

Achim Stephan

PHÄNOMENALER PESSIMISMUS

Es soll eingefleischte Pessimisten geben, die bezweifeln, daß wir jemals einen Platz für unsere phänomenalen Erlebnisse in der physischen Welt finden werden. Da die Neuro- und Kognitionswissenschaften bislang wenig Erhellendes über das phänomenale Bewußtsein beisteuern konnten, schließen einige von ihnen, daß *nichts* von dem, was diese Wissenschaften noch herausfinden werden, zu einem wirklichen Verständnis qualitativer Zustände wird führen können. Patricia Churchland, die man wohl besser nicht den Pessimisten zurechnen sollte, hat diese Schlußweise „the argument from ignorance“ genannt und gezeigt, daß sie verfehlt ist (vgl. 1996, 406).

Eine Optimistin hat jedoch mehr zu leisten. Sie sollte uns wenigstens überzeugen können, daß es Hoffnung auf ein tieferes Verständnis gibt. Patricia Churchland stützt sich dazu auf das folgende induktive Argument: In der Geschichte der Wissenschaften wurden schon viele Probleme als schwierig oder unlösbar angesehen. Sie ist aber eine Geschichte grandioser Erfolge. Sie lehrt uns, daß wir in jedem einzelnen Fall immense Fortschritte gemacht haben: Erfolgreiche chemische Erklärungen für Verbrennungsvorgänge lösten die Phlogiston-Theorie ab. Elektrische Phänomene, die sich zunächst mechanistischen Erklärungen widersetzten, fanden später plausible Erklärungen in der Elektrodynamik, und selbst die lange Zeit für unzugänglich gehaltenen Phänomene des Lebendigen wurden nach und nach entschlüsselt. Man denke nur an die Leistungen, die seit den fünfziger Jahren auf dem Gebiet der Gentrifugation und Gentechnik vollbracht wurden.

Schön und gut, mag man denken: Aber induktives Schließen ist ein seltsames Geschäft – und das selbst dann, wenn es sich auf eine stattliche Induktionsbasis stützen kann. Im Falle wissenschaftlicher Erfolge verfügen wir aber noch nicht einmal über eine breite Induktionsbasis. Schlimmer noch, die Wissenschaftsgeschichte enthält auch Beispiele vergeblichen Bemühens. Es gab nicht immer Erfolg und Fortschritt.

I. KREIS-QUADRATUREN UND ANDERE SCHWIERIGKEITEN

Schon vor mehr als 2000 Jahren formulierten griechische Mathematiker einige Fragen, die zwar leicht zu verstehen, aber keineswegs leicht zu beantworten

sind. Eines ihrer klassischen Konstruktionsprobleme befaßt sich mit der Quadratur des Kreises: Ist es möglich, mit Zirkel und Lineal ein Quadrat zu konstruieren, das denselben Flächeninhalt wie ein vorgegebener Kreis hat? Ein weiteres altherwürdiges Problem bezieht sich auf die Lösung von Polynomgleichungen. Bereits die Griechen wußten, wie man in allgemeiner Weise quadratische Gleichungen löst. Wie das funktioniert, lernen Jugendliche heute in der neunten Klasse. Komplizierter ist es mit Gleichungen höheren Grades: Können auch für sie allgemeine Lösungen gefunden werden?

Es mußten einige Jahre vergehen, bis um 1500 die italienischen Mathematiker Scipione del Ferro und Niccolò Tartaglia eine allgemeine Lösung für Gleichungen dritten Grades formulieren konnten. Nur vierzig Jahre später präsentierte Girolamo Cardanos Student Ludovico Ferrari eine allgemeine Lösung für Gleichungen vierten Grades. Man kann sich vorstellen, wie groß die Erwartungen waren, nun auch allgemeine Lösungen für Gleichungen noch höheren Grades zu finden. Ganze Mathematikerschulen arbeiteten an diesem Problem. Zweihundert Jahre lang wurde viel geistige Energie in diese Probleme investiert. Es stellte sich heraus, daß die Suche nach allgemeinen Lösungen für Gleichungen 5ten und höheren Grades ein wirklich *hartes* Problem war; die zuvor bewältigte Aufgabe, allgemeine Lösungen für Gleichungen 3ten und 4ten Grades zu finden, war zwar anspruchsvoll, zählte aber doch zu den „*easy problems*“.

Es sollte bis ins beginnende neunzehnte Jahrhundert dauern, ehe Nils Henrik Abel und Evariste Galois zeigen konnten, daß es *keine* allgemeine Lösung für Gleichungen 5ten und höheren Grades gibt. Jede weitere Suche würde vergebens sein. Die nach Galois benannte algebraische Theorie zeigt darüber hinaus, daß und weshalb einige der klassischen Konstruktionsprobleme wie das Verdoppeln des Würfels oder die Quadratur des Kreises ebenfalls *unlösbar* sind.

DAS PHILOSOPHISCHE PROBLEM PHÄNOMENALER QUALITÄTEN

Doch wozu erwähne ich das alles? Zunächst zeigen diese Beispiele, daß es zwischen den Optimisten und den Pessimisten zumindest unentschieden steht. Die Wissenschaftsgeschichte kennt Beispiele großer Erfolge. Sie kennt aber auch einige Probleme, um die jahrzehntelang *vergeblich* gerungen wurde – und zwar deshalb, weil diese Probleme, wie sich in der Regel erst später zeigte, prinzipiell *unlösbar* waren. Dieses Ergebnis wäre für sich allein nicht besonders bemerkenswert. Spannend wird es, wenn wir uns fragen, ob das philosophische Problem des phänomenalen Erlebens eher den (unlösbaren) mathematischen oder eher den erfolgreich gelösten – und von Churchland angeführten – empirischen Problemen ähnelt. Deshalb sollten wir zunächst prüfen, um was es wirklich geht: Was genau ist das philosophische Problem

phänomenaler Erlebnisse? Was heißt es, einen Platz für phänomenale Qualitäten in der physischen Welt zu finden?

Nun, gefordert ist das, was David Chalmers (1996, 42–51) und Jaegwon Kim (1998a, 198–201) eine *reduktive Erklärung* phänomenaler Erlebnisse genannt haben und was Joseph Levine (1993, 127–132) einige Jahre zuvor als *explanatorische Reduktion* bezeichnete.¹ Ist man nämlich in der Lage, phänomenale Qualitäten reduktiv zu erklären, so ist damit die *explanatorische Lücke* (vgl. Levine 1983) geschlossen, und man hat diesen Entitäten unbestreitbar einen Platz innerhalb der physischen Welt zugewiesen.

Leugnet man, daß die Angabe einer reduktiven Erklärung das *philosophische* Qualia-Problem ist, so sucht man nach etwas anderem. Worum es beispielsweise nicht geht, ist die empirische Frage nach neuronalen Korrelaten phänomenaler Erlebnisse oder um Brückengesetze vom Nagel-Typ.² Ebenso wenig geht es jedoch um die *Verabschiedung* des Problems durch die Identifizierung phänomenaler Erlebnisse mit ihren vermeintlichen neuronalen Korrelaten.³ Aber was heißt es, eine reduktive Erklärung anzugeben? Gemäß Joseph Levine erfolgt eine Reduktion, die explanatorisch ist, in zwei Schritten:

Stage 1 involves the (relatively? quasi?) *a priori* process of working the concept of the property to be reduced „into shape“ for reduction by identifying the causal role for which we are seeking the underlying mechanisms. Stage 2 involves the empirical work of discovering just what those underlying mechanisms are (Levine 1993, 132).

Jaegwon Kim nannte kürzlich die erste Stufe die sogenannte „Präparations-Prozedur“ für die zu reduzierende Eigenschaft:

To reduce a property *M* to a domain of base properties we must first „prime“ *M* for reduction by construing, or reconstruing, it relationally or extrinsically. This turns *M* into a relational/extrinsic property. For functional reduction we construe *M* as a second-order property defined by its causal role – that is, by a causal specification *H* describing its (typical) causes and effects (Kim 1998b, 98).

Die zentrale Frage sei daher, ob Qualia in der für reduktive Erklärungen erforderlichen Weise funktionalisiert werden könnten, oder ob sie sich solchen Funktionalisierungen widersetzen. Kim selbst ist eher skeptisch: „To get to the point without fuss, it seems to me that the felt, phenomenal qualities of

¹ Levines Terminologie bin ich selbst in früheren Arbeiten (vgl. 1997, 307–311 und 1999, 172–174) gefolgt. Derzeit scheint sich jedoch der Ausdruck „reduktive Erklärung“ durchzusetzen.

² Vgl. dazu Kim (2001, 153–160).

