hat diese Person die mentale Eigenschaft *F*

können sehr unterschiedlichen Charakter haben – sie können *emergent*, d.h. selbst *nicht weiter ableitbar* sein; sie können sich aber auch *zwangsläufig* aus den *allgemeinen Naturgesetzen* ergeben. Im ersten Fall heißt auch *F* selbst *emergent*. Im zweiten Fall heißt *F* *reduktiv erklärbar*.

Der Unterschied zwischen emergenten und reduktiv erklärbaren Eigenschaften wurde in aller Klarheit zum ersten Mal von Charles D. Broad in seinem Buch *The Mind and Its Place in Nature* entwickelt.⁵ Diese Überlegungen entstanden im Rahmen der *Vitalismusdebatte*, die zu Beginn des 20. Jahrhunderts die Gemüter vieler Wissenschaftler genauso erregte wie heute die Debatte um das Leib-Seele-Problem. Ein entscheidender Schritt in diesen Überlegungen war, daß Broad klar machte, daß das Vitalismusproblem nur der Spezialfall eines sehr allgemeinen Problems ist – des Problems, wie sich die *Makro-eigenschaften* eines Systems zu seiner *Mikrostruktur* verhalten, d.h. zu den *Eigenschaften* der Teile, aus denen das System besteht, und der *Anordnung* dieser Teile. Was diese Frage betrifft, so sind im Prinzip genau zwei Antworten möglich: Man kann die Auffassung vertreten, daß eine Makroeigenschaft *F* *nicht* durch die physischen Teile des Systems und deren Anordnung, sondern nur durch eine weitere nichtphysische Substanz erklärt werden kann (wer diese Auffassung vertritt, ist nach Broad Vertreter einer *Komponententheorie*); man kann aber auch der entgegengesetzten Auffassung sein, daß die Eigenschaft *F* sehr wohl durch die physischen Teile des Systems und deren Anordnung erklärbar ist. Im letzteren Fall muß man Broad zufolge jedoch zwei weitere Möglichkeiten unterscheiden. Denn wenn *F* durch die physischen Teile des Systems und deren Anordnung erklärbar ist, kann *F* immer noch *reduktiv erklärbar* oder aber *emergent* sein. Broad selbst charakterisiert diesen Unterschied so:

“Put in abstract terms the emergent theory asserts that there are certain wholes, composed (say) of constituents A, B, and C in a relation R to each other; that all

⁵ Broad spricht zwar nicht von reduktiver, sondern von *mechanischer* Erklärbarkeit. Da es jedoch auch Broad nicht um mechanische Erklärungen im Wortsinn geht, scheint mir der erste Ausdruck angemessener. Eine ausführliche Rekonstruktion der Überlegungen Broads findet sich in Beckermann (2001b).

wholes composed of constituents of the same kind as A, B, and C in relations of the same kind as R have certain characteristic properties; that A, B, and C are capable of occurring in other kinds of complex where the relation is not of the same kind as R; and that the characteristic properties of the whole R(A,B,C) cannot, even in theory, be deduced from the most complete knowledge of the properties of A, B, and C in isolation or in other wholes which are not of the form R(A,B,C). The mechanistic theory rejects the last clause of this assertion.”
(Broad 1925, 61)

Zwei Dinge gilt es also festzuhalten:

1. Sowohl reduktiv erklärbare als auch emergente Makroeigenschaften hängen Broads Terminologie zufolge *in gesetzmäßiger Weise* von den Mikrostrukturen der entsprechenden Systeme ab. Wenn wir die Makroeigenschaft F eines Systems S betrachten, das über die Mikrostruktur $[C_1, \dots, C_n; R]$ verfügt, d.h. das aus den Teilen C_1, \dots, C_n besteht, die in der Weise R angeordnet sind, ist nach Broad der Satz

„Für alle x : wenn x die Mikrostruktur $[C_1, \dots, C_n; R]$ hat, dann hat x die Makroeigenschaft F “

also ein *wahres Naturgesetz* – unabhängig davon, ob F reduktiv erklärbar oder emergent ist.

2. Reduktiv erklärbare unterscheiden sich von emergenten Eigenschaften jedoch darin, daß die ersteren (wenigstens im Prinzip) aus der vollständigen Kenntnis all der Eigenschaften abgeleitet werden können, die die Komponenten isoliert oder in anderen Anordnungen haben, während dies für die letzteren nicht gilt.

