

Wolfgang Prinz

Wie das Bewusstsein erfunden wurde

In der Wissenschaft geht es immer um zweierlei: um gute Fragen und um gute Antworten. Wenn man gute Fragen hat, besteht die Kunst darin, durch methodische Anstrengung zu guten Antworten zu kommen – das ist also die *Kunst des Findens*. Sie ist der sichtbare und manchmal auch spektakuläre Teil der Forschung. Schwieriger ist die Lage, wenn man keine guten Fragen hat. Dann besteht die Kunst darin, gute Fragen erst einmal zu entwickeln – Fragen nämlich, die sich überhaupt mit wissenschaftlichen Mitteln traktieren lassen. Das ist die *Kunst des Suchens*, und sie macht den weniger sichtbaren, aber nicht minder wichtigen Teil der Forschung aus.

Vor dem Rätsel, von dem ich im Folgenden sprechen will – wie nämlich Subjektivität in die Welt kommt und was sie in einem im Übrigen subjektlosen Universum eigentlich verloren hat –, hat selbst diese Kunst bisher versagt. Nicht, dass wir überhaupt keine Fragen hätten – im Gegenteil. Sie bedrängen uns nachhaltig, weil sie die Grundlagen unseres Selbstverständnisses berühren. Aber wir haben keine *guten* Fragen, denn bisher ist es uns nicht gelungen, sie so zu disziplinieren, dass sie traktierbar wären – von guten Antworten ganz zu schweigen.

Dennoch glaube ich, dass die Zeit für einen solchen Disziplinierungsversuch gekommen ist. Ich will ihn im Folgenden in drei Schritten beschreiben. Zuerst wird es darum gehen, die Ausgangsfrage nach dem Status von Subjektivität *in traktierbare Form* zu überführen. Im zweiten Schritt wird es dann darum gehen, die *Konturen einer Antwort* auf einige dieser Fragen zu skizzieren. Dabei werde ich, wie es sich für einen Psychologen gehört, zwar die *repräsentationalen Grundlagen* von Subjektivität in den Mittelpunkt stellen, daneben aber auch auf ihre naturgeschichtlichen und kulturgeschichtlichen Wurzeln eingehen, also auch biologischen und historischen Perspektiven Rechnung tragen. In einem dritten Schritt betrachte ich schließlich gesellschaftliche Diskurse, in denen Subjektivität erzeugt und aufrechterhalten wird.

Subjektivität und Bewusstsein

Wieso ist Subjektivität ein Rätsel? Was unterscheidet eigentlich das Subjektive vom Objektiven? Wenn wir über Sterne und Steine, über Ozeane und Flüsse reden, wenn es um Viren, Algen oder Pilze, ja selbst um Farne, Gräser oder Bäume geht, glauben wir im Allgemeinen, dass alles, was über diese Dinge überhaupt zu sagen ist, mit den Mitteln von Physik, Chemie und Biologie gesagt werden kann. Das Gleiche mag auch noch für Quallen, Würmer und Schwämme gelten, aber dann wird es schwierig – vielleicht bei den Insekten, dann aber bei den Wirbeltieren und gewiss bei den Säugern, besonders den Primaten und, erst recht natürlich, beim *homo sapiens*. Von diesen Tieren glauben wir nämlich, dass mit den Mitteln dieser Wissenschaften *nicht* alles gesagt werden kann, was über sie zu sagen ist. Wir schreiben ihnen ein darüber hinausgehendes *Innenleben* zu, abgestuft vielleicht, aber im Prinzip ähnlich und verwandt mit unserem eigenen bewussten Erleben.

Subjektivität oder Bewusstsein – was soll man mit dieser neuen Qualität anfangen, die da irgendwo zwischen Fischen und Menschen entstanden zu sein scheint? *Eine* Möglichkeit besteht darin, sie zu leugnen – entweder ihre Existenz oder jedenfalls ihre wissenschaftliche Dignität. Bewusstseinserscheinungen, so wird oft argumentiert, sind keine beobachtbaren Erscheinungen, die man messen kann wie Farben, Längen oder elektrische Ladungen, sondern verborgene Eigenschaften, die man Lebewesen allenfalls *zuschreiben* kann, ohne dass man gleich glauben muss, dass sie diese Eigenschaften auch *haben* – mit der Folge, dass sie nicht Gegenstand objektiver Wissenschaften sein können.

Wer aber nicht leugnen mag, muss anerkennen – und handelt sich damit unangenehme Fragen ein, die schwer zu beantworten sind: Wie kommt Subjektivität in die Welt? Wozu ist sie gut? Und vor allem: In welcher Beziehung stehen die psychischen Qualitäten der subjektiven Welt zu den physischen Qualitäten der objektiven Welt, die uns aus Physik, Chemie und Biologie vertraut sind?

Diese Fragen haben die abendländische Philosophie seit über zweieinhalb Jahrtausenden in Atem gehalten. Das Rätsel Subjektivität lässt uns nicht los, und dass das so ist, liegt sicher nicht nur an den Philosophen, die es niemals von ihrer Agenda gestrichen haben, sondern auch an uns selbst: daran, dass wir selbst die Systeme *sind*, die wir *untersuchen*. Dieser Umstand bringt es mit sich, dass wir nicht davon lassen können, uns immer wieder Gedanken darüber zu machen, wie es mit dem Bewusstsein bestellt ist.

