

Das Bewußtsein und seine neurobiologische Erklärung

VON GÜNTER RAGER

Auf Grund der großen Fortschritte in den Neurowissenschaften während der letzten Jahre hat das Interesse an Fragen nach dem Bewußtsein stark zugenommen. Es werden immer neue und immer speziellere Strukturen und Funktionen im Gehirn entdeckt, die mit Bewußtseinsvorgängen in besonderer Weise verknüpft sind. Es liegt deshalb nahe anzunehmen, das Bewußtsein könne schließlich neurobiologisch erklärt werden, auch wenn wir heute im einzelnen dazu noch nicht in der Lage sind. Diese Perspektive wiederum ermutigt eine Reihe von Philosophen der naturalistischen Richtung, Modelle für die naturalistische Erklärung des Bewußtseins zu entwickeln. Einer der prominentesten Vertreter dieser Richtung, nämlich Daniel Dennett¹, hat die These „Consciousness Explained“ sogar zum Titel eines seiner Bücher erhoben. Auf der anderen Seite wird weiterhin mit guten Gründen die Auffassung vertreten, das Bewußtsein lasse sich nicht mit den Mitteln der Neurowissenschaften erklären.

Um in dieser kontroversen Situation Stellung beziehen zu können, müssen zuerst einige Begriffe geklärt werden. Sodann werden wir an Beispielen aus der neurobiologischen Literatur aufzeigen, in welche Richtung die gegenwärtigen neurobiologischen Erklärungen von Bewußtsein weisen. Nach einer kritischen Beurteilung dieser Erklärungen, werde ich im vierten und letzten Teil einen Vorschlag vorlegen, wie wir neurobiologische Erkenntnisse zum Bewußtsein mit unserer Selbsterfahrung in Verbindung bringen können.

1. Begriffsklärung

Zuerst sollen also die drei Begriffe geklärt werden, welche im Titel dieser Arbeit vorkommen: Was soll hier unter Bewußtsein verstanden werden? Was verstehen wir unter Neurobiologie? Und schließlich: Was heißt neurobiologisch erklären?

Jeder von uns meint zu wissen, was *Bewußtsein* ist, „solange uns keiner danach fragt, es zu definieren“². Für Thomas Metzinger ist das Bewußtsein „alles andere als klar“; es gebe noch „keinen einheitlichen theoretischen Hintergrund, vor dem wirklich so etwas wie eine Wissenschaft des Bewußtseins entstehen könnte“³. David Chalmers⁴ sieht keine Möglichkeit der Definition, sondern allenfalls die Möglichkeit, mit Umschreibungen wie *Erfahrung* oder *Erleben* zur Klärung beizutragen. Da Bewußtsein nicht de-

¹ D. C. Dennett, *Consciousness explained*, Boston 1991.

² W. James, *Principles of psychology*, New York 1890, 225: „We know so long as no one asks us to define it“.

³ T. Metzinger, Einleitung. Das Problem des Bewußtseins. In: T. Metzinger (Hrsg.): *Bewußtsein. Beiträge aus der Gegenwartsphilosophie*, Paderborn 1996, 15–53.

⁴ D. J. Chalmers, *The conscious mind*. In: *Search of a fundamental theory*, New York 1996.

finiert werden kann, kann es auch nicht erklärt werden. Deshalb müßten wir eigentlich bereits hier unsere Überlegungen abbrechen. Damit würden wir aber nicht nur die umfangreiche Literatur zu diesem Thema, sondern auch das große Problem ignorieren, das heute viele Menschen sehr beschäftigt. Deshalb sei im folgenden trotz mangelnder Definition eine Annäherung an die Bewußtseinsproblematik versucht.

Mentale Zustände gelten dann als bewußt, wenn sie Erfahrungsqualitäten haben, welche auch als *phänomenale Qualitäten* oder „*Qualia*“ bezeichnet werden. Ein Bewußtsein, welches durch subjektive Erfahrung gekennzeichnet ist, wird *phänomenales Bewußtsein* genannt⁵. Wenn die bewußten Akte jedoch als Ursache oder Erklärung von Verhalten verstanden werden, dann spricht man von *psychologischem Bewußtsein*⁶.

Neurobiologie sei definiert als eine Naturwissenschaft, welche die Struktur und die Funktion der Nervensysteme des Menschen und der Tiere aufklären will. Sie benutzt dazu eine Vielzahl verschiedener Methoden und Disziplinen. Allen neurowissenschaftlichen Disziplinen ist gemeinsam, daß sie das Nervensystem letztlich unter quantitativ-mathematischen Gesichtspunkten untersuchen wollen.

Das Bewußtsein *erklären* ist heute eine vielgebrauchte Redeweise. Häufig geben jedoch die neurobiologischen Autoren nicht an, was sie unter „erklären“ verstehen. Meist ergibt sich erst aus dem Kontext, daß eine reduktive Erklärung gemeint ist, d. h. das Bewußtsein soll auf einfachere Entitäten zurückgeführt werden. Demnach hätten wir das Bewußtsein dann reduktiv erklärt, wenn wir die neuronalen Prozesse angeben könnten, welche die bewußten Akte nicht nur begleiten, sondern auch hervorbringen.

Von einer derartigen Erklärung muß man erwarten, daß für das *Explanandum* die notwendigen und die hinreichenden Bedingungen angegeben werden. Wenn wir neurobiologisch faßbare Prozesse angeben, dann sollten diese Prozesse nicht nur notwendige Bedingungen für bewußte Akte sein, sondern diese Akte auch hinreichend bestimmen. Wir könnten diese Forderung in einem Satz formulieren und sagen: Das Bewußtsein ist nichts anderes als das Ergebnis der Prozesse des Nervensystems. Und noch einmal präziser formuliert: Es besteht eine Äquivalenz zwischen neuronalen Prozessen und Bewußtsein⁷.

⁵ Ebd. 4; Chalmers spricht von einem „qualitativen Gefühl“ („a qualitative feel“).

⁶ Ebd. 11 und 26–28.

⁷ In der deduktiv-nomologischen Erklärung (Hempel-Oppenheim) sind die Antecedens-Bedingungen notwendige Bedingungen. Wenn sie zusammen mit den Gesetzesaussagen vollständig sind, dann sind sie auch hinreichend für die Erklärung des Explanandum. Die Gesetzesaussagen würden dann lauten: Wenn die Antecedensbedingungen hinreichend sind, ist Bewußtsein erklärt. Mit dem Hempel-Oppenheim-Schema wäre zu prüfen, ob das oben formulierte Explanandum wirklich hergeleitet werden kann, ob sich tatsächlich nicht nur einige, sondern alle notwendigen Bedingungen für den behaupteten Sachverhalt finden lassen und damit die Menge der Bedingungen auch hinreichend wäre für die Erklärung. Siehe C. G. Hempel, *Aspects of scientific explanation*, New York 1965; W. Stegmüller, *Wissenschaftliche Erklärung und Begründung*, Berlin 1969.

2. Neurobiologische Erklärungsversuche

Wie sieht es nun mit den neurobiologischen Erklärungen tatsächlich aus? Was wird mit dem Nachweis von neuronalen Funktionskreisen wirklich gezeigt? Zur Beantwortung dieser Frage wollen wir aus der Vielzahl von Veröffentlichungen vier Gruppen von Publikationen auswählen, welche die Erklärung des Bewußtseins ausdrücklich anzielen. Bei den ersten drei Publikationsgruppen handelt es sich um Veröffentlichungen in wissenschaftlichen Zeitschriften, bei der vierten Gruppe – den Arbeiten von Gerald Edelman – um umfangreiche Buchveröffentlichungen, für deren Besprechung auch in diesem Rahmen wesentlich mehr Raum erforderlich wird.

2.1 *Das visuelle Bewußtsein*

Die Arbeit von Francis Crick und Christof Koch⁸ trägt den Titel: „Auf dem Weg zu einer neurobiologischen Theorie des Bewußtseins“. Sie wird viel zitiert, weil sie Hypothesen zur neuronalen Implementierung von Bewußtseinsakten liefert, welche an vorhandenen Daten anknüpfen und konkrete Vorschläge liefern, wie sie experimentell getestet werden können. Sie ist aber auch eine der wenigen neurobiologischen Arbeiten, in welchen einige der Begriffe definiert werden, die in den Hypothesen auftreten. Eine Definition des Bewußtseins selbst wird jedoch ausdrücklich ausgeklammert. Die Autoren befürchten, eine derartige Definition sei zur jetzigen Zeit unreif und könnte deshalb leicht zu Mißverständnissen führen. An Stelle einer Definition berufen sie sich auf die grobe Vorstellung von Bewußtsein, die jeder von uns hat.

