

Wolf Singer

Vom Gehirn zur Psyche

Vor etwa 150 Jahren wurden hier in Berlin folgende Sätze gesprochen:

„Dies neue Unbegreifliche ist das Bewußtsein. Ich werde jetzt, wie ich glaube, in sehr zwingender Weise dartun, daß nicht allein bei dem heutigen Stand unserer Kenntnis das Bewußtsein aus seinen materiellen Bedingungen nicht erklärbar ist, was wohl jeder zugibt, sondern daß es auch der Natur der Dinge nach aus diesen Bedingungen nicht erklärbar sein wird. (...) Welche denkbare Verbindung besteht zwischen bestimmten Bewegungen bestimmter Atome in meinem Gehirn einerseits, andererseits den für mich ursprünglichen, nicht weiter definierbaren, nicht wegzuleugnenden Tatsachen: ‚ich fühle Schmerz, fühle Lust; ich schmecke Süßes, rieche Rosenduft, höre Orgelton, sehe Rot‘ und der ebenso unmittelbar daraus schließenden Gewißheit: ‚Also bin ich? (...) Es ist in keiner Weise einzusehen, wie aus ihrem – gemeint ist der Atome – „Zusammensein Bewußtsein entstehen könne. Sollte ihre Lagerungs- und Bewegungsweise ihnen nicht gleichgültig sein, so müßte man sie sich nach Art der Monaden schon einzeln mit Bewußtsein ausgestattet denken. Weder wäre damit das Bewußtsein überhaupt erklärt, noch für die Erklärung des einheitlichen Bewußtseins das Mindeste gewonnen.“

So Emil Du Bois-Reymond in seinem Vortrag „Über die Grenzen des Naturerkennens“, den er 1872 auf der Tagung der Naturforscher und Ärzte gehalten hat. Emil Du Bois-Reymond war Mitglied der Preußischen Akademie der Wissenschaften, die jetzt die Berlin-Brandenburgische

heißt. Er äußert, wie ich glaube, begründete und nachvollziehbare Zweifel im Hinblick auf die Möglichkeit einer reduktionistischen Erklärung mentaler Phänomene, unserer subjektiven Empfindungen, unserer Möglichkeit zur freien Entscheidung und unserer Erfahrung, ein autonomes Selbst zu sein, das zwar in einem biologisch begründeten Organismus residiert, von diesem aber als ontologisch verschieden empfunden wird. Diese mentalen Phänomene, so die über Jahrhunderte unveränderte Position, verschlossen sich einer reduktionistischen Erklärung im Rahmen naturwissenschaftlicher Beschreibungssysteme. Und je überzeugender die Beweise dafür werden, daß wir unser Dasein und unser Sosein einem kontinuierlichen evolutionären Prozeß verdanken, in dessen Verlauf es keinerlei Hinweise auf ontologische Sprünge gibt, um so zwingender wird natürlich die Notwendigkeit, sich erneut mit dem Phänomen der Emergenz mentaler Qualitäten auseinanderzusetzen. Da die Phänomene, die wir gemeinhin unter Bewußtsein subsumieren, unzweifelhaft auf kognitiven Funktionen unserer Gehirne beruhen, möchte ich das Phänomen des Bewußtseins im Lichte dessen erneut kommentieren, was wir heute über die Evolution unserer Gehirne und über deren Funktionsweise zu wissen glauben.

Ein epistemologisches Caveat

Bevor ich mich dem Gehirn selbst zuwende, möchte ich ein erkenntnistheoretisches Problem ansprechen, das jemandem, der Hirnforschung be-

treibt, besonders oft und eindringlich begegnet. Bei der Erforschung des Gehirns betrachtet sich ein kognitives System im Spiegel seiner selbst. Es verschmelzen also Erklärendes und das zu Erklärende. Und es stellt sich die Frage, inwieweit wir überhaupt in der Lage sind, das, was uns ausmacht, selbst zu erkennen. Natürlich ist dies ein Problem, dem sich auch die anderen Wissenschaften stellen müssen, denn erkennbar ist ja nur, was unser kognitiver Apparat, unser Gehirn, zu denken und zu erkennen vermag. Betrachtet man die evolutionären Prozesse, die dieses Organ hervorgebracht haben, drängt sich unweigerlich der Schluß auf, daß die während der Evolution wirksamen Selektionsmechanismen vermutlich nicht dazu angetan waren, kognitive Strukturen auszubilden, die für die Erfassung dessen optimiert sind, was hinter den Dingen möglicherweise sich verbirgt. Unser Gehirn ist einzig und allein an den funktionalen Kriterien gemessen worden, den Organismus, der es trägt, so lange am Leben zu erhalten, bis dieser sich reproduzieren kann – so zumindest die klassische Auffassung. Unsere kognitiven Funktionen sind deshalb an eine makroskopische Welt angepaßt, und nicht an die Welt, in der die Quantenmechanik relevant ist, oder an die Welt kosmischer Dimensionen. Bedeutsam ist für uns die Welt, die im Zentimeter- bis Meterraum sich ereignet, und vornehmste Aufgabe unseres kognitiven Systems ist es, Regelmäßigkeiten dieser Welt zu begreifen. Daher rühren denn auch die Schwierigkeiten, die wir mit der Vorstellung von Welten haben, die uns Beckwith und von Klitzing (siehe Beiträge in diesem Band) vorgestellt haben. Prozesse im Bereich von Nanometern und Lichtjahren sind zwar berechenbar, aber sie verwehren die Anschaulichkeit und sind kaum nachempfindbar.

Evolution und Emergenz neuer Qualitäten

Ich möchte zunächst auf die Evolution unseres kognitiven Organs, des Gehirns, eingehen, dann an einigen Beispielen verdeutlichen, was wir heute über die funktionelle Organisation dieses Organs wissen, und zum Schluß noch kurz über die höchsten kognitiven Funktionen des Gehirns sprechen, die sich im Bewußt-Sein ausdrücken. Ich schicke voraus, um keine falschen Erwartungen zu wecken, daß ich der Überzeugung bin, daß diese höchsten Hervorbringungen unserer Gehirne, jene, die uns die Erfahrung vermitteln, autonome, selbstbestimmte Agenten zu sein, vermutlich kulturelle Konstrukte und deshalb der neurobiologischen Erklärung nicht direkt zugänglich sind.

Bei der Betrachtung der Evolution des Gehirns fasziniert die ungeheure Beständigkeit, mit der frühe Erfindungen über Jahrtausende hinweg konserviert wurden. Nervenstrukturen, die bereits zu Beginn der Evolution von Nervennetzen, also schon von Invertebraten entwickelt wurden, finden sich nahezu unverändert in den Nervensystemen der spät hinzugekommenen Säugetiere wieder. Die charakteristischen Merkmale von Nervenzellen, die Ausbildung von Dendritenbäumen, über die sie Information von anderen Nervenzellen empfangen, und von Axonen, mit denen sie Kontakt zu nachgeschalteten Nervenzellen aufnehmen, diese Polarisierung in einen Empfänger- und Senderbereich, ist seit Jahrtausenden unangetastet erhalten geblieben. Unverändert geblieben sind auch fast alle biochemischen Bestandteile dieser Zellen. Etwa 90 % der Gene, die in menschlichen Nervenzellen exprimiert sind, finden sich, abgesehen von kleinen, funktionell wenig relevanten Modifikationen, auch schon in Nervenzellen von Schnecken. Was an

diesen Weichtieren über zelluläre Eigenschaften zu lernen ist, läßt sich in der Regel direkt auf höhere Säuger und den Menschen übertragen. Konserviert sind erstaunlicherweise auch bis ins Detail die chemischen Überträger-substanzen, über die Nervenzellen miteinander kommunizieren. **Abbildung 1** zeigt eines der synaptischen Endknöpfchen, mit dem eine Nervenzelle über ihr Axon eine nachgeschaltete Nervenzelle kontaktiert. Hier wird durch Freisetzung einer chemischen Überträger-substanz die elektrische Aktivität der sendenden Zelle in ein chemisches Signal umgesetzt, das dann seinerseits über Rezeptoren und gekoppelte Ionenkanäle in der nachgeschalteten Zelle wiederum elektrische Potentiale erzeugt. Es gibt fast keine Überträger-substanzen im Säugetiergehirn, die nicht auch schon in einfachen Organismen, wie Insekten und Schnecken zu finden wären. Konserviert worden ist auch der allgemeine Bauplan von Gehirnen, vor allem der von Chordaten, also jenen Spezies, die über ein Rückenmark verfügen.