Was eine Theorie des Bewusstseins leisten muss

In den letzten Jahren sind eine Reihe von Büchern erschienen, die uns versprechen, das letzte noch verbliebene Menschheitsrätsel mit naturwissenschaftlichen Mitteln zu lösen: das Rätsel der Subjektivität, namentlich das Rätsel, *wie Bewusstsein aus Gehirntätigkeit entsteht*.¹ Vieles lesen wir darüber, was die moderne Neurobiologie über Gehirnfunktionen weiß: über die Chemie, die Physik, die Physiologie und nicht zuletzt die neuroinformatische Modellierung neuronaler Aktivität auf verschiedenen Ebenen. Weiteres lesen wir darüber, was die moderne Psychologie über Informationsverarbeitungsprozesse weiß: über funktionale Systemarchitekturen und Operationen, die auf ihnen arbeiten. Und zuletzt finden wir schließlich Mutmaßungen darüber, *welche* der so beschriebenen Strukturen und Prozesse es denn sind, die *Bewusstsein hervorbringen* oder *mit Bewusstsein verbunden sind*.

Aber: So faszinierend diese Lektüre auch sein mag, bleibt doch meist ein ausgesprochen schales Gefühl zurück. Man hat nämlich nicht den Eindruck, dass man dadurch besser versteht, warum subjektive Phänomene so sind, wie sie sind, d.h. warum sie gerade jenen Charakter tragen, der sie auszeichnet: bewusst zu sein. Und man wird das Gefühl nicht los, dass man selbst dann, wenn man schon genau wüsste, worin die neurobiologischen Korrelate und/oder die psychologischen Funktionen von Bewusstseinserscheinungen bestehen, noch lange nicht verstünde, warum diese physiologischen Prozesse *just diese besondere Qualität hervorbringen* bzw. die psychologischen Funktionen *just diese besondere Qualität erfordern* – und nicht irgendeine andere. Und ebenso wenig verstünde man umgekehrt, warum *bestimmte* physiologische Prozesse diese besondere Qualität hervorbringen, *andere* dagegen nicht. Was unterscheidet die Prozesse, die über dieses Potential verfügen, von denen, die

1 Vgl. z.B. William H. Calvin, *The cerebral's symphony: Seashore reflections on the structure of consciousness*, 1989, (dt.: *Die Symphonie des Denkens: Wie aus Neuronen Bewußtsein entsteht*, München/Wien 1993); Gerald M. Edelman, *The remembered present: A biological theory of consciousness*, New York 1989; William H. Calvin und George E. Ojemann, *Conversation with Neil's brain*, New York 1994, (dt.: *Einsicht ins Gehirn: Wie Denken und Sprache entstehen*, München/Wien 1995); Francis Crick, *The astonishing hypothesis – The scientific search for the soul*, New York 1994, (dt.: *Was die Seele wirklich ist – Die naturwissenschaftliche Erforschung des Bewußtseins*, München 1994); António R. Damásio, *Descartes' error: Emotion, reason and the human brain*, New York 1994, (dt.: *Descartes' Irrtum: Fühlen, Denken und das menschliche Gehirn*, München 1995); Gerhard Roth, *Das Gehirn und seine Wirklichkeit, Kognitive Neurobiologie und ihre philosophischen Konsequenzen*. Frankfurt 1994; Gerhard Roth und Wolfgang Prinz, *Kopf-Arbeit. Gehirnfunktionen und kognitive Leistungen*, Heidelberg 1996.

nicht darüber verfügen? Keine der derzeit angebotenen Theorien über den Zusammenhang zwischen Hirnfunktionen und Bewusstsein hat eine Antwort auf derartige Fragen anzubieten. Was sie anbieten, sind *formale Korrelationszusammenhänge*, die man mit Staunen zur Kenntnis nehmen, nicht aber *inhaltliche Fundierungszusammenhänge*, die man wirklich verstehen kann. Man könnte auch sagen, dass sie bestenfalls das Dasein, nicht das Sosein von Subjektivität erklären.

Was aber brauchen wir, um das Sosein von Subjektivität zu verstehen? Dazu müssen wir zweierlei klären: erstens die Natur des Explanandums: Was heißt eigentlich *bewusst*? Und zweitens: Wie hätte eine befriedigende Erklärung von Bewusstseinserscheinungen denn auszusehen? Was muss eine Fundierungstheorie des Bewusstseins eigentlich leisten?

Bewusstsein beschreiben

Was genau heißt eigentlich *bewusst*? Wer diese Frage stellt, betritt vermintes Gelände. Alle Termini, alle Begriffe und alle Ideen, die jemals zur Eingrenzung und Abgrenzung in die Diskussion gebracht worden sind, haben eine verworrene Geschichte und eine verwirrende Gegenwart.

Wenn ich im Folgenden von Bewusstsein rede, meine ich damit nicht den Zustand von Personen, d. h. dass jemand bei Bewusstsein ist, oder sich im Zustand von Bewusstlosigkeit befindet. Hier soll dagegen ausschließlich vom bewussten Charakter *mentaler Inhalte* die Rede sein. Was ist damit gemeint? Zunächst weist die Rede vom bewussten Charakter mentaler Inhalte darauf hin, dass wir niemals in reiner, inhaltsleerer Form bei Bewusstsein sein können. Immer sind es bestimmte Inhalte, die uns bewusst sind, und am *bewusst-sein dieser Inhalte* erkennen wir, dass wir *selbst bei Bewusstsein* sind. Ähnlich wie Personen bewusst oder bewusstlos sein können, können auch mentale Inhalte bewusst oder nicht-bewusst sein. Wir bezeichnen sie dann, je nach theoretischem Geschmack, z. B. als *unbewusste, vorbewusste*, oder *nicht bewusstseinsfähige* Inhalte.