Crick und Koch gehen davon aus, daß es einen elementaren Mechanismus für Bewußtsein gibt, welcher in den verschiedenen Subsystemen des Gehirns ziemlich ähnlich abläuft. Wenn man den Mechanismus für eine Form des Bewußtseins kennt, dann hat man vielleicht den Schlüssel zum Verständnis für alle Bewußtseinsformen gefunden. Deshalb schränken sie sich vorläufig auf das visuelle System ein, welches ihnen aus verschiedenen Gründen besonders geeignet erscheint. Dabei klammern sie die Frage nach den Qualia, welche als charakteristisch für das phänomenale Bewußtsein angesehen werden, ausdrücklich aus. Für den Einbezug quantenmechanischer Aspekte, an die heute wieder verschiedentlich angeknüpft wird⁹, sehen sie keinen Anlaß. Ganz entscheidend für den Bewußtseinsprozeß ist das Vorhandensein eines ‚Kurzzeitgedächtnisses‘. Die Autoren unterscheiden zwei Formen des Kurzzeitgedächtnisses: ein sehr schnelles (Größenord-

⁸ F. Crick, F. and C. Koch, Towards a neurobiological theory of consciousness, *Semin. Neurosci.* 2 (1990) 263–275.

⁹ J. C. Eccles, *How the self controls its brain*, Berlin 1994; R. Penrose, *The emperor's new mind*. Oxford 1989; R. Penrose, *Shadows of the mind: A search for the missing science of consciousness*, Oxford 1994.

nung ein Bruchteil einer Sekunde) oder bildhaftes (ikonisches) Gedächtnis und ein langsames (Größenordnung einige Sekunden) oder Arbeitsgedächtnis (working memory). Das bildhafte oder ikonische Gedächtnis kann sehr schnell elementare sensorische Informationen aufnehmen, so z. B. Orientierung oder Bewegung eines Objekts. Es hat eine sehr große, aber nur kurz dauernde Kapazität. Das Arbeitsgedächtnis hingegen hat eine begrenzte Kapazität, kann aber dafür verschiedene Muster miteinander kombinieren und sie für mehrere Sekunden speichern¹⁰.

Da die Information über ein bestimmtes Objekt im Gehirn verteilt ist, wird ein Mechanismus postuliert, der für eine bestimmte Zeit die Aktivität aller in diesem Moment relevanten Neurone koordiniert. Dies geschieht unter dem Einfluß der *visuellen Aufmerksamkeit* („visual attention“). Die visuelle Aufmerksamkeit wählt ein bestimmtes Objekt an einem bestimmten Ort im Gesichtsfeld aus. Sie benutzt dazu eine besondere topographische Karte, in welcher die Objekte im Gehirn repräsentiert sind und durch bestimmte Eigenschaften hervorstechen. Sobald in der topographischen Karte ein hervorstechendes Objekt ausgewählt ist, wird die topographische Information verbunden mit der Information über die individuellen Merkmale des Objekts. Die Aktivität der verschiedenen relevanten Neuronengruppen wird dann für eine bestimmte Zeit zusammengebunden („binding problem“). Das Zusammenbinden erfolgt durch Synchronisierung der Aktionspotentiale dieser Neuronengruppen, welche dann gemeinsam in 40 Hz Oszillationen feuern. Genauer gesagt bewegt sich die Oszillationsfrequenz in dem sogenannte g-Band, welches etwa von 40 bis 70 Hz reicht. Die in den einzelnen Neuronengruppen schon vorhandene Information wird so zu einer kohärenten Perzeption zusammengefaßt. Ein Objekt, für welches das Zusammenbinden neuronaler Aktivitäten gelungen ist, wird in das Arbeitsgedächtnis gelegt. Damit wird es automatisch für eine kurze Zeit erinnert, und somit – nach diesen Autoren – bewußt.

Die Hypothese von Crick und Koch kann sich auf eine Reihe von experimentellen Befunden stützen. Das Auftreten von 40 Hz Oszillationen als Funktion von besonderen Reizsituationen ist in verschiedenen kortikalen Regionen gut abgesichert¹¹. Auch zur Hypothese über das Arbeitsgedächtnis (working memory) gibt es neue Befunde. Insbesondere hat die Arbeitsgruppe von Patricia Goldman-Rakic¹² gezeigt, daß im präfrontalen Cortex

¹⁰ Es wird beobachtet, daß die Neurone in einem bestimmten neuronalen Schaltkreis (des präfrontalen Cortex) während einiger Sekunden nach dem Anstoß durch einen sensorischen Eingang eine hohe Feuerrate beibehalten.

¹¹ Hierzu gibt es inzwischen eine umfangreiche Liste von Veröffentlichungen. Ohne Anspruch auf Vorrang seien nur die beiden folgenden Arbeiten genannt: A. K. Engel, P. König, W. Singer, Bildung repräsentationaler Zustände im Gehirn. In: Gehirn und Bewußtsein, Heidelberg 1993, 42–46; G. Rager, W. Singer, The response of cat visual cortex to flicker stimuli of variable frequency, In: Europ. J. Neurosci. 10 (1998), 101–122.

¹² Hauptvortrag von P. Goldman-Rakic auf dem Annual Meeting of the American Society for Neuroscience in New Orleans 1997.

(Area 46) von Affen Neuronenpopulationen identifiziert werden können, welche Informationen aus den visuellen Areae erhalten und dann in der Größenordnung von einer oder mehreren Sekunden feuern. Dies spricht für ihre Beteiligung an den Prozessen des Arbeitsgedächtnisses. In einer neueren Arbeit kommen Crick und Koch¹³ auf Grund experimenteller Evidenz zu dem Schluß, daß das visuelle Gewahrsein (awareness) sich nicht direkt auf die neuronale Aktivität im primären visuellen Cortex (V1) bezieht, sondern auf hierarchisch höher gelegene corticale Areale, die ihrerseits in enger Beziehung zum präfrontalen und prämotorischen Cortex stehen. Ein funktionsfähiger primärer visueller Cortex (V1) bildet aber die Voraussetzung für bewußtes Sehen¹⁴. Problematisch wird es jedoch an jener Stelle, wo Crick und Koch den neurobiologischen Kontext verlassen und zu der allgemeinen Aussage vorstoßen: „Wir glauben, daß das Problem des Bewußtseins langfristig nur durch Erklärungen auf der neuronalen Ebene gelöst werden kann.“¹⁵ Damit wird zum Ausdruck gebracht, daß es sich um eine reduktive Erklärung handeln soll. Es wird jedoch nicht gesagt, was das Problem des Bewußtseins überhaupt ist, um das es hier geht¹⁶.

2.2 Die cortico-thalamischen Schleifen

Die Arbeiten aus der Gruppe von Llinás sind der zentralen Bedeutung der cortico-thalamischen Schleifen für das Bewußtsein gewidmet. Während sich Crick und Koch auf die visuelle Wahrnehmung beschränken, ist die Hypothese von Llinás¹⁷ allgemeiner und umfassender. Ein wesentliches Element dieser Hypothese ist wiederum der Befund der 40 Hz Oszillationen, welche beim temporären Zusammenbinden von Neuronengruppen eine wichtige Rolle spielen.

Es werden zwei cortico-thalamische Schleifen unterschieden, ein schon

¹³ F. Crick and C. Koch, Are we aware of neural activity in primary visual cortex? *Nature* 375 (1995) 121–123.

¹⁴ P. Stoerig, Varieties of vision: from blind responses to conscious recognition, in: *Trends Neurosci.* 19 (1996) 401–406.

¹⁵ F. Crick and C. Koch (1990) ebd. 263: „We believe that the problem of consciousness can, in the long run, be solved only by explanations at the neural level.“

¹⁶ Während Crick und Koch in ihrem Papier von 1990 sachlich und argumentativ vorgehen, verläßt F. Crick (1995) in seinem Buch „Was die Seele wirklich ist“ (amerikanische Originalausgabe „The astonishing hypothesis“ 1994) den Boden der wissenschaftlichen Argumentation und wird zum Prediger. In dem Kapitel „Dr. Cricks Wort zum Sonntag“ schreibt er, es sei „entscheidend, daß man bei dem Nachdenken über diese Fragen von den Neuronen ausgehe“ (ebd. 313). Nach einem Generalangriff auf Religionen und Philosophen postuliert er, daß alles menschliche Verhalten auf neuronale Aktivität zurückzuführen sei. Das sei eben gerade die „Erstaunliche Hypothese“.