Was meint Broad aber damit, wenn er davon spricht, daß emergente Eigenschaften „aus der vollständigen Kenntnis all der Eigenschaften abgeleitet werden können, die die Komponenten isoliert oder in anderen Anordnungen ha-

ben“? Soweit ich sehen kann, soll dies folgendes heißen: Ist F emergent, dann folgt aus den *allgemeinen* Naturgesetzen, die für die Teile C_1, \dots, C_n gelten, daß ein System mit der Mikrostruktur $[C_1, \dots, C_n; R]$ alle *Merkmale* besitzt, die für F charakteristisch sind. Broads Überlegungen lassen sich daher in den folgenden beiden Definitionen zusammenfassen:

Reduktive Erklärbarkeit

Die Makroeigenschaft F eines komplexen Systems mit der Mikrostruktur $[C_1, \dots, C_n; R]$ ist genau dann *reduktiv erklärbar*, wenn aus den *allgemeinen* für die Komponenten C_1, \dots, C_n geltenden Naturgesetzen folgt, daß Systeme mit der Mikrostruktur $[C_1, \dots, C_n; R]$ alle für die Eigenschaft F charakteristischen Merkmale besitzen.

Emergenz

Die Makroeigenschaft F eines komplexen Systems mit der Mikrostruktur $[C_1, \dots, C_n; R]$ ist genau dann *emergent*, wenn

- (a) auf der einen Seite der Satz „Für alle x : wenn x die Mikrostruktur $[C_1, \dots, C_n; R]$ hat, dann hat x die Eigenschaft F “ ein wahres Naturgesetz ist, wenn aber
- (b) aus den *allgemeinen* für die Komponenten C_1, \dots, C_n geltenden Naturgesetzen *nicht* folgt, daß Systeme mit der Mikrostruktur $[C_1, \dots, C_n; R]$ alle für die Eigenschaft F charakteristischen Merkmale besitzen.

Ohne Zweifel sind diese Definitionen kompliziert, und ich will deshalb gleich versuchen, sie mit Hilfe einiger Beispiele zu veranschaulichen; vorab läßt

sich der Unterschied zwischen emergenten und reduktiv erklärbaren Eigenschaften aber auch schon mit Hilfe einer theologischen Metapher erläutern.⁶ Wenn die grundlegenden Eigenschaften aller Dinge festliegen, ergeben sich die reduktiv erklärbaren Eigenschaften von selbst. Nachdem Gott die physikalischen Komponenten, aus denen alle Gegenstände aufgebaut sind, erschaffen und die allgemeinen für diese Komponenten geltenden Naturgesetze in Kraft gesetzt hatte, brauchte er sich um die reduktiv erklärbaren Makroeigenschaften daher nicht mehr zu kümmern. Sie ergaben sich von allein. Die emergenten Eigenschaften bedurften jedoch eines eigenen Schöpfungsaktes. Gott mußte erstens diese Eigenschaften erschaffen und er mußte zweitens all die Gesetze eigens in Kraft setzen, die physikalische Mikrostrukturen mit emergenten Eigenschaften verbinden.

An dieser Stelle sollte auch noch ein weiterer wichtiger Unterschied zwischen emergenten und reduktiv erklärbaren Eigenschaften angesprochen werden. Broad selbst betont an verschiedenen Stellen, daß das Gesetz „Alle Systeme mit der Mikrostruktur $[C_1, \dots, C_n; R]$ haben die Makroeigenschaft F “ bei reduktiv erklärbaren Eigenschaften einen ganz anderen Status hat als bei emergenten. Wenn F emergent ist, ist dieses Gesetz, wie Broad sagt, ein nicht weiter ableitbares Gesetz („a *unique* and *ultimate* law“). D.h., dieses Gesetz ist (a) kein Spezialfall, der aus einem allgemeinen Gesetz durch Einsetzung bestimmter Werte für bestimmbare Variablen gewonnen werden kann. Es ist (b) kein Gesetz, das durch Kombination zweier oder mehrerer allgemeiner Gesetze gewonnen werden kann. Und was vielleicht am wichtigsten ist: Wenn F emergent ist, dann kann dieses Gesetz (c) *nur* dadurch *entdeckt* werden, daß man erstens eine Reihe von Systemen mit der Mikrostruktur $[C_1, \dots, C_n; R]$ untersucht, zweitens dabei feststellt, daß alle diese Systeme die Eigenschaft F haben, und drittens

⁶ Vgl. dazu z.B. Kripke (1972) und Chalmers (1996, 38, 41).

dieses Ergebnis induktiv auf alle Systeme mit dieser Mikrostruktur überträgt.⁷ Bei reduktiv erklärbaren Eigenschaften liegen die Dinge dagegen ganz anders.