Wie aber sind *bewusste* mentale Inhalte möglich und worin genau besteht ihr bewusster Charakter? Bei der Beantwortung dieser Frage will ich mir eine Charakterisierung zu eigen machen, die vor über 100 Jahren der Philosoph Franz Brentano, einer der Wegbereiter der modernen Phänomenologie, gegeben hat und der sie in einem entscheidenden Punkt erweitern.² Brentano

² Vgl. hierzu Franz Brentano, *Psychologie vom empirischen Standpunkt* (Bd. 1), Leipzig 1924 (Erstausgabe 1874).

erörtert die Natur psychischer Akte an einem denkbar einfachen Beispiel: Was geschieht eigentlich, wenn wir einen Ton hören? Was ist es, das den bewussten Charakter dieses Ereignisses ausmacht? Nach Brentano sind in diesem psychischen Akt zwei Inhalte miteinander verwoben: der *Ton*, den wir hören und die Tatsache, dass wir ihn *hören*. Allerdings sind diese beiden mentalen Inhalte nicht in gleicher Weise repräsentiert. Der Ton ist das primäre Objekt des Hörens; ihn können wir im psychischen Akt direkt beobachten. Das Hören selbst ist das sekundäre Objekt des psychischen Aktes. Von ihm sagt Brentano, dass es nicht direkt beobachtet werden kann, sondern in einer anderen, indirekten Form zu Bewusstsein gelangt: »Die Töne, die wir hören, können wir beobachten, das Hören der Töne können wir nicht beobachten; denn nur im Hören der Töne wird das Hören selbst mit erfasst.«

Soweit Brentano. Will man die Struktur psychischer Akte allerdings erschöpfend charakterisieren, muss man noch einen Schritt weitergehen. Wenn nämlich zutrifft, dass im Hören des Tons nicht nur der Ton selbst, sondern implizit auch sein Hören enthalten ist, dann muss das *Subjekt*, das da hört, in abermaliger Verschachtelung auch in dem Akt enthalten sein. Denn ebenso wenig wie ein Ton vorstellbar ist ohne ein Hören, das auf ihn gerichtet ist, ist ein Hören vorstellbar ohne ein mentales Ich oder Subjekt, das da hört. Bewusste mentale Akte sind also dadurch charakterisiert, dass in ihnen das mentale Ich implizit gegenwärtig ist.³

Bewusstsein erklären

Mit diesen Bestimmungen haben wir den *Gegenstand*, um den es geht, soweit eingegrenzt, dass wir uns jetzt auf die Suche nach geeigneten *Erklärungen* machen können. Wenn nämlich zutrifft, dass die Bezogenheit mentaler Inhalte auf

³ Dem entspricht, dass die bewusste Repräsentanz einer Situation genau dann endet, wenn sich das Ich aus ihr verabschiedet. Wenn wir etwa während eines Spaziergangs in ein Gespräch verwickelt sind, das unsere ganze Aufmerksamkeit beansprucht, ist unsere bewusste Wahrnehmung auf den Inhalt des Gesprächs und auf die Gesprächssituation selbst konzentriert. Dies sind, um noch einmal mit Brentano zu sprechen, die primären Objekte, auf die das implizit anwesende Ich sich richtet. Nur sie sind es, die wir bewusst wahrnehmen. Andere Merkmale der Situation – nämlich die Szenerie, die wir durchschreiten – nehmen wir mit Bewusstsein nicht zur Kenntnis. Natürlich kann kein Zweifel darin bestehen, dass diese Informationen *verarbeitet* werden, denn andernfalls wäre nicht zu erklären, dass wir, obwohl ins Gespräch vertieft, voll in der Lage sind, unsere Schritte umgebungsgerecht zu steuern. Die Verarbeitung erzeugt aber keine *bewusste* Repräsentation, die auf das implizit anwesende Ich bezogen ist.

ein implizit anwesendes Ich die entscheidende Grundlage für die Ausbildung ihres bewussten Charakters ist, ist damit das Erklärungsproblem auf die Frage der Konstitution des mentalen Ich und seiner impliziten Repräsentation verschoben.

Benötigt werden also Theorien, die die Rolle des implizit anwesenden mentalen Ich aufklären. Hätte man diese Rolle verstanden, hätte man zugleich verstanden, wie der bewusste Charakter mentaler Inhalte zustande kommt. Da nämlich die Qualität der Bewusstheit nicht nur *dann entsteht, wenn* die Bedingung der impliziten Gegenwart des Ich erfüllt ist, sondern auch *darin besteht, dass* genau diese Bedingung erfüllt ist, hätten wir nämlich nicht nur verstanden, *unter welchen Bedingungen* bewusste mentale Inhalte entstehen, sondern auch, *warum* sie unter diesen Bedingungen *eben diese Qualität* annehmen und nicht irgendeine andere. Dann hätte man genau das, was man von einer Fundierungstheorie der Subjektivität verlangen muss.

Was aber könnte das heißen: die Rolle des mentalen Ich erklären? Was heißt es überhaupt, *eine Rolle zu erklären*? Rollen- oder Funktionserklärungen verlangen immer zweierlei – die Spezifikation von *Leistungen* und von *Mechanismen*, die diese Leistungen erbringen – ultimate und proximate Erklärungen also. Was das mentale Ich betrifft, müssen wir auf ultimativer Ebene klären, *was* es leistet, und auf proximativer Ebene, *wie* es diese Leistungen realisiert.

Wozu ist also das mentale Ich gut und wie wird es realisiert? Antworten auf diese Frage werden oft in Form von Geschichten angeboten, die erklären, wie es in grauer Vorzeit dazu kam, dass das mentale Ich erfunden wurde und welche Vorteile damit verbunden waren. Mit anderen Worten: Es werden in diesen Geschichten hypothetische Szenarien konstruiert, die es plausibel machen, dass bestimmte Lebewesen, die zunächst noch nicht über ein mentales Ich verfügen, *fitness*-Vorteile dadurch gewinnen, dass sie eine derartige Instanz ausbilden.

