

# In: Science+Fiction. Begleitpublikation zur Ausstellung. VolkswagenStiftung 2002

ANDREAS K. ENGEL

DIE BIOLOGISCHEN WURZELN DES GEISTES: HERAUSFORDERUNGEN DER MODERNEN HIRNFORSCHUNG

Die moderne Hirnforschung ist auf dem besten Wege, eine der spannendsten Fragen zu beantworten, die man an die Naturwissenschaft stellen kann: die Frage nach den biologischen Wurzeln der Intelligenz. Ganz offensichtlich ist intelligentes Handeln eine Leistung unseres Nervensystems. Als Organ der Informationsspeicherung und Informationsverrechnung ermöglicht das Gehirn es uns, die Welt und uns selbst wahrzunehmen und zu erkennen, Neues zu lernen und Vergangenes zu erinnern, Pläne zu machen und Probleme zu lösen. Besonders wichtig für intelligentes Verhalten ist die Fähigkeit, über reflexhafte Reizantworten hinauszukommen, sich von der aktuellen Situation durch Generalisierung und Abstraktion zu lösen und Strategien zur Problemlösung zu entwickeln, die sich auch auf andere, neue Situationen anwenden lassen.

Welche Eigenschaften aber sind es, die Nervensysteme so intelligent und damit ihre Träger so erfolgreich machen? Mit der Frage nach den Mechanismen der natürlichen Intelligenz stellt sich auch die nach der Möglichkeit künstlicher Intelligenz: Warum eigentlich sind die heute herstellbaren Rechner und Roboter noch nicht zu intelligentem Handeln fähig? Was unterscheidet biologische Organismen von Maschinen und Gehirne von Computern? Gibt es hier prinzipielle Grenzen oder ließen sich die Mechanismen der natürlichen Intelligenz – wenn sie denn identifiziert werden könnten – in Artefakten eines Tages implementieren? Hier ergeben sich aus der aktuellen neurobiologischen Forschung eine Reihe neuer und interessanter Antworten. Rasante Fortschritte sind in der Hirnforschung einerseits aufgrund neuer Methoden möglich, mit denen das lebende Gehirn in seinen Funktionen untersucht werden kann, andererseits aber auch aufgrund neuer Konzepte und Denkweisen, mit denen die Forscher ihren Erkenntnisgegenstand betrachten. Beides will die Ausstellung in ihrem neurowissenschaftlichen Schwerpunkt in Ausschnitten deutlich machen.

## Was macht Gehirne leistungsfähig?

Wie sich aus neuen Erkenntnissen der Hirnforschung lernen lässt, sind für die komplexen Leistungen der höher entwickelten Nervensysteme eine ganze Reihe von Faktoren ausschlaggebend. Interessanterweise ist für die meisten der nachstehend genannten Merkmale heute noch gar nicht abzusehen, wie man solche Eigenschaften in Artefakten überhaupt verwirklichen könnte. Zu erwarten ist jedoch, dass in künstlichen Systemen erst dann komplexe kognitive Verarbeitungsprozesse auftreten, wenn sie einige dieser architektonischen Schlüsseleigenschaften mit den biologischen kognitiven Systemen teilen.

Wie wir heute wissen, sind Parallelität und eine verteilte Architektur hervorstechende Merkmale natürlicher intelligenter Systeme. Dies bedeutet, dass die Verarbeitung nicht durch einen einzelnen Prozessor geleistet wird, sondern auf verschiedene Subsysteme oder Module verteilt ist, was Geschwindigkeit und Robustheit der Systemleistung enorm steigert. Ein sehr gut untersuchtes Beispiel hierfür bietet das Sehsystem der höheren Säugetiere, für das inzwischen über 30 Hirnrindenareale beschrieben wurden, die unterschiedliche Arten von Informationen gleichzeitig verarbeiten. Dies zeigt, dass Repräsentation und Speicherung von Sinnesinformation auf ausgedehnten neuronalen Netzwerken beruhen. Charakteristisch für natürliche kognitive Systeme ist zudem, dass es keine zentralen Steuerungsprozesse gibt, die Verhalten hervorbringen. Auch die Handlungssteuerung basiert auf Prozessen der Selbstorganisation, in denen sich in biologischen Netzwerken nach lokalen Regeln globale Ordnung herstellt.

