

Thomas Metzinger

Die Selbstmodell-Theorie der Subjektivität:

Eine Kurzdarstellung in sechs Schritten*

1. SMT: Was ist die Selbstmodell-Theorie der Subjektivität?

Das Ziel dieses Beitrags besteht darin, eine sehr kurze Darstellung der „Selbstmodell-Theorie der Subjektivität“ anzubieten, die auch für solche Leute verständlich ist, die keine Berufsphilosophen sind. Die Selbstmodell-Theorie der Subjektivität ist eine philosophische Theorie darüber, was ein Selbst ist, eine Theorie darüber, was es eigentlich bedeutet, dass geistige Zustände „subjektive“ Zustände sind und auch darüber, was es heißt, dass ein bestimmtes System eine „phänomenale Erste-Person-Perspektive“ besitzt.¹ Eine der Kernaussagen dieser Theorie ist, dass es so etwas wie Selbste in der Welt nicht gibt: Selbste und Subjekte gehören nicht zu den irreduziblen Grundbestandteilen der Wirklichkeit. Was es gibt, ist das erlebte Ichgefühl und die verschiedenen, ständig wechselnden Inhalte unseres Selbstbewusstseins – das, was Philosophen das „phänomenale Selbst“ nennen. Dieses bewusste *Erleben* eines Selbst wird als Resultat von Informationsverarbeitungs- und Darstellungsvorgängen im zentralen Nervensystem analysiert. Natürlich gibt es auch höherstufige, begrifflich vermittelte Formen des phänomenalen Selbstbewusstseins, die nicht nur neuronale, sondern auch *soziale* Korrelate besitzen.² Der Fokus der Theorie liegt jedoch zunächst auf der Frage nach den minimalen repräsentationalen und funktionalen Eigenschaften, die ein informationsverarbeitendes System wie der Mensch besitzen muss, um die

* Dieser Text ist die vollständig überarbeitete und stark erweiterte Version eines Buchbeitrags (Metzinger 2000a), der zuerst unter dem Titel *Philosophische Perspektiven auf das Selbstbewusstsein: Die Selbstmodell-Theorie der Subjektivität* erschienen ist, in W. Greve (Hrsg.), *Psychologie des Selbst*. Psychologie Verlagsunion.

¹ Eine allererste deutschsprachige (und überholte) Formulierung der Theorie findet sich in Metzinger 1993. Eine umfassende Darstellung ist Metzinger 2003a, eine frei verfügbare elektronische Zusammenfassung 2005b.

² Die Beziehung zwischen begrifflichen und nicht-begrifflichen Inhalten des Selbstbewusstseins habe ich in Metzinger 2003d genauer analysiert, eine frühere deutsche Version dieses Texts ist Metzinger 2003b. Eine Hypothese zur Rolle des unbewussten Selbstmodells bei der Entstehung nicht begrifflich vermittelter Formen der sozialen Kognition findet sich in Metzinger & Gallese 2003.

Möglichkeitsbedingungen für diese höherstufigen Varianten des Selbstbewusstseins zu realisieren. Die erste Frage lautet: Was sind die minimal hinreichenden Bedingungen dafür, dass *überhaupt* ein bewusstes Selbst entsteht?

Die Selbstmodell-Theorie geht davon aus, dass die gesuchten Eigenschaften repräsentationale und funktionale Eigenschaften des Gehirns sind. Diejenige psychologische Eigenschaft, die uns überhaupt erst zu Personen macht, wird also mit den begrifflichen Mitteln *subpersonaler* Beschreibungsebenen analysiert. In der Philosophie des Geistes nennt man ein solches Verfahren manchmal auch eine „Naturalisierungsstrategie“: Ein schwer verständliches Phänomen – etwa das Entstehen von phänomenalem Bewusstsein mit einer subjektiven Innenperspektive – wird auf begrifflich eine Weise analysiert, die es empirisch behandelbar machen soll. Naturalistische Philosophen versuchen, über eine interdisziplinäre Öffnung klassische Probleme ihrer eigenen Disziplin für die Naturwissenschaften traktabel zu machen, zum Beispiel für die Neuro- und Kognitionswissenschaften. Naturalismus und Reduktionismus sind für solche Philosophen aber keine szientistische Ideologie, sondern einfach eine rationale Forschungsheuristik: Wenn es sich zum Beispiel zeigen sollte, dass es – wie viele glauben³ - etwas am menschlichen Selbstbewusstsein gibt, dass sich dem naturwissenschaftlichen Zugriff aus *prinzipiellen* Gründen entzieht, dann werden sie auch damit zufrieden sein. Sie haben dann das erreicht, was von Anfang an ihr Ziel war: Philosophen nennen es gerne einen „epistemischen Fortschritt“. Ein Erkenntnisfortschritt könnte nämlich auch darin bestehen, dass man hinterher auf wesentlich präzisere und gehaltvollere Weise beschreiben kann, *warum* es auf bestimmte Fragen prinzipiell keine befriedigende wissenschaftliche Antwort geben kann. Seriöse philosophische Anti-Naturalisten und Anti-Reduktionisten kann man also immer daran erkennen, dass gerade sie es sind, die ein besonders großes Interesse an neuen empirischen Erkenntnissen und an ernsthaften interdisziplinären Dialogen haben.