Bausteine für solche Szenarien sind in den letzten Jahren verschiedentlich angeboten worden, z. B. von Daniel Dennett, Julian Jaynes oder Thomas Metzinger.⁴ Bei allen Unterschieden konvergieren sie in der Idee, dass die Mechanismen, die bewusste Inhalte hervorbringen, eine *metarepräsentationale Struktur* aufweisen. Das heißt, dass sie (mindestens) zwei Repräsentations-

4 Vgl. Daniel C. Dennett, *The origin of selves*, Bericht Nr. 14/1990 der Forschungsgruppe »Kognition und Gehirn« am ZiF, Universität Bielefeld; ders., »The self as the center of narrative gravity«, in F. S. Kessel, P. M. Cole und D. L. Johnson, Hg., *Self and consciousness: Multiple perspectives*, Hillsdale 1992, S. 103–115; Edelman, *The remembered present* (s. Fn. 1); Julian Jaynes, *The origin of consciousness in the breakdown of the bicameral mind*, Boston, MA 1976; Thomas Metzinger, *Subjekt und Selbstmodell. Die Perspektive phänomenalen Bewußtsein vor dem Hintergrund einer naturalistischen Theorie mentaler Repräsentationen*, Paderborn 1993.

ebenen enthalten, die in Brentanos Analyse vorgezeichnet sind: eine für die Repräsentation der Inhalte selbst und eine weitere für die Repräsentation ihrer Bezogenheit auf das mentale Ich.

Im Folgenden skizziere ich ein psychohistorisches Szenario, das an diese Ideen anknüpft und das dazu bestimmt ist, die Konstitution des mentalen Ich zu rekonstruieren – gewissermaßen ein Steinzeitmärchen, das davon handelt, wie aus bewusstlosen Zombies bewusste Akteure werden.

Konstitution

Denken wir uns als Ausgangspunkt ein bewusstloses, zombieartiges Lebewesen. Obwohl es bewusstlos ist, ist es mit hochentwickelten kognitiven Fähigkeiten ausgestattet. Allerdings gibt es eine charakteristische Einschränkung: Es soll sich um ein *vorsprachliches* Lebewesen handeln, bei dem symbolische Kommunikation und Repräsentation nicht vorkommen.

Ein ungefähres Bild der kognitiven Leistungsfähigkeit unseres Zombies, zugleich aber auch der Grenzen seiner Leistungsmöglichkeiten, lässt sich wie folgt zeichnen: Auf der Haben-Seite können wir ihm – verkürzt gesagt – die Fähigkeit zuschreiben, die verhaltensrelevanten Implikationen der jeweils aktuellen Reizsituation zu bewerten und in entsprechendes Verhalten umzusetzen. Diese Bewertung mag auf der Grundlage komplexer Algorithmen erfolgen, die teils zu seinem genetisch angelegten Verhaltensrepertoire gehören, teils in Lernprozessen entstanden sein mögen. Weitere Algorithmen sorgen dafür, dass die Ergebnisse dieser Bewertungen gegen die aktuellen Prioritäten des Lebewesens abgeglichen werden und dass dieser Abgleich in Handlungsentscheidungen umgesetzt wird.

So komplex die Berechnungen auch sein mögen, die der Verhaltenssteuerung zugrunde liegen, unterliegen sie doch der prinzipiellen Beschränkung der *Kopplung an die jeweils aktuelle Situation*. Sie nehmen ihren Ausgang von der aktuellen Reizinformation, und sie bewerten Handlungsoptionen, die sich auf die aktuelle Situation beziehen. Keine Rolle spielen demgegenüber Prozesse, die symbolische Repräsentation voraussetzen, wie z. B. die Vergegenwärtigung vergangener oder die Planung zukünftiger Ereignisse. Unser Zombie ist sklavisch an die Gegenwart gebunden.

Wie kann nun aus solcher Bewusstlosigkeit Bewusstsein entstehen? Bewusstsein – so die Antwort, die ich vorschlage – kann sich aus dieser Ausgangssituation dann und nur dann entwickeln, wenn zwei Entwicklungsschritte aufeinanderfolgen. In einem ersten Schritt muss die Fähigkeit zur *Vergegenwärtigung abwesender Sachverhalte* entwickelt werden, verbunden mit der Fähig-

keit zur *getrennten Repräsentation von Wahrnehmung und Vergegenwärtigung*. Diese Voraussetzung bezeichne ich im Folgenden als *duale Repräsentation*. In einem zweiten Schritt muss dann hinzukommen, dass auch selbsterzeugte Vergegenwärtigungen *als Kommunikationshandlungen interpretiert* werden. Diese Voraussetzung bezeichne ich als *personale Attribution*.

Die erste dieser beiden Voraussetzungen betrifft die *Naturgeschichte* der Verhaltensorganisation und ihrer Realisierung durch das Gehirn. Die zweite betrifft dagegen die *Kulturgeschichte* der Spezies *homo sapiens*. Obwohl sich die beiden Entwicklungsschritte systematisch unterscheiden lassen, sind sie doch historisch so eng miteinander verknüpft, dass man sie nur aufeinander bezogen darstellen kann.

Duale Repräsentation

Wir erweitern jetzt den Lebenshorizont unseres fiktiven Lebewesens um einen entscheidenden Schritt, indem wir annehmen, dass der soziale Verband, in dem es lebt, einfache Formen *symbolischer Kommunikation* entwickelt.