Ein weiteres entscheidendes Merkmal ist die Aktivität und Selektivität natürlicher kognitiver Systeme, die nicht einfach passiv Information aus ihrer Umgebung aufnehmen, sondern selbst – bezogen auf ihre Bedürfnisse – Hypothesen über die Welt generieren und diese gezielt im Wahrnehmungsprozess abfragen. Klassische Modelle nahmen an, dass kognitive Systeme weitgehend als serielle Reiz-Reaktions-Maschinen verstanden werden können. Derzeit treten neue Konzepte stark in den Vordergrund, die die Rolle von intrinsisch generierter Aktivität betonen

und nach den neuronalen Korrelaten von antizipativen Prozessen suchen. Das Bild, das sich aus aktuellen Forschungsergebnissen ergibt, ist das eines selektiven und aktiven Gehirns, das auf der Basis von intern generierten Erwartungshaltungen gezielt die Umwelt »abfragt« und so sehr schnell und effizient zu Problemlösungen kommt. Wichtige Faktoren, die diese intern erzeugte Aktivität anstoßen, sind neben den Gedächtnissystemen vor allem auch Emotion und Motivation. Diese beiden Faktoren, die zur Bewertung von Umweltereignissen führen, spiegeln die vitalen Interessen eines Organismus wider und sind damit entscheidend für sensomotorische Auswahlprozesse.

Eines der wichtigsten Arbeitsfelder der neueren Hirnforschung ist die Untersuchung der Mechanismen neuronaler Plastizität, also der Veränderbarkeit der neuronalen Netze. Die Tatsache, dass sich neuronale Systeme in ihrer organischen Struktur verändern können, markiert einen wichtigen Unterschied zu technischen Systemen. Ein entscheidender Punkt hierbei ist, dass die plastischen Veränderungen erfahrungs- und aktivitätsabhängig sind und in Auseinandersetzung mit den von der Umwelt gelieferten Reizen stattfinden. Man kann sagen, dass der »Gebrauch« hier die Struktur prägt, während in den heutigen Rechnern »Software« bislang noch nicht fähig ist, die »Hardware« zu verändern. Nach dem, was wir heute wissen, ist kaum vorstellbar, dass Intelligenz und Bewusstsein ohne solche Lernvorgänge in einem Organismus auftreten können. Diese Plastizität verleiht dem Gehirn ein hohes Maß an Anpassungsfähigkeit in Zeiträumen von Stunden, Tagen und Jahren. Die Adaptivität auf kürzeren Zeitskalen verdankt das Gehirn der Tatsache, dass es ein dynamisches System ist. Dies heißt zum einen, dass der Fluss der neuronalen Aktivität – also die Art und Weise, wie im Muster elektrischer Impulse Informationen verrechnet werden – rasch veränderbar ist; diese Veränderungen ergeben sich in direkter Kopplung mit den Veränderungen in der Umwelt. Zum anderen bedeutet dies, dass die Zeit als Kodierungsdimension systemintern für die Verarbeitung von Information nutzbar gemacht wird. Eine der vielleicht wichtigsten Quellen der Intelligenz besteht schließlich in der Sozialisierbarkeit von Gehirnen, also der Fähigkeit, Teil einer Gemeinschaft von Individuen zu werden. Die Einbettung in intersubjektive Kontexte ist ganz offensichtlich für kognitive Verarbeitungsprozesse, für die Entstehung von Intelligenz und Bewusstsein von entscheidender Bedeutung. Insbesondere für die Herausbildung von Ichbewusstsein (die Entstehung eines »Selbstmodells«) ist die Interaktion von Gehirnen eminent wichtig, sie ermöglicht eine wechselseitige Repräsentation verschiedener Individuen und eine interne »Simulation« von Ko-Subjekten. Diesen Sachverhalt bezeichnen die Psychologen als »theory of mind« – also die Fähigkeit sich vorzustellen, was in einem anderen Menschen vorgeht, und daher auch Vorhersagen über die Handlungen eines anderen zu machen.

Diese Schlüsselmerkmale intelligenter Systeme, die von der aktuellen Forschung in den Vordergrund gerückt werden, definieren auch die Arbeit eines interdisziplinären Schwerpunktprogramms, mit dem die VolkswagenStiftung unter dem Titel »Dynamik und Adaptivität neuronaler Systeme« neurowissenschaftliche Forschung fördert. Der neurowissenschaftliche Bereich der Ausstellung greift diese Themen auf drei Beschreibungsebenen auf, die sich methodisch unterscheiden, inhaltlich aber eng miteinander verzahnt sind: erstens die Ebene der Bildgebung, die die Frage der Lokalisation und Modularität von Hirnfunktionen untersucht; zweitens, die zelluläre und molekulare Ebene, auf der Funktion und Veränderbarkeit einzelner Nervenzellen dargestellt werden; und drittens, die Ebene der neuronalen Netze, die die Dynamik größerer Nervenzellverbände thematisiert und damit zwischen den beiden anderen Beschreibungsebenen vermittelt.