³ Dies ist die Strategie der *New-Wave Materialisten*, auf die ich weiter unten eingehen werde.

our experiences, or qualia, are intrinsic properties if anything is“ (1998b, 102). Qualia scheinen demnach nicht zu den Entitäten zu gehören, die erfolgreich für reduktive Erklärungen präpariert werden können. Offenbar können sie nicht über kausale Rollen individuiert werden; wenigstens ist es das, was Levine als Grund angibt:

What seems to be responsible for the explanatory gap, then, is the fact that our concepts of qualitative character do not represent, at least in terms of their psychological contents, causal roles. ... Thus, to the extent that there is an element in our concept of qualitative character that is not captured by features of its causal role, to that extent it will escape the explanatory net of a physicalistic reduction (Levine 1993, 134).

Ähnliche Überlegungen hatte Charles Dunbar Broad bereits vor mehr als siebzig Jahren in seinen Tarnner Lectures *The Mind and Its Place in Nature* (1925) angestellt. Da weder sekundäre Qualitäten noch phänomenale Qualitäten adäquat in Begriffen des mikro- oder makroskopischen *Verhaltens* der Systeme, die sie haben, analysiert werden könnten, seien diese Entitäten irreduzibel. Denn wenn wir sagten, daß ein bestimmter Gegenstand rot sei oder eine chemische Substanz wie Salmiakgeist rieche, dann meinten wir damit einfach nicht, daß sich die Bestandteile der entsprechenden Systeme auf eine bestimmte Weise *verhielten* oder *bewegten*. Und ebensowenig meinten wir, wenn wir von unseren bewußten Empfindungen sprechen, daß sich Teile unseres Körpers oder Gehirns auf eine bestimmte Weise verhielten oder bewegten.⁴ Kein Fortschritt in den Wissenschaften könne daran etwas ändern. Deshalb bestehe auch keine Aussicht, daß durch einen Zuwachs an Wissen die Lücke zwischen physischen Vorgängen und sekundären Qualitäten bzw. zwischen neuronalen Vorgängen und phänomenalen Qualitäten geschlossen werden kann.⁵

Das Vorkommen von Eigenschaften, die nicht reduktiv erklärbar sind, impliziert im übrigen nicht, wie mitunter gedacht wird, ein abwärts gerichtetes Verursachen. Weder folgt, daß die *Systeme*, die Eigenschaften besitzen, die sich gegenüber einer funktionalen Analyse als sperrig erweisen, einen abwärts gerichteten Einfluß auf ihre Konstituenten besitzen, noch daß die unanalysierbaren *Eigenschaften* selbst abwärts auf die Konstituenten des fraglichen Systems wirken. Eher wäre zu fragen, inwiefern Eigenschaften, die nicht adäquat

⁴ „Neither colour nor consciousness can be exhibited without residue in physical and chemical terms in the sense in which life can, since to be coloured or to be conscious does not mean to move in certain peculiar ways“ (Broad 1921, 145f.).

⁵ Eine ausführliche Darstellung von Broads Argumenten für die Irreduzibilität phänomenaler Qualitäten habe ich in Stephan (1999, 32–44, 51 und 185–196) gegeben. Man beachte auch Ansgar Beckermans Beitrag „Die reduktive Erklärbarkeit phänomenalen Bewußtseins – C.D. Broad zur Erklärungslücke“ (2001).

über zuschreibbare kausale Rollen individuiert werden können, überhaupt einen kausalen Einfluß haben können.

EINE AUSWEGLOSE SITUATION?

Betrachten wir noch einmal die theoretische Situation, in der wir uns befinden. Offenbar haben wir es mit drei Behauptungen zu tun:

- (1) Das *philosophische* Problem, um das es geht, ist die Frage, ob phänomenale Erlebnisse reduktiv erklärt werden können.
- (2) Reduktive Erklärungen erfordern eine Präparierung der zu reduzierenden Eigenschaften über kausale oder funktionale Rollen.
- (3) Phänomenale Qualitäten entziehen sich jedweder Form der erforderlichen Präparierung.

Insofern sich phänomenale Erlebnisse der erforderlichen Präparierung entziehen, fügen sie sich *nicht* nahtlos in die physische Welt ein: Es klappt eine explanatorische Lücke – und zwar aus begrifflichen Gründen. Somit scheint das philosophische Qualia-Problem eher den von den Pessimisten angeführten unlösbaren mathematischen Problemen zu ähneln als den erfolgreich gelösten Problemen empirischer Wissenschaften, auf die Patricia Churchland so gerne verweist.

Wo stehen wir damit? Der theoretische Rahmen, in dem es zur Erklärungslücke kommt, ist durch die Annahme ausgezeichnet, daß die Träger qualitativer Phänomene komplexe materielle Systeme sind. Damit scheint es zwei Möglichkeiten ihrer Realisierung zu geben: (i) Sie sind mit spezifischen physischen Eigenschaften des Systems typ-identisch, oder (ii) sie sind durch verschiedene physische Zustände, d.h. auf mehrfache Weise instantiierbar. Gegen Variante (i) spricht neben der mit hoher Wahrscheinlichkeit anzunehmenden mehrfachen Realisierbarkeit qualitativer Phänomene die ursprünglich von Levine (1983) diagnostizierte Erklärungslücke für psychophysische Identitätsbehauptungen. Gegen Variante (ii) spricht die behauptete Erklärungslücke für reduktive Erklärungen mehrfach realisierbarer Eigenschaften. Welche Optionen bleiben dann aber?

Nun, man könnte den vorgegebenen theoretischen Rahmen generell in Zweifel ziehen. Vor etwas längerer Zeit hatte Leibniz ähnliche Überlegungen unter der zusätzlichen Annahme, daß ein materielles komplexes System nur die Eigenschaften wirklich haben kann, die sich aus dem Zusammenwirken seiner Bestandteile mechanisch erklären lassen, als *reductio ad absurdum* eingesetzt. Da nämlich „die Perzeption und das, was von ihr abhängt, aus mechanischen Gründen, d.h. aus Gestalt und Bewegung, nicht erklärbar“ sei,

könne der Träger der mentalen Eigenschaften kein zusammengesetzter materieller sein; in Frage komme nur etwas Einfaches, Nicht-Zusammengesetztes – eine Monade (Leibniz, *Monadologie*, § 17). Leibniz verläßt damit den von uns vorausgesetzten theoretischen Rahmen; er favorisiert statt dessen eine dualistische Position. Ebenso verliesse man den theoretischen Rahmen, wenn man die Existenz qualitativer Phänomene bezweifelte und eine eliminativ materialistische Theorie bezüglich phänomenaler Eigenschaften vertreten würde.

Will man den theoretischen Rahmen beibehalten, so stehen derzeit vor allem drei Strategien zur Verfügung:⁶ Erstens kann man die oben skizzierte Situation, d.h. die Behauptungen (1)–(3), akzeptieren und einräumen, daß die explanatorische Lücke nicht zu schließen ist. Dann vertritt man eine *emergentistische* Position bezüglich der phänomenalen Eigenschaften. In diesem Fall wird es allerdings schwierig, für die qualitativen Phänomene einen angemessenen Platz im Kausalgefüge der Natur zu finden.⁷ Zweitens kann man die Behauptungen (1) und (2) akzeptieren, aber Behauptung (3) zurückweisen, indem man zu zeigen versucht, daß sich phänomenale Qualitäten doch adäquat über ihre funktionale Rolle präparieren und damit reduktiv erklären lassen. Diesen Weg beschreitet Michael Pauen (2001). Und schließlich kann man drittens dafür argumentieren, daß phänomenale Qualitäten als Eigenschaften materieller Systeme konzipiert werden können, ohne daß sich dabei eine explanatorische Lücke auftut, obwohl man zugesteht, daß qualitative Phänomene nicht adäquat über ihre funktionale Rolle erfaßt werden können. Diesen Weg beschreiten derzeit einige New-Wave Materialisten, indem sie der Typen-Identitätstheorie zu einer Neuauflage verhelfen (vgl. McLaughlin, 2001). Interessanterweise stimmen die New-Wave Materialisten mit den Emergentisten darin überein, daß sie die Erklärungslücke für unvermeidlich halten, wenn man versucht, qualitative Eigenschaften reduktiv zu erklären. Beide akzeptieren demnach die Behauptungen (2) und (3). Die New-Wave Materialisten bestreiten jedoch Behauptung (1) und favorisieren statt dessen einen identitätstheoretischen Ansatz. Diesem wende ich mich nun zu.