Was ist dafür erforderlich? Betrachten wir z. B. den Fall einer Mitteilung, die sich auf einen Sachverhalt bezieht, der außerhalb des aktuellen Wahrnehmungshorizonts ihres Rezipienten liegt. Damit er eine derartige Mitteilung verstehen kann, muss er über die Fähigkeit verfügen, *Vergegenwärtigungen* auszubilden, d. h. *Repräsentationen von Sachverhalten, die gegenwärtig nicht wahrnehmbar sind*. Dabei muss er Vergegenwärtigungen abwesender Sachverhalte von Wahrnehmungen gegebener Sachverhalte unterscheiden können, denn es muss sichergestellt sein, dass diese Vergegenwärtigungen für das Handeln in der aktuellen Situation unschädlich sind.

Die gleichzeitige Repräsentation von vergegenwärtigten *neben* wahrgenommenen Inhalten macht eine tiefgreifende Erweiterung der kognitiven Verarbeitungsarchitektur erforderlich. Erforderlich wird jetzt eine Architektur, die zwischen Vordergrund- und Hintergrundverarbeitung unterscheidet und die es erlaubt, vorübergehend vergegenwärtigte Information im Vordergrund zu verarbeiten und gleichzeitig im Hintergrund die Verarbeitung der aktuellen Wahrnehmungsinformation fortzusetzen – jedenfalls so weit, dass elementare Grundfunktionen intakt bleiben, wie z. B. die Bewegungssteuerung oder auch Orientierungsreaktionen, mit denen der Organismus auf überraschende Reize reagiert.

Den Repräsentationsmodus, der mit dieser neuen Organisation der Informationsverarbeitung verbunden ist, bezeichne ich als *duale Repräsentation*. Ich verstehe darunter die Fähigkeit, wahrgenommene Inhalte und vergegen-

wärtigte Inhalte *nebeneinander* und *funktional getrennt* zu unterhalten – mit der Folge, dass zwischen Wahrgenommenem und Vergegenwärtigtem jederzeit unterschieden werden kann.

Personale Attribution

Wir haben bisher nur solche Vergegenwärtigungen betrachtet, die durch die Rezeption sprachlicher Mitteilungen angestoßen werden und insofern von außen induziert sind. Wenn eine duale Repräsentationsarchitektur ausgebildet ist, bietet sie aber auch Raum für die Induzierung von Vergegenwärtigungen von innen her, wie z. B. Gedanken, Erinnerungen oder Phantasien. Der Kürze halber verwende ich im Folgenden den Ausdruck *Gedanken* stellvertretend für alle Formen intern induzierter Vergegenwärtigung.

Intern erzeugte Gedanken unterscheiden sich aber in einem wichtigen Merkmal von extern induzierten Mitteilungen. Das Auftreten von Vergegenwärtigungen, die von außen durch sprachliche Mitteilungen angestoßen werden, ist stets von der *Wahrnehmung einer Kommunikationshandlung* begleitet, die selbst in der aktuellen Wahrnehmungssituation stattfindet. D.h. es gibt stets eine Person in der Umgebung des Rezipienten, und diese Person ist die wahrnehmbare Quelle der Mitteilung. *Gedanken* sind dagegen intern erzeugte Vergegenwärtigungen, die nicht von der Wahrnehmung einer Kommunikationshandlung begleitet sind. Sie können also nicht auf eine in der aktuellen Wahrnehmungssituation anwesende Person zurückgeführt werden.

Woher kommen also die Gedanken? Wer oder was erzeugt sie und wie können sie mit der aktuellen Wahrnehmungssituation verknüpft werden? Eine naheliegende Lösung für dieses Problem besteht darin, das Schema für die Interpretation extern induzierter Mitteilungen auch auf intern induzierte Gedanken zu übertragen. Das bedeutet, dass auch Gedanken auf personale Quellen zurückgeführt werden, und zwar auf Quellen, die in der aktuellen Wahrnehmungssituation gegenwärtig sein müssen.

Wie kann das geschehen? *Eine* mögliche Lösung dieses Attributionsproblems besteht darin, das Auftreten von Gedanken auf *Stimmen* zurückzuführen – Stimmen von *Göttern, Priestern, Königen oder Ahnen*, personalen Autoritäten also, von denen man glaubt, dass sie in der aktuellen Wahrnehmungsumgebung unsichtbar gegenwärtig sind. Eine *andere* Lösung lokalisiert die Quelle der Gedanken dagegen in einer eigenständigen personalen Instanz, die an den Körper des Akteurs gebunden ist: dem *eigenen Ich*.

Diese beiden Lösungen des Attributionsproblems unterscheiden sich in vielerlei Hinsicht: historisch, politisch und psychologisch. *Historisch* dürfte die

erste älter als die zweite Lösung sein. Der Übergang zwischen den beiden Lösungen und der mit ihnen verbundenen Mentalitäten ist Gegenstand der spekulativen Bewusstseinstheorie von Julian Jaynes. Folgt man Jaynes, ist dieser Übergang sogar in historischer Zeit geschehen: zwischen Ilias und Odyssee. In der Ilias ist nach Jaynes die Geistesverfassung der Protagonisten durchweg noch so strukturiert, dass sie Gedanken, Gefühle und Absichten nicht als Produkte eines eigenen Ich erleben, sondern als Eingebungen göttlicher Stimmen. Anders in der Odyssee: Odysseus verfügt über ein Ich, und dieses Ich ist es, das denkt und handelt. Jaynes ist der Überzeugung, dass das moderne Bewusstsein des Odysseus nur dadurch entstehen konnte, dass das Ich sich an die Stelle der Götter setzte und insofern ihre Nachfolge antrat.⁵

Dass ferner die *politischen* Implikationen der beiden Lösungen verschieden sind, liegt auf der Hand. Gesellschaften, deren Akteure ihre Handlungen auf Stimmen weltlicher oder überweltlicher Autoritäten zurückführen, werden Priester- und Adelseliten ausbilden, die für sich die Rolle natürlicher Autoritäten bzw. authentischer Interpreten solcher Autoritäten in Anspruch nehmen und daraus die Legitimation zur Ausübung von Herrschaft ableiten. Erst wenn das Ich an die Stelle der Götter tritt, werden diese Eliten obsolet, und autoritäre Konstruktionen werden durch Organisationsformen abgelöst, die die Quellen des Handelns in den Akteuren selbst verankern.