“In order to predict the behaviour of a clock a man need never have seen a clock in his life. Provided he is told how it is constructed, and that he has learnt from the study of other material systems the general rules about motion and about the mechanical properties of springs and of rigid bodies, he can foretell exactly how a system constructed like a clock must behave.” (Broad 1925, 65)

Wenn die Makroeigenschaft F eines Systems S mit der Mikrostruktur $[C_1, \dots, C_n; R]$ reduktiv erklärbar ist, kann man also *ohne je ein System mit dieser Mikrostruktur untersucht zu haben* wissen, daß S – genauso wie alle Systeme mit dieser Mikrostruktur – F besitzt. In diesem Fall folgt das einfach aus den allgemeinen für die Komponenten C_1, \dots, C_n geltenden Naturgesetzen. Bei reduktiv erklärbaren Eigenschaften ist es daher in diesem Sinne *undenkbar*, daß ein System zwar die Mikrostruktur $[C_1, \dots, C_n; R]$, aber nicht die Eigenschaft F besitzt. Wenn aus den allgemeinen Naturgesetzen folgt, daß alle Systeme mit dieser Mikrostruktur F besitzen, ist es – zumindest *relativ zu diesen Naturgesetzen* – unmöglich, daß ein System die Mikrostruktur $[C_1, \dots, C_n; R]$ hat, die Eigenschaft F aber nicht besitzt. Es gibt also einen eindeutigen *Test*, um herauszufinden, ob die Makroeigenschaft F eines Systems reduktiv erklärbar ist. Man muß nur fragen, ob diese Eigenschaft *vor dem ersten Auftreten* von Systemen mit der Mikrostruktur $[C_1, \dots, C_n; R]$ hätte prognostiziert werden können bzw. ob es – relativ zu den grundlegenden Naturgesetzen – *undenkbar* ist, daß ein System mit der Mikrostruktur $[C_1, \dots, C_n; R]$ die Eigenschaft F nicht besitzt.

⁷ Vgl. Broad (1925, 64f.).

Gibt es gute Beispiele für emergente und für reduktiv erklärbare Eigenschaften? Broad selbst war der Auffassung, daß das Verhalten fast aller chemischen Verbindungen in dem von ihm erläuterten Sinne emergent ist. Zumindest war er der Meinung,

„that, so far as we know at present, the characteristic behaviour of Common Salt cannot be deduced from the most complete knowledge of the properties of Sodium in isolation; or of Chlorine in isolation; or of other compounds of Sodium, such as Sodium Sulphate, and of other compounds of Chlorine, such as Silver Chloride.“ (Broad 1925, 59).

Allerdings war er auch der Meinung, daß diese Auffassung durch den Fortschritt der Wissenschaft überholt werden könne. Und in der Tat sehen die meisten Wissenschaftler dies heute anders. Nehmen wir als Beispiel die *Festigkeit* und die *Wasserlöslichkeit* von Kochsalz.

Kochsalz ist bekanntlich Natriumchlorid; es besteht aus Natrium- und Chloratomen, d.h. genauer aus Natrium- und Chlorionen. Natriumatome besitzen in ihrer äußersten Schale nur ein einziges Elektron, das leicht abgespalten werden kann. Chloratome dagegen besitzen in ihrer äußersten Schale sieben Elektronen; diese Atome sind daher „bestrebt“, ihre äußerste Schale mit einem weiteren Elektron aufzufüllen, um so auf die Idealzahl von acht Elektronen zu kommen. Wenn Natrium- und Chloratome miteinander reagieren, geschieht deshalb folgendes. Das Natriumatom gibt sein äußerstes Elektron ab, und dieses Elektron wird vom Chloratom aufgenommen. So entstehen positiv geladene Natrium- und negativ geladene Chlorionen, die sich aufgrund der zwischen ihnen bestehenden elektromagnetischen Anziehungskräfte in einer Gitterstruktur anordnen. Entscheidend sind hier zunächst diese Anziehungskräfte. Sie erklären, warum Kochsalz unter normalen Bedingungen fest ist. Diese Kräfte sind nämlich so groß, daß die einzelnen Ionen an ihren relativen Positionen „festgezurr“ sind. Wenn sich ein Stück Kochsalz bewegt, bewegt sich daher immer das ganze