¹⁷ R. Llinás and D. Paré, Commentary of dreaming and wakefulness. *Neuroscience* 44 (1991) 521–535; R. Llinás and U. Ribary, Coherent 40-Hz oscillation characterizes dream state in humans, In: *Proc. Natl. Acad. Sci. USA* 90 (1993) 2078–2081; D. Paré and R. Llinás, Conscious and pre-conscious processes as seen from the standpoint of sleep-waking cycle neurophysiology, in: *Neuropsychologia* 33 (1995) 1155–1168.

seit längerem bekannter Befund¹⁸. Neurone der spezifischen Kerne des Thalamus dorsalis, einer Kerngruppe im Zwischenhirn (Diencephalon), projizieren vorzugsweise in die Schicht IV des zugehörigen spezifischen (z. B. visuellen) Cortex. Kollateralen der Axone dieser Neurone gehen an den retikulären Kern des Thalamus. Neurone der Schicht VI desselben kortikalen Moduls projizieren wieder zurück in den Thalamus und schließen so die *erste cortico-thalamische Schleife*. Eine *zweite cortico-thalamische Schleife* entsteht dadurch, daß Neurone der unspezifischen, intralaminären Kerne des Thalamus diffus in die Schicht I des ganzen Cortex cerebri projizieren und gleichzeitig Kollateralen an den retikulären Kern des Thalamus abgeben. Neurone der Schicht V projizieren wieder zurück zu den Neuronen in den intralaminären Kernen und auch zum retikulären Kern, wodurch diese zweite Schleife geschlossen wird. Querverbindungen zwischen den beiden Schleifen sollen vor allem in Schicht V des Cortex erfolgen. In beiden Schleifen werden 40 Hz Oszillatoren beobachtet, die miteinander gekoppelt werden. In der ersten oder spezifischen Schleife werden sinnesspezifische Informationen (z. B. visuelle Informationen) verarbeitet. Die Kopplung mit der zweiten oder unspezifischen Schleife soll bewirken, daß diese sinnesspezifischen Informationen in den Gesamtzusammenhang der neuronalen Informationsverarbeitung gestellt werden. Die über die Oszillatoren gekoppelten Schleifen bilden die Organisation, welche dem Bewußtsein zugrunde liegt.

Diese gekoppelten Systeme verhalten sich unterschiedlich im Wachen, Schlafen und Träumen. Im Tiefschlaf sollen die Schleifen entkoppelt sein; Signale aus der Umwelt haben dann praktisch keinen Einfluß auf die Informationsverarbeitung im Gehirn. Beim Träumen sind die Schleifen wahrscheinlich gekoppelt, reagieren aber kaum auf Signale aus der Umwelt, sondern arbeiten vorwiegend mit intern hervorgebrachten Wahrnehmungen. Im Wachzustand sind sie gekoppelt und reagieren auf externe Reize.

2.3 Neurobiologie der Qualia

Wie wir schon bei der Klärung der Begriffe gesehen haben, ist das phänomenale Bewußtsein charakterisiert durch die Empfindungen, die man als Qualia bezeichnet. Nach Ansicht von Thomas Nagel¹⁹, David Chalmers²⁰ und anderen verleihen die Qualia den Empfindungen einen qualitativen Charakter, der nicht auf neurobiologische Prozesse reduziert werden kann. Ramachandran und Hirstein (1997) versuchen nun, gerade diese Qualia, und nicht nur die Qualia, sondern auch das Selbst neurobiologisch zu erklä-

¹⁸ O. Creutzfeldt, *Cortex cerebri. Leistung, strukturelle und funktionelle Organisation der Hirnrinde*, Berlin 1983.

¹⁹ T. Nagel, *What is it like to be a bat?* PhRev 83 (1974): 435–450.

²⁰ D. J. Chalmers, *Facing up to the problem of consciousness*. J. Consc.Studies 2 (1995), 200–219; D. J. Chalmers, 1996.

ren. Sie stützen sich auf Untersuchungen an neurologischen Patienten und auf Phänomene der Wahrnehmungspsychologie. Es geht ihnen darum, „das Problem des Bewußtseins zu lösen“²¹. Wie man aus dem Kontext ersehen kann, soll eine reduktive Erklärung die Lösung des Problems sein.

Nach Ramachandran und Hirstein sind die Prozesse der bewußten Erfahrung und die Qualia vorwiegend im Temporallappen, im Gyrus cinguli, in der Sprachregion von Wernicke und in anderen damit assoziierten Strukturen lokalisiert. Sie stellen ein Bindeglied zwischen Perzeption und Aktion dar²².

Dem Einwand, daß Qualia gerade dadurch charakterisiert seien, daß sie privat sind und nicht auf neurobiologische Vorgänge reduziert werden können, begegnen Ramachandran und Hirstein mit einem Argument aus dem Bereich des Farbensehens. Wie könnte ein Farbenblinder verstehen, was eine mit normalem Farbsehen befähigte Person erlebt, wenn sie Farben wahrnimmt? Solange man versucht, die Farbempfindung der farbenblinden Person über das Medium der Sprache zu vermitteln, wird die Farbempfindung privat bleiben. Die farbenblinde Person wird nicht nachvollziehen können, wie Farbe sich „anföhlt“. Ramachandran und Hirstein schlagen deshalb vor, ein Kabel herzustellen, das aus Neuronen der für die Wahrnehmung der Farben zuständigen Area V4 besteht. Mit diesem Neuronen-Kabel solle sowohl die sprachliche Vermittlung als auch der ganze Weg der Übertragung von der Retina bis zum Cortex umgangen werden. Die V4 des Farbtüchtigen könnte auf diese Weise direkt mit der V4 des Farbblinden verbunden werden. Mit dieser Verbindung würde der Farbenblinde vielleicht verstehen, was der Farbentüchtige meint, wenn er eine Farbe als rot empfindet. Damit wäre die von Philosophen für unüberwindbar angesehene Schranke eingerissen und eine freie Passage für das Verständnis des Quale Farbe zwischen verschiedenen Personen ermöglicht.

Wenn so die Qualia nicht mehr einem privaten Selbst angehören, sondern mitteilbar werden, dann wird auch eine Erklärung des Selbst möglich, weil das Selbst eng verknüpft ist mit diesen Qualia. „Die Vorstellung eines einheitlichen, dauerhaften Selbst dürfte sich als eine Form von adaptiver Selbsttäuschung herausstellen.“²³ Die Illusion über das Selbst entstehe dadurch, daß man meint, die Qualia seien für jemanden da, der dann als Homunculus oder als Selbst vorgestellt wird. Sie sind aber nicht für jemand,

²¹ V. S. Ramachandran and W. Hirstein, Three laws of qualia: What neurology tells us about the biological functions of consciousness, *J. Consc.Studies* 4 (1997) 429–457, besonders 429.

²² V. S. Ramachandran und W. Hirstein formulieren drei „Gesetze“ für die Qualia, die in ihrer Bedeutung in die Nähe der drei Gesetze der klassischen Mechanik von Newton gerückt werden. Qualia sind erstens unwiderruflich (irrevocable), d. h. ich kann nicht einfach entscheiden, einen grünen Sonnenuntergang zu sehen. Sie haben zweitens eine potentiell unendliche Menge von möglichen Verhaltensweisen zur Folge („flexibility on the output side“, ebd. 438) und benötigen drittens das Kurzzeitgedächtnis, d. h. in Zuständen der Unbewußtheit gibt es keine Qualia.

²³ Ebd. 454: „Even though the notion of a unitary, enduring self may turn out to be a form of adaptive self-deception or delusion ... we must consider why the illusion arises.“

sondern für etwas da, nämlich für einen Kontrollprozeß. „Ein Kontrollprozeß ist nicht etwas, was alle Eigenschaften eines vollen menschlichen Wesens hat ... er ist kein Homunculus. ... Er wird angeführt durch bestimmte Hirnregionen (d. h. Regionen der Perzeption und der Motivation), welche die Aktivitäten von anderen Regionen kontrollieren (d. h. Regionen der Motorik und der Planung)“²⁴. Von daher sei es nicht schwierig zu sehen, wie solche Kontrollprozesse „Anlaß geben können zu der Mythologie eines Selbst als einer aktiven Gegenwart im Gehirn – einem ‚Geist in der Maschine‘“²⁵.

2.4 Die biologische Theorie des Bewußtseins von Edelman

Eine umfassende biologische Theorie des Bewußtseins versucht Gerald Edelman²⁶ vorzulegen. Er weist die funktionalistische Vorstellung vom Gehirn als einer Turing Maschine zurück²⁷. Mit Putnam²⁸ hält er fest, daß psychologische Zustände nicht durch ein Computermodell beschrieben werden können. Begriffe und Überzeugungen können nicht individuiert werden ohne Bezug auf die Umwelt²⁹. Mit Searle³⁰ argumentiert er, daß Algorithmen keine hinreichenden Bedingungen für das Denken und für intentionale Zustände liefern können. Computerprogramme seien strikt durch ihre formale syntaktische Struktur definiert; die Syntax aber sei unzureichend, um die Semantik hervorzubringen. Das Mentale sei aber gerade durch seine semantischen Inhalte charakterisiert. Kein Organismus könne intentionale Zustände haben ohne eine subjektive Erfahrung³¹.