Ein wichtiger *psychologischer* Unterschied besteht schließlich darin, dass die Entwicklung eines Ich-Konzepts die Voraussetzung dafür schafft, dass Individuen sich als Personen mit kohärenter Biografie verstehen können. Einmal konstituiert, ist das Ich in *jedem* Vergegenwärtigungsvorgang als implizite personale Quelle gegenwärtig, und ähnlich wie es immer der gleiche Körper ist, der in jeder Wahrnehmungssituation anwesend ist, ist es auch das gleiche mentale Ich, das über Zeit und Situationen hinweg identisch bleibt.

Wenn dieses psychohistorische Szenario zutrifft, mutet es uns einiges zu. Wir müssen uns dann nämlich von der verbreiteten Ansicht verabschieden, dass das Ich eine naturgegebene, fundamentale *mentale Instanz* ist. Das Ich wäre dann vielmehr nichts weiter als ein *mentaler Inhalt*, gebildet in Lernprozessen und ausgeformt in sozialer Interaktion – nicht grundsätzlich verschieden von mentalen Repräsentationen von Erscheinungen der Außenwelt. Verschieden wäre es von diesen allerdings durch seine repräsentationale Sonderstellung als eine Quellinstanz, die andere mentale Inhalte erzeugt und ihnen insofern metarepräsentational gegenübersteht.

5 Vgl. Jaynes, *The origin of consciousness* (s. Fn. 4).

Subjektivitätsdiskurse

Aus all dem folgt: Das Ich ist eine Erfindung zur Lösung eines Attributionsproblems. Es wird zunächst als Quelle intern induzierter Vergegenwärtigungen konstituiert. Nachdem es einmal konstituiert ist, bildet seine implizite Gegenwart in allen psychischen Akten die funktionale und auch inhaltliche Grundlage für den bewussten Charakter ihrer Repräsentanz.

Allerdings darf man sich diese Erfindung nicht als heroische Tat einzelner Individuen vorstellen. Das Ich wird vielmehr sozial konstruiert. Es wird im konkreten sozialen Austausch erzeugt – in Praktiken und Diskursen, die die Sozialisation der Individuen steuern und ihnen eine ich-förmige mentale Organisation zuweisen. Zum Abschluss skizziere ich drei Spielarten solcher Subjektivitätsdiskurse: Diskurse der Attribution, der Reflexion und der Abgrenzung.

Attributionsdiskurse

Zunächst: Wie hat man sich den sozialen Austausch in den Attributionsdiskursen des Alltagslebens vorzustellen? Die elementarsten Vermittlungsmechanismen stützen sich auf unmittelbare *face-to-face*-Interaktionen im *mikrosozialen Bereich* und sind nicht einmal notwendigerweise an sprachliche Kommunikation gebunden. Wenn in einer sozialen Gruppierung sämtliche Akteure den Umgang miteinander so organisieren, dass sie ich-förmige Organisation bei allen Kommunikationspartnern voraussetzen, trifft jeder Akteur – auch jeder neu hinzutretende – auf eine Situation, in der durch das Handeln der anderen eine ich-förmige Rolle für ihn bereitgehalten wird. Fremdzuschreibung von Eigenschaften erzeugt dann Selbstzuschreibung, und der Akteur macht sich schließlich die ihm zugeschriebene Ich-Rolle zu eigen.⁶

Komplexere Vermittlungsmechanismen stützen sich auf sprachlich gebundene Attributionsdiskurse im *makrosozialen Bereich*. Zu nennen ist hier zunächst der Diskurs des psychologischen Common Sense, d. h. der alltagspsychologischen Konstrukte, die Kulturen oder Sprachgemeinschaften verwenden, um das Handeln ihrer Akteure zu erklären.⁷ So operiert z. B. die moderne

⁶ Die frühkindliche Entwicklung von Subjektivität und Subjektivitätsverständnis ist in den letzten 20 Jahren ausführlich untersucht worden. Vgl. z. B. Josef Perner, *Understanding the representational mind*, Cambridge, MA 1991; Henry M. Wellman, *The child's theory of mind*, Cambridge, MA 1990.

⁷ Vgl. Fritz Heider, *The psychology of interpersonal relations*, New York 1958; Wolfgang Prinz, »Explaining voluntary action: The role of mental content«, in M. Carrier und

Alltagspsychologie mit einem Begriff von menschlichen Personen, dessen Kern ein explizites, lebenslang identisches Ich bildet, das gleichsam als Brennpunkt aller Erfahrungen und Handlungen des Individuums fungiert. Ebenso einschlägig sind die Diskurse der Moral und des Rechts. In diesen Diskursen wird persönliche Verantwortung von Akteuren für ihr Handeln daraus abgeleitet, dass ihr Ich als autonome Quelle von Handlungsentscheidungen verstanden wird. Das Gleiche gilt natürlich für die narrativen Diskurse der Literatur, die uns vorführen, was Personen sind und wie sie funktionieren bzw. funktionieren sollen.