II. DIE VERABSCHIEDUNG DES QUALIA-PROBLEMS IM NEW-WAVE-MATERIALISMUS DAVID PAPINEAUS

In den letzten drei bis vier Jahren ist es zu einer umfassenden Neubewertung der „klassischen“ Argumente von Kripke und Levine gegen die (alte) Typen-Identitätstheorie gekommen. Während sich Chris Hill (1997) insbesondere mit Saul Kripkes metaphysischem Argument auseinandersetzt und zu zeigen

⁶Theoretisch gibt es zwar noch andere Optionen; diese werden jedoch nicht vertreten.

⁷Vgl. dazu Kim (1993, 1998b) und Stephan (1997).

versucht, daß sich die scheinbare Kontingenz psychophysischer Identitätsaussagen doch wegerklären läßt, versuchen Brian Loar (1997), Ned Block und Robert Stalnaker (1999) sowie David Papineau (1998) auch dem „Explanatory gap“-Argument den Wind aus den Segeln zu nehmen. Im folgenden konzentriere ich mich auf Papineaus Überlegungen, der in *Mind the Gap* eine ganz schlicht erscheinende Antwort anbietet: „Ich denke, daß der Physikalismus überhaupt keine Probleme mit der Erklärung dafür hat, warum bewußte Zustände mit Gehirnzuständen einhergehen. Mein Hauptgrund dafür ist, daß der Physikalismus am besten als eine These über die Identität von bewußten und materiellen Eigenschaften aufzufassen ist und daß Identitätsbehauptungen keine Erklärung benötigen“ (1998, 373; 2001, 224).⁸

Wenden wir uns zunächst den beiden Hauptthesen Papineaus zu: Was spricht für die Identitätsthese, und was spricht dafür, daß Identitätsbehauptungen keine Erklärungslücken entstehen lassen? Zunächst sollte man sich klar machen, daß Papineau unter der Identitätsthese nicht die Typen-Identität alter Prägung versteht, sondern unter den „materiellen Eigenschaften“ sowohl physische Eigenschaften im herkömmlichen Sinne (z.B. neurophysiologische Eigenschaften) als auch physisch realisierte funktionale-Rollen-Eigenschaften subsumiert.

Papineaus Argument für die Identitätsthese basiert auf drei Prämissen:

- (1) Bewußte geistige Vorkommnisse haben physische Wirkungen.
- (2) Alle physischen Wirkungen haben hinreichende physische Ursachen.
- (3) Die physischen Wirkungen bewußter Ursachen sind nicht immer überdeterminiert. (Papineau 1998, 375; 2001, 227-228)

Der eleganteste und einfachste Weg, alle drei Prämissen gleichzeitig behaupten zu können, sei es, bewußte Vorkommnisse mit physischen Vorkommnissen zu identifizieren. Sollten bewußte mentale Eigenschaften mit physischen Eigenschaften jedoch nicht streng identisch sein, dann müßten sie zumindest mit physisch realisierten Rollen-Eigenschaften identifiziert werden. Andernfalls sei nicht zu sehen, wie bewußte Vorkommnisse überhaupt physische Wirkungen hervorbringen könnten. Die weiteren Schritte gehen dann ganz schnell: Wenn bewußte mentale Eigenschaften mit materiellen Eigenschaften identisch sind, dann gebe es nichts weiter zu erklären. Es gebe „kein Rätsel hinsichtlich der Frage, weshalb materielle Eigenschaften bewußte Eigenschaften ‚hervorbringen‘. Das ist so, weil Identitäten keiner Erklärung

⁸ Brian McLaughlin sieht das ganz ähnlich: Wenn die Korrelationsthese, derzufolge jedem Typ eines sensorischen Zustandes ein Typ eines physischen oder funktionalen Zustandes eineindeutig mit nomologischer Notwendigkeit zugeordnet ist, empirische Bestätigung findet, dann bietet die Typen-Identitätstheorie die beste Erklärung für das Bestehen jener Korrelationsthese (vgl. McLaughlin 2001, 210 f.).

bedürfen. Wenn die ‚zwei‘ Eigenschaften in Wirklichkeit eine sind, dann ist es nicht so, daß die materielle Eigenschaft die bewußte Eigenschaft ‚hervorbringt‘ – sie *ist* die bewußte Eigenschaft. Und wenn sie es ist, dann gibt es kein Rätsel, warum sie ist, was sie ist“ (1998, 379; 2001, 233). Ebenso unsinnig wäre es, wenn man fragte, weshalb Pascal Mercier mit Peter Bieri identisch ist, wenn es sich *de facto* um ein und dieselbe Person handelt. Man mag vielleicht wissen wollen, was *zeige*, daß sie miteinander identisch sind, und die Antwort wäre vermutlich, daß sie stets dieselbe Stelle im Kausalgefüge des Weltlaufs einnehmen. Aber es wäre unsinnig zu fragen, *warum* sie identisch seien. Wenn sie es sind, dann sind sie es. Sie können nicht zwei sein.

Da Papineau allerdings selbst – und wohl zu recht – glaubt, daß seine Antwort nicht alle Leser überzeugen dürfte – „Ich vermute, daß viele Leser durch das Argument des letzten Abschnittes wenig befriedigt sein werden“ (1998, 379; 2001, 234) –, geht er abschließend der Frage nach, weshalb viele Philosophen trotz allem zu dem Eindruck gelangten, daß eine Erklärungslücke bestehe.

Vielleicht meinten sie erklären zu müssen, daß bestimmte physische Zustände diejenigen Beschreibungen erfüllen, die wir *a priori* mit unseren bewußten Zuständen verbänden. Über entsprechende Erklärungen verfügten wir im Falle bewußter Vorkommnisse allerdings nicht. Der Grund sei schlicht, daß es gar keine Beschreibungen gebe, die wir *a priori* mit unseren alltagssprachlichen Begriffen bewußter Zustände verbänden (1998, 381; 2001, 236). Papineau zufolge referieren Begriffe von bewußten Zuständen nämlich nicht mit Hilfe von Beschreibungen, vielmehr lösen sie solche Zustände simulationsartig aus. „Phänomenale Begriffe referieren durch eine Art der Simulation, ohne dabei irgendwelche Beschreibungen heranzuziehen“ (1998, 383; 2001, 239). Dieses besondere Merkmal der phänomenalen Begriffe sei auch der eigentliche Grund für die weitverbreitete und machtvolle Illusion der Verschiedenheit von Geist und Gehirn, die der explanatorischen Lücke zugrunde liege und dazu führe, daß „die meisten Menschen einfach nicht bereit sind zu akzeptieren, daß phänomenale Begriffe von vornherein auf materielle Eigenschaften referieren“ (1998, 383; 2001, 240). Papineau hat diese Illusion als *antipathetischen* Fehlschluß charakterisiert und an verschiedenen Beispielen, u.a. auch dem folgenden, illustriert, wie er seines Erachtens zustande kommt: „Wir betrachten die linke Seite der [Identitäts-]Behauptung, wenden unseren phänomenalen Begriff des Schmerzes an (*jene* Empfindung) und fühlen ein kleines Stechen. Dann betrachten wir die rechte Seite, wenden unseren Begriff des Feuerns von C-Fasern an und fühlen nichts (oder wenigstens nichts aus dem Bereich der Schmerzen – wir mögen uns Nervenzellen oder ähnliches vorstellen). So folgern wir, daß die rechte Seite die Schmerz-Empfindung selbst *auslöst* ...“ (1998, 384–385; 2001, 242).

ZWEIFEL BLEIBEN

Hat nun Papineau die Lücke schließen können? Hat er wirklich die Lücke erfaßt, auf die Levine vor zwanzig Jahren zu zeigen meinte? Hat er die richtigen Gründe genannt, die viele von uns eine reduktive Erklärung fordern und eine explanatorische Lücke „sehen“ lassen? Kann er die von ihm vorgeschlagene Identifizierung hinreichend begründen? Ist nun das „Rätsel der phänomenalen Qualitäten“ gelöst? Mir scheint, daß auf keine dieser Fragen mit einem klaren „Ja“ geantwortet werden kann.