³ Ein gutes Beispiel ist hier der Philosoph Thomas Nagel. Vgl. Nagel 1992, besonders Kapitel 4, dazu auch Metzinger 1995a.

2. Der erste Schritt: Was genau ist das Problem?

Was wir in alltagspsychologischen Zusammenhängen als „das Ich“ bezeichnen, ist das phänomenale Selbst: Der im subjektiven Erleben unmittelbar gegebene Inhalt des Selbstbewusstseins. Das phänomenale Selbst ist vielleicht die interessanteste Form phänomenalen Gehalts überhaupt - unter anderem dadurch, dass es unserem Bewusstseinsraum zwei äußerst interessante *strukturelle* Merkmale verleiht: Zentriertheit und Perspektivität. Solange es ein phänomenales Selbst gibt, ist unser Bewusstsein ein *zentriertes* Bewusstsein und an das gebunden, was in der Philosophie als die „Perspektive der ersten Person“ bezeichnet wird. Zustände, die sich innerhalb dieses Bewusstseinszentrums befinden, sind dem Erleben nach meine *eigenen* Zustände, denn der Mittelpunkt meines Bewusstseinsraums bin immer *ich selbst*. Dadurch, dass ich dann im Erleben und im Handeln ständig wechselnde Beziehungen zu meiner Umwelt und meinen eigenen geistigen Zuständen aufnehme, entsteht die subjektive Innenperspektive. Die Tatsache, dass ich eine solche Innenperspektive besitze, ist mir selbst wiederum kognitiv verfügbar.⁴ Das bedeutet: Von anderen Tieren unterscheiden wir Menschen uns auch dadurch, dass wir eine subjektiv erlebte Innenperspektive nicht nur *haben*, sondern dass wir sogar einen Begriff von uns selbst bilden können *als eines Wesens mit einer solchen Innenperspektive*.

Das Problem besteht nun darin, dass wir eigentlich gar nicht genau wissen, was wir da sagen, wenn wir so reden. Weder sind wir in der Lage, Begriffe wie „Ich“, „Selbst“ oder „Subjekt“ zu definieren, noch gibt es irgendwelche beobachtbaren Gegenstände in der Welt, auf die diese Begriffe sich beziehen könnten. Was wir deshalb zuallererst verstehen müssen, sind die strukturellen Merkmale unseres inneren Erlebens, die dazu führen, dass wir so reden. Um die Logik der Selbstzuschreibung psychologischer Eigenschaften zu analysieren und um zu verstehen, *worauf* sie sich in Wirklichkeit beziehen, muss man zuerst die repräsentationale Tiefenstruktur des bewussten Erlebens selbst

⁴ Zumindest ist sie das ab einem gewissen *Stadium* in meiner psychologischen Entwicklung. Als erste Einführung in das Problem der kognitiven Selbstbezugnahme als einer möglichen Schwierigkeit für den philosophischen Naturalismus eignet sich Baker 1998. Eine gekürzte deutsche Übersetzung ist verfügbar als Baker 2000. Siehe dazu auch Metzinger 2003a (Abschnitt 6.4.4) und insbesondere Metzinger 2003d (eine frühere deutsche Version ist Metzinger 2003b).

untersuchen. Es gibt drei phänomenale Eigenschaften höherer Ordnung, die in diesem Zusammenhang das Zentrum des Interesses bilden:

- **„Meinigkeit“:** Dies ist eine höherstufige Eigenschaft einzelner Formen von phänomenalem Gehalt. Hier sind Beispiele dafür, wie wir sprachlich auf diese phänomenale Eigenschaft Bezug nehmen: „Ich erlebe *mein* Bein subjektiv als immer schon zu mir gehörend“; „Ich erlebe *meine* Gedanken und *meine* Gefühle immer als Teil *meines* Bewusstseins“; „Meine Willensakte werden von *mir selbst* initiiert“.
- **Selbstheit, „präreflexive Selbstvertrautheit“:** Dies ist die phänomenale Kerneigenschaft, das erlebnismässig unhintergehbare „Ichgefühl“. Wieder einige Beispiele dafür, wie wir von außen auf dieses Merkmal unseres inneren Erlebens hinweisen: „Ich bin *jemand*“; „Ich erlebe mich selbst als *identisch* durch die Zeit hinweg.“; „Die Inhalte meines Selbstbewusstseins bilden eine zusammenhängende *Ganzheit*.“; „Mit dem Inhalt meines Selbstbewusstseins bin ich vor allen gedanklichen Operationen „immer schon“ vertraut.“
- **„Perspektivität“:** In unserem Zusammenhang ist Perspektivität das dominante Strukturmerkmal des Bewusstseinsraums als Ganzem: Er wird durch ein handelndes und erlebendes Subjekt zentriert, durch ein Selbst, das Beziehungen zu sich selbst und zur Welt aufbaut. Beispiele: „Meine Welt besitzt einen unverrückbaren Mittelpunkt und dieser Mittelpunkt bin ich *selbst*.“; „Bewusstsein zu haben bedeutet, eine *individuelle Innenperspektive* zu besitzen“; „Im Erleben nehme ich diese Ich-Perspektive sowohl auf Personen und Dinge in der Welt, als auch auf meine eigenen geistigen Zustände ein“.

Was jetzt geleistet werden muss, ist eine repräsentationale und eine funktionale Analyse dieser Eigenschaften. Man muss fragen: Was sind die funktionalen und repräsentationalen Eigenschaften, die ein informationsverarbeitendes System mindestens besitzen muss, um die fragliche *phänomenale* Eigenschaft zu instantiieren? Welche dieser Eigenschaften sind hinreichend, welche notwendig? Was *genau* bedeutet es für ein solches System, eine Erste-Person-Perspektive auf die Welt und auf seine eigenen mentalen Zustände einzunehmen? Benötigt wird ein konsistenter begrifflicher Hintergrund, der flexibel genug für eine kontinuierliche Integration neuer empirischer Erkenntnisse ist und gleichzeitig dem Reichtum, der Vielfalt und der Subtilität des phänomenalen Erlebens Rechnung trägt. Ich werde deshalb jetzt versuchen, in fünf weiteren, kurzen Schritten die Grundlinien eines solchen Begriffsrahmens skizzieren.

3. Der zweite Schritt: Das Selbstmodell

Der zweite Schritt besteht darin, eine neue theoretische Entität einzuführen: Das phänomenale *Selbstmodell*. Es bildet den wichtigsten Teil der repräsentationalen Instantiierungsbasis⁵ der zu erklärenden phänomenalen Eigenschaften. Eine unserer Kernfragen war: Was ist die minimal hinreichende Menge an *repräsentationalen* Eigenschaften, die ein System entwickeln muss, um die Zieleigenschaften zu besitzen? Eine erste, vorläufige Antwort auf diese Frage lautet jetzt: Das System muss in jedem Fall eine kohärente Selbstrepräsentation besitzen, ein zusammenhängendes inneres Modell von sich selbst als einer Ganzheit. Ein Selbstmodell ist in unserem eigenen Fall eine nur episodisch aktive repräsentationale Entität, deren Gehalt durch Eigenschaften des Systems selbst gebildet wird. Immer dann, wenn eine solche Selbstrepräsentation gebraucht wird um die Interaktion mit der Umwelt zu regulieren, wird sie vorübergehend vom System aktiviert - zum Beispiel dann, wenn wir am Morgen aufwachen.