Stück. Seine Teile verändern ihre relativen Positionen nicht, und deshalb behält das Stück Kochsalz seine Form. Die starken Anziehungskräfte sind auch dafür verantwortlich, daß es immer eines gewissen Kraftaufwands bedarf, um ein Stück Kochsalz zu zerteilen. Mit anderen Worten: Aus den allgemeinen Naturgesetzen ergibt sich, welche Kräfte zwischen den Natrium- und Chlorionen wirken, aus denen Kochsalz besteht. Und aus diesen Kräften ergibt sich, daß sich Ionengitter aus Natrium- und Chlorionen genau so verhalten, wie dies für feste Körper charakteristisch ist. Und warum ist Kochsalz wasserlöslich? Das liegt zum einen wieder daran, daß Kochsalz aus einem Gitter von positiv und negativ geladenen Ionen besteht. Zum anderen liegt es an der Dipolstruktur der H_2O -Moleküle. Denn aufgrund dieser Struktur können H_2O -Moleküle die einzelnen Ionen aus ihrer Position im Gitter herauslösen.

Gibt es auch Beispiele für emergente Eigenschaften? Diese Frage ist nicht so leicht zu beantworten. Denn die wenigen Beispiele von Eigenschaften, bei denen man vermuten könnte, sie seien emergent, erweisen sich bei näherem Hinsehen in aller Regel doch als reduktiv erklärbar. Deshalb muß an dieser Stelle ein fiktives Beispiel genügen – ein Beispiel, das auf das schöne Buch *Vehicles* von Valentin Braitenberg zurückgeht. Braitenbergs Grundidee ist, daß schon sehr einfach konstruierte künstliche Wesen ein Verhalten zeigen, das man versucht ist, mit psychologischen Begriffen zu beschreiben. Nehmen wir zwei Wesen a) und b), die im wesentlichen aus je zwei Sensoren und Motoren bestehen. Der einzige Unterschied ist, daß die Sensoren beim Wesen a) mit den Motoren auf der gleichen Seite, beim Wesen b) mit den Motoren auf der gegenüberliegenden Seite verbunden sind. Dies führt dazu, daß sich das Wesen a) von einer Lichtquelle, auf die es zufährt, abwendet, während das Wesen b) sich dieser Lichtquelle annähert. Das Wesen b) „mag“ Licht, während das Wesen a) Licht „meidet“ (Abb. 1).

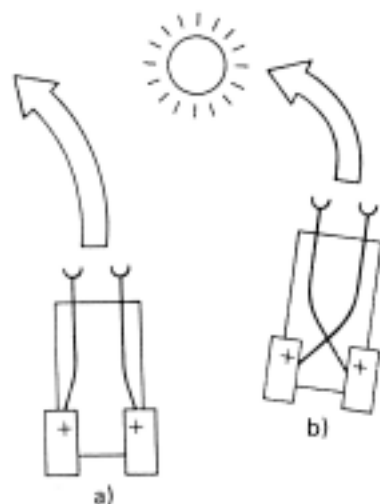


Abbildung 1

Stellen wir uns nun folgendes vor: Im allgemeinen verhalten sich Wesen des Typs a) genau wie gerade beschrieben – sie wenden sich von einer Lichtquelle ab, auf die sie zufahren. Aber unter ganz bestimmten Umständen, nämlich dann, wenn die Sensoren mit den Motoren durch eine bestimmte Art von Draht – sagen wir Silberdraht – verbunden sind, verhalten sich diese Wesen ganz anders. Sie fahren wie Wesen des Typs b) auf die Lichtquelle zu, ohne daß dies aufgrund der normalen Materialeigenschaften zu erwarten wäre (Abb. 2). Dieses Verhalten wäre emergent. Denn so, wie das Beispiel konstruiert ist, ist es auf der einen Seite ein Naturgesetz, daß jedes Wesen auf eine Lichtquelle zufährt, in dem Sensoren und Motoren wie beim Wesen a), aber mit Silberdraht verbunden sind; auf der anderen Seite ergibt sich dieses Verhalten jedoch *nicht* aus den allgemeinen für Sensoren, Silberdrähte und Motoren geltenden Naturgesetzen.