Natürlich wäre es unsinnig, nach den Gründen für eine *Identität* fragen zu wollen. Handelt es sich bei einer Entität, die uns als zweierlei „Ding“ erscheint, *de facto* nur um eins, so gibt es in der Tat für ihr „Eins-Sein“ keinen Grund zu nennen. Levine hat aber gerade nicht gefragt, *warum* Schmerzen mit dem Feuern von C-Fasern identisch sind, *wenn* sie miteinander identisch sind, sondern *warum* sich das Feuern der C-Fasern so anfühlt, wie es sich anfühlt, *während* sich das Feuern anderer Neuronenverbände ganz anders anfühlt, bzw. *warum* es sich überhaupt irgendwie anfühlt, *während* sich das Feuern anderer Neuronenverbände nicht irgendwie anfühlt. D.h., es geht hier vor allem um den *systematischen* Zusammenhang zwischen bestimmten neuronalen und bestimmten phänomenalen Zuständen. Betrachten wir zum Vergleich die chemischen Stoffen zukommende Eigenschaft, flüssig zu sein. Diese ist mit einer bestimmten Kräftekonstellation unter den Molekülen identisch, aus denen die Stoffe bestehen. Im Unterschied zum Qualia-Problem ist hier der systematische Kontext jedoch völlig transparent. Es läßt sich genau zeigen, *wie* mit spezifischen Veränderungen der Bindungskräfte spezifische Modifikationen des Charakters der Flüssigkeiten einhergehen, *weshalb* sie dünn- oder dickflüssig sind und *warum* dies so ist. Dagegen scheint eine entsprechende Transparenz bezüglich des systematischen Zusammenhangs zwischen neuronalen und phänomenalen Zuständen aufgrund des intrinsischen Charakters der phänomenalen Qualitäten gerade nicht erreichbar zu sein. Und diese Kluft im Hinblick auf reduktive Erklärungen scheint prinzipiell nicht schließbar zu sein.

Darüber hinaus zeigen sich fundamentale Unterschiede zwischen der angestrebten Identifizierung von phänomenalen und materiellen Eigenschaften auf der einen Seite und den Identifizierungen, die Papineau als Beispiele anführt, auf der anderen. Papineau verweist nämlich hauptsächlich auf Entitäten, die ohnehin zur selben Kategorie gehören wie z.B. Berge, Himmelskörper oder Personen; d.h., es geht dann nur noch um die Frage, ob zwei Namen oder Beschreibungen, die auf verschiedene Exemplare eines bestimmten Typs zu referieren scheinen, *de facto* auf ein und dasselbe Exemplar referieren. So ist es aber im Falle qualitativer Phänomene und neuronaler Vorgänge gerade nicht. Hier ist noch nicht einmal in Ansätzen klar, was die genauen Individualisierungsbedingungen für die „Exemplare“ der neuronalen Zustände oder Vorgänge sind, die mit bestimmten Exemplaren eines qualitativen Zustandes

identifiziert werden sollen: Von welchem physischen Zustand läßt sich denn begründet sagen, daß er genau den Platz innerhalb des „Kausalgefüges der Welt“ einnehme, den Papineau dem entsprechenden qualitativen Zustand zuschreiben gewillt ist? Auch hier bleiben Verständnis- und Begründungslücken. Somit könnte Papineau vor allem solche explanatorischen Lücken geschlossen haben, von denen niemand annahm, daß sie offen seien.

Betrachten wir deshalb die alternative Strategie, über vermeintlich adäquate Funktionalisierungen die explanatorische Lücke zu schließen.

III. QUALIA-FUNKTIONALISIERUNGEN ALS LÜCKENFÜLLER

Michael Pauen verfolgt in seinem Beitrag *Invertierte Schmerzen* (in diesem Band) ein ambitioniertes Projekt. Anders als etwa David Papineau und einige weitere New-Wave Materialisten versucht er nicht, das Problem der Erklärungslücke als ein Scheinproblem zu entlarven. Vielmehr akzeptiert er es als ein echtes Problem: Wäre es uns aus prinzipiellen Gründen unmöglich, reduktive Erklärungen für phänomenale Eigenschaften anzugeben, dann würde auch nach Pauen die Erklärungslücke Bestand haben und der Physikalismus müßte sich mit einem wohl unlösbaren Rätsel auseinandersetzen. Im Unterschied zu Levine und Chalmers ist Pauen jedoch ein verhaltener Optimist. Sein Ziel besteht darin zu zeigen, daß schlicht nicht auszuschließen ist, daß man auch für phänomenale Eigenschaften reduktive Erklärungen im Prinzip anbieten können. Darüber hinaus skizziert er einen Weg, wie wir vielleicht reduktiven Erklärungen einiger phänomenaler Eigenschaften näher kommen könnten.

Um Pauens Vorschläge und seine Argumentation besser einschätzen zu können, ist es ratsam, sich noch einmal zu vergewissern, worin die Erklärungslücke bestehen soll. Im Anschluß an Levine, Chalmers und Kim läßt sie sich wie folgt diagnostizieren:

- (EL) Phänomenale Eigenschaften entziehen sich prinzipiell einer reduktiven Erklärung, da sie sich nicht funktional konstruieren oder rekonstruieren lassen.

Will man diese These bestreiten, so kann dies in einer sehr starken, aber auch in einer eher moderaten Form geschehen; die starke Gegenthese könnte lauten:

- (GT₁) Von *keiner* phänomenalen Eigenschaft kann ausgeschlossen werden, daß sie wenigstens im Prinzip reduktiv erklärt werden kann; d.h. wenigstens im Prinzip könnte sich jede phänomenale Eigenschaft funktional konstruieren oder rekonstruieren lassen.

Eine schwächere Variante lautet:

- (GT₂) Zumindest von *einigen* phänomenalen Eigenschaften kann nicht ausgeschlossen werden, daß sie sich wenigstens im Prinzip reduktiv erklären lassen; d.h. einige phänomenale Eigenschaften könnten sich wenigstens im Prinzip funktional konstruieren oder rekonstruieren lassen.

Erwiese sich (GT₁) als zu stark, (GT₂) hingegen als plausibel, so wäre das Erklärungslückenproblem zwar nicht vom Tisch, es bestünde jedoch nicht mehr in seiner allgemeinen Form; insbesondere bestünde es nicht mehr für alle phänomenalen Eigenschaften. Eine unüberwindbare „Lücke“ klaffte dann nicht generell zwischen phänomenalen und physisch/funktionalen Zuständen, sondern innerhalb des phänomenalen Bereiches gäbe es einen Bruch zwischen reduktiv erklärbaren Eigenschaften und solchen, die sich einer Erklärung prinzipiell widersetzen.

Wie wir noch sehen werden, argumentiert Pauen in *Invertierte Schmerzen* nur für die schwächere Gegenthese (GT₂). Gerade darin liegt jedoch die Brisanz für Anhänger des Erklärungslückenargumentes. Sollte Pausens Argumentation erfolgreich sein, so wäre das Erklärungslückenargument in seinem umfassenden Anspruch erschüttert. Zugleich zwingt er durch seinen eher moderaten Einspruch die Verteidiger des Erklärungslückenargumentes zu einer sehr subtilen Auseinandersetzung, da er nur behauptet, daß nicht ausgeschlossen werden kann, daß unter noch näher zu beschreibenden Bedingungen in ferner Zukunft eine reduktive Erklärung einiger phänomenaler Eigenschaften wenigstens im Prinzip gelingen könnte. Will man dieser Ansicht widersprechen, so bedarf es einer sorgfältigen Vorgehensweise.

Wenden wir uns daher zunächst der Argumentation Pausens etwas genauer zu; sie vollzieht sich in zwei aufeinander aufbauenden Teilen. Im ersten Teil geht es darum, ein Argument auszuhöhlen, auf das sich einige Vertreter des Erklärungslückenargumentes wie z.B. David Chalmers (1996, 94 ff.) gestützt haben, wenn sie für die prinzipielle Unüberwindbarkeit der Erklärungslücke argumentierten. Im zweiten Teil versucht Pauen dann zu zeigen, unter welchen Umständen es zumindest bei einigen phänomenalen Eigenschaften zu reduktiven Erklärungen kommen könnte.

DAS UNMÖGLICHE LEBEN DER TEILZEIT-ZOMBIES

Um im positiven Sinne aufzeigen zu können, wie eine reduktive Erklärung aussehen könnte, müssen Pauen zufolge zunächst einige psychophysische Dissoziationsmöglichkeiten ausgeschlossen werden (vgl. Pauen 2001, 270 ff.);

denn wenn es theoretisch möglich wäre, daß sich physisch/funktionale Duplikate in phänomenaler Hinsicht von ihren „Urbildern“ unterscheiden, dann – so Pauen – wäre in der Tat ausgeschlossen, daß es zu einer adäquaten funktionalen Analyse phänomenaler Eigenschaften kommen kann. Es muß also zunächst gezeigt werden, daß bestimmte psychophysische Dissoziationen weder begrifflich noch intuitiv möglich sind.