Abbildung 2 zeigt den Stammbaum von Wirbeltieren mit den entsprechenden Gehirnen. Die Ähnlichkeiten sind unverkennbar. Bei allen Gehirnen, ob von Fischen, Reptilien oder Säugern, läßt sich die gleiche Unterteilung in Vorderhirn, Riechhirn, Zwischenhirn, Mittelhirn, Kleinhirn und Hirnstamm vornehmen. Diese Unterteilungen ergeben sich aufgrund der Konnektivität der verschiedenen Zentren und der regionalen Expressionsmuster hirnspezifischer Gene. Besonders auffällig sind diese Ähnlichkeiten natürlich zwischen den Gehirnen von Säugetieren. Bemerkenswert ist dabei die enorme Volumenzunahme der Großhirnrinde in den hochentwickelten Gehirnen von Primaten. Es ist jedoch keineswegs so, daß wir Menschen das größte Gehirn haben; Größe ist eine not-

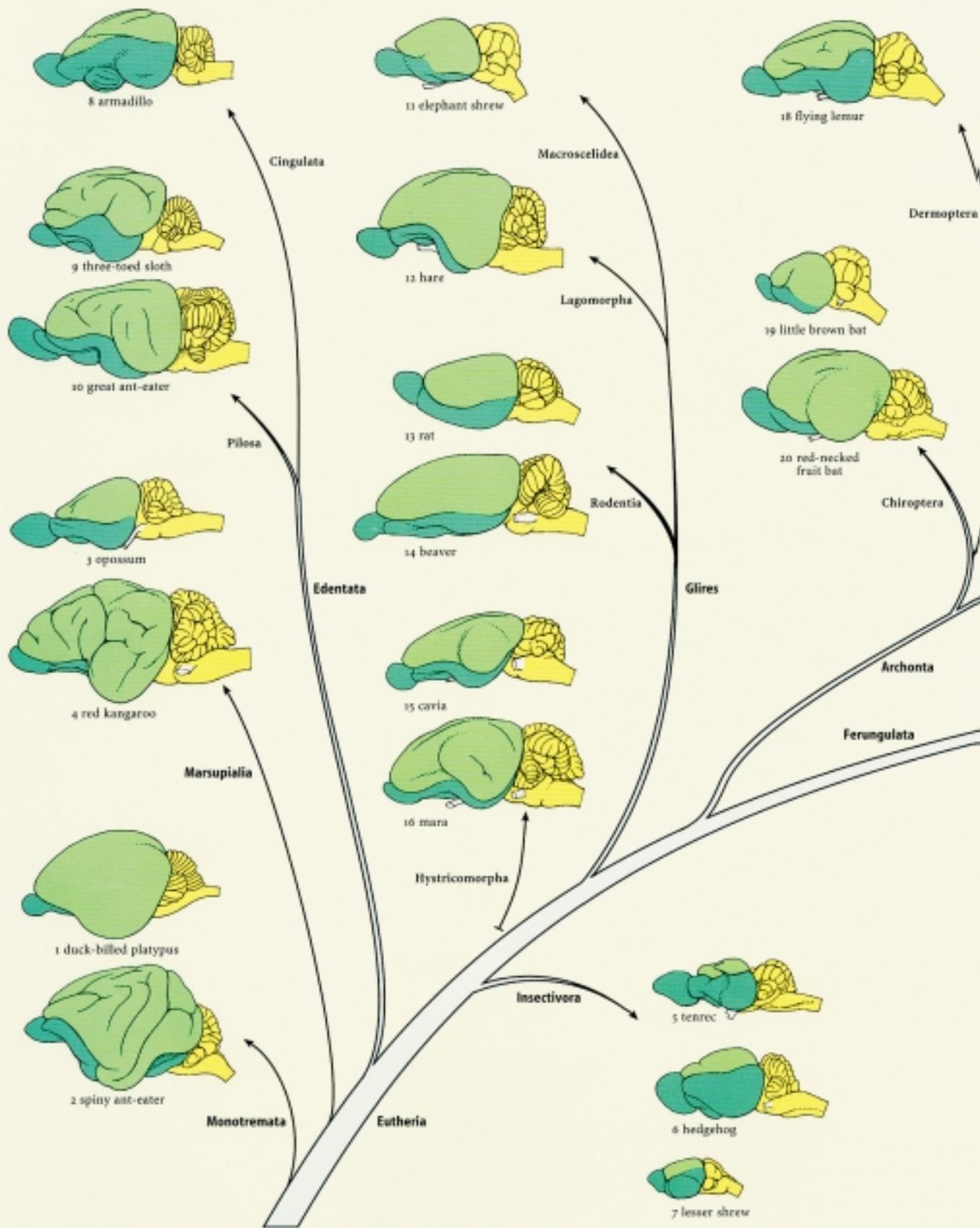


Abb. 1: Rekonstruktion einer Synapse. Die gelben Sphären enthalten chemische Überträgerstoffe, die bei elektrischer Erregung freigesetzt werden. (Quelle: Archiv des MPI für Hirnforschung)

wendige, aber nicht eine hinreichende Voraussetzung für Komplexität und Leistung; es kommt auch auf die Verschaltungsweise an. Dennoch gilt, daß all die kognitiven Eigenschaften, die Säugetiere voneinander und den Menschen von diesen unterscheiden, einzig und allein auf einer Volumenzunahme der Großhirnrinde beruhen. Abgesehen von diesem quantitativen Unterschied läßt sich keine wesentliche Veränderung im Aufbau der verschiedenen Gehirne ausmachen.

Bei der Großhirnrinde handelt es sich um eine etwa 2 mm dünne gefaltete Schicht von dicht gepackten Nervenzellen, die gemeinhin als graue Substanz bezeichnet wird, im Gegensatz zu der darunter liegenden wei-

Abb. 2 (nächste Doppelseite): Evolution des Gehirns von Wirbeltieren. Die Farbmarkierungen beziehen sich auf homologe Hirnstrukturen. (Quelle: Nieuwenhuys, ten Donkelaar, Nicholson: The Central Nervous System of Vertebrates, Springer 1998. Abdruck mit freundlicher Genehmigung des Springer-Verlages, Heidelberg.)

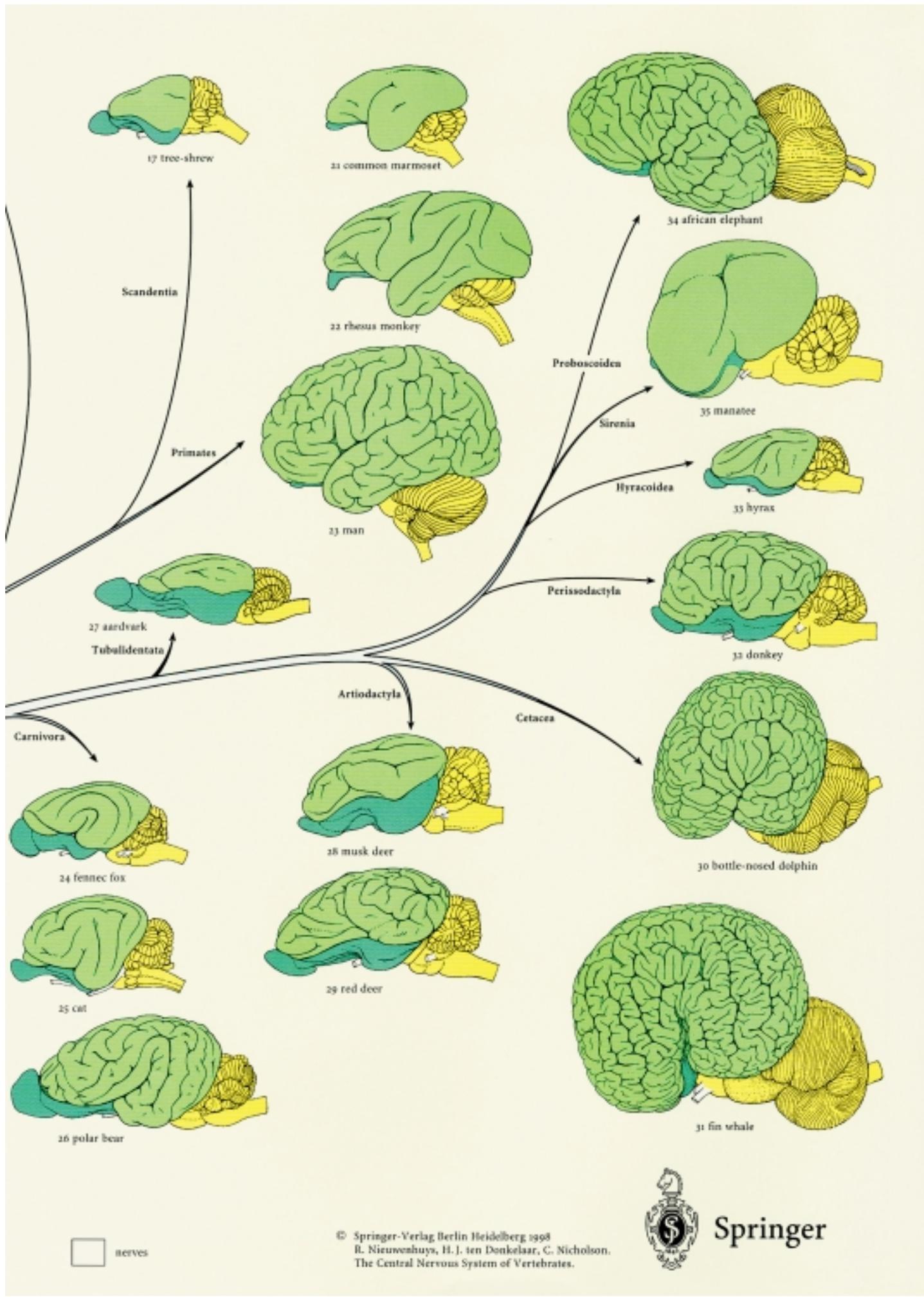


Poster 1.
Lateral views of the brains of selected mammals.