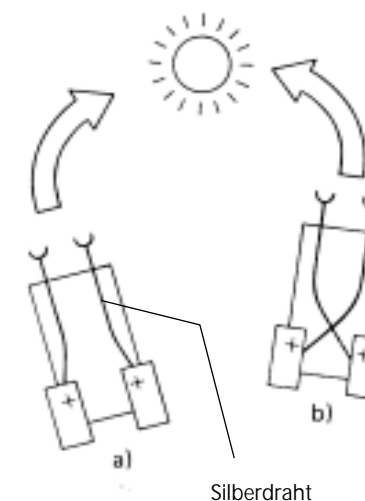


Abbildung 2

Sicher, diese Geschichte scheint irgendwie absurd. Denn einerseits scheint es einfach undenkbar, daß sich ein Wesen vom Typ a) nur deshalb wie ein Wesen vom Typ b) verhalten soll, weil in ihm Sensoren und Motoren durch Silberdrähte verbunden sind. Silberdrähte erweisen sich schließlich in allen anderen Zusammenhängen als ganz normale Leiter. Und was passiert in diesen Drähten – oder in den durch sie verbundenen Sensoren und Motoren – eigentlich, wenn sich Wesen vom Typ a) auf eine Lichtquelle zu bewegen? Falls der geschilderte Fall allerdings trotzdem eintreten sollte, neigen wir auf der anderen Seite dazu anzunehmen, daß Silberdraht eine grundlegende, für das beobachtete Verhalten verantwortliche Eigenschaft hat, die uns bisher nur entgangen ist. D.h., auf die eine oder andere Weise versuchen wir zu zeigen, daß das vermeintlich emergente Verhalten letzten Endes doch reduktiv erklärbar ist. Wir können uns einfach nicht vorstellen, daß auf der grundlegenden physikalischen Ebene etwas nicht mit rechten Dingen zugeht. Und genau deshalb haben viele eine intuitive

Abneigung gegen emergente Eigenschaften. Denn die Existenz dieser Eigenschaften ist offenbar mit der allgemeinen Geltung der grundlegenden Naturgesetze unvereinbar.⁸ Und wenn überhaupt etwas, sind es doch offenbar diese grundlegenden Naturgesetze, die allgemein und ausnahmslos gelten. Trotzdem ist es natürlich *möglich*, daß die grundlegenden Naturgesetze nicht allgemein gelten. Und dementsprechend ist es auch durchaus möglich, daß es emergente Eigenschaften gibt. Selbst wenn dies in gewisser Weise sehr unschön wäre.

3.

Nachdem wir soweit geklärt haben, wie der Begriff der reduktiven Erklärbarkeit zu verstehen ist, können wir zu der Frage zurückkommen, ob Erlebnisse wie Schmerzen reduktiv erklärbar sind oder nicht. Ich hatte schon gesagt, daß viele Philosophinnen und Philosophen gerade im qualitativen Charakter von Erlebnissen eine unüberwindbare Hürde für die reduktive Erklärbarkeit dieser mentalen Zustände sehen. Ich möchte das am Beispiel der Überlegungen Joseph Levines⁹ erläutern, die unter dem Stichwort „Argument der Erklärungslücke“ bekannt geworden sind.

Betrachten wir die beiden Aussagen

(3) Schmerzen sind auf das Feuern von XYZ-Neuronen im Gehirn zurückführbar

und

⁸ Vgl. hierzu Beckermann (2000; 2001a, 240-245).

⁹ Vgl. besonders Levine (1983; 1993).

(4) Die Temperatur eines Gases ist auf die mittlere kinetische Energie seiner Moleküle zurückführbar.¹⁰

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 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 in meinem Gehirn die XYZ-Neuronen feuern. Das kann man auch so ausdrücken: Wenn man alle Gesetze kennt, die für die Bewegungen der Moleküle eines Gases gelten, kann man auch, *ohne entsprechende empirische Untersuchungen anzustellen*, wissen, daß ein Gas, in dem die mittlere kinetische Energie der Moleküle $6.21 \cdot 10^{21}$ Joule beträgt, eine Temperatur von 300 K besitzt. Ohne entsprechende empirische Untersuchungen kann man aber *nicht* wissen, daß ich Schmerzen fühle, wenn in meinem Gehirn die XYZ-Neuronen feuern. Worauf beruht dieser Unterschied?