2.4.1 Die Basistheorie

Während der Funktionalismus als Theorie des Bewußtseins versagt habe, soll nach Edelman die Theorie der Selektion von neuronalen Gruppen („Theory of Neuronal Group Selection (TNGS)“³²) zureichend sein, um Bewußtsein biologisch zu erklären. Edelman hat diese Theorie bereits

²⁴ Ebd. 455: „A control process is not something which has all the properties of a full human being, of course – it is not at all a homunculus. All the notion of a control process entails, as we are employing it, is that control processes are guided by some brain areas (i. e. perceptual areas and motivational areas) as they control the activities of other brain areas (i. e. motor and planning areas).“

²⁵ Ebd. 455: „It is not difficult to see how such processes could give rise to the mythology of a self as an active presence in the brain – a ‚ghost in the machine‘.“

²⁶ G. M. Edelman, *The remembered present. A biological theory of consciousness*, New York 1989.

²⁷ Ebd. 28, 258.

²⁸ H. Putnam, *Repräsentation und Realität*, Frankfurt 1991. Die englische Originalausgabe „Representation and Reality“ ist 1988 erschienen. Wir zitieren jedoch nach der deutschen Ausgabe von 1991.

²⁹ G. M. Edelman, 1989, 29, 42.

³⁰ J. R. Searle, *Minds, brains and science*, Cambridge, Mass. 1984.

³¹ G. M. Edelman, 1989, 29, 42.

³² Ebd. 10.

1978³³ grundgelegt, 1987 hat er sie in ihren phylogenetischen³⁴ und 1988 in ihren ontogenetischen³⁵ Aspekten weiterentwickelt. „The Remembered Present“³⁶ faßt die schon vorhandenen Elemente zusammen und liefert eine erweiterte biologische Theorie des Bewußtseins („extended TNGS“). Die TNGS baut auf drei Mechanismen auf: Selektion während der Entwicklung, Selektion durch Erfahrung und die für das Gehirn so charakteristische Struktur der reziproken Verbindungen zwischen verschiedenen Hirngebieten („reentry“, „reentrant mapping“). Es gilt als gesichert, daß weder die Zahl noch die Variabilität der neuronalen Strukturen durch den genetischen Code festgelegt werden. Die strukturelle Mannigfaltigkeit, die sich während der Entwicklung ergibt, entsteht durch eine epigenetische Regulation von Zellteilung, Zellwanderung, Zelltod, Ausbildung von Zellfortsätzen und neuronaler Aktivität (Selektion durch Entwicklung, *erster Mechanismus*). Insbesondere spielen hierfür die Kontrollschleifen eine wichtige Rolle, welche die Expression von Zelladhäsionsmolekülen (CAM) und Substratadhäsionsmolekülen (SAM) regulieren. Es entsteht so eine Mannigfaltigkeit von neuronalen Gruppen und lokalen Schaltkreisen, welche ein primäres Repertoire für die weitere Differenzierung darstellen.

Auf der Selektion durch Entwicklung baut die Selektion durch Erfahrung auf (*zweiter Mechanismus*). Erfahrung und Verhalten führen zur Abschwächung oder Verstärkung synaptischer Verbindungen und damit zur Variation der Schaltkreise selbst. Es entsteht ein sekundäres Repertoire von neuronalen Gruppen. Der *dritte Mechanismus* arbeitet mit reziproken Abbildungen zwischen Hirngebieten („reentrant mapping“). Die Signale aus unserer Umwelt treffen auf Sinnesoberflächen (visuelle, akustische, somatosensorische etc.). Die Sinnesoberflächen werden wiederum so auf bestimmte Hirnregionen projiziert, daß dort eine topographisch organisierte Abbildung („map“) entsteht³⁷ (Beispiel: Projektion der Retina auf den primären visuellen Cortex). Diese Abbildung steht wieder in reziproker Verbindung mit anderen Abbildungen. Unter „reentry“ versteht Edelman nicht bloß eine Rückkopplungsschleife³⁸, sondern den wiederholten Austausch von Signalen zwischen zwei verschiedenen reziproken Abbildungen³⁹ („maps“). Beispiele solcher reziproker Verbindungen zwischen Abbildungen dieser Art sind in erster Linie die kortiko-kortikalen Verbindungen,

³³ G. M. Edelman and V. B. Mountcastle, *The mindful brain: Cortical organization and the group selective theory of higher brain function*, Cambridge 1978.

³⁴ G. M. Edelman, *Neural darwinism: The theory of neuronal group selection*, New York (1987).

³⁵ G. M. Edelman, *Topobiology: An introduction to molecular embryology*, New York 1988.

³⁶ G. M. Edelman, 1989.

³⁷ Sofern solche Abbildungen mit einer Sinnesoberfläche verbunden sind, muß wenigstens eine von ihnen topographisch organisiert sein (ebd. 68).

³⁸ Ebd. 48.

³⁹ „Exchange of signals between different repertoires in a parallel and recursive manner“ (ebd. 243).

aber auch die kortiko-thalamischen Schleifen, wie wir sie weiter oben besprochen haben, gehören dazu. Der Mechanismus des wiederholten Austauschs von Signalen („reentry“) bewirkt, daß Neuronengruppen in verschiedenen Hirngebieten, die unterschiedliche Eingänge („inputs“) erhalten, ausgewählt und in ihrer Aktivität korreliert werden. Dies wiederum liefert die Grundlage für die Kategorisierung der Eingänge auf der Ebene der Perzeption.

Wenn zwei verschiedene Abbildungen wechselseitig auf einander bezogen und wenn zwischen ihnen wiederholt Signale ausgetauscht werden, dann handelt es sich noch um ein lokales Geschehen, eine lokale wechselseitige Abbildung („reentrant local maps“). Die lokalen Abbildungen sind wiederum eingebettet in ein globales Abbildungssystem. Das globale Abbildungssystem („global mapping“) umfaßt multiple lokale Abbildungen, und zwar sowohl auf der sensorischen wie auch auf der motorischen Seite. „Das globale Abbildungssystem wird erstellt durch eine Serie lokaler Abbildungen, die reziprok miteinander verbunden sind und ständig in Gang gehalten werden durch motorisches und sensorisches Erproben der Umwelt im Verhalten und durch beständige synaptische Selektion ihrer neuronalen Gruppen.“⁴⁰ Daß die Perzeption wesentlich auch von kontinuierlicher motorischer Aktivität abhängt, wird besonders deutlich beim Sehen: um Objekte wahrzunehmen, sind beständige Augenbewegungen nötig, mit welchen wir diese Objekte nach ihren bestimmenden Elementen abrastern und erfassen.

Das Erproben der Umwelt über verschiedene Sinneskanäle führt zu Klassifikationen in zwei oder mehreren verschiedenen (n-tupel) Sinnesmodalitäten, die wiederum die Grundlagen der „Perzeptionskategorien“ („perceptual categorization“) liefern⁴¹. Die Perzeptionskategorien sind jedoch mehrdeutig: mehrere verschiedene Kombinationen von neuronalen Gruppen können zu demselben Ergebnis („output“) führen⁴². Das Geschehen hat wesentlich eine Zeitstruktur, welche im *Gedächtnis* festgehalten wird. Im Gedächtnis werden die gebildeten Kategorien wieder umgeformt („recategorization“) in Abhängigkeit von den synaptischen Änderungen, die sich in den lokalen Abbildungen ergeben⁴³.

2.4.2 Die erweiterte Theorie

Die Selektion neuronaler Gruppen in den reziproken Abbildungen des Gehirns ist der eine Vorläufer des Bewußtseins; der andere Vorläufer ist das

⁴⁰ „A series of such local maps reentrantly connected to each other and driven continuously by motor and sensory sampling of the environment through behavior and undergoing continuous synaptic selection of its neuronal groups constitutes a global mapping“ (ebd. 243).

⁴¹ Ebd. 243. Philosophisch würde man hier eher von Anschauungsformen als von Kategorien sprechen.

⁴² G. M. Edelman, 1989, bezeichnet diese mehrdeutige Beziehung als „degeneracy“ oder Entartung. „The neuronal group selection ... is degenerate“ (ebd. 243). Weitere Ausführungen über die „degeneracy“ siehe ebd. 50.