rhinencephalon

neopallium

remaining parts



Ben Substanz, die aus Leitungsbahnen besteht. In einem Kubikmillimeter Hirnrinde drängen sich etwa vierzigtausend Nervenzellen, die untereinander aufs innigste in Verbindung stehen. Eine Nervenzelle kontaktiert etwa zwanzigtausend andere und empfängt von ebenso vielen ihre Eingangssignale. Dabei kommunizieren sowohl Nervenzellen miteinander, die in unmittelbarer Nachbarschaft angeordnet sind, als auch Zellen, die weit entfernt in verschiedenen Hirnstrukturen liegen. Über Einzelheiten dieser Verbindungsarchitekturen wird noch zu sprechen sein.

Die Evolution höherer kognitiver Leistungen scheint also ganz vorwiegend auf der Vergrößerung dieses dünnen Mantels von Hirnrindenzellen zu beruhen. Bestehend ist dabei, daß diese Struktur im Laufe der Evolution ihre interne Organisation nahezu unverändert beibehalten hat. Die Großhirnrinde der Maus ist von der des Menschen kaum zu unterscheiden. Dies hat wichtige Implikationen hinsichtlich der Evolution neu-

er Funktionen. Anders als in technischen Systemen ist im Gehirn keine Trennung zwischen Hard- und Software möglich. Im Gehirn wird das Programm für Funktionsabläufe ausschließlich durch die Verschaltungsmuster der Nervenzellen festgelegt. Die Netzstruktur ist das Programm. Die Algorithmen, nach denen die Großhirnrinde arbeitet, haben sich somit im Laufe der Evolution kaum verändert. Es sind lediglich mehr Areale hinzugekommen. Dies bedeutet, erstens, daß die von der Großhirnrinde erbrachten Verarbeitungsleistungen sehr allgemeiner Natur sein müssen und, zweitens, daß die Iteration von im Prinzip gleichen Prozessen neue, qualitativ verschiedene Funktionen hervorbringen kann.

Wie **Abbildung 3** zeigt, läßt sich die Hirnrinde aufgrund anatomischer und funktioneller Kriterien in Regionen einteilen. Im parietalen und temporalen Bereich liegen Areale, die sich mit der Verarbeitung visueller Signale befassen, dazwischen finden sich Areale, die akustische Aktivität

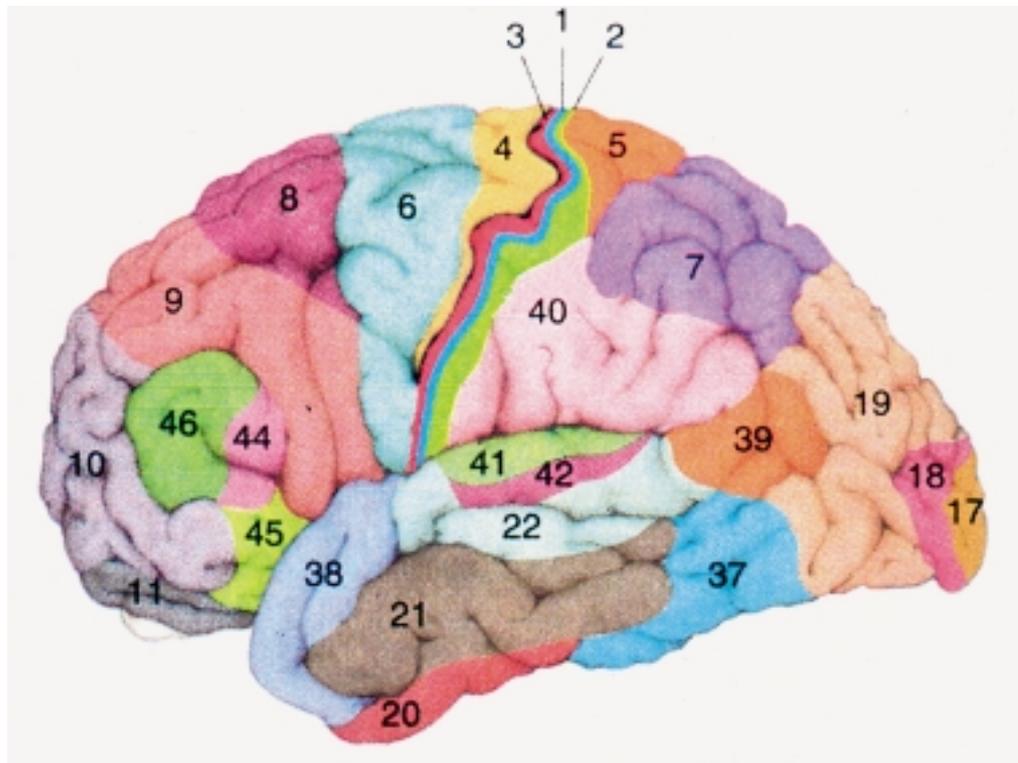


Abb. 3: Brodmanns Topologie der Rindenareale des menschlichen Gehirns. (Quelle: Archiv des MPI für Hirnforschung.)

vermitteln, und wenn es sich um die sprachdominante Hirnhälfte handelt, liegen hier auch Areale, die sich mit der sensorischen Verarbeitung von Sprachmaterial befassen. Ferner gibt es Areale, die sich mit der Körperfühlsphäre auseinandersetzen, also mit den Signalen, die von den Rezeptoren im Körper vermittelt werden. In frontalen Rindengebieten werden Bewegungsprogramme erstellt und in der dominanten Hemisphäre wird hier zusätzlich die Sprachproduktion verwaltet. Schließlich sind da die stammesgeschichtlich relativ rezenten präfrontalen Areale, die für die Handlungsplanung und vermutlich auch für die Einbindung in soziale Gefüge zuständig sind. Hier findet sich auch der Kurzzeitspeicher, der es uns ermöglicht, Reaktionen auf Reize aufzuschieben und Handlungsentwürfe gegeneinander abzuwägen.

Das Bestechende an dieser funktionellen Unterteilung ist, daß die interne Struktur der verschiedenen Hirnrindenareale praktisch identisch ist, obgleich sie doch offensichtlich ganz verschiedene Funktionen wahrnehmen. Nur der Spezialist ist in der Lage, ein histologisches Präparat, das von der Sehrinde entnommen wurde, von einem zu unterscheiden, das von der Sprachregion stammt. Es gibt feine Unterschiede, aber die generelle Organisation, die Verschaltung, ist nahezu identisch.

Dies legt die Schlußfolgerung nahe, daß in der Hirnrinde ein Verarbeitungsalgorithmus realisiert wird, der zur Behandlung unterschiedlichster Inhalte taugt und dessen Iteration alleine offenbar zu immer höheren kognitiven Leistungen führen kann.

Welches nun sind die Leistungen, die in der Hirnrinde erbracht werden, oder allgemeiner gefragt, welches sind die grundlegenden Funktionsprinzipien, nach denen Gehirne organisiert sind?

Das Bindungsproblem

Bis vor kurzem, und wohl schon seit geraumer Zeit, sind Fachleute wie Laien gleichermaßen, der Intuition folgend, davon ausgegangen, daß irgendwo im Gehirn ein Konvergenzzentrum existieren müsse, wo alle Signale, die über die Sinnesorgane gesammelt werden, konvergieren, um dort einer einheitlichen Interpretation zugeführt zu werden. Es wäre dies dann auch der Ort, wo Handlungsentwürfe erarbeitet und Entscheidungen gefällt werden; und für die, die dualistische Positionen bevorzugen, wäre dies auch der Ort, wo der mit mentalen Eigenschaften ausgestattete Homunkulus wirkt, der über alle Hirnfunktionen wacht und koordinierend tätig ist. Aber selbst wer monistischen Positionen zuneigt, ist versucht, wenn er seiner Intuition folgt, ein hierarchisches oder pyramidales Ordnungsprinzip zu postulieren – ganz so, wie es Descartes natürlich und unvermeidlich schien.