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

(4') 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.

¹⁰ Levine spricht in diesem Zusammenhang zwar nicht von „reduktiver Rückführbarkeit“, sondern von „Identität“; seine Überlegungen lassen sich aber ohne weiteres übertragen.

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:

“... our knowledge of chemistry and physics makes intelligible how it is that something like the motion of molecules could play the causal role we associate with heat. Furthermore, antecedent to our discovery of the essential nature of heat, its causal role, captured in statements like [(4)], exhausts our notion of it. Once we understand how this causal role is carried out there is nothing more we need to understand.” (Levine 1983, 357)

Für den explanatorischen Charakter der Aussage (4) 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.

Mit anderen Worten: Aus den grundlegenden Gesetzen der Physik und Chemie folgt, daß in Gasen die mittlere kinetische Energie der Moleküle genau die kausale Rolle spielt, die für die Temperatur charakteristisch ist. D.h., die Temperatur von Gasen ist *reduktiv erklärbar*.

Warum ist dann aber die Aussage (3) 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 Rückführung von Schmerzen auf das Feuern von XYZ-Neuronen im Gehirn den Mechanismus erklärt, auf dem diese kausale Rolle beruht. Dennoch gibt es seiner Meinung nach einen entscheidenden Unterschied.

“However, there is more to our concept of pain than its causal role, there is its qualitative character, how it feels; and what is left unexplained by the discovery of C-fiber firing is why pain should feel the way it does! For there seems to be nothing about C-fiber firing which makes it naturally ‘fit’ the phenomenal properties of pain, any more than it would fit some other set of phenomenal properties. Unlike its functional role, the identification of the qualitative side of pain with C-fiber firing ... leaves the connection between it and what we identify it with completely mysterious. One might say, it makes the way pain feels into merely a brute fact.” (Levine 1983, 357)

Levines erster Grund für die These, daß die Aussage (3) 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 (3) könnte immer noch vollständig explanatorisch sein, wenn die Neurobiologie nur verständlich machen könnte, daß sich das Feuern von XYZ-Neuronen im Gehirn 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 XYZ-Neuronen im Gehirn auf die für Schmerzen charakteristische Weise – nämlich schmerzhaft – anfühlt.

Könnte das nicht aber daran liegen, daß die Neurobiologie noch nicht weit genug fortgeschritten ist? Könnte es nicht sein, daß die Neurobiologie eines Tages doch zeigen wird, daß sich das Feuern von XYZ-Neuronen im Gehirn notwendigerweise schmerzhaft anfühlt – in dem Sinne, in dem die Chemie heute schon zeigen kann, daß Natriumchlorid notwendigerweise fest und wasserlöslich ist?

Levine meint, daß dies nicht möglich sei – aus Gründen, die er besonders in seinem Aufsatz “On Leaving Out What It’s Like” erläutert. Jede Reduktion, so Levine, muß zu einer *Erklärung* des reduzierten Phänomens führen. Und daß eine solche Erklärung gelungen ist, zeigt sich in seinen Augen daran, daß es hinterher in einem epistemischen Sinn unmöglich ist, sich vorzustellen, daß das Explanans ohne das Explanandum vorliegt.

“The basic idea is that a reduction should explain what is reduced, and the way we tell whether this has been accomplished is to see whether the phenomenon to be reduced is epistemologically necessitated by the reducing phenomenon, i.e. whether we can see why, given the facts cited in the reduction, things must be the way they seem on the surface.” (Levine 1993, 129)

Denken wir noch einmal an die reduktive Erklärung der Festigkeit von Salz. Diese Erklärung beruhte zum einen auf der Tatsache, daß aus den allgemeinen Naturgesetzen folgt, daß zwischen Natrium- und Chlorionen so starke anziehende Kräfte wirken, daß diese Ionen an ihren relativen Positionen festgezurt sind; zum anderen ergab sie sich aus einer Analyse der Eigenschaft, fest zu sein, der zufolge feste Körper formstabil sind und nur mit erheblichem Kraftaufwand zerteilt werden können. Genau genommen klafft in dieser Erklärung aber noch eine Lücke. Denn aus den allgemeinen Naturgesetzen folgt nur, wie sich die *einzelnen Ionen* verhalten, aus denen Salz besteht. Sie sagen uns nicht, wie sich *ein Stück Salz als Ganzes* verhält. Offenbar benötigen wir daher zusätzlich *Brückenprinzipien*,¹¹ aus denen hervorgeht, wie das Verhalten des komplexen Gegenstandes Salz mit dem Verhalten seiner Teile zusammenhängt – Prinzipien wie