⁴³ Vgl. ebd. Kap. 6, 244.

Auftauchen von zwei sehr verschiedenen neuronalen Strukturen und Funktionen während der Evolution. Die erste dieser beiden Strukturen ist mit der interozeptiv wahrgenommenen homöostatischen Regulation, die andere mit der exterozeptiven Wahrnehmung der Umwelt befaßt⁴⁴. Die interozeptiven Signale werden kategorisiert und verglichen mit den Kategorisierungen der exterozeptiven Signale, die in der schnell wechselnden Auseinandersetzung mit der Umwelt aufgenommen werden. Dieser Vergleich von inneren und äußeren Signalen führt bereits zur Unterscheidung von Selbst und Nicht-Selbst in einem elementaren, vorindividuellen Sinn. Bewußtsein ist schließlich für Edelman das „Ergebnis eines rekursiv vergleichenden Gedächtnisses, in welchem vorausgegangene Kategorisierungen von Selbst und Nicht-Selbst beständig bezogen werden auf aktuelle gegenwärtige Kategorisierungen“⁴⁵.

Edelman unterscheidet *zwei Formen des Bewußtseins*, ein primäres („primary“) und ein höheres („higher-order“) Bewußtsein. Das *primäre Bewußtsein* ergibt sich aus dem Gedächtnis von Selbst und Nicht-Selbst. Auf der Basis dieses Gedächtnisses werden fortlaufend perzeptuelle Kategorisierungen vorgenommen und unterschieden. Die aktuellen Perzeptionen werden kontinuierlich zu den bereits im Gedächtnis vorhanden hinzugefügt⁴⁶. Zugleich werden Beziehungen hergestellt zu den biologischen Wertsetzungen, die sich aus den selbstregulatorischen Funktionen wie Homöostase, Sättigung, Anpassung, Übereinstimmung mit den Lustzentren („hedonic centers“), Erfüllung der Antriebe und Befriedigtsein ergeben. Tiere mit primärem Bewußtsein sind aber mit ihrem Kurzzeitgedächtnis gebunden an kleine Zeitintervalle und haben keinen Begriff von der Vergangenheit.

Das *höhere Bewußtsein* taucht erst mit der Entwicklung der Sprache auf. Die Sprache ermöglicht eine symbolische Kategorienbildung. Damit werden Begriffe von der Vergangenheit und von der Zukunft möglich. Zugleich mit der Entwicklung der Vorstellungen von Vergangenheit und Zukunft können zeitunabhängige Modelle des Selbst und der Welt entstehen⁴⁷. Es bildet sich das konzeptuelle Selbst, welches das elementare, bloß biologische Selbst überschreitet⁴⁸ und sich aus der Gebundenheit in die Gegenwart befreien kann. Hatte sich das elementare, biologische Selbst auf Grund seiner inneren Zustände vom Nicht-Selbst der Umgebung unterschieden, so verfügt das konzeptuelle Selbst, welches über die Sprache mit anderen Individuen kommunizieren kann, über einen Begriff von seinem Selbstsein und

⁴⁴ Ebd. 151.

⁴⁵ Ebd. 155: „... consciousness is an outcome of a recursively comparative memory in which previous self-nonself categorizations are continually related to ongoing present perceptual categorizations ...“

⁴⁶ Ebd. 102.

⁴⁷ Ebd. 246.

⁴⁸ Ebd. 187.

grenzt sich als subjektive Identität von den Artgenossen ab⁴⁹. Höheres Bewußtsein ist gleichbedeutend mit direktem Gewahrsein („direct awareness“⁵⁰) und Bewußtsein von Bewußtsein („consciousness of consciousness“⁵¹). Höheres Bewußtsein ist auch die Voraussetzung für die Planung von Handlungen⁵². Es kommt wahrscheinlich ansatzweise schon bei Schimpansen vor. Richtig ausgebildet ist es aber erst beim Menschen. Für die Entstehung des höheren Bewußtseins spielt sowohl die Sprache als auch das soziale Gefüge eine wesentliche Rolle⁵³.

Edelman benennt eine Reihe von Strukturen und Funktionskreisen im Gehirn, von denen man mit einigen Gründen annehmen kann, daß sie in die Prozesse des primären Bewußtseins eingebunden sind. Die neuronalen Strukturen für das höhere Bewußtsein anzugeben, bezeichnet er jedoch selbst als spekulativ. Sicher gehören hierher die Sprachzentren von Broca und Wernicke. Weitere wichtige Kandidaten sind die langen kortikalen Assoziationsfasersysteme, der mediale Kern des Pulvinar (Thalamus dorsalis, Diencephalon), welcher reziprok mit dem präfrontalen, cingulären und parietalen Cortex in Verbindung steht und deren Aktivitäten korrelieren kann, und schließlich die kortiko-thalamische Schleife, über deren Bedeutung wir uns weiter oben schon Rechenschaft abgelegt hatten⁵⁴.

2.4.3 Die Bedeutung von „Erklärung“

Die erweiterte TNGS versteht sich als biologische Theorie des primären und des höheren Bewußtseins. Sie versucht, sich auf Strukturen und Zuständen des Gehirns abzustützen⁵⁵ und „das Bewußtsein mit den materiellen Strukturen und Funktionen des Gehirns zu erklären“⁵⁶. Die Theorie soll einen globalen Charakter haben und mit Physik und Evolutionstheorie kompatibel sein⁵⁷. Dies scheint wiederum nur dann möglich, wenn man metaphysisch einen – obgleich subtilen – materialistischen Standpunkt einnimmt⁵⁸. So macht es auf den ersten Blick den Anschein, als sei diese Theorie nur eine Variante der streng reduktionistischen Erklärungen, welche die Identität von Hirnzuständen und Bewußtsein behaupten. Bei genauerem Zusehen erweist sich die erweiterte TNGS jedoch als nicht nur sehr breit angelegt, sondern auch als sehr differenziert. Die bewußte oder phänomenale Erfahrung wird als eigene Realität anerkannt⁵⁹ und ein extremer Re-

⁴⁹ Ebd. 247.

⁵⁰ Ebd. 104.

⁵¹ Ebd. 194, 263.

⁵² Ebd. 192.

⁵³ Ebd. 187. Siehe auch den Hinweis auf die Sprachgemeinschaft, ebd.247.

⁵⁴ Ebd. 163.

⁵⁵ Ebd. 249.

⁵⁶ „... to explain consciousness in terms of the material structures and functions of the brain“ (ebd. 273, ähnlich ebd. 8–14).

⁵⁷ Ebd. 239.

⁵⁸ Ebd. 264.

⁵⁹ Ebd. 15, 20–24, 167–168.

duktionismus als „leer“ („empty“⁶⁰) bezeichnet. Während Ramachandran und Hirstein⁶¹ meinen, auch die Qualia reduktionistisch erklären zu können, verwirft Edelman diesen Versuch. Er stellt ihn auf die gleiche Ebene wie den Versuch, wissenschaftlich zu erklären, warum Materie existiert⁶². „Als Wissenschaftler können wir nicht betroffen sein von ontologischen Mysterien, bei denen es um die Frage geht, warum etwas ist und nicht vielmehr nichts oder warum warm sich ‚warm‘ anfühlt.“⁶³

Wägt man nun die verschiedenen Aussagen zur biologischen Theorie des Bewußtseins (erweiterte TNGS) sorgfältig ab, kann man sie nicht als Reduktionismus im strengen Sinn bezeichnen. Eher geht es um eine Korrelation von mentalen und neuronalen Zuständen⁶⁴. Die Forschung beginnt mit einem Alltagsverständnis („common sense“⁶⁵) oder einer philosophischen Sicht und versucht, dieses Wissen auf Vorgänge im Gehirn zu beziehen und damit zu verifizieren⁶⁶. Obwohl die Wissenschaft die größte Errungenschaft der Menschheit ist, bleibt sie „nur eine partielle Erfahrung des Bewußtseins“ („partial experience of the consciousness“⁶⁷). Die „subjektive persönliche Erfahrung“ („subjective personal experience“⁶⁸) behält den Vorrang.

3. Kritische Prüfung der neurobiologischen Erklärungen

Bevor wir in die kritische Prüfung der vorgestellten neurobiologischen Theorien und Erklärungen eintreten, sei bemerkt, daß nur wenige Neurobiologen solchen Extrempositionen zustimmen, wie sie vor allem von Ramachandran und Hirstein⁶⁹ vertreten werden. Viele Neurobiologen meinen, die ganze Debatte um die Neurobiologie des Bewußtseins sei verfrüht. Andere sind skeptisch, ob diese Debatte überhaupt zu sinnvollen Resultaten führen kann. Nur relativ wenige Veröffentlichungen liefern einen substantiellen Beitrag zur neuronalen Deutung des Bewußtseins. Das Werk von Edelman hingegen markiert einen wichtigen Schritt auf dem Weg zu einem besseren Verständnis der neuronalen Korrelate des Bewußtseins. Im folgenden sei nun versucht, die vorgelegten Modelle und Theorien kritisch zu analysieren und zu beurteilen.