Nun hat uns die moderne Neurobiologie belehrt, daß wir alle, Descartes eingeschlossen, irrten, daß die tatsächliche Organisation des Nervensystems auf dramatische Weise verschieden ist. Es trifft zwar immer noch zu, und ich will dies am Beispiel des Sehsystems illustrieren, daß die ersten Schritte der Informationsverarbeitung dem seriellen Prinzip folgen. Licht wird im Auge durch Photorezeptoren in neuronale Aktivität umgewandelt, und diese elektrischen Signale gelangen über Fasersysteme zum Thalamus und dann zur primären Sehrinde. Aber dann beginnt das große Verwirrspiel. **Abbildung 4** zeigt, wie wir uns heute das Sehsystem von Primaten vorzustellen haben. Die verschiedenfarbigen Kästchen stehen für Hirnrindenareale, die sich alle direkt mit der Verarbeitung visueller Signale befassen. Bis zur

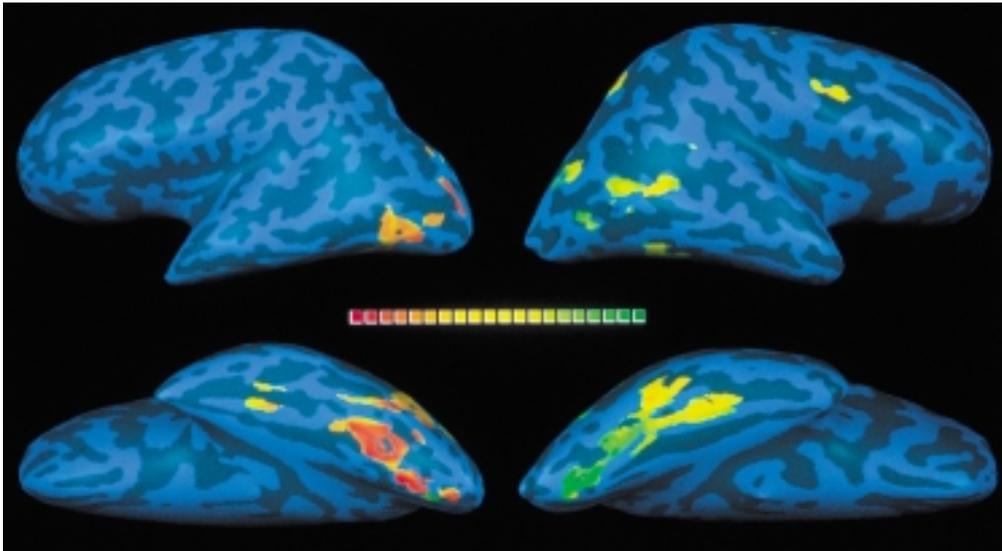


Abb. 5: Darstellung von Hirnrindenarealen, die bei der Betrachtung eines einfachen visuellen Reizes aktiv werden. Diese Aktivitätskarte wurde mit Hilfe der funktionellen Kernspintomographie erstellt (Dr. R. Goebel, MPI Hirnforschung).

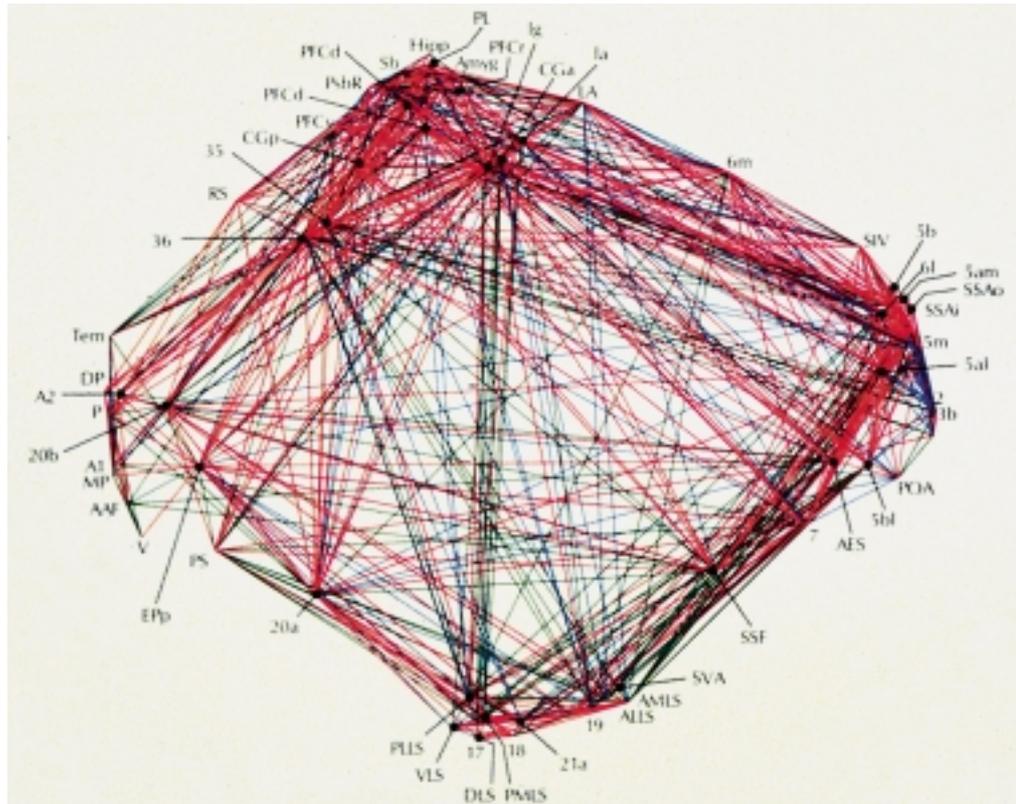
kennen eines dreidimensionalen Objektes, zur gleichzeitigen Aktivierung zahlreicher Hirnrindenareale. **Abbildung 5** illustriert das Ergebnis einer Untersuchung mit der funktionellen Kernspintomographie und zeigt die Aktivitätsverteilungen im Gehirn einer gesunden Versuchsperson während eines einfachen Wahrnehmungsaktes. Die Großhirnhemisphären sehen hier etwas ungewöhnlich aus, weil hier durch Computerrekonstruktion die Hirnrinde geglättet wurde. Die Hemisphären wurden wie ein Ballon aufgeblasen, bis die Faltungen der Hirnrinde verstrichen waren und auch die Bereiche in der Tiefe der Furchen sichtbar wurden.

Nun könnte man einwenden, daß es vielleicht doch Konvergenzzentren geben könnte, wenn man mehrere sensorische Modalitäten zusammen betrachtet. Doch auch diese Hoffnung trägt, wie das Schaltdiagramm in **Abbildung 6** zeigt. Hier stehen die schwarzen Punkte für Hirnrindenareale der sensorischen Systeme der Katze, und die farbigen Striche symbolisieren die Verbindungen, die zwischen den Hirnrindenarealen ausgespannt sind. Unten liegt das visuelle System – es hat weniger Areale als das des Primaten – rechts das soma-

to-sensorische System, also das System, das sich mit der Körperfühlsphäre befaßt und links das auditorische System. Zusätzlich eingezeichnet ist noch ein Teil des limbischen Systems (oben), das sich mit der Zuordnung von emotionalen Beiwerten für die jeweiligen Wahrnehmungsinhalte befaßt. Wir können Gesichter ja nicht nur identifizieren, sondern auch deren Gestimmtheit ablesen und meist lösen Wahrnehmungen auch in uns bestimmte Emotionen aus.

Also selbst dann, wenn man mehrere Sinnesmodalitäten zusammenfaßt und deren Verbindungen untersucht, lassen sich keine Konvergenzzentren identifizieren. Man sieht sich vielmehr einem hoch distributiv und parallel organisierten System gegenüber, das auf außerordentlich komplexe Weise reziprok vernetzt ist. Und dies wirft die kritische Frage auf, wie diese vielen gleichzeitig ablaufenden Verarbeitungsprozesse so koordiniert werden können, daß eine kohärente Interpretation der Welt möglich wird, daß sinnvolle Entscheidungen getroffen werden können und daß gezielte Handlungsentwürfe realisierbar sind. Es gibt hier keinen Agenten, der interpretiert, kontrol-

Abb. 6: Vernetzungsdiagramm von Hirnrindenarealen des visuellen, auditorischen, somatosensorischen und limbischen Systems der Katze (von M. Young).



liert und befiehlt. Koordiniertes Verhalten und kohärente Wahrnehmung müssen als emergente Qualitäten oder Leistungen eines Selbstorganisationsprozesses verstanden werden, der alle diese eng vernetzten Zentren gleichermaßen einbezieht. Zu klären, wie diese Koordination erfolgt, ist eine der großen Herausforderungen, mit der sich die Neurobiologie im Augenblick beschäftigt. Wir bezeichnen dieses Problem als das Bindungsproblem. Ich will hier nicht ins Detail gehen, weil im Max-Planck-Spiegel in Heft 4/1998 über dieses Problem ausführlich berichtet wurde.