Ein wesentlicher Teil von Pauens Argumentation dient dem Nachweis, daß *intrapyschische* phänomenale Dissoziationen aus begrifflichen Gründen ausgeschlossen werden können; dies ist deshalb für Pauen von großer Bedeutung, weil er darüber hinaus die These vertritt, daß aus der theoretischen Möglichkeit *interpersoneller* psychophysischer Dissoziationen die theoretische Möglichkeit intrapsychischer Dissoziationen folgt.⁹ Sei letztere unmöglich, so liefere der Umkehrschluß das gewünschte Ergebnis: die Unmöglichkeit auch interpersoneller psychophysischer Dissoziationen. Damit wäre einer der „Königswege“ zur Begründung der explanatorischen Lücke unpassierbar geworden. Betrachten wir also zunächst die Argumentation für die begriffliche Unmöglichkeit intrapsychischer (psychophysischer) Dissoziationen:

- (1) Einer Person *P* (oder allgemeiner: einem Organismus, einem System) können nur diejenigen Zustände als Instantiierungen *phänomenaler* Eigenschaften zugeschrieben werden, die von *P* in phänomenaler Hinsicht individuiert werden können; d.h. *P* muß in der Lage sein, extreme Unterschiede zwischen den instantiierten Eigenschaften (bzw. deren völlige Abwesenheit) aus der Perspektive der ersten Person zu erkennen.
- (2) Von der Erkenntnis eines Unterschiedes zwischen verschiedenen Instantiierungen phänomenaler Eigenschaften aus der Perspektive der ersten Person kann nur dann gesprochen werden, wenn sich diese Erkenntnis in den entsprechenden Urteilen und Erinnerungen von *P* niederschlagen kann. (Diese Urteile und Erinnerungen müssen freilich nicht öffentlich mitgeteilt werden.)
- (3) Sind intrapsychische psychophysische Dissoziationen in *P* möglich, dann können sich diese Unterschiede nicht in den Urteilen und Erinnerungen von *P* niederschlagen.
- (4) Sind intrapsychische psychophysische Dissoziationen in *P* möglich, dann können *P* diejenigen Zustände, die in ihr mit unerkennbaren psy-

⁹ Intrapyschische psychophysische Dissoziationen liegen dann vor, wenn die phänomenalen Zustände einer Person bei gleichbleibenden physisch/funktionalen Zuständen variieren. Unterscheiden sich die phänomenalen Zustände von zwei in physisch/funktionaler Hinsicht identischen Personen, so spreche ich von interpersonellen Dissoziationen.

chophysischen Variationen einhergehen, nicht als Instantiierungen phänomenaler Eigenschaften zugeschrieben werden. (aus 1–3)

Die drei Prämissen (1)–(3) dieses Argumentes werden von Pauen plausibel begrifflich begründet (vgl. 2001, 273–279), die Konklusion ergibt sich zwingend. Festzuhalten ist jedoch, daß Prämisse (1) phänomenalen Zuständen implizit eine minimale kausale Rolle zuschreibt: Extreme Unterschiede zwischen ihnen müssen erkennbar sein; ein durchgehender phänomenaler Epiphänomenalismus wird mit dieser These ausgeschlossen.

Doch betrachten wir nun die Anwendung dieser Ergebnisse auf die in der philosophischen Literatur so beliebten *Gedankenschöpfe*: Zombies, invertierte Zwillinge usw. Pauen erweitert diesen Zoo für seine Argumentation durch die Einführung des sogenannten „Teilzeit-Zombies“, eines physisch/funktionalen Duplikats bewußtseinsfähiger Personen. Anders als „normale“ Zombies haben Teilzeit-Zombies jedoch nur zeitweilig keine bewußten Erfahrungen. Auf diese Weise, so Pauen, „bieten die Teilzeit-Zombies die Möglichkeit, psychophysische Dissoziationen nicht nur im Vergleich unterschiedlicher Individuen, sondern auch bei *einem* Individuum zu untersuchen“. Weil „die theoretische Möglichkeit von Teilzeit-Zombies ebenso wie die von normalen Zombies nur davon abhängt, daß phänomenale Eigenschaften nicht von physischen Eigenschaften determiniert werden, kann man sie nur bestreiten, wenn man auch die theoretische Möglichkeit ‚normaler‘ Zombies und damit auch die Möglichkeit psychophysischer Dissoziationen ablehnt“ (2001, 277).

Nach Voraussetzung kann der Teilzeit-Zombie nicht zwischen der An- und Abwesenheit phänomenaler Zustände (mindestens eines Typs) unterscheiden und deshalb können ihm auf der Basis des zuvor skizzierten Argumentes diejenigen Zustände, die mit für ihn unerkennbaren psychophysischen Dissoziationen einhergehen, überhaupt nicht als phänomenale Zustände zugeschrieben werden. Im Hinblick auf phänomenale Zustände dieses Typs ist der Teilzeit-Zombie folglich wie ein normaler Zombie anzusehen. Damit scheint freilich nicht ausgeschlossen zu sein, daß ihm phänomenale Zustände eines anderen Typs zugeschrieben werden könnten. Im Hinblick auf phänomenale Zustände eines *prima facie* zuschreibbaren Typs unterschiede sich der Teilzeit-Zombie dann jedoch nicht von seinem „Urbild“. Der Begriff eines Teilzeit-Zombies erweist sich folglich als instabil. (Zu dem gleichen Ergebnis kommt man bei Teilzeit-Inversionszwillingen; auch diesen können diejenigen Zustände, die mit für sie unerkennbaren psychophysischen Dissoziationen einhergehen, nicht als phänomenale Zustände zugeschrieben werden. Im Hinblick auf phänomenale Zustände des unzuschreibbaren Typs sind folglich auch Teilzeit-Inversionszwillinge wie normale Zombies anzusehen.)

Und was ist mit diesem Ergebnis gewonnen? Die Begriffe von Teilzeit-Zombies und Teilzeit-Inversionszwillingen erweisen sich als inkohärent. Aber ist damit auch der Begriff eines normalen Zombies und eines normalen Inversionszwillings inkohärent? Ich bin noch nicht davon überzeugt, daß dies aus

Pauens Argument folgt, auch wenn ein solches Ergebnis durch den oben angeführten Umkehrschluß auf den ersten Blick nahegelegt wird. Da der Nachweis der Unmöglichkeit der beiden fiktiven Teilzeit-Geschwister wesentlich auf der Forderung basiert, daß intrapsychische Dissoziationen aus der Perspektive der ersten Person erkennbar sein müssen, sich eine vergleichbare Forderung aber gerade nicht für interpersonelle psychophysische Dissoziationen erheben läßt – aus wessen Perspektive sollten diese Dissoziationen erkennbar sein? –, gelten nicht die gleichen Bedingungen der Möglichkeit für normale und für Teilzeit-Zombies. Der Übergang von der Unmöglichkeit intrapsychischer Dissoziationen zur Unmöglichkeit interpersoneller Dissoziationen scheint also noch nicht etabliert.

Aber nehmen wir einmal an, es wäre Pauen gelungen, auch die Inkohärenz von „normalen“ Zombies und Inversionszwillingen zu zeigen. Wo stünden wir damit? Nun, ein wichtiges *a priori* Argument für die Existenz der Erklärungslücke stünde dann nicht mehr zur Verfügung – ein Argument allerdings, dem keineswegs alle an der Debatte Beteiligten hohe Priorität eingeräumt haben. So schreibt beispielsweise Chalmers: „Still, some people find conceivability arguments difficult to adjudicate, particularly where strange ideas ... are concerned. It is therefore fortunate that every point made by using zombies can also be made in other ways ...“ (1996, 99). Dennoch tat Pauen gut daran, sich mit dem *a priori* Argument auseinanderzusetzen; denn gelänge es nicht, dieses Argument aus dem Weg zu räumen, so wäre sein weiterführendes Vorhaben, aufzeigen zu wollen, wie die Erklärungslücke vielleicht geschlossen werden könnte, von vornherein zum Scheitern verurteilt.

Meines Erachtens werden wir jedoch durch die Konzentration auf Dissoziationsbehauptungen auf der Basis *physisch/funktionaler* Duplikate eher auf eine falsche Fährte geführt. Insbesondere entspricht dies nicht der ursprünglichen Einführung der Erklärungslücke, soweit sie ihre Wurzeln in der Emergenzdebatte hat. Broad zum Beispiel hielt es, wie u.a. aus dem Beitrag von Ansgar Beckermann (in diesem Band) hervorgeht, für durchaus plausibel, daß Korrelationsbeziehungen zwischen bestimmten Mikrostrukturen und *bestimmten* phänomenalen Zuständen einer Person einen nomologischen Charakter besitzen. Zugleich war er jedoch der Auffassung, daß Empfindungen notwendig emergent sind, weil es sich bei den Gesetzen, die Hirnzustände mit Empfindungen verbinden, um nicht ableitbare Gesetze handelt, um Gesetze also, die nicht den Status von Brückenprinzipien¹⁰ besitzen. Seine eigenen Intuitionen hinsichtlich möglicher phänomenaler Dissoziationen beschränken sich freilich auf behaviorale (oder funktionale) Übereinstimmungen.