⁶⁰ Ebd. 254.

⁶¹ V. S. Ramachandran and W. Hirstein, 1997.

⁶² G. M. Edelman, 1989, 167–168.

⁶³ Ebd. 168: „As scientists, we can have no concern with ontological mysteries concerned with why there is something and not nothing, or why warm feels ‚warm‘“. Die Beurteilung des Buches durch D. J. Chalmers, 1996, wird der Sache nicht gerecht.

⁶⁴ Ebd. 260.

⁶⁵ Ebd. 268.

⁶⁶ Ebd. 268.

⁶⁷ Ebd. 270.

⁶⁸ Ebd. 270.

⁶⁹ V. S. Ramachandran and W. Hirstein, 1997.

3.1 Neuronale Korrelate des Bewußtseins

Die Befunde, auf welche in den Theorien und Modellen Bezug genommen wird, sind sicher keine Einzelphänomene. Man wird vermutlich die neurobiologischen Aussagen in dem Sinne verallgemeinern dürfen, daß immer dann, wenn visuelle Wahrnehmung erfolgt, Gruppen von Neuronen in verschiedenen Hirnregionen in Oszillationen synchronisiert werden und vorzugsweise mit einer Frequenz von 40–70 Hz (g-Band) feuern. Es gibt schon gute Evidenz für Neurone im präfrontalen Cortex, die in Abhängigkeit vom sensorischen Eingang in einer Weise aktiv sind, welche den Annahmen über das Arbeitsgedächtnis entspricht. Auch für die Aussagen über die doppelte cortico-thalamische Schleife gibt es experimentell schon gute Evidenz. Während früher das ganze Gehirn als Sitz des Bewußtseins betrachtet wurde, werden jetzt immer kleinere und spezifischere Bereiche dafür verantwortlich gemacht. Es ist möglich, daß vor allem der Temporalappen und das limbische System unter Einschluß der Amygdala eine wichtige Rolle bei bewußten Akten spielen und beim Erleben von Farben oder anderen Qualia besonders involviert sind. Auch wenn wir uns klar darüber sind, daß viele Elemente der vorgebrachten Theorien experimentell noch nicht belegt sind, so spricht einiges dafür, daß die Vorgänge im Gehirn so sein könnten, wie sie in den Theorien und Modellen postuliert werden. Trotz der bestehenden Unsicherheiten könnten wir wahrscheinlich so weit gehen zu behaupten, daß die neuronalen Prozesse mit Bewußtseinsakten *korreliert* sind. Wir könnten sogar zu der These vordringen, daß alle unsere mentalen Akte zugleich neuronale Prozesse sind⁷⁰. Wir behaupten damit eine Korrelation von neuronalen Prozessen mit den Akten des Bewußtseins. Wenn dem so ist, dann liefern uns die Neurowissenschaften wichtige Befunde, Modelle und Theorien über diese Korrelate.

3.2 Faktisch notwendige Bedingungen

Die genannten neuronalen Prozesse sind nicht nur Korrelate, sondern auch faktisch notwendige Bedingungen des Bewußtseins. Wenn diese Prozesse gestört sind, dann sind die bewußten Akte nicht mehr möglich. Mit den faktisch notwendigen Bedingung ist folgendes gemeint: Bei der gegebenen Struktur und Funktion des Nervensystems erscheint es als notwendig, daß die Aktivitäten von Neuronengruppen durch 40 Hz Oszillationen miteinander gekoppelt und synchronisiert werden, daß Elemente des Arbeits-

⁷⁰ G. Rager, Das Leib-Seele-Problem. Begegnung von Hirnforschung und Philosophie. FZPhTh (1982), 443–464; G. Rager, Der Mensch als er selbst und als Objekt der Wissenschaft, Universitas Friburgensis 3 (1993), 7–9; G. Rager, Menschsein zwischen Lebensanfang und Lebensende. Grundzüge einer medizinischen Anthropologie, in: L. Homfeldler und G. Rager (Hrsg.): Ärztliches Urteilen und Handeln. Zur Grundlegung einer medizinischen Ethik, Frankfurt 1994, 53–103; G. Rager, Bewußtsein und Hirnforschung: Befunde und Theorien, in: J. Quitterer und E. Runggaldier (Hrsg.): Der neue Naturalismus, Stuttgart 1999, im Druck.

gedächtnisses im präfrontalen Cortex realisiert sind oder daß das Gehirn über zahlreiche Rückkopplungsschleifen verfügt. Besonders eindrücklich sind hier die Evidenzen, die sich aus der Ontogenese ergeben. Solange bestimmte Strukturen des Gehirns noch nicht entwickelt sind, sind auch die entsprechenden Bewußtseinsakte nicht möglich.

Was mit faktisch notwendigen Bedingungen gemeint ist, läßt sich mit dem Hempel-Oppenheim Schema wie folgt ausdrücken. V sei die gegebene Struktur und Funktion des Nervensystems. Für V sollen die Gesetze G_0 und die Antezedensbedingungen A_0 gelten. Unter der Voraussetzung von V wird nun das Explanandum E durch das Explanans G_1, \dots, A_n *bedingt* erklärt

$$\begin{array}{l} V \\ G_1, \dots, G_m \\ \hline A_1, \dots, A_n \\ E \end{array} \qquad \text{also:} \qquad \begin{array}{l} G_0, G_1, \dots, G_m \\ \hline A_0, A_1, \dots, A_n \\ E \end{array}$$

Faktisch notwendige Bedingungen sind in diesem Sinne notwendige Bedingungen, aber unter einer bestimmten Voraussetzung. Sie können nicht als hinreichende Bedingungen gelten ohne Berücksichtigung der Anwendbarkeit auf Bereiche, in denen die Voraussetzung erfüllt ist.

3.3 Logisch notwendige Bedingungen

Die neuronalen Prozesse sind zwar faktisch notwendig, sie sind aber nicht logisch notwendig. Aus der Struktur des Bewußtseins folgt nicht, daß es nur in der Weise neuronal verwirklicht werden kann, in der es tatsächlich verwirklicht ist⁷¹. Man kann sich auch andere Strategien und Systeme vorstellen, welche die gleiche Funktion hätten. Solche Argumente sind schon von Putnam in der Funktionalismusdebatte vorgebracht und mit Beispielen aus der Computerwelt erläutert worden. Man kann bestimmte Funktionen auf verschiedene Weisen physisch realisieren. Die Computer können verschieden sein in ihrer Hardwarearchitektur. Man kann auf diesen Computern unterschiedliche Algorithmen für die gleiche Funktion einsetzen und diese Algorithmen in verschiedenen Sprachen programmieren. Ein gegebener mentaler Zustand kann also physisch multipel realisiert werden⁷².

3.4 Hinreichende Bedingungen

Bleibt noch die Frage, ob die neuronalen Prozesse hinreichend sind für die Erklärung des Bewußtseins. Bereits Levine⁷³ hat am Beispiel der Schmerz-

⁷¹ Vgl. auch G. M. Edelman, 1989, 260.

⁷² Siehe auch die Diskussion dieser Frage bei G. Brüntrup, *Das Leib-Seele-Problem. Eine Einführung*, Stuttgart 1996.

⁷³ J. Levine, *Materialism and qualia: The explanatory gap*. *Pacific Philosophical Quarterly* 64 (1983) 354–361.

empfindung gezeigt, daß man die neuronalen Prozesse nicht als hinreichende Bedingungen bezeichnen kann. Man weiß, daß bei einer Form der Schmerzempfindung C-Fasern Signale vom Ort der Schmerzauslösung zum Zentralnervensystem weiterleiten. Es wird aber damit keine Antwort auf die Frage gegeben, warum der Schmerz sich so anfühlt, wie er dies tut. Das Gleiche gilt für alle anderen Qualitäten des Bewußtseins. Aus neuronalen Prozessen oder aus mehr oder weniger umfassenden neuronalen Modellen läßt sich nicht herleiten, wie daraus die bewußte Erfahrung hervorgeht⁷⁴. Das ist genau die heute viel diskutierte Erklärungslücke (explanatory gap), die weiterhin offen ist und aus grundsätzlichen Überlegungen offen bleiben wird.