Die Struktur von Bindungsproblemen, die in solch distributiv organisierten Systemen gelöst werden müssen, läßt sich auch an scheinbar einfachen Wahrnehmungsakten veranschaulichen. Wenn man die Szene in **Abbildung 7** betrachtet und sich dabei vergegenwärtigt, daß sie auf der Netzhaut lediglich eine zweidimensionale Helligkeitsverteilung er-

zeugt, wird deutlich, welche immense Leistung das Sehsystem erbringen muß, um die dargestellten Figuren vom Hintergrund abzugrenzen und als Pferde erkennen zu können.

Unsere Sehzentren müssen von den vielen Konturen und Helligkeitsunterschieden jene herausfinden, die konstitutiv für eine bestimmte Figur sind, diese perzeptuell binden und dann gemeinsam interpretieren. Es muß also wieder ein Bindungsproblem gelöst werden. Würde dieses Bindungsproblem falsch gelöst, würden z. B. die dunklen Flecken der Pferde als zur Wiese gehörig interpretiert, wäre es natürlich unmöglich, die Tiere zu erkennen. Die Segmentierung muß folglich dem Erkennungsprozeß vorausgehen. Erst wenn richtig segmentiert wurde, kann erkannt werden. Dies bedeutet aber, daß der Segmentierungsprozeß sehr allgemeinen Regeln folgen muß, die auf beliebige Szenen gleichermaßen angewandt werden können. Wir ge-



Abb. 7: Pferde auf aus-
spernder Wiese. (Zeich-
nung der Künstlerin
Bev Doolittle. Abdruck
mit freundlicher Ge-
nehmigung von „The
Greenwich Workshop“,
Shelton, CT/USA.)

hen heute davon aus, daß die Regeln, denen solche Segmentierungsleistungen gehorchen, zum großen Teil angeboren sind, also auf Wissen beruhen, das im Laufe der Evolution erworben und in den Genen gespeichert wurde; Wissen über zweckmäßige Gruppierungen, das sich in genetisch determinierten Verschaltungsmustern ausdrückt, die ihrerseits das Programm für die Gruppierungsoperationen darstellen. Gruppierungsregeln können natürlich auch gelernt werden, und dies dürfte vor allem für solche zutreffen, die auf komplexen Gestaltkriterien beruhen. Auch dieses durch Erfahrung erworbene Wissen muß aber letztlich über Änderungen der funktionellen Koppelung von Neuronen abgespeichert werden.

Wie sollen wir uns die Realisierung solcher Bindungsoperationen im

Gehirn vorstellen? Wieder ist da das klassische Konzept, das unsere Forschung über Dekaden hinweg motiviert hat und das aus methodischen und konzeptionellen Gründen am nächsten lag. Man postulierte hierarchisch aufgebaute Verarbeitungsstrukturen, in denen die Bindung von Merkmalen über die Konvergenz von anatomischen Bahnen auf spezielle Bindungsneurone erreicht werden sollte. Und einige Befunde sprachen auch für diese Annahme. In der Peripherie des Systems dominieren Nervenzellen, die selektiv auf elementare Merkmale ansprechen, senkrechte oder horizontale Konturen, einfache Texturen und Farbkontraste. Hubel und Wiesel wurden mit dem Nobelpreis bedacht, nachdem sie vor inzwischen fast 30 Jahren entdeckt hatten, daß Nervenzellen in der primären

Sehrinde, also auf einer sehr frühen Verarbeitungsstufe, selektiv auf die Orientierung von Lichtbalken ansprechen. Wenn ein Lichtbalken geringfügig von der Vorzugsorientierung des rezeptiven Feldes abweicht, verstummt die Zelle. Diese Beobachtung legte nahe, daß die Zellen als Merkmalsdetektoren arbeiten und die Orientierung einer Kante signalisieren. Folgerichtig haben sich die Neurophysiologen dann zu höheren Verarbeitungsstrukturen vorangetastet. **Abbildung 8** zeigt, was sie dort entdeckten. Hier sind die Reize abgebildet, die Neuronen auf einer höheren Verarbeitungsstufe bevorzugen. Dabei handelt es sich um Zellen in einem Areal, das zum ventralen Pfad gehört, dem Verarbeitungsweg, dem die Objektidentifikation obliegt. Weil diese Nervenzellen bereits auf recht komplexe Konstellationen von Merkmalen ansprechen, stand zu erwarten, daß sich schließlich Zellen finden würden, die selektiv auf reale Objekte

der Sehwelt reagieren. Doch die Suche war vergebens. Es fanden sich keine Nervenzellen, die selektiv durch reale Objekte wie Bananen oder Bäume aktiviert wurden. Theoretiker hatten überdies darauf hingewiesen, daß dies auch nicht zu erwarten sei. Computerwissenschaftler hatten versucht, auf der Basis solcher konvergenter Architekturen mustererkennende Maschinen zu entwickeln, und mußten erkennen, daß Merkmalsbindung über Konvergenz alleine nicht zu realisieren ist. Wir können bekannte Objekte auch dann wiedererkennen, wenn sie im Raum gedreht sind. Dies führt jedesmal zu völlig anderen Merkmalskonstellationen. Man bräuchte also für jedes Objekt einen ganzen Satz von Bindungsneuronen, die sich auf die verschiedenen Ansichten eines bestimmten Objekts spezialisiert haben. Es bedürfte also einer viel zu großen Zahl von Nervenzellen, wollte man für jedes erkennbare und unterscheidbare Ob-

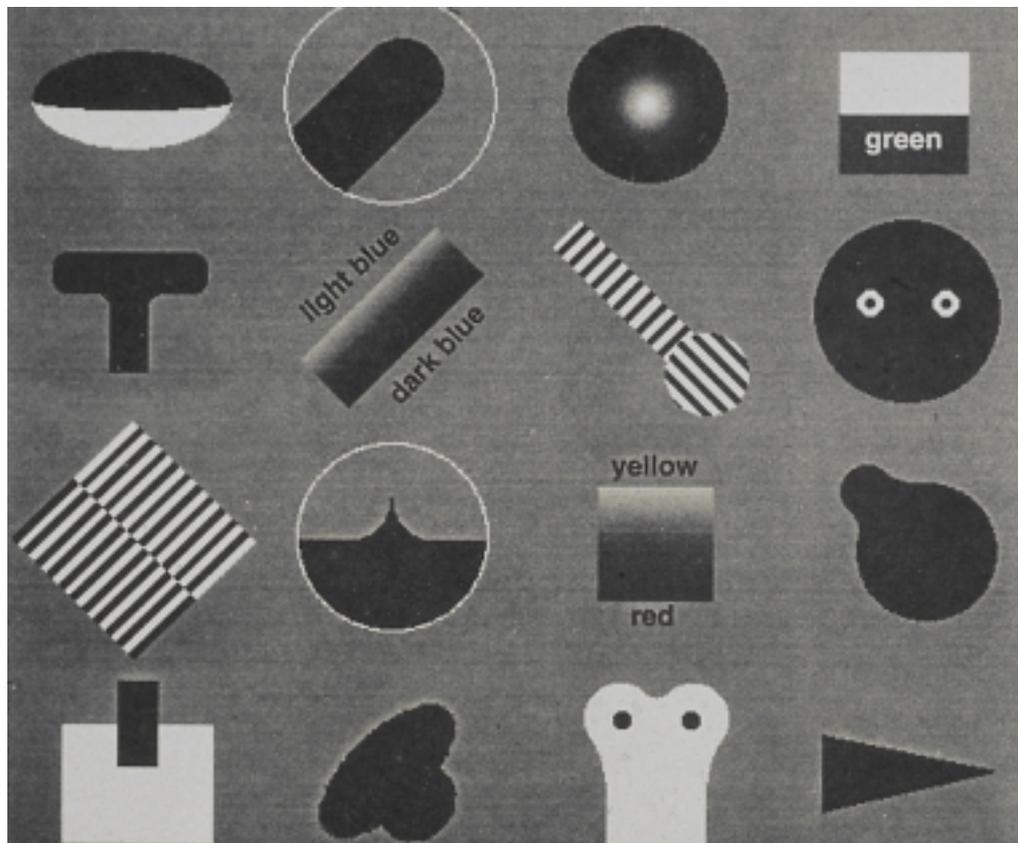


Abb. 8: Muster, auf die Neuronen in höheren Hirnrindengebieten des Sehsystems bevorzugt ansprechen (nach K. Tanaka).

jekt Bindungsneurone einrichten. Ferner bräuchte man ein riesiges Reservoir von nicht festgelegten Neuronen, um dem Umstand Rechnung zu tragen, daß wir neue Objekte sofort repräsentieren können, sobald wir sie gesehen haben.