¹¹ Vgl. Levine (1993, 131). Brückenprinzipien dieser Art sind nicht zu verwechseln mit den Brückengesetzen, die nach Nagel bei der Reduktion zweier Theorien benötigt werden, um die verschiedenen Vokabulare in Beziehung zu setzen.

(P1) Wenn die Anziehungskräfte zwischen den Teilen eines Gegenstandes so stark sind, daß sie an ihren relativen Positionen „festgezurt“ sind, dann ist der Gegenstand formstabil.

(P2) Wenn die Anziehungskräfte zwischen den Teilen eines Gegenstandes ziemlich stark sind, dann läßt sich dieser Gegenstand nur mit erheblichem Kraftaufwand zerteilen.

Offenbar haben diese Brückenprinzipien einen speziellen epistemischen Status; man kann sich einfach nicht vorstellen, daß sie falsch sind. Genauso wenig wie man sich vorstellen kann, daß eine Scheibe stillsteht, deren Teile alle mit derselben Winkelgeschwindigkeit in derselben Ebene in dieselbe Richtung um einen bestimmten Punkt kreisen. Und genauso wenig, wie man sich vorstellen kann, daß ein Pfeil nach links fliegt, wenn sich alle seine Teile nach rechts bewegen.

Damit ergibt sich folgende Antwort auf die Frage, warum wir uns – nach der gegebenen Erklärung – nicht mehr vorstellen können, daß Salz (unter normalen Bedingungen) *nicht* fest ist. Der erste Grund dafür ist, daß aus den allgemeinen Naturgesetzen folgt, daß zwischen den Natrium- und Chlorionen, aus denen Salz besteht, so starke anziehende Kräfte wirken, daß diese Ionen an ihren relativen Positionen festgezurt sind. Mindestens ebenso wichtig ist jedoch der zweite Grund, der sich aus dem speziellen Status der Brückenprinzipien (P1) und (P2) ergibt. Denn offenbar ist dieser Status dafür verantwortlich, daß es für uns *nicht* denkbar ist, daß die Teile eines Gegenstandes an ihren relativen Positionen festgezurt sind, dieser Gegenstand aber nicht formstabil ist bzw. daß zwischen den Teilen eines Gegenstands starke anziehende Kräfte herrschen, dieser Gegenstand aber schon mit minimalem Kraftaufwand zerteilt werden kann.

Wenn wir das Verhältnis zwischen Schmerzen und dem Feuern von XYZ-Neuronen im Gehirn betrachten, liegen die Dinge jedoch anders. Selbst wenn wir bis ins letzte Detail darüber informiert sind, welche neurophysiologischen

Prozesse (oder welche Informationsverarbeitungsprozesse) im Gehirn ablaufen, ist es immer noch denkbar, daß die Person, in deren Gehirn diese Prozesse ablaufen, keine Schmerzen empfindet. Warum?

Wenn wir die Erklärung, die dazu führt, daß wir uns nicht vorstellen können, daß Salz unter normalen Bedingungen nicht fest ist, im Detail analysieren, ergeben sich drei zentrale Punkte:

1. Die charakteristischen Merkmale der Eigenschaft, fest zu sein, bestehen *alle* darin, daß sich feste Stoffe unter bestimmten Bedingungen auf eine bestimmte Art und Weise *verhalten*.
2. Aus den allgemeinen Naturgesetzen folgt, daß zwischen den Natrium- und Chlorionen, aus denen Salz besteht, so starke anziehende Kräfte wirken, daß diese Ionen an ihren relativen Positionen festgezurr sind.
3. Es gibt Brückenprinzipien, aus denen sich ergibt, daß ein Stoff, zwischen dessen Molekülen diese Kräfte bestehen, genau das Verhalten zeigt, das für die Eigenschaft, fest zu sein, charakteristisch ist.