Reflexionsdiskurse

Die Attributionsdiskurse des Alltagslebens werden in unserem Kulturkreis durch wissenschaftliche Diskurse der Reflexion über die Rolle und den Status von Bewusstsein ergänzt und erweitert, und wenn wir eine umfassende Perspektive einnehmen wollen, müssen wir auch diese Reflexionsdiskurse als Bestandteile des kulturellen Gesamtdiskurses sehen, der Subjektivität konstituiert und sichert. Von besonderer Bedeutung ist hier die Philosophie, die in diesem Diskurs seit jeher federführend ist. Oft neigt der philosophische Diskurs dazu, dem Bewusstsein einen besonderen, privilegierten Status zuzuschreiben.⁸ Eine besonders prägnante Formulierung hat diese Auffassung durch Descartes erfahren. Descartes war der Überzeugung, dass der Zugang zu den eigenen Bewusstseinsinstanzen ein Prozess von viel einfacherer Struktur ist als der Zugang zur Außenwelt. Beim Nachdenken über das eigene Bewusstsein ist der Geist gleichsam bei sich selbst – statt irgendwelchen materiellen Sachverhalten gegenüberzustehen, die ihm wesensfremd sind. Deshalb muss das, was wir über unsere psychischen Vorgänge wissen, stets notwendigerweise wahr sein: Es beruht nämlich auf einem Vorgang des Innewerdens des wirklichen Sachverhalts selbst – und nicht auf einem Abbildungsvorgang, bei dem man die Frage nach der Beziehung zwischen wirklichem und wahrgenommenem Sachverhalt sinnvoll stellen könnte. Bewusstseinserscheinungen sind für die Erkenntnis primär und fundamental, Erscheinungen der Außenwelt sekundär und abgeleitet. Diese Doktrin bildet den analytisch reflektierten ideologischen Kern des modernen Subjektivitätsdiskurses.

P. Machamer, Hg., *Mindscales: Philosophy, science, and the mind*, Konstanz 1997, S. 153–175.

⁸ Vgl. Wolfgang Prinz, »Bewußtsein und Ich-Konstitution«, in Roth und Prinz, Hg., *Kopf-Arbeit* (s. Fn. 1), S. 451–467.

Nebenbei sei bemerkt, dass sie inzwischen als ebenso ehrwürdig wie fragwürdig gilt. Neben der analytischen Philosophie haben nicht zuletzt entwicklungspsychologische Studien zu ihrer Erschütterung beigetragen. Hier konnte nämlich gezeigt werden, dass Kinder zu keinem Zeitpunkt ihrer Entwicklung ihre eigenen mentalen Zustände besser verstehen als die mentalen Zustände anderer Personen. Oft scheint sogar das Umgekehrte der Fall zu sein: dass das Selbstverstehen dem Fremdverstehen *folgt* – ein Befund, der mit der Doktrin eines direkten und privilegierten Zugangs zu den eigenen mentalen Zuständen nicht zu vereinbaren ist.⁹ Aus diesem und einer Reihe weiterer Gründe lässt sich das bewusstseinsfundamentalistische Dogma kaum aufrechterhalten – so sehr wir es vielleicht auch aus allerlei weltanschaulichen Gründen in unser Herz geschlossen haben mögen.

Abgrenzungsdiskurse

Zuletzt will ich noch einen kurzen Blick auf Abgrenzungsdiskurse werfen – Diskurse von teils attributivem, teils reflexivem Charakter, die dazu dienen, die Grenzen von Subjektivität zu bestimmen. Sie grenzen ein, was in das Reich normal entwickelter Subjektivität gehört und grenzen aus, was jenseits liegt. So weiß die Psychopathologie seit langem, dass personale Attribution und Ich-Konstitution im Einzelfall auch andere als normale Wege gehen können – Wege, die als *pathologisch* gelten.

Ein Beispiel ist die Ausbildung von Wahnsymptomen bei psychotischen Erkrankungen, besonders bei Schizophrenien. Folgt man den in den letzten Jahren entwickelten kognitiven Schizophrenietheorien¹⁰, lassen sich diese Symptome nach dem gleichen Grundmuster erklären, das nach der Theorie von Julian Jaynes für die mentale Organisation der Protagonisten der Ilias charakteristisch ist. Wahnpatienten leiden darunter, dass ihnen das genormte Attributionsschema nicht zur Verfügung steht, welches die Quellen der Gedanken im Ich lokalisiert. Deshalb sind sie darauf angewiesen, den Ursprung ihrer Gedanken, Vorstellungen und Wünsche auf andere Weise zu erklären. Sie führen sie dann auf personale Quellen zurück, die unsichtbar anwesend sind – wie

9 Vgl. Alison Gopnik, »How we know our mind: The illusion of first-person knowledge of intentionality«, in *Behavioral and Brain Sciences* 16, 1993, S. 1–14.

10 Vgl. z. B. Elena Daprati u. a., »Looking for the agent: An investigation into consciousness of action and self-consciousness in schizophrenic patients«, in *Cognition* 65, 1997, S. 71–86; Christopher Frith, *The cognitive neuropsychology of schizophrenia*, Hillsdale 1992.

Angehörige, Ärzte, berühmte Personen oder Außerirdische. Oft konstruieren sie auch Wirkungsmechanismen, die plausibel machen, wie die von diesen Quellen ausgehenden Gedanken übertragen werden, etwa durch Stimmen oder Bilder, die über Strahlen oder Drähte weitergeleitet werden, neuerdings auch oft über Telefone, Funkgeräte oder Computer.

Ein anderes Beispiel nicht normgerechter Ich-Konstitution liefert das in letzter Zeit in Verruf geratene Syndrom der multiplen Persönlichkeit.¹¹ Von multiplen Persönlichkeiten ist die Rede, wenn in einem Individuum zwei oder mehr unabhängige Persönlichkeiten ausgebildet sind, die jede ihr eigenes Leben führen. Auch wenn einige besonders spektakuläre Fallbeschreibungen sich inzwischen als überzogen, z. T. auch als Fälschungen herausgestellt haben, bleibt es dabei, dass das Auftreten derartiger Persönlichkeitsspaltungen gut dokumentiert ist. Sie bilden gewissermaßen das Gegenstück zur Wahnsymptomatik: Dort fehlt das Ich, hier stehen gleich mehrere zur Verfügung.