Es wurden deshalb andere Hypothesen erdacht. Eine, die derzeit favorisiert wird und auch experimentellen Überprüfungen zugänglich ist, geht davon aus, daß die Repräsentation von Inhalten nicht über einzelne hochspezialisierte Nervenzellen erfolgt, sondern über ganze Ensembles von Nervenzellen, die über große Bereiche der Großhirnrinde verteilt sein können und sich ad hoc aufgrund der vorhandenen Kopplungen zusammenschließen. Jede einzelne dieser Zellen würde dann nur Teilmerkmale eines bestimmten kognitiven Objektes repräsentieren. In ihrer Gesamtheit aber wären die Antworten der Zellen, die sich an einem Ensemble beteiligen, die nicht weiter reduzierbare Beschreibung eines bestimmten Inhaltes. Der große Vorteil dieser Repräsentationsstrategie ist natürlich, daß die gleiche Nervenzelle zu verschiedenen Zeitpunkten benutzt werden kann, um ganz verschiedene Inhalte mit zu repräsentieren, indem sie einfach in verschiedene Ensembles eingebunden wird. Eine Zelle, die auf vertikale Konturen anspricht, kann dann für die Kodierung aller Objekte benutzt werden, die vertikale Konturen enthalten usw. Dies löst jedoch noch nicht das Bindungsproblem. In der Regel sind sehr viele Nervenzellen gleichzeitig aktiv und es muß für die nachfolgenden Verarbeitungsstrukturen geklärt werden, welche Zellen jeweils zu einem bestimmten Ensemble gehören und gemeinsam einen bestimmten Inhalt kodieren. Für die Lösung dieses Bindungsproblems wurden verschiedene Mechanismen vorgeschlagen. Wir favorisieren aufgrund experimenteller

Hinweise die Hypothese, daß Neuronen in der Hirnrinde, die sich mit der Repräsentation des gleichen Objekts befassen, sich dadurch als zusammengehörig zu erkennen geben, daß sie ihre Aktivität synchronisieren. Die Signatur eines Ensembles wäre demnach die zeitliche Kohärenz der Aktivität der jeweils teilhabenden Neuronen. Die zeitliche Auflösung, mit der diese Signatur definiert wird, liegt dabei im Bereich von Millisekunden. Entsprechend hoch ist die Taktfrequenz, mit der verschiedene Ensembles aufeinander folgen können. Experimentelle Befunde legen nahe, daß die Synchronisationsprozesse auf der Basis von Oszillationen im 40 Hz-Bereich erfolgen.

Von Repräsentationen zum Bewußtsein

Ich will der faszinierenden Frage nach dem neuronalen Code von Objektrepräsentationen nicht weiter nachgehen, sondern mich wieder dem eingangs von Du Bois-Reymond angesprochenen Problem zuwenden. Wie kommt es, daß wir nicht nur das in unserem Gehirn repräsentieren können, was in der Umwelt vorhanden ist, sondern daß wir uns dessen auch bewußt sein können, daß wir uns gewahr sind, Wahrnehmungen und Empfindungen zu haben? – ein Phänomen, das die Angelsachsen als *phenomenal awareness* ansprechen.

Voraussetzung für diese Fähigkeit ist, daß es im Gehirn kognitive Strukturen gibt, die die Repräsentation des Draußen noch einmal reflektieren, noch einmal auf die gleiche Weise verarbeiten wie die peripheren Areale die sensorischen Signale aus der Umwelt und dem Körper. Die Funktion des „inneren Auges“ läßt sich denken als die Iteration, als die wiederholte Anwendung auf sich selbst, der gleichen kognitiven Opera-

tionen, die den unreflektierten Primärrepräsentationen des Draußen zu Grunde liegen.

Nun gibt es tatsächlich Hinweise, daß die in der Evolution später hinzutretenden Hirnrindenareale ihre Eingangssignale nicht mehr direkt von den Sinnesorganen beziehen, sondern von den bereits vorhandenen stammesgeschichtlich älteren Arealen, die ihrerseits mit den Sinnesorganen verbunden sind. Die neuen Areale scheinen die Signale, die sie von den alten, von den primären Arealen bekommen, auf die gleiche Weise zu verarbeiten wie letztere die Signale, die sie von den Sinnesorganen erhalten. So lassen sich durch Iteration der immer gleichen Repräsentationsprozesse Metarepräsentationen aufbauen – Repräsentationen von Repräsentationen – die hirnhinterne Prozesse abbilden anstatt die Welt draußen.

Metarepräsentationen aufbauen zu können, bringt Vorteile. Gehirne,

die dies vermögen, können Reaktionen auf Reize zurückstellen und Handlungsentscheidungen abwägen, sie können interne Modelle aufbauen und den erwarteten Erfolg von Aktionen an diesen messen. Sie können mit den Inhalten der Metarepräsentationen spielen und prüfen, was die Konsequenzen bestimmter Reaktionen wären. Die Möglichkeit, Metarepräsentationen aufzubauen, befähigt zu umsichtigem Handeln und erlaubt damit, Gefahren präventiv aus dem Weg zu gehen. Wie bedeutend die Rolle dieser internen Mustererzeugung, dieser internen Modellbildung ist, läßt sich mit der funktionellen Kernspintomographie demonstrieren. **Abbildung 9** illustriert, welche Hirnrindenareale aktiviert werden, wenn man sich etwas vorstellt. Zwei Bedingungen wurden verglichen: In einem Fall sah die Versuchsperson eine rotierende Scheibe, im anderen hatte sie die Augen geschlossen und stellte

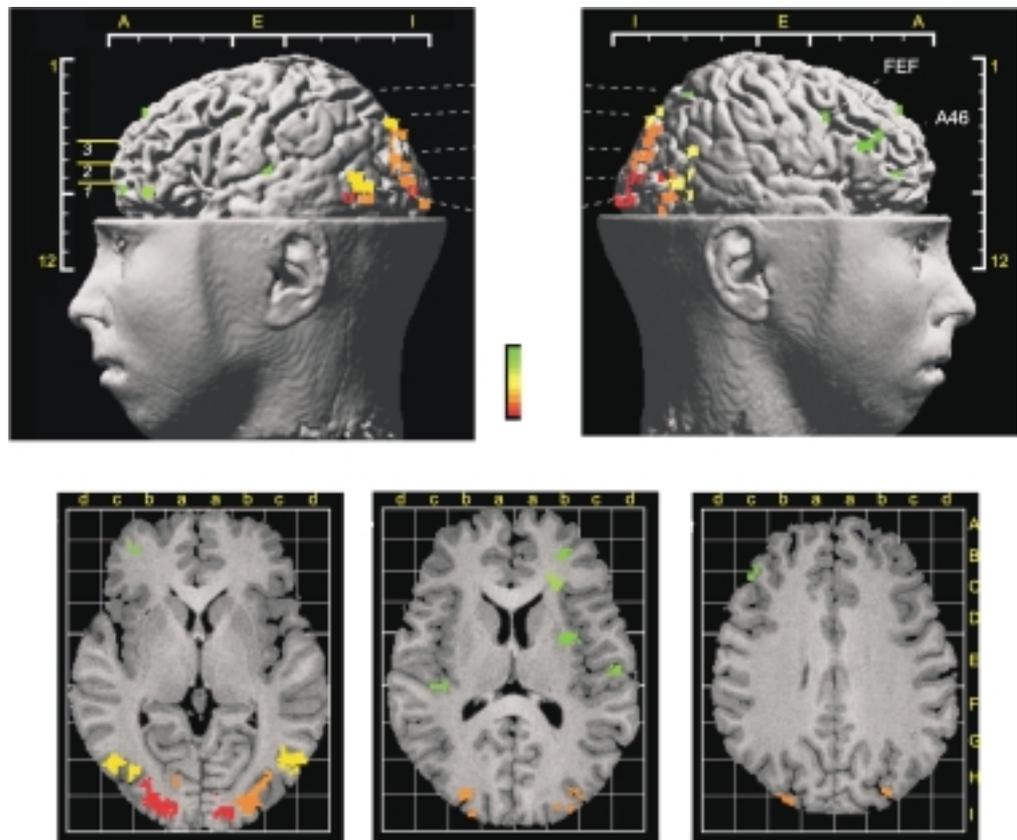


Abb. 9: Vergleich der Aktivierungsmuster der Hirnrinde bei visueller Wahrnehmung und der Vorstellung desselben Musters. Rot: Areale, die nur bei der Wahrnehmung realer Objekte aktiv werden. Orange und gelb: Areale, die sowohl bei der Wahrnehmung als auch bei der bloßen Vorstellung aktiv werden. Grün: Areale, die nur bei der Vorstellung aktiv werden. (Weitere Erläuterungen im Text, aus Goebel et al., Eur. J. Neurosci. 10, 1563-1573 (1998).)