### Module des Geistes

Die modernen Methoden der anatomischen und funktionellen Bildgebung erlauben neue Einblicke in die materiellen Grundlagen des menschlichen Geistes. Es ist abzusehen, dass diese Erkenntnisse unser Selbstverständnis tief greifend verändern werden, da die Hirnforschung hier auch zu den Wurzeln von bislang kaum erforschten und spezifisch menschlichen Fähigkeiten vordringt.

Den Hirnforschern steht heute ein großes Spektrum von bildgebenden Methoden zur Verfügung. Diese liefern sehr unterschiedliche Daten, die sich für die Kartierung von Hirnfunktionen hervorragend ergänzen. Die Positronen-Emissions-Tomographie (PET) erfordert die Injektion schwach radioaktiver Substanzen in die Blutbahn, deren Zerfall zur Erzeugung eines Bildes der Aktivität im Hirngewebe genutzt werden kann. Die beste räumliche Auflösung liefert heute die Magnet-Resonanz-Tomographie (MRT), bei der die Anregung durch starke Magnetfelder zur Darstellung unterschiedlicher Gewebeschichten im Hirn genutzt wird, wie etwa der grauen und der weißen Substanz in den Großhirnhemisphären. Mit einer Variante dieser Methode, der funktionellen MRT (fMRT), können Erhöhungen der regionalen Durchblutung im Gehirn gemessen werden, die mit der Aktivierung der entsprechenden Regionen einhergehen. Mit Hilfe der Elektro- bzw. Magnetoencephalographie (EEG bzw. MEG) können schließlich elektrische und magnetische Felder registriert werden, die bei Hirnaktivität entstehen. Diese Methoden ergänzen die PET bzw. fMRT durch ihre hervorragende zeitliche Auflösung. Bereits in den 1980er Jahren konnten mit Hilfe der PET geistige Leistungen umschriebenen Hirnarealen zugeordnet werden. Damit flammte auch die alte Debatte zwischen Lokalisationismus und Holismus wieder auf. Während die Verfechter der ersten Auffassung von einer strikten Ortsgebundenheit geistiger Aktivität ausgehen, sind die Vertreter des Holismus der Meinung, es werde stets das ganze Gehirn beteiligt. Neue Ergebnisse weisen hier den Weg zu einem Kompromiss. Zwar besitzt jede Hirnregion eine oder mehrere spezifische Funktionen, jedoch ist für alle komplexen kognitiven Abläufe das Zusammenwirken verschiedener Areale in einem sehr ausgedehnten Netzwerk erforderlich. Dies gilt insbesondere für Prozesse wie Wahrnehmung, Bewusstsein, Emotion und Sprache, die im Brennpunkt der aktuellen Forschung stehen. Die Ausstellung zeigt hierzu exemplarisch Ergebnisse aus Forschungsarbeiten. Die Beispiele illustrieren dabei nicht nur experimentell gewonnene Daten, sondern veranschaulichen auch die »Bildersprache«, in der sich dieser Bereich der Hirnforschung heute bewegt.

#### Wandelbare Bausteine

Die Nervenzellen (Neurone) sind als elementare Bausteine des Gehirns für die Signalverarbeitung zuständig, die allen geistigen Abläufen und allen Verhaltensweisen von Menschen und Tieren zugrunde liegt. Der Aufbau der Neurone spiegelt ihre Funktion wider. Vom Zellkörper, der lebenswichtige Stoffwechselorganellen enthält, gehen zwei Arten von Fortsätzen aus. Die baumartig verzweigten Dendriten stellen den Eingangsbereich dar, in dem das Neuron Signale von anderen Zellen empfängt. Die von der Zelle verrechnete Information wird über das Axon weitergegeben. Durch Kontakte, so genannte Synapsen, sind die Neurone untereinander verbunden. In den Synapsen werden Botenstoffe (Neurotransmitter) freigesetzt, die erregend oder hemmend auf die nachgeschaltete Zelle wirken. Etwa 10000 Synapsen münden auf jede der 100 Milliarden Zellen des menschlichen Gehirns. Die Ausformung der neuronalen Verzweigungen und ihrer Kontakte ist von Lernvorgängen abhängig und damit im Detail von Zelle zu Zelle verschieden. So entsteht die enorme, bislang für technische Systeme unerreichbare Komplexität der Rechenleistungen, die im Nervensystem auf kleinstem Raum stattfinden.