Was wir im Grunde brauchen, ist also eine umfassende Theorie des Selbstmodells von *Homo sapiens*.⁶ Ich selbst gehe davon aus, dass eine solche Theorie in wesentlichen Teilen eine neurokomputationale Theorie sein wird.⁷ Das bedeutet, dass das Selbstmodell des Menschen nicht nur eine wahre repräsentationale und eine wahre funktionale Beschreibung besitzt, sondern auch eine wahre neurobiologische Beschreibung - zum Beispiel als ein komplexes Aktivierungsmuster im menschlichen Gehirn.⁸ Das *phänomenale* Selbstmodell ist aber immer nur derjenige Teil des *mentalen* Selbstmodells der gegenwärtig in die höchststufige, integrierte Struktur eingebettet ist, in das globale Modell der Welt.⁹ Es kann also durchaus unbewusste, aber funktional aktive Teile des Selbstmodells geben. Das phänomenale Selbstmodell ist eine

⁵ Vgl. Cummins 1983.

⁶ Psychologie kann man – wenn diese metatheoretische Bemerkung eines philosophischen Außenseiters erlaubt ist – in ihrem methodologischen Kern und auf heuristisch sehr fruchtbare Weise als *Selbstmodellforschung* analysieren: Sie ist eine wissenschaftliche Disziplin, die sich mit dem repräsentationalen Gehalt, dem funktionalen Profil und der neurobiologischen Realisierung des menschlichen Selbstmodells beschäftigt

⁷ Vgl. Churchland 1989.

⁸ Vgl. Damasio 1999.

⁹ Vgl. Yates 1975, Baars 1988. Eine ausführliche Analyse der Kriterien, die den Grad der *Bewusstheit* bestimmen findet sich in Metzinger 2003a, Kapitel 3.

kohärente multimodale Struktur, die wahrscheinlich auf einem teilweise angeborenen und „fest verdrahteten“ Modell der räumlichen Eigenschaften des Systems beruht (davon später mehr).¹⁰ Bei diesem Typ von Analyse wird der selbstbewusste Mensch also als eine ganz bestimmte Art von Informationsverarbeitungssystem betrachtet: Der subjektiv erlebte Gehalt des phänomenalen Selbst ist der repräsentationale Inhalt einer jetzt gerade aktiven Datenstruktur in seinem zentralen Nervensystem.

Man kann parallel zur repräsentationalen Beschreibungsebene auch eine *funktionale* Analyse des Selbstmodells entwickeln. Ein aktives Selbstmodell ist dann ein subpersonaler funktionaler Zustand: Eine – unter Umständen sehr komplexe – Menge von Kausalbeziehungen, die realisiert sein können oder auch nicht. Dadurch, dass dieser funktionale Zustand eine konkrete neurobiologische Realisierung besitzt, spielt er eine bestimmte kausale Rolle im System. Man kann sich diesen Gedanken verdeutlichen, indem man die Perspektive der klassischen Kognitionswissenschaft einnimmt und sagt: Das Selbstmodell ist ein *transientes computationales Modul*, das vom System vorübergehend aktiviert wird, um seine Interaktion mit der Umwelt zu regulieren. Der Besitz von immer besseren Selbstmodellen als einer neuen Art von „virtuellen Organen“ ermöglichte - diesen Punkt darf man nicht übersehen - überhaupt erst die Bildung von Gesellschaften. Plastische und immer komplexere Selbstmodelle erlaubten nicht nur eine fortlaufende Optimierung somatomotorischer, perzeptiver und kognitiver Funktionen, sondern später auch *soziale* Kognition und damit die Entwicklung von kooperativem Verhalten. Mit ihnen entstanden die fundamentalen repräsentationalen Ressourcen für Perspektivenübernahmen, Empathie und Schuldbewusstsein, später auch für metakognitive Leistungen wie die Entwicklung eines Selbstkonzepts und einer *theory of mind*.¹¹

Der offensichtlichen Tatsache, dass die Entwicklung unseres Selbstmodells eine lange evolutionsbiologische und eine (etwas kürzere) soziale Geschichte besitzt, kann man nun Rechnung tragen, indem man im nächsten Schritt das einführt,

¹⁰ Vgl. hierzu den fünften Abschnitt und z.B. den Begriff eines „Long-term body image“ bei O’Shaughnessy 1995; siehe dazu insbesondere auch Damasio 1994, 1999, Metzinger 1993, 1996, 1997.