¹⁰ Brückenprinzipien sind Prinzipien, aus denen hervorgeht, wie sich ein System als Ganzes verhält, wenn sich seine Teile auf eine bestimmte Weise verhalten. Man sollte sie nicht mit den „Brückengesetzen“ Ernest Nagels verwechseln (vgl. Beckermann 2001, 131).

Wie wenig jedoch mit dem bloßen Zurückweisen des *a priori* Argumentes gewonnen ist, sieht man vor allem, wenn man einmal die molekülidentischen Zombies und Inversionszwillinge beiseite läßt und stattdessen fragt, welche Aussagen wir über „Näherungszwillinge“ treffen können. Wenn auch nur ein neuronales Netz in meinem Näherungszwilling anders strukturiert ist als in mir, was folgt aus der bisherigen Argumentation Pauens über dessen phänomenale Zustände? Nun, es folgt, daß auch diesem nur diejenigen Zustände als phänomenale zugeschrieben werden können, die er in phänomenaler Hinsicht individuieren kann. Ich habe jedoch kein zwingendes Argument mehr dafür, ihm dieselben oder ähnliche phänomenale Zustände wie mir zuzuschreiben. Noch schwieriger gestalten sich Aussagen über phänomenale Zustände von Organismen und Systemen, die mit uns nur in grobkörniger Weise funktional übereinstimmen (etwa so, wie Röhrenverstärker, Transistorverstärker und Mikrochipverstärker miteinander übereinstimmen).

Natürlich ist sich auch Pauen dieser Probleme bewußt. Und diese zu lösen, ist an dieser Stelle seiner Argumentation, wie ich bereits andeutete, auch nicht sein Anspruch. Uneingeschränkte Dissoziationsbehauptungen zurückzuweisen ist für ihn lediglich eine unverzichtbare Vorarbeit für seine Skizze, wie man der Erklärungslücke beikommen könnte. Betrachten wir nun diese Strategie etwas näher.

LÜCKENLOSE EMOTIONEN

Im zweiten Teil seiner Argumentation stellt sich Michael Pauen die Aufgabe, am Beispiel von paradigmatischen Zuständen des phänomenalen Bewußtseins (wie Furcht- und Schmerzzuständen) deutlich zu machen, daß das Erklärungslückenargument sein Ziel insofern verfehlt, als nicht *ausgeschlossen* werden kann, daß jene Zustände adäquat in funktionalen Beschreibungen erfaßt werden können. Folglich gäbe es einige phänomenale Eigenschaften, die zumindest im Prinzip reduktiv erklärt werden könnten. Seine Argumentation vollzieht sich in drei Schritten:

- (1) Bei bestimmten Typen phänomenaler Zustände wie Emotionen und Schmerzen haben bereits unsere Alltagsbegriffe konkrete funktionale Implikationen, auch wenn diese noch ungenau und wahrscheinlich korrekturbedürftig sind (2001, 280–282, 289–289).
- (2) Entwicklungen in den empirischen Wissenschaften wie der Emotionspsychologie geben zu der Hoffnung Anlaß, daß jene funktionalen Implikationen angemessen korrigiert und konkretisiert werden können (2001, 290–292).

- (3) Die korrigierten und konkretisierten Beschreibungen können zugleich eine Basis zur Bestimmung derjenigen neuronalen Prozesse liefern, die diese Funktionen und damit die phänomenalen Zustände selbst realisieren (2001, 292–293).

Pauens Gedankengang stützt sich u.a. auf den Nachweis, daß bereits unsere alltagspsychologischen phänomenalen Begriffe in ein funktionales Netz eingebettet sind. Dazu imaginiert er in Analogie zu den Gedankenexperimenten, bei denen die erlebten Farbspektren als vertauscht angenommen werden, Situationen, in denen Empfindungen des Verlangens und der Abneigung miteinander vertauscht sind, und fragt, ob unsere Alltagsbegriffe *generell* die Möglichkeit einer substantiellen Veränderung auf der phänomenalen Ebene zulassen, die keine Auswirkungen auf der funktionalen Ebene hat, so wie es das Argument des invertierten Spektrums zu zeigen scheint.

Nach Pauens Ansicht ist leicht zu erkennen, daß das folgende Beispiel mit unseren Alltagsintuitionen kollidieren würde. Nehmen wir an, zwei Kinder, ein „normales“ und eines, bei dem Schmerzen und angenehme Empfindungen invertiert sind, würden nach der Geburt von ihrer Mutter gestreichelt: „Während das eine diese Berührungen als angenehm empfindet, sind sie für das zweite schmerzhaft und damit unangenehm. Es ist nicht zu sehen, warum dieses Kind nicht das typische Schmerzverhalten zeigen sollte. Genausowenig scheint die Erwartung berechtigt, daß leichte Verletzungen bei diesem Kind das übliche Schmerzverhalten auslösen sollten, schließlich erlebt es diese Verletzungen ja als angenehm“ (2001, 283).¹¹

Vermutlich kann sich kaum jemand der suggestiven Kraft von Pauens Überlegung entziehen. Sie zeigt, daß es zumindest in einigen Fällen unplausibel ist anzunehmen, es könne auch extreme Qualia-Inversionen geben, ohne daß damit zugleich funktionale Konsequenzen einhergehen. Dennoch möchte ich die Aussagekraft dieses Gedankenexperiments in Frage stellen. Es könnte nämlich sein, daß die fiktiven Inversionen zwischen Empfindungen der *Abneigung* und des *Angenehmen* weniger einer Inversion zwischen Rot- und Grünempfindungen als einer zwischen z.B. Farb- und Tonempfindungen entsprechen. So weit ich weiß, hat jedoch niemand, und das räumt auch Pauen ein, behauptet, daß es plausibel wäre, Qualia-Inversionen des zuletzt genannten Typs anzunehmen, ohne daß damit zugleich funktionale Konsequenzen einhergehen müßten. Umgekehrt ließe sich fragen, ob sich fiktive Qualia-Inversionen innerhalb des Spektrums der aversiven Emotionen (bzw. innerhalb des Spektrums der angenehmen Emotionen) ähnlich zwingend in Verhaltensänderungen niederschlagen müßten. Es kommt also entscheidend darauf

¹¹ Dieses Beispiel verwendet Pauen eigentlich zur Zurückweisung eines weiteren, intuitiven Arguments für die Möglichkeit psychophysischer Dissoziationen; ich diskutiere es hier unter anderen systematischen Gesichtspunkten in Verbindung mit seinem „Furchtbeispiel“.

an, welche Kandidaten als adäquate Inversionspaare angesehen werden können.

Die von Pauen diskutierte fiktive Inversion zwischen Empfindungen des Angenehmen und der Abneigung bezieht sich ja auf Emotionen, die derart basal sind, daß man sie bei nahezu allen Lebewesen antrifft, ohne daß damit notwendigerweise phänomenale Erlebnisse einhergehen müssen. Es scheint sich geradezu um Konstituenten tierlichen Verhaltens zu handeln, die, soweit wir wissen, auch je unterschiedlich implementiert sind: Aversive Emotionen scheinen vor allem mit Vorgängen in einem „festverdrahteten“ System mit der Amygdala als zentraler Struktur einherzugehen, dagegen sind Gefühle des Angenehmen besonders mit der Ausschüttung von Neuropeptiden (Oxytocin) und Neuromodulatoren (Dopamin, Serotonin) verbunden. Es ist daher fraglich, ob wir Emotionen des Angenehmen überhaupt mit Emotionen der Abneigung zu einer einzigen Kategorie verschmelzen sollten, oder ob wir sie nicht besser zwei verschiedenen Kategorien zurechnen. Dann wäre aber gerade zu bezweifeln, ob es sich bei Inversionen zwischen Empfindungen des Angenehmen und der Abneigung um adäquate Inversionen handelt. Es stellt sich somit die Frage, was eigentlich genau gezeigt ist, wenn wir extreme Qualia-Inversionen wie die von Schmerz- und Lustempfindungen ausschließen können.