Da weder logisch notwendige noch hinreichende neurobiologische Bedingungen für Bewußtsein aufgezeigt werden, wird der Anspruch der Autoren, das Bewußtsein reduktiv zu erklären, nicht eingelöst. Die beschriebenen neuronalen Prozesse und Modelle ergeben wertvolle Korrelationen der Bewußtseinsakte, liefern aber keine hinreichenden Bedingungen für diese. Eine Äquivalenz zwischen neuronalen Prozessen und Bewußtsein läßt sich deshalb nicht herstellen.

3.5 Grundsätzliche Grenzen der Reduktion und der Formalisierung

Der Versuch, Bewußtsein mit neuronalen Prozessen zu identifizieren, stößt auf grundsätzliche Grenzen. Dies gilt gleichermaßen für neurobiologische wie für funktionalistische Ansätze. Um wirklich das Mentale auf das Funktionale und das Funktionale auf das Neuronale reduzieren zu können, müßten wir – so Putnam – über einen „General-Algorithmus“ der ganzen menschlichen Rationalität verfügen. Wir müßten Gott sein, um „die Gesamtheit der möglichen Denk- und Begriffsbildungsweisen des Menschen“⁷⁵ überblicken zu können. „Wollte man einen Menschen in einer zeitgebundenen menschlichen Kultur dazu auffordern, einen Überblick zu geben über alle Weisen der sprachlichen Existenz des Menschen ..., so wäre

⁷⁴ Vgl. D. J. Chalmers, 1996, 117. Selbst die den naturalistischen Standpunkt vertretende P. Churchland spricht dem Modell mit den cortico-thalamischen Schleifen ab, daß es die hinreichenden Bedingungen aufzeigt (P. Churchland, Die Neurobiologie des Bewußtseins. Was können wir von ihr lernen? In: T. Metzinger (Hrsg.): Bewußtsein. Beiträge aus der Gegenwartsphilosophie, Paderborn, 1996: 463–490, besonders 486). Nach Prinz (W. Prinz, Bewußtsein und Ich-Konstitution, in: G. Roth and W. Prinz (Hrsg.): Kopf-Arbeit. Gehirnfunktionen und kognitive Leistungen, Heidelberg 1996, 451–467) kann die Neurobiologie keine hinreichenden Erklärungen geben. Vgl. auch Bieri (P. Bieri, Was macht Bewußtsein zu einem Rätsel? In: Gehirn und Bewußtsein. Heidelberg 1994, 172–180), Beckermann (A. Beckermann, Können mentale Phänomene neurobiologisch erklärt werden? In: G. Roth and W. Prinz (Hrsg.): Kopf-Arbeit. Gehirnfunktionen und kognitive Leistungen, Heidelberg 1996, 413–425) und D. J. Chalmers, 1996, welche zeigen, daß subjektive Erfahrungen nicht naturalistisch erklärt werden können.

⁷⁵ H. Putnam, 1991, 164; ebenso: H. Putnam, Why functionalism didn't work. In J. Earman (Hrsg.): Inference, explanation and other frustrations. Essays in the philosophy of science, Berkeley 1992, 255–270, besonders 268.

das gleichbedeutend mit der Suche nach einem unmöglichen archimedischen Punkt.“⁷⁶

Schließlich argumentiert Putnam gegen die funktionalistische Formalisierung und greift dabei auf das Gödelsche Theorem⁷⁷ zurück: „Was Gödel gezeigt hat, ist sozusagen dies: daß wir unsere eigene mathematische Fähigkeit nicht vollständig formalisieren können, weil es mit zu eben dieser mathematischen Fähigkeit gehört, daß sie über alles, was sie zu formalisieren vermag, hinausgehen kann.“⁷⁸ Was hier von der mathematischen Fähigkeit gesagt wurde, gilt für die Tätigkeit der Vernunft überhaupt. Die Vernunft kann über alles hinausgehen, was sie formalisieren kann⁷⁹; sie kann alles transzendieren, was sie überblicken kann⁸⁰. Die Unmöglichkeit der vollständigen Formalisierung ist die eine Grenze für das Programm von Funktionalismus und Reduktionismus, die Intentionalität ist die andere Grenze. Es ist gezeigt worden, daß es nicht möglich ist, mentale oder intentionale Zustände „so auf funktionale Zusammenhänge zurückzuführen, daß für ihre Beschreibung keine intentionalen Ausdrücke verwendet werden“⁸¹. Die Intentionalität selbst aber kann nicht reduktionistisch erklärt werden⁸².

4. Rückbesinnung auf unseren Entwurf von Wissenschaft

Auf der einen Seite stehen also neurobiologische Fakten und Modelle, welche mit Bewußtseinsprozessen korrelieren, auf der anderen Seite befinden sich die Basiserfahrungen von Bewußtsein, Qualia, Freiheit und Verantwortung. Es bleibt also bei dem Gegensatz zwischen neurobiologisch beschreibbaren Vorgängen im Gehirn und subjektiver Erfahrung. Die menschliche Vernunft will sich jedoch mit diesem Gegensatz nicht dauerhaft abfinden, sondern sucht eine letzte Einheit, in welcher diese Gegensätze aufgehoben sind. Wo liegt nun ein möglicher Einigungsgrund, der

⁷⁶ H. Putnam, 1991, 164.

⁷⁷ Das Gödel-Theorem oder der Unvollständigkeitssatz (*K. Gödel, Über formal unentscheidbare Sätze der Principia mathematica und verwandter Systeme I, Mh. Math. Phys.* 38 (1931), 173–198) zeigt die Undurchführbarkeit des Hilbert-Programms: Jedes für die Darstellung der elementaren Zahlentheorie ausreichende und zugleich widerspruchsfreie formale System ist unvollständig und seine Widerspruchsfreiheit kann nicht mit den in ihm formalisierten Mitteln bewiesen werden. Damit wird die Hoffnung auf eine rein axiomatisch fundierte Mathematik zerstört, die alle mathematischen Probleme lösen kann. Die Mathematik läßt sich also nicht aus sich selbst heraus begründen. Über die Gesetze der Logik läßt nicht zugleich vollständig und korrekt verfügen. In der ‚Philosophy of Mind‘ vertritt Gödel (*K. Gödel, Basic theorems on the foundations of mathematics and their implications*, in: S. Feferman et al. (Hrsg.): *Gödel, Collected works III*, New York 1951, 304–323) die Auffassung, daß entweder der menschliche Geist jeder endlichen Maschine unendlich überlegen ist oder daß es absolut unentscheidbare Probleme gibt.

⁷⁸ Ebd. 208.

⁷⁹ Ebd. 208.

⁸⁰ Ebd. 209.

⁸¹ E. Runggaldier, Was sind Handlungen? Eine philosophische Auseinandersetzung mit dem Naturalismus, Stuttgart 1996, 76.

⁸² H. Putnam, 1991, 189f.

sowohl den Anliegen der Naturwissenschaft als auch den Realitäten der subjektiven Erfahrung Rechnung trägt?

4.1 Entwurf der Neurowissenschaften

Um auf diese Frage eine Antwort zu finden, kann uns die Reflexion auf die Neurobiologie als Wissenschaft weiterhelfen. Die Neurobiologie ist nicht etwa eine Gegebenheit, die wir unabhängig von uns vorfinden. Sie ist vielmehr das Ergebnis des Entwurfs der Gemeinschaft der Forschenden. In der Gemeinschaft der Forschenden haben wir uns zum Ziel gesetzt, uns selbst in unserer Leiblichkeit besser zu verstehen und die notwendigen physischen Bedingungen unseres Bewußtseins zu entdecken⁸³. Für die Erreichung dieses Ziels machen wir uns selbst zum Objekt dieser Wissenschaften und entwickeln Instrumente und Methoden, um dieses Objekt zu erforschen. Mit diesen gemeinsam festgelegten Methoden gelangen wir zu sogenannten objektiven Ergebnissen, worunter wir verstehen, daß sie von jedermann nachprüfbar sind, sofern er sich an die vereinbarte Methode hält. Wir meinen aber damit nicht eine von unserer Subjektivität völlig unabhängige Wirklichkeit⁸⁴. Der gesuchte Einigungsgrund ist also die Gemeinschaft der Personen, aus welcher der Entwurf von Wissenschaft hervorgeht⁸⁵.