So bizarr sich diese Syndrome auf dem Hintergrund unserer Standardvorstellung von Subjektivität und Persönlichkeit ausmachen, so nahtlos fügen sie sich in die theoretische Vorstellung ein, dass die Ich-Förmigkeit unserer mentalen Organisation kein Naturphänomen ist, sondern ein kulturelles Artefakt, das in Attributionsprozessen konstituiert wird.¹² Einheitlichkeit und Konsistenz

11 Vgl. William N. Coner und Billie S. Ables, *Multiple Personality: Etiology, diagnosis, and treatment*, New York 1983.

12 Weitere Beobachtungen über die Störung der gewohnten Beziehung zwischen Akteur und Handlung sind von split-brain Patienten und von gesunden Personen unter posthypnotischer Suggestion bekannt. Von Patienten, bei denen die Verbindung zwischen den beiden Großhirnhemisphären operativ durchtrennt worden ist, wird berichtet, dass sie manchmal Äußerungen tätigen oder Handlungen ausführen, ohne sich dessen bewusst zu sein, die sie dann in einem zweiten Schritt nachträglich begründen und rationalisieren. Dabei werden die unbewusste Ausführung und bewusste Rationalisierung je einer der beiden getrennten Hirnhemisphären zugeschrieben (vgl. z. B. Michael S. Gazzaniga und Charlotte S. Smylie, »What does language do for a right hemisphere?«, in M. S. Gazzaniga Hg., *Handbook of cognitive neuroscience*, New York / London 1984; Eran Zaidel, »Language functions in the two hemispheres following complete cerebral commissurotomy and hemispherectomy«, in F. Boller und J. Grafman, Hg., *Handbook of neuropsychology (Vol. 4)*, Amsterdam 1990, S. 115–150). Entsprechendes wird auch regelmäßig bei der Ausführung posthypnotischer Aufträge beobachtet: Wenn Probanden Aufträge ausführen, die ihnen unter Hypnose aufgegeben wurden, liefern sie oft zugleich Rationalisierungen für ihre (z. T. recht skurrilen) Handlungen. Offenbar geht hier das Tun dem Wollen voraus: Es sieht so aus, als würden Patienten und Probanden nicht etwa tun, was sie wollen, sondern vielmehr (auch noch) wollen, was sie (ohnehin schon) tun bzw. getan haben (vgl. Wolfgang Prinz, »Freiheit oder Wissenschaft?«, in K. Foppa und M. von Cranach, Hg., *Freiheit des Entscheidens und Handelns*, Heidelberg 1996, S. 86–103).

des Ich sind keine natürliche Notwendigkeit, sondern eine kulturelle Üblichkeit, und wenn Individuen besonderen Entwicklungs- und Lebensbedingungen ausgesetzt sind, mag es sein, dass sie andere als die üblichen Attributionsmuster entwickeln. Ob diese Abweichungen auf Störungen der *personalen Attribution* zurückgehen oder auf Störungen der *dualen Repräsentationsarchitektur*, ist ohne weiteres nicht zu entscheiden. Es sind eben biologische *und* gesellschaftliche Bedingungen, die in die Ich-Konstitution eingehen, und wenn sie anders als gewohnt verläuft, können die Ursachen dafür in beiden Bereichen liegen.

Lassen Sie mich enden mit einem doppelten Gedankenspiel. Es bezieht sich auf eine beliebte Frage, die im Zentrum des Abgrenzungsdiskurses steht – die Frage nämlich, ob und wie weit andere Lebewesen als Menschen über ich-förmige mentale Organisation und Bewusstsein verfügen und ob es auch sein kann, dass es Menschen gibt, die nicht darüber verfügen:

Erstens: Kann sich bei Tieren Bewusstsein entwickeln, wenn wir es ihnen zuschreiben? Würde z. B. mein Hund eine ich-förmige Organisation ausbilden, wenn er ausschließlich mit menschlichen Akteuren Umgang hätte, die ihn so behandeln als hätte er eine? Hinter dieser Überlegung verbirgt sich die Frage, ob das soziale Angebot von personalen Attributionsmustern *hinreichend* ist für die Ausbildung einer ich-förmigen mentalen Organisation. Aufgrund der psychohistorischen Skizze, die ich entworfen habe, muss ich diese Frage mit Nein beantworten, jedenfalls so lange ich nicht annehmen will, dass auch die zweite hierfür notwendige Voraussetzung bei Hunden ausgebildet ist, nämlich die Fähigkeit zur dualen Repräsentation. Personale Attribution, so sehen wir an diesem Beispiel, ist zwar *notwendig*, aber nicht *hinreichend* für die Ausbildung von Bewusstsein.

Zweitens: Können Menschen zu bewusstlosen Zombies werden, wenn ihnen alle Interaktionen und Diskurse vorenthalten werden, die Angebote für personale Attribution enthalten? Wäre z. B. denkbar, dass Kaspar Hauser völlig ich-los und somit bewusst-los war? Diese Frage muss unsere Theorie leider mit Ja beantworten, denn sie nimmt an, dass ohne sozial vermittelte Attributionen ich-förmige Organisation und Bewusstsein nicht entstehen können.

Danach scheint es, als wäre es wohl möglich, dass Menschen bewusstlos leben, nicht aber, dass Tiere Bewusstsein entwickeln: Zombies mag es geben, aber Bambi und Lassie und Fury bleiben wohl auf immer eine schöne Illusion.