Zur Erörterung dieser Frage ist die von Pauen zu Beginn seiner Untersuchung eingeführte Unterscheidung zwischen starken und schwachen funktionalen Implikationen hilfreich: „Schwache funktionale Implikationen liegen dann vor, wenn eine höherstufige Eigenschaft eine funktionale Rolle besitzt, die sie von *einigen*, nicht jedoch von *allen* anderen höherstufigen Eigenschaften desselben Typs unterscheidet“ (2001, 272). Typen von phänomenalen Eigenschaften seien beispielsweise Farbempfindungen, Geruchsempfindungen oder Emotionen.¹² Von starken funktionalen Implikationen könne dann die Rede sein, „wenn eine höherstufige Eigenschaft oder eine Gruppe eng verwandter höherstufiger Eigenschaften, beispielsweise Schmerzigenschaften, eine funktionale Rolle besitzt, die *nur* für sie charakteristisch ist.“ Starke funktionale Implikationen würden uns daher die Möglichkeit geben, „wesentliche phänomenale Unterscheidungen, im Idealfall sogar *alle* phänomenalen Unterscheidungen auf das funktionale Vokabular abzubilden“ (2001, 272).

Das gerade vorgestellte Beispiel invertierter Schmerzen habe zumindest schwache funktionale Implikationen, und diese reichen Pauen zufolge bereits aus, um einen weiteren wichtigen Einwand *gegen* die funktionale Analysierbarkeit phänomenaler Begriffe auszuräumen, da es zeige, daß auch die vermeintliche intuitive Möglichkeit beliebiger psychophysischer Dissoziationen ausgeschlossen werden kann (vgl. 2001, 284). Allerdings ist damit, und auch

¹² Wie die Diskussion im vorangegangenen Abschnitt zeigt, ist es jedoch fraglich, ob man alle emotionalen Erlebnisse einem einzigen Typ phänomenaler Eigenschaften zurechnen sollte.

das räumt Pauen ein, noch nicht *positiv* gezeigt, daß phänomenale Eigenschaften in funktionalen Begriffen adäquat erfaßt werden können. „Hierzu wäre zumindest eine deutliche Annäherung an *starke* funktionale Implikationen erforderlich“ (2001, 272).

Pauen stellt sich auch dieser Aufgabe, wenn er zu zeigen versucht, daß zumindest auf längere Sicht eine reelle Chance besteht, zu funktionalen Beschreibungen und damit letztlich auch zu zufriedenstellenden Erklärungen konkreter phänomenaler Zustände zu kommen. Als Beispiel dient ihm hierbei die Emotion der Furcht.

Obzwar unsere derzeitigen alltagspsychologischen Kriterien bezüglich der Ursachen und Wirkungen von Furcht noch so rudimentär seien, „daß man bei einer Person, die alle diese Kriterien erfüllt, immer noch zweifeln kann, ob sie tatsächlich Furcht empfindet – auch wenn man möglicherweise auszuschließen vermag, daß sich die Person in einem diametral entgegengesetzten Zustand befindet“ (2001, 290), hält Pauen adäquate Präzisierungen im Prinzip für möglich. Sollte es nämlich neue und weiterreichende wissenschaftliche Erkenntnisse über die kausale Rolle phänomenaler Zustände geben, dann wäre zu erwarten, daß diese Erkenntnisse auf die Dauer auch unsere alltagspsychologischen Überzeugungen beeinflussen (vgl. 2001, 285). Als positives Beispiel dienen ihm evolutionspsychologische Modelle der Emotionen wie das von Robert Plutchik (vgl. 1993), die die funktionalen Implikationen unseres Alltagsbegriffs der Furcht weitgehend bestätigen und weiter konkretisieren:

Den Erkenntnissen der Emotionspsychologie zufolge zählen zu den kognitiven Bestandteilen von Furchtzuständen zum einen die Bewertung der eigenen Situation, zum anderen die Planung von Flucht- und Abwehrreaktionen. Solche Reaktionen bilden gleichzeitig die Handlungstendenzen, die für Furchtzustände konstitutiv sind. ... Außerdem gehören auch spezifische Körperreaktionen zu den Konstituenten von Furchtzuständen. Diese Reaktionen bereiten den Organismus auf eine mögliche Flucht vor. Zu nennen sind hier insbesondere die Erhöhung von Puls, Blutdruck und Atemfrequenz (2001, 291).

Auch wenn Pauen ausdrücklich darauf verweist, daß er diese empirischen Untersuchungen nicht anführt, weil sie schon die verlangte reduktive Erklärung lieferten, sondern weil es nicht unrealistisch wäre anzunehmen, „daß die dabei anfallenden Erkenntnisse über funktionale Merkmale phänomenaler Zustände auch Einfluß auf unsere phänomenalen Begriffe gewinnen“ könnten (vgl. 2001, 290), stellt sich die Frage, was sich im Hinblick auf das *philosophische* Problem phänomenaler Qualitäten mit diesen Erkenntnissen gewinnen läßt? Inwiefern sind wir damit einer relationalen oder funktionalen „Präparierung“ des *qualitativen Gehaltes* eines emotionalen Zustandes näher gekommen? Was Pauen zeigen müßte, ist, daß sich nach, sagen wir, idealen Präzisierungen, die denen von Plutchik weit überlegen wären, nicht mehr trefflich streiten ließe,

ob diejenigen, auf die die angeführten Beschreibungen zutreffen, tatsächlich das *empfinden*, was als eine genuine Furchtempfindung gilt.

Nun hat man in Tierversuchen jedoch herausgefunden, daß Tiere trotz massiver Ablationen bei der emotionalen Reaktionsfähigkeit ein auffällig normales Verhalten zeigen (vgl. LeDoux 1998, 87). So waren bei Katzen auch nach Entfernung der gesamten Hirnrinde noch alle typischen Anzeichen emotionaler Erregung zu beobachten (z.B. Fellsträuben, Erhöhung des Blutdrucks und der Herzfrequenz). Da aber nicht anzunehmen ist, daß es auch ohne Beteiligung des Kortex zu *bewußten* emotionalen Erlebnissen kommt, kann man schließen, daß die funktionale Einbettung von emotionalen Zuständen allein nicht ausreicht, verlässliche Auskunft über das Vorhandensein und die Qualität eines phänomenalen Erlebnisses zu geben.

Was Plutchik und anderen also wahrscheinlich gelingen wird, ist, die funktionale Einbettung eines *emotionalen Zustandes* weiter voranzutreiben. Nicht zu gelingen scheint jedoch die funktionale Individuierung des spezifischen *qualitativen Gehaltes* des entsprechenden emotionalen Zustandes. Doch Pauen ist da anderer Ansicht. Er hält die Verbindung zwischen der phänomenalen und der physiologischen Ebene – entgegen früheren Annahmen – für so eng, „daß es nicht nur möglich ist, einen emotionalen Zustand anhand seiner physiologischen Konsequenzen zu identifizieren, sondern umgekehrt die phänomenale Erfahrung entstehen zu lassen, indem man die zugehörigen physiologischen Konstituenten aktiviert“ (2001, 291). Auch das will ich natürlich nicht bestreiten. Doch die Fähigkeit, einen phänomenalen Zustand (willkürlich) entstehen zu lassen, impliziert nicht die Fähigkeit, den qualitativen Gehalt dieses Zustandes *reduktiv* erklären zu können. Für Pauen ist hingegen entscheidend,

daß die skizzierten kognitiven, dispositionalen und körperlichen Reaktionen nicht einfach nur lose mit Furchtzuständen oder ihren phänomenalen Eigenschaften assoziiert sind, vielmehr muß es gute Gründe für die Annahme geben, daß es sich hier um *konstitutive Bestandteile* dieser Zustände handelt. Solche Gründe gibt es. Es spricht nämlich bereits heute einiges dafür, daß Herzpochen nicht einfach nur in einem kontingenten Zusammenhang mit der davon unabhängigen phänomenalen Qualität steht, vielmehr handelt es sich offenbar um einen direkten Bestandteil der phänomenalen Erfahrung selbst. Die phänomenale Qualität eines Furchtzustandes ändert sich also unmittelbar dadurch, daß das Herz mehr oder weniger stark pocht, daß die Fluchttendenz mehr oder minder groß ist oder aber die eigenen Gedanken mehr oder minder stark darauf konzentriert sind, einen Ausweg aus der Situation zu finden (2001, 291).

Betrachten wir Pauens zentrale Behauptungen: (1) Das mit einem Furchtzustand korrelierte Herzpochen ist nicht bloß lose mit diesem Zustand assoziiert, sondern ein *konstitutiver* Bestandteil desselben. Und mehr noch: (2) Es ist nicht nur ein konstitutiver Bestandteil des *emotionalen* Zustandes, sondern auch Bestandteil der mit ihm einhergehenden *phänomenalen Qualität*. (3) Deshalb ließen sich die phänomenalen und die funktionalen Eigenschaften eines emotionalen Zustandes nicht wirklich voneinander trennen.