sich die Scheibe nur vor. Ein robustes Ergebnis solcher Untersuchungen ist, daß eine Vielzahl von Arealen in gleicher Weise aktiv werden, unabhängig davon, ob die Muster tatsächlich gesehen oder nur vorgestellt werden. Insbesondere die höheren Areale, also jene, denen die Erstellung von Metarepräsentationen obliegt, werden auch aktiv, wenn sich die Probanden bestimmte Inhalte nur vorstellen – und diese interne Aktivierung ist modalitätsspezifisch. Bei visuellen Vorstellungen werden visuelle Areale aktiv und beim stummen Sprechen die Sprachareale. Aber es gibt auch Areale, die *nur* bei der Vorstellung aktiv werden und nicht bei der Wahrnehmung realer Inhalte. Diesen Arealen fällt die Aufgabe zu, die Aktivität in den spezifischen Arealen zu orchestrieren, in denen die zur Vorstellung erforderlichen Repräsentationen gespeichert liegen. Schließlich fallen einige Areale auf, die tatsächlich nur bei der Wahrnehmung realer Inhalte aktiviert werden. Es sind dies die phylogenetisch alten, primären, sensorischen Areale, die ihre Eingangssignale vorwiegend von den Sinnesorganen beziehen.

Unter bestimmten pathologischen Bedingungen, z. B. bei Halluzinationen, werden diese intern generierten Aktivitätsmuster als von draußen kommend wahrgenommen. In solchen Fällen ändern sich dann die Verteilungsmuster. Das Beispiel in **Abbildung 10** zeigt das Ergebnis einer Messung mit der funktionellen Kernspintomographie bei einem schizophrenen Patienten, der verbale Halluzinationen hatte. Dieser Patient vernahm zu genau angebbaren Zeitpunkten eine Stimme, die von einem realen Sprecher zu kommen schien und Schmähreden hielt. Diese Beschimpfungen wurden als sehr unangenehm empfunden und führten zur Aktivierung von Zentren im limbischen System (den Mandelkernen), von denen bekannt ist, daß sie bei negativen Empfindungen aktiviert werden. Hier also erzeugt sich das Gehirn Erregungsmuster, die als real erlebt und entsprechend emotional bewertet werden. Anders als beim Gesunden, der sich etwas vorstellt oder stumme Sprache spricht, werden bei halluzinierenden Patienten jedoch auch die primären Sinnesareale mit aktiviert. Bei akustischen Halluzina-

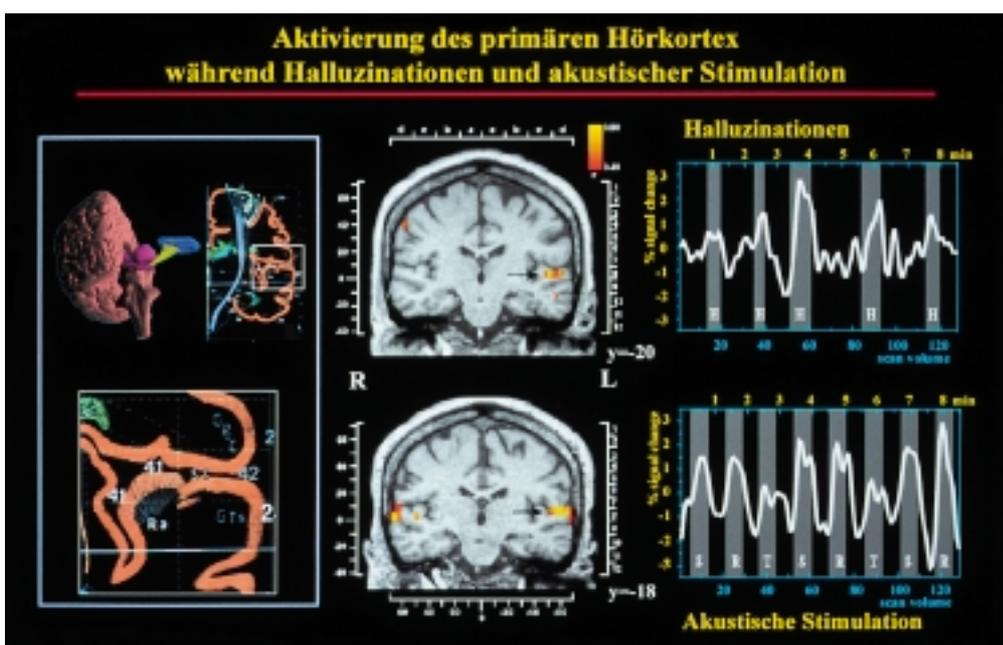


Abb. 10: Räumliche und zeitliche Verteilung von Aktivitätsmustern in der Hirnrinde eines halluzinierenden Patienten während Halluzinationen (obere Diagramme) und während akustischer Reizung (untere Diagramme). Die Kurven stellen Aktivitätsschwankungen in der primären Hörrinde (in den Hirnschnitten farblich markiert) dar, und zwar bei Halluzinationen (oberes Diagramm, schattierte Episoden) und bei akustischer Reizung (unteres Diagramm). (Aus Dierks et al., *Neuron* 22, 615–621 (1999).)

tionen betrifft dies die primäre Hörrinde in der Heschelschen Querwindung der linken sprachdominanten Hemisphäre. Jedesmal, wenn der haluzinierte Sprecher spricht, und die Patienten können das genau angeben, läßt sich eine Zunahme der Hirnaktivität messen, hier indirekt erschlossen über die Zunahme der Durchblutung in den entsprechenden Arealen. Die Aktivierung des primären sensorischen Areals erfolgt vermutlich über Rückkopplungsschleifen, die von höheren Hirnrindenarealen kommen. Wenn dieses primäre Areal in der sprachkompetenten Hemisphäre mitaktiviert wird, werden die selbsterzeugten Erregungsmuster offenbar so wahrgenommen, als kämen sie von draußen. Werden diese primären Areale nicht mitaktiviert, wie es bei Gesunden der Fall ist, wenn sie stumme Sprache sprechen, bleibt die Wahrnehmung des Gesprochenen als selbst Erzeugtes erhalten.

Diese Beispiele sollten deutlich machen, wie groß bei Wahrnehmungsprozessen der Anteil selbstgenerierter Aktivität sein kann. Es bestätigt dies auf eindrucksvolle Weise, was wahrnehmungsphysiologische Untersuchungen nahelegen: daß Wahrnehmung nicht als passive Abbildung von Wirklichkeit verstanden werden darf, sondern als das Ergebnis eines außerordentlich aktiven, konstruktivistischen Prozesses gesehen werden muß, bei dem das Gehirn die Initiative hat. Das Gehirn bildet ständig Hypothesen darüber, wie die Welt sein sollte, und vergleicht die Signale von den Sinnesorganen mit diesen Hypothesen. Finden sich die Voraussagen bestätigt, erfolgt die Wahrnehmung nach sehr kurzen Verarbeitungszeiten. Treffen sie nicht zu, muß das Gehirn seine Hypothesen korrigieren, was die Reaktionszeiten verlängert. In den meisten Fällen dürfte sich der Wahrnehmungsakt jedoch auf das Bestätigen bereits formulierter Hypothesen beschränken.

Somit erscheint, zumindest im Prinzip, nachvollziehbar, wie die Funktion des inneren Auges neuronal realisiert sein kann, wie das Sich-Gewahr-Werden seiner eigenen Wahrnehmungen und Empfindungen über die Etablierung von Metarepräsentationen erreicht werden kann, ohne daß es ontologischer Diskontinuität in der Evolution bedarf. Offenbar genügt es zum Aufbau von Metarepräsentationen, Areale hinzuzufügen, die auf hirninterne Prozesse genauso „schauen“ wie die bereits vorhandenen Areale auf die Peripherie.

Das Subjekt als kulturelles Konstrukt

Zum Schluß nun will ich mich noch kurz einer der schwierigsten Fragen zuwenden, die gegenwärtig im Grenzgebiet zwischen Neurobiologie und Philosophie verhandelt werden – der Frage, ob wir innerhalb neurobiologischer Beschreibungssysteme angeben können, wie unsere Selbstkonzepte entstehen, unser Ichbewußtsein und unsere Erfahrung, ein autonomes Agens zu sein, das frei ist zu entscheiden. Es geht um die Frage, wie es möglich ist, daß unser Ich, das wir als eine mentale Entität erleben, losgelöst von allen materiellen Bindungen, etwas beschließen kann, das dann, um ausgeführt zu werden, in neuronale Aktivität übersetzt werden muß. Behandelt werden soll also die Frage nach unserem Selbstbewußtsein, nach unserer Erfahrung, ein autonomes freies Ich zu sein.