4.2 *Petitio principii*: Erklärung und Erklärungsgrund

Das Ergebnis dieser Reflexion auf den Entwurf der Neurobiologie wirft ein neues Licht auf den Problembereich neurobiologischer Erklärungen. Wenn wir gemeinschaftlich Instrumente und Methoden entwickelt haben und immer weiter entwickeln, um uns selbst in unserer neuronalen Konstitution⁸⁶ zu verstehen, dann sind wir als bewußte Personen selbst der Grund, von dem her erklärt wird. Wird aber auf diesen Prozeß vergessen und die Neurobiologie zu einer von uns unabhängigen Gegebenheit gemacht und dann versucht, vom Entworfenen her den Entwurf zu erklären, dann kommt es zu dem Trugschluß einer *Petitio principii*. Das zu Erklärende ist selbst Voraussetzung oder Bedingung der Möglichkeit des Erklärens⁸⁷. Diese Überlegung macht auch verständlich, warum wir nicht definieren

⁸³ G. Rager, 1982; G. Rager, 1994.

⁸⁴ Vgl. hierzu auch die Diskussion der objektiven und der indexikalischen Wirklichkeit bei E. Runggaldier, 1996, 154–168.

⁸⁵ Die Gemeinschaft der Personen gründet im Kommerzium der Freiheit. „Das Kommerzium der Freiheit ist transzendental früher als die Subjektivität, und im Begriffe des Subjekts ist der Begriff der Intersubjektivität als der transzendentallogisch frühere Begriff schon enthalten“ (H. Krings, System und Freiheit, Freiburg 1980). – Der pragmatische Realismus von Putnam kommt zu einem ähnlichen Ergebnis.

⁸⁶ „Neuronale Konstitution“ ist ebenfalls ein von uns entwickelter operativer Begriff, der aber auf Realität abzielt.

⁸⁷ G. Rager, 1982.

können, was Bewußtsein ist. Das Bewußtsein selbst ist der Grund, von dem her wir überhaupt erst definieren und ein System aufbauen können.

4.3 *Selbst-Erfahrung und Wissen vom Gehirn: Dieselbe Wirklichkeit*

Wie verhalten sich nun Selbst-Erfahrung und Neurobiologie zueinander hinsichtlich der Wirklichkeit, auf die sie sich richten? Sind diese Wirklichkeiten von einander verschieden? Schon der Vorgang des Entwurfs von Wissenschaft zeigt, daß wir als bewußte Personen es sind, die sich in ihrer Leiblichkeit besser verstehen wollen. Es handelt sich also sowohl in unserer Selbsterfahrung als auch in den Neurowissenschaften um uns selbst. In der Selbst-Erfahrung geht es um uns in einem umfassenden Horizont. In den Neurowissenschaften geht es um uns unter eingeschränkten, sehr spezifischen Hinsichten. Für die methodisch differenzierten Programme werden je neue Begriffssysteme entwickelt, welche den ganzen Reichtum und die Mannigfaltigkeit der Neurowissenschaften ausmachen⁸⁸.

4.4 *Neurowissenschaften und Materialismus*

Bis zu diesem Punkt ist ein Konsens mit einigen neurobiologischen Theoretikern des Bewußtseins möglich. Die bewußte Erfahrung beansprucht zu Recht ihren Vorrang und kann nicht auf materielle Gegebenheiten reduziert werden. „Die epistemologische Position der modernen Naturwissenschaft ist also ... implizit dualistisch.“⁸⁹ Jenseits von diesem Punkt aber betritt Edelman den epistemologischen Zirkel: Obwohl er eigentlich bei der Konstruktion einer globalen Hirntheorie aus historischen und kreativen Gründen mit dem Bewußtsein beginnen müßte, um den Prozeß nachzuvollziehen, der zur Entwicklung der Hirnforschung geführt hat, sieht er sich verpflichtet, der Materie den ersten Platz für die Erklärung einzuräumen. „Die Materie existiert vor dem Geist“ („matter exists prior to mind“⁹⁰). So ergibt sich eine Aufeinanderfolge von „Materie → Evolution → Vertebraten → primäres Bewußtsein → Mensch → soziale Kommunikation und Sprache → Naturwissenschaft“⁹¹, welche eine materialistische Metaphysik erfordert. Der Naturwissenschaftler müsse deshalb metaphysisch ein „subtiler Materialist“ („the physical scientist as observer is a subtle materialist“⁹²) sein. Die von Edelman beschriebene Aufeinanderfolge von Materie über Evolution bis zur Naturwissenschaft ist aber schon Ergebnis

⁸⁸ Zum Verhältnis der verschiedenen Einzelwissenschaften zueinander und zur Philosophie siehe O. Muck, *Methodologie und Metaphysik*, in: E. Coreth (Hrsg.), *Aufgaben der Philosophie*, Innsbruck 1958, 99–157.

⁸⁹ G. M. Edelman, 1989, 256: „The epistemological position of modern science ... is thus either noncommittal or implicitly dualistic“.

⁹⁰ Ebd. 270, ähnlich 254, 263.

⁹¹ Ebd. 254: „matter → evolution → vertebrates → primary consciousness → man → social communication and language → natural science“.

⁹² Ebd. 264.

unseres Entwurfs von Wissenschaft. Erst in der Wissenschaft entwickeln wir Vorstellungen über Materie, Selektion und Gehirn. Diese Vorstellungen richten sich auf Realität, sind aber zugleich auch unsere Interpretation von Realität, weil wir die Realität unter einer bestimmten Hinsicht betrachten und ohne eine solche Hinsicht kein Ergebnis erzielt werden kann. In unseren wissenschaftlichen Begriffssystemen operieren wir mit dem elementaren Begriff Materie, wobei aber niemand sagen kann, was Materie eigentlich ist⁹³. Von daher erscheint es als nicht sinnvoll, den operativ entwickelten Begriff Materie zu einem Begriff von Wirklichkeit zu machen und eine materialistische Metaphysik einzufordern.

4.5 Rekonstruktion der Neurowissenschaften

Das Bewußtsein geht also allen Entwürfen von Wissenschaft historisch und kreativ voraus. Es richtet sich auf die Wirklichkeit als solche. Indem es versucht zu verstehen, wie sich seine Akte in ihrer Leiblichkeit vollziehen, entwirft es die verschiedenen Disziplinen der Neurowissenschaften mit ihren je besonderen Verstehenshorizonten, die aber dieselbe Wirklichkeit zum Gegenstand haben wie die bewußte Selbsterfahrung. Mit den Mitteln der Neurowissenschaften versuchen wir uns in unserer neuronalen Konstitution zu verstehen. Alle unsere neurobiologischen Hypothesen und Theorien beinhalten schon ein Vorauswissen, das wir wissenschaftlich gleichsam an Hand von Hirnzuständen verifizieren⁹⁴. Die Prüfung auf Konsistenz, auf Übereinstimmung mit der bewußten Erfahrung und auf Entsprechung mit beschriebenen und deutbaren Hirnzuständen, setzt immer ein Vorwissen voraus, welches uns bei der Weiterentwicklung der neurowissenschaftlichen Zugänge zum Verständnis des Gehirns leitet. Auf dieser Basis planen wir auch unsere Experimente, mit welchen wir unsere Hypothesen testen und falsifizieren⁹⁵.

4.6 Bedeutung der Neurowissenschaften

Obleich die Neurowissenschaften nicht in der Lage sind, das Bewußtsein im oben definierten Sinne zu erklären, liefern sie uns doch Systeme von Korrelationen zwischen bewußten Akten und neuronalen Vorgängen. Diesen Korrelationen verdanken wir äußerst wertvolle Einsichten in die faktischen Bedingungen unserer Leiblichkeit. Sie können uns vor unrealistischen

⁹³ Die theoretische Physik beschreibt nur Zustände und Relationen, macht aber keine ontologischen Aussagen über Materie (W. Heisenberg, *Der Teil und das Ganze. Gespräche im Umkreis der Atomphysik*, München, 1969; C. F. v. Weizsäcker, *Die Einheit der Natur*, München 1971).

⁹⁴ Vgl. auch G. M. Edelman, 1989, 268–9.

⁹⁵ Mit der Planung und Durchführung von Experimenten durchbrechen wir die Kette der Ereigniskausalität und setzen einen je neuen Anfang im Sinne der Agenskausalität. Zur Gegenüberstellung von Ereignis- und Agenskausalität siehe E. Runggaldier (1996) und U. Meixner (*U. Meixner, Ereignis und Substanz. Die Metaphysik von Realität und Realisation*, Paderborn 1997).

und falschen Vorstellungen über uns selbst als leibliche Wesen bewahren. Zugleich geben sie uns auch die Mittel für den sorgsamsten Umgang mit unserer Leiblichkeit in die Hand: Wir könnten das als Neurohygiene bezeichnen. Schließlich führen die immer genaueren Kenntnisse der neuronalen Prozesse dazu, daß wir die auftretenden Störungen besser erkennen und immer präzisere therapeutische Mittel einsetzen.