Zwar handele es sich dabei zunächst nur um empirische Befunde, die in dieser Form noch keine Erklärungen der phänomenalen Eigenschaften emotionaler Zustände durch deren funktionale Eigenschaften lieferten. Dazu sei die Integration dieser Ergebnisse in unsere phänomenalen Alltagsbegriffe nötig. Es zeige sich jedoch bereits, daß es möglich sei, „den Zusammenhang zwischen der funktionalen und der phänomenalen Ebene aus der Perspektive der ersten Person nachzuvollziehen“ (2001, 291).

Ich bestreite nicht, daß man die mit einem emotionalen Zustand verbundenen physiologischen Zustände als konstitutive Bestandteile dieser Zustände ansehen kann. Ebenso wenig will ich bestreiten, daß man das phänomenale Erleben eines pochenden oder rasenden Herzens als einen konstitutiven Bestandteil des phänomenal erlebten Furchtzustandes ansehen kann. Was ich allerdings bestreite, ist, daß man den *physiologischen* Zustand eines pochenden Herzens als einen konstitutiven Bestandteil des *phänomenal* erlebten Furchtzustandes ansehen kann. Aber selbst wenn es einen Grund für diese Annahme gäbe, folgte daraus nicht, daß eine Erklärung der funktionalen Implikationen eines Furchtzustandes eine reduktive Erklärung der mit ihm einhergehenden phänomenalen Qualitäten überflüssig macht. Auch das Erleben der phänomenalen Qualität eines pochenden Herzens bedarf selbst noch einer reduktiven Erklärung: Wie fühlt sich ein stark pochendes Herz aus der Perspektive der ersten Person an? Fühlt es sich im Zustand der Furcht genau so an wie im Zustand freudiger Erregung? Oder wie nach einem anstrengenden Dauerlauf? Muß es sich so anfühlen?

Es scheint, als würde man beim angestrebten Versuch der reduktiven Erklärung phänomenaler Eigenschaften vor einem ähnlichen Problem stehen wie die Behavioristen und semantischen Physikalisten, als sie Mitte des letzten Jahrhunderts versuchten, mentales Vokabular durch ausschließlich physikalisches Vokabular zu ersetzen. Sicher, es mag gelingen, phänomenale Konstituenten eines Furchtgefühls herauszuarbeiten. Diese selbst bleiben jedoch weiterhin erklärungsbedürftig. Ich sehe nicht, wie solch basale Konstituenten wie die der Empfindung eines klopfenden Herzens in eine funktionale Analyse unserer Alltagsbegriffe phänomenaler Zustände sollten Eingang finden können.

Ein wirklich erfolgreich durchgeführtes Programm reduktiven Erklärens phänomenaler Zustände müßte darüber hinaus natürlich auch auf *andere* erlebnisfähige Organismen und Systeme anwendbar sein. Es müßte transparent machen, wie etwa ein Klaustrophobiker seine Furcht empfindet, wenn er sich

in einem engen Aufzug befindet, wie ein Alzheimer-Patient Emotionen erlebt, und es müßte zeigen, auf welche Weise ein Hase, der mit pochendem Herzen und Haken schlagend einem Fuchs zu entkommen sucht, seine Hasen-Angst empfindet. Darüber hinaus müßte es möglich werden, Auskunft darüber zu geben, ob bzw. unter welchen Bedingungen ein künstliches System – etwa ein sehr komplexer Roboter – zu Erlebnissen fähig ist.

Hilft der von Pauen vorgeschlagene Ansatz in diesen Fragen weiter? Ich meine, nein. Trotz des zweifellos zu erwartenden Fortschritts in unserem Verständnis emotionaler Zustände durch die in Aussicht gestellten funktionalen Präzisierungen habe ich nicht den Eindruck, daß wir hinsichtlich der explanatorischen Lücke entscheidend vorankommen können. Was Pauen erhält, wenn sich sein Vorhaben einigermaßen erfolgreich umsetzen läßt, sind Revisionen verschiedener emotionaler Begriffe in funktionaler Hinsicht. Was er nicht erhält, ist eine funktionale Einbettung der phänomenalen Qualitäten emotionaler Zustände.

IV. FAZIT

Die neueren Vorschläge, die explanatorische Lücke als geschlossen oder doch wenigstens als schließbar zu betrachten, die ich an den Beiträgen von David Papineau und Michael Pauen etwas ausführlicher diskutiert habe, haben mich nicht überzeugen können. Phänomenale Qualitäten scheinen nach wie vor sperrig gegenüber reduktiven Erklärungen zu sein, und das aus prinzipiellen Gründen.

Wenn man behauptet, daß phänomenale Qualitäten *prinzipiell* nicht reduktiv erklärt werden können, so ist damit jedoch kein Mißlingen der empirischen Schritte verbunden. Impliziert wird vielmehr ein Mißlingen der begrifflichen Schritte. Das bedeutet natürlich nicht, daß Neurobiologen oder Kognitionswissenschaftler aufhören sollten, qualitative Zustände zu erforschen. Aber wir sollten anerkennen: Was immer sie entdecken mögen, wird nicht das *philosophische* Problem phänomenalen Erlebens lösen. Vergleichen wir erneut mit den mathematischen Problemen: Natürlich hörten Mathematiker nicht auf, nach Lösungen von Gleichungen fünften Grades zu suchen. Sie entwickelten andere, numerische Verfahren. Aber sie lösten nicht das *alte* Problem. Das alte Problem verlor, nachdem bekannt war, daß es nicht nur ein schwieriges, sondern ein unlösbares Problem darstellt, seine Anziehungskraft. Es wurde zu einem Teil der Mathematikgeschichte. Niemand würde das Problem heute noch als solches in Angriff nehmen wollen. Das hat durchaus Vorteile: Man kann seine Kraft auf solche Probleme verwenden, die beantwortbar erscheinen.

V. LITERATUR

- Beckermann, A. 2001: Die reduktive Erklärbarkeit phänomenalen Bewußtseins – C.D. Broad zur Erklärungslücke. In diesem Band.
- Block, N. und Stalnaker, R. 1999: Conceptual Analysis, Dualism, and the Explanatory Gap. *Philosophical Review* 108: 1–46.
- Broad, C.D. 1921: Prof. Alexander's Gifford Lectures. *Mind* 30 : 25–39 und 129–150.
– 1925: *The Mind and Its Place in Nature*. London.
- Chalmers, D. 1996: *The Conscious Mind*. Oxford.
- Churchland, P. 1996: The Hornswoggle Problem. *Journal of Consciousness Studies* 3: 402–408.
- Hill, C. 1997: Imaginability, Conceivability, Possibility and the Mind-Body Problem. *Philosophical Studies* 87: 61–85 (dt. Übersetzung in diesem Band).
- Kim, J. 1993: The Nonreductivist's Troubles with Mental Causation. In ders.: *Supervenience and Mind*. Cambridge.
– 1998a: *Philosophie des Geistes*. Wien.
– 1998b: *Mind in a Physical World*. Cambridge.
– 2001: Emergenz, Reduktionsmodelle und das Mentale. In diesem Band,
- LeDoux, J. 1998: *Das Netz der Gefühle. Wie Emotionen entstehen*. München.
- Leibniz, G.W. 1714: *Monadologie*. Hamburg 1969.
- Levine, J. 1983: Materialism and Qualia: The Explanatory Gap. *Pacific Philosophical Quarterly* 64: 354–361.
– 1993 On Leaving Out What It's Like. In: M. Davies und G.W. Humphreys (Hrsg.): *Consciousness*. Oxford.
- Loar, B. 1997: Phenomenal States. In: N. Block, O. Flanagan, und G. Güzeldere (Hrsg.): *The Nature of Consciousness*. Cambridge MA.
- McLaughlin, B. 2001: In Defense of New Wave Materialism. In: B. Loewer und C. Gillett (Hrsg.): *Physicalism and Its Discontents*. Cambridge (dt. Übersetzung in diesem Band).
- Papineau, D. 1998: Mind the Gap. *Philosophical Perspectives* 12: 373–388 (dt. Übersetzung in diesem Band).
- Pauen, M. 2001: Invertierte Schmerzen. In diesem Band.
- Plutchik, R. 1993: Emotions and their Vicissitudes. In: M. Lewis und J.M. Haviland (Hrsg.): *Handbook of Emotions*. New York London.
- Stephan, A. 1997: Armchair Arguments Against Emergentism. *Erkenntnis* 46: 305–314.
– 1999: *Emergenz. Von der Unvorhersagbarkeit zur Selbstorganisation*. Dresden.