Die bereits hervorgehobene Plastizität des Gehirns erklärt sich daraus, dass sich in Abhängigkeit von der Intensität des Signalfusses die Verknüpfung der Neurone untereinander verändert. Diese Veränderbarkeit liefert die Grundlage für Lernen, Gedächtnis und die Ausprägung von Individualität. Das Ausmaß der Plastizität unterscheidet sich dabei beträchtlich zwischen heranwachsendem und ausgereiftem Gehirn. Während der Entwicklung wachsen neue Faserbahnen, nicht gebrauchte werden selektiv abgebaut. Im erwachsenen Gehirn verändert sich dagegen im Wesentlichen nur die funktionelle Wirksamkeit der einzelnen Synapsen. Die Stärke der synaptischen Aktivierung ist dabei vom Muster der eingehenden Impulse abhängig: Wenn mehrere Impulse kurz nacheinander über dasselbe Axon oder gleichzeitig über verschiedene Fortsätze eintreffen, wird wesentlich mehr Neurotransmitter freigesetzt. Findet wiederholt eine solche starke Aktivierung statt, so kommt es zu einem Phänomen, das als Langzeitpotenzierung bezeichnet wird. Hierbei handelt es sich um eine langanhaltende Erhöhung der synaptischen Übertragungs-

stärke. Sie geht neben dem erhöhten Transmitterausstoß auch mit anatomischen Veränderungen einher, durch die sich die Fläche vergrößert, die für die synaptische Übertragung auf die nachgeschaltete Zelle zur Verfügung steht. Dies wurde erst kürzlich mit der Methode der Zwei-Photonen-Lasermikroskopie gezeigt, mit deren Hilfe die Neurowissenschaftler auch am lebenden Gehirn zu immer feineren Details vordringen können.

### Dynamische Netze

Auf der Ebene der neuronalen Netze versucht die Hirnforschung zu erklären, wie Kognition und Verhalten aus der Wechselwirkung der einzelnen Nervenzellen entstehen können. Wie bereits erwähnt, zeigen die Teilsysteme des Gehirns eine hochgradig parallele und arbeitsteilige Struktur, in der viele Areale untereinander zu einem komplexen Netzwerk verknüpft sind. Dies wirft die Frage auf, wie die durch zahlreiche Verarbeitungskanäle geleiteten Informationen wieder zusammengefasst werden können – ein Problem, das im Bereich der Wahrnehmung seit den Gestaltpsychologen bekannt ist und heute als »Bindungsproblem« bezeichnet wird. Am Beispiel dieses Problems lassen sich aktuelle Forschungsansätze im Bereich der netzwerkorientierten Hirnforschung verdeutlichen.

Einer der gegenwärtig diskutierten Lösungsvorschläge für dieses Problem besagt, dass die bereits angesprochene neuronale Dynamik der Schlüssel zum Verständnis solcher Netzwerkfunktionen sein könnte. Diese Hypothese geht davon aus, dass zusammengehörige Informationen dadurch integriert werden, dass die jeweils aktiven Neurone sich durch eine Synchronisation ihrer Impulse zu so genannte Assemblies (also kohärent aktiven Zellverbänden) zusammenschließen. Trifft diese Hypothese zu, so wäre das synchrone Feuern der Hirnrindenneurone beispielsweise die Ursache für die ganzheitliche Struktur unserer Wahrnehmungseindrücke. Die zeitlichen Korrelationen würden nämlich die Zusammengehörigkeit der Merkmale eines Objektes repräsentieren und wären auf diese Weise für die Erzeugung eines kohärenten Perzepts von entscheidender Bedeutung.

In zahlreichen Arbeiten wurde inzwischen nachgewiesen, dass die Neurone in sensorischen und motorischen Systemen ihre elektrischen Impulse sehr präzise synchronisieren können. Viele Forschungsergebnisse weisen zudem darauf hin, dass diese zeitlichen Muster tatsächlich bedeutsam für die perzeptive Integration und auch für die Verknüpfung von sensorischen und motorischen Signalen sind. Beim Menschen lassen sich solche Synchronisationsphänomene durch die Messung von Hirnströmen mit Hilfe von EEG und MEG nachweisen. Diese Untersuchungen zeigen, dass Wahrnehmung, Aufmerksamkeit und Bewusstsein mit erhöhter neuronaler Synchronisation einhergehen. So wurde zum Beispiel beschrieben, dass schnelle Hirnwellen (so genannte Gamma-Wellen), die eine präzise neuronale Synchronisation reflektieren und im Wachzustand gut nachweisbar sind, in tiefer Narkose verschwinden. EEG-Studien haben darüber hinaus bestätigt, dass erhöhte Wachheit und Aufmerksamkeit mit einer verstärkten Synchronisation im Gamma-Band einhergehen, und neuere Untersuchungen zeigen, dass eine Gamma-Synchronisation mit der bewussten Verarbeitung sensorischer Reize korreliert. Diese Ergebnisse machen es sehr wahrscheinlich, dass die Synchronisation eine wichtige Grundlage für dynamische Integrationsprozesse in neuronalen Netzen darstellt.