Nach meinem Dafürhalten läßt sich diese Frage nicht mehr allein innerhalb neurobiologischer Beschreibungssysteme fassen, da diese sich ausschließlich an der naturwissenschaftlichen Analyse einzelner Gehirne orientieren, die Ich-Erfahrung bzw. die subjektiven Konnotationen von Bewußtsein jedoch kulturelle

Konstrukte sind, soziale Zuschreibungen, die dem Dialog zwischen Gehirnen erwachsen und deshalb aus der Betrachtung einzelner Gehirne nicht erklärbar sind. Die Hypothese, die ich diskutieren möchte, ist, daß das Konstrukt des autonomen, subjektiven Ichs nur hat entstehen können, weil die Evolution Gehirne hervorbrachte, die zwei Eigenschaften aufweisen: Erstens, ein inneres Auge zu haben, also über die Möglichkeit zu verfügen, Protokoll zu führen über hirninterne Prozesse, diese in Metarepräsentationen zu fassen und deren Inhalt über Gestik, Mimik und Sprache anderen Gehirnen mitzuteilen; und, zweitens, die Fähigkeit, mentale Modelle von den Zuständen der je anderen Gehirne zu erstellen, eine „theory of mind“ aufzubauen, wie die Angelsachsen sagen. Diese Fähigkeit ist dem Menschen vorbehalten und fehlt dem Tier. Allenfalls Schimpansen haben eine begrenzte Möglichkeit, sich vorzustellen, was im anderen vorgeht, wenn er bestimmten Situationen ausgesetzt ist. Wir Menschen können dies in hervorragender Weise und sind deshalb in der Lage, in einen Dialog einzutreten der Art „ich weiß, daß du weißt, wie ich fühle“ oder „ich weiß, daß du weißt, daß ich weiß, wie du fühlst“ usw. Interaktionen dieser Art führen also zu einer iterativen wechselseitigen Be Spiegelung im je anderen. Diese Reflexion wiederum ist, wie ich glaube, die Voraussetzung dafür, daß der Individuationsprozeß einsetzen kann, daß die Erfahrung, ein Selbst zu sein, das autonom und frei agieren kann, überhaupt möglich wird.

Warum nun erscheinen uns die subjektiven Konnotationen von Bewußtsein von so ganz anderer Art als die üblichen Erfahrungen? Ich vermute, daß dies eine entwicklungspsychologische Begründung hat. Der Dialog, der den Individuationsprozeß erst möglich macht, vollzieht sich be-

reits in der frühen Kindheit und erlaubt erste Ich-Identifikationen schon nach den ersten paar Lebensjahren. Dieser frühe Dialog zwischen Bezugsperson und Kind vermittelt diesem in sehr prägnanter und asymmetrischer Weise die Erfahrung, offenbar ein autonomes, frei agierendes, verantwortliches Selbst zu sein, hört es doch ohne Unterlaß: „Tu nicht dies, sondern tu das, laß das, sonst –“, oder „Mach das, andernfalls –!“ Diese Hinweise sind in idealer Weise dazu angetan, dem Kind klar zu machen, daß es offensichtlich frei ist, nicht dies, aber das zu tun, und daß es für seine Entscheidung zur Verantwortung gezogen, belohnt oder bestraft werden kann. Wichtig für mein Argument ist nun, daß dieser frühe Lernprozeß in einer Phase sich ereignet, in dem die Kinder noch kein episodisches Gedächtnis aufbauen können. Wir erinnern uns nicht an die ersten zwei bis drei Lebensjahre, weil in dieser frühen Entwicklungsphase die Hirnstrukturen noch nicht ausgebildet sind, die zum Aufbau eines episodischen Gedächtnisses erforderlich sind. Es geht dabei um das Vermögen, Erlebtes in raum-zeitliche Bezüge einzubetten und den gesamten Kontext zu erinnern. Zwar kann auch ohne episodisches Gedächtnis gelernt werden, es fehlt aber dann die kontextuelle Einbettung des Gelernten: Man weiß das Gelernte, spürt das Erfahrene, aber weiß nicht, woher das Wissen, woher die Erfahrung kommt. Was Kleinkinder wissen, wissen sie *an sich*. Fragt man sie, woher sie dies oder jenes wissen, dann werden sie sagen, dies sei halt so, selbst wenn ihnen das Abgefragte erst vor kurzem beigebracht wurde. Diese frühkindliche Amnesie scheint mir dafür verantwortlich, daß die subjektiven Konnotationen von Bewußtsein für uns eine ganz andere Qualität haben als die Erfahrungen mit anderen sozialen Konstrukten. Vielleicht erleben wir

diese Aspekte unseres Selbst deshalb auf so eigentümliche Weise als von ganz anderer Qualität, als aus Bekanntem nicht herleitbar, weil die Erfahrung, so zu sein, in einer Entwicklungsphase installiert worden ist, an die wir uns nicht erinnern können. Wir haben an den Verursachungsprozeß keine Erinnerung. Und deshalb erscheinen uns die subjektiven Aspekte von Bewußtsein als immer schon dagewesen, als von aller Gebundenheit losgelöst, als alles Materielle transzendierende Entitäten, die jeder Verursachung entzogen sind und jedem reduktionistischen Erklärungsansatz trotzen.

Aus neurobiologischer Sicht liegt somit der Schluß nahe, daß auch die höheren Konnotationen von Bewußtsein, die wir mit unseren Konzepten von Freiheit, Identität und Verantwortlichkeit verbinden, Produkt eines evolutionären Prozesses sind, der zunächst Gehirne hervorgebracht hat, die in der Lage waren, eine Theorie des Geistes zu erstellen, mentale Modelle der Befindlichkeit des je anderen zu entwerfen. Dies und die Herausbildung differenzierter Sprachen ermöglichte die Entwicklung von Kommunikationsprozessen, die schließlich zur Evolution menschlicher Kulturen führte und zur Emergenz der nur den Menschen eigenen subjektiven Aspekte von Bewußtsein. Wenn dem so ist, wenn also die subjektiven Konnotationen von Bewußtsein Zuschreibungen sind, die auf Dialogen zwischen sich wechselseitig spiegelnden Menschen gründen, dann ist zu erwarten, daß die Selbsterfahrung von Menschen kulturspezifische Unterschiede aufweist. Auch kann nicht ausgeschlossen werden, daß bestimmte Inhalte dieser Selbsterfahrung, z. B. die Überzeugung, frei entscheiden zu können, illusionäre Komponenten haben. Im Bezugssystem neurobiologischer Beschreibungen gibt es keinen Raum für sub-

jektive Freiheit, weil die je nächste Handlung, der je nächste Zustand des Gehirns immer determiniert wäre durch das je unmittelbar Vorausgegangene. Variationen wären allenfalls denkbar als Folge zufälliger Fluktuationen. Innerhalb neurobiologischer Beschreibungssysteme wäre das, was wir als freie Entscheidung erfahren, nichts anderes als eine nachträgliche Begründung von Zustandsänderungen, die ohnehin erfolgt wären, deren tatsächliche Verursachungen für uns aber in der Regel nicht in ihrer Gesamtheit faßbar sind. Nur ein Bruchteil der im Gehirn ständig ablaufenden Prozesse ist für das innere Auge sichtbar und gelangt ins Bewußtsein. Unsere Handlungsbegründungen können folglich nur unvollständig sein und müssen a posteriori-Erklärungen miteinschließen.

Hier also haben wir ein weiteres Beispiel dafür – die moderne Physik hält weitere bereit – daß naturwissenschaftliche Erklärungsmodelle mit subjektiven Erfahrungen und auf Intuition beruhenden Überzeugungen in krassem Widerspruch stehen können. Die Rezeptionsgeschichte der heliozentrischen Kosmologielehre und der Darwinschen Evolutionstheorie legen nahe, daß sich schließlich die naturwissenschaftlichen Beschreibungen gegen Überzeugungen durchsetzen, die auf unmittelbarer Wirklichkeitserfahrung beruhen und daß wir uns letztlich an die neuen Sichtweisen gewöhnen. Ob dies auch der Fall sein wird für Erkenntnisse, die unser Selbstverständnis noch nachhaltiger verändern als die vorangegangenen wissenschaftlichen Revolutionen, muß die Zukunft beantworten. Unaufschiebbar werden jedoch schon jetzt Überlegungen über die Beurteilung von Fehlverhalten, über unsere Zuschreibungen von Schuld und unsere Begründungen von Strafe.