Wie sieht es nun mit der vermeintlichen Erklärung von Schmerzen durch das Feuern von XYZ-Neuronen im Gehirn aus? Offenbar gibt es hier schon im ersten Punkt einen wesentlichen Unterschied:

1. Unser Begriff von Schmerzen erschöpft sich nicht in einer kausalen Rolle, und Schmerzen sind auch nicht allein durch ein bestimmtes Verhalten charakterisiert; vielmehr umfaßt unser Begriff von Schmerzen einen qualitativen Aspekt – die Art, wie es sich anfühlt, Schmerzen zu haben.

Doch dieser Punkt ist nicht entscheidend. Denn Schmerzen könnten immer noch auf das Feuern von XYZ-Neuronen im Gehirn zurückgeführt werden, *wenn* es nur Brückenprinzipien gäbe, aus denen hervorginge, daß sich das Feuern von

XYZ-Neuronen im Gehirn auf die für Schmerzen charakteristische Weise anfühlt. Entscheidend sind daher die folgenden beiden Punkte:

2. Aus den Gesetzen der Neurobiologie folgt nur, unter welchen Bedingungen welche Neuronen mit welcher Geschwindigkeit feuern.

Und:

3. Es gibt keinerlei Brückenprinzipien, die das Feuern von Neuronen mit bestimmten Erlebnisqualitäten verbinden.

Entscheidend für Levines Argument der Erklärungslücke ist also die These, daß es keine Brückenprinzipien gibt, die das Feuern von Neuronen mit bestimmten Erlebnisqualitäten verbinden. Damit meint er nicht, daß Sätze wie

- (1) Immer wenn im Gehirn einer Person die XYZ-Neuronen feuern, fühlt diese Person Schmerzen

keine wahren Naturgesetze sein können; er meint nur, daß sie selbst dann, wenn sie wahr sind, niemals den speziellen Status von Brückenprinzipien haben. Wenn das richtig ist, scheint der Schluß aber tatsächlich unausweichlich, daß Erlebnisse *nicht reduktiv erklärbar*, sondern *emergent* sind.

Literatur

- Beckermann, Ansgar (2000) Ein Argument für den Physikalismus. In: G. Keil und H. Schnädelbach (Hg.) *Naturalismus*. Frankfurt/M.: Suhrkamp, 128-143.
- Beckermann, Ansgar (2001a) *Analytische Einführung in die Philosophie des Geistes*. 2. überarbeitete Aufl., Berlin/New York: Walter de Gruyter.
- Beckermann, Ansgar (2001b) Die reduktive Erklärbarkeit phänomenalen Bewußtseins – C.D. Broad zur Erklärungslücke. In: M. Pauen und A. Stephan (Hg.) *Phänomenales Bewußtsein*, Paderborn: Mentis Verlag, 122-147.
- Braitenberg, Valentin (1984) *Vehicles*. 7th ed. Cambridge MA: MIT-Press 2000. (Dt.: *Künstliche Wesen*. Braunschweig: Vieweg&Sohn 1986)
- Broad, Charles Dunbar (1925) *The Mind and Its Place In Nature*. London: Kegan Paul, Trench, Turbner, & Co.
- Chalmers, David (1995) Facing Up to the Problem of Consciousness. *Journal of Consciousness Studies* 2, 200-219. Wiederabdruck in: J. Shear (ed.) *Explaining Consciousness – The ‘Hard Problem’*. Cambridge MA: MIT Press 1997, 9-30. (Dt. in: F. Esken und D. Heckmann (Hg.) (1998) *Bewußtsein und Repräsentation*. Paderborn: Schöningh, 221-253)
- Chalmers, David (1996) *The Conscious Mind*. Oxford: Oxford University Press.
- Kripke, Saul (1972) Naming and Necessity. In: D. Davidson and G. Harman (eds.) *Semantics of Natural Language*. 2nd edition. Dordrecht: Reidel Publishing Company, 253-355. Rev. Fassung erschienen als Monographie: *Naming and Necessity*. Cambridge MA: Harvard University Press 1980. (Dt.: *Name und Notwendigkeit*. Frankfurt am Main: Suhrkamp 1981)
- Levine, Joseph (1983) Materialism and Qualia: The explanatory gap. *Pacific Philosophical Quarterly* 64, 354-361.
- Levine, Joseph (1993) On Leaving Out What It’s Like. In M. Davies and G.W. Humphreys, eds., *Consciousness: Psychological and Philosophical Essays*. Oxford: Basil Blackwell, 121-136.