### Ignorabimus?

Die neuen Ergebnisse der Hirnforschung und die rasche Expansion des hier entstehenden Faktenwissens werfen die Frage nach der Erklärungsmächtigkeit der neurowissenschaftlichen Ansätze auf. Ihrem Selbstverständnis zufolge versucht sich die Kognitionswissenschaft an einer naturalistischen Theorie mentaler Prozesse, die sie aus Vorgängen im Gehirn erklären möchte. Wie eingangs dargestellt, könnte es durchaus möglich sein, die biologischen Faktoren aufzudecken, die zu intelligentem Handeln führen. Aber kann die Hirnforschung tatsächlich auf empirischem Wege auch Bewusstsein, Selbst und Subjektivität erklären? Oder tritt spätestens hier eine fundamentale Erklärungslücke auf, die aus einer unüberbrückbaren Kluft zwischen dem Objektiven und dem Subjektiven resultiert, aus einer mangelnden Übersetzbarkeit der

Erlebnisperspektive der »1. Person« in die »3. Person«-Perspektive, mit der die Wissenschaft zwangsläufig operiert?

Aus einer Reihe von Gründen scheint es zweifelhaft, dass die Hirnforschung eine wirklich umfassende Theorie kognitiver Prozesse liefern könnte, dass sich also in neurobiologischen Beschreibungen alles Relevante über Geistiges sagen ließe. Zum einen gibt es hier möglicherweise ein »Kontext-Problem«. Selbst bei vollständiger Kenntnis der Hirnzustände einer Person könnte man – nur aus der Kenntnis dieser Daten – nicht sagen, was diese denkt oder fühlt. Der Grund hierfür liegt darin, dass die Bedeutung dieser Zustände nur mit Bezug auf die Umgebung dieser Person und relativ zu der Situation, in der sie sich befindet, definiert ist. Hirnzustände per se haben keinen Gehalt, da sie nur mit Bezug auf den situativen Kontext individuierbar sind. Hieraus ergibt sich, dass mentale Prozesse nicht »individualistisch« erklärt werden können, d.h. nicht ausschließlich durch Bezug auf innere Zustände eines kognitiven Systems, wie es die Neurobiologie in der Regel versucht.

Darüber hinaus kann eingewendet werden, dass kognitive Prozesse nicht ausschließlich durch Bezug auf die sub-personale Ebene adäquat beschrieben werden können. Die Tatsache, dass kognitive Akte von Personen vollzogen werden, wird in der kognitiven Neurobiologie gerne übersehen, die durchaus davon spricht, dass das Gehirn Objekte »erkenne« und das Sehsystem Szenen »interpretiere«. Dies ist ein unzulässiger Sprachgebrauch (in der philosophischen Debatte gelegentlich als »homunculus fallacy« bezeichnet), der durch Kategorienfehler Pseudolösungen vortäuscht. Schließlich steht einem konsequent reduktionistischen Ansatz das ungelöste Problem des subjektiven Erlebens (das »Qualia-Problem«) im Wege: Nur aus der Kenntnis der Hirnzustände, die einen Schmerz begleiten, kann man nicht entnehmen, wie Schmerz sich anfühlt. Es könnte also etwas irreduzibel Subjektives geben, was möglicherweise nicht in eine naturwissenschaftliche Hirntheorie abgebildet werden kann.

Obwohl die methodischen und inhaltlichen Fortschritte der Kognitionswissenschaften sicherlich berechtigten Anlass zum Optimismus geben, wäre also von der Anmeldung voreiliger Hegemonialansprüche gegenüber den Geistes- und Sozialwissenschaften abzuraten. Dennoch darf das »Ignorabimus«, das der Berliner Neurophysiologe Emil Du Bois-Reymond Ende des 19. Jahrhunderts in einer berühmt gewordenen Rede konstatierte, heute mit einem Fragezeichen versehen werden. Möglicherweise liefern Alleingänge der Naturwissenschaften keine Erklärung des Mentalen, daraus folgt jedoch keineswegs, dass Bewusstsein und Subjektivität sich wissenschaftlicher Theoriebildung im Gesamten entzögen.