¹¹ Vgl. hierzu etwa Bischof-Köhler 1996, 1989. Bezüglich möglicher neurobiologischer Korrelate solcher basalen sozialen Leistungen, die gut in den hier skizzierten Rahmen passen, vgl. Gallese & Goldman 1999, Metzinger & Gallese 2003. Eine neuere deutsche Textsammlung ist Newen & Vogeley 2000.

was in der Philosophie des Geistes als eine *teleofunktionalistische Zusatzannahme* bezeichnet wird.¹² Die Entwicklung und Aktivierung dieses komputationalen Moduls spielt eine Rolle *für* das System: Das funktionale Selbstmodell besitzt eine wahre evolutionsbiologische Beschreibung, d.h. es war eine *Waffe*, die im Verlauf eines „kognitiven Wettrüstens“ erfunden und immer weiter optimiert wurde. Die funktionale Instantiierungsbasis der phänomenalen Erste-Person-Perspektive ist somit eine spezifische kognitive Leistung: Die Fähigkeit, zentrierte Darstellungsräume zu öffnen. Phänomenale Subjektivität (im Sinne des Entstehens einer subsymbolischen, nicht-begrifflichen Erste-Person-Perspektive) ist also eine Eigenschaft, die nur dann instantiiert wird, wenn das betreffende System ein kohärentes Selbstmodell aktiviert und dieses in sein globales Weltmodell integriert.

Mit dem Vorhandensein eines stabilen Selbstmodells kann das entstehen, was in der Philosophie des Geistes als die „Perspektivität des Bewusstseins“ bezeichnet wird: Die Existenz eines einzigen, kohärenten und zeitlich stabilen Modells der Wirklichkeit, welches repräsentational um oder „auf“ ein einziges, kohärentes und zeitlich stabiles phänomenales Subjekt zentriert ist, d.h. um ein Modell des Systems *als erlebend*. Dieses strukturelle Merkmal des globalen Darstellungsraums führt episodisch zur Instantiierung einer zeitlich ausgedehnten und nicht-begrifflichen Erste-Person-Perspektive. Wenn diese globale repräsentationale Eigenschaft verloren geht, verändert sich auch die Phänomenologie, und verschiedene neuropsychologische Störungsbilder oder veränderte Bewusstseinszustände treten hervor. Vielleicht klingen diese Überlegungen jetzt in den Ohren einiger meiner Leser sehr abstrakt. Ein Selbstmodell ist jedoch nichts Abstraktes, sondern etwas ganz und gar Konkretes. Ein erstes Beispiel soll deshalb an dieser Stelle verdeutlichen, was ich – unter vielem anderen – mit dem Begriff „Selbstmodell“ meine.

Was ein phänomenales Selbstmodell ist, hat der indische Neuropsychologe Vilayanur Ramachandran in einer Serie von faszinierenden Experimenten gezeigt, bei denen er mit Hilfe von einfachen Spiegeln Synästhesien und Bewegungszusammenhänge in Phantomgliedern auslöste.¹³ Phantomglieder sind subjektiv erlebte Gliedmaßen, die typischerweise nach dem Verlust eines Arms

¹² Vgl. etwa Bieri 1987, Dennett 1987, Dretske 1988, 1998, Lycan 1996, Millikan 1984, 1993.

¹³ Vgl. Ramachandran & Rogers-Ramachandran 1996, eine populäre Darstellung findet sich in Ramachandran & Blakeslee 1998: 46ff. Ich bin Ramachandran für die Überlassung der Abbildung im Text zu Dank verpflichtet.

oder einer Hand oder nach chirurgisch durchgeführten Amputationen auftreten. In manchen Fällen, zum Beispiel nach einer nicht-traumatischen Amputation durch einen Chirurgen, sind die Patienten subjektiv in der Lage, ihr Phantomglied willentlich zu kontrollieren und zu bewegen. Das neurofunktionale Korrelat dieser phänomenalen Konfiguration könnte darin bestehen, dass – da es keine widersprechende Rückmeldung aus dem amputierten Arm gibt – Motorbefehle, die im motorischen Kortex entstehen, immer noch kontinuierlich durch Teile des Parietallappens überwacht und dabei in denjenigen Teil des Selbstmodells integriert werden, der als ein *Motoremulator*¹⁴ dient. In anderen Situationen dagegen kann die subjektiv erlebte Beweglichkeit und Kontrolle über das Phantomglied verloren gehen. Solche alternativen Konfigurationen könnten etwa durch eine präamputationale Lähmung als Folge peripherer Nervenschädigungen oder durch ein längeres Fehlen einer die Beweglichkeit bestätigenden „Rückmeldung“ durch propriozeptives und kinästhetisches Feedback entstehen. Das Resultat auf der phänomenalen Darstellungsebene ist dann ein paralysiertes Phantomglied.

Ramachandran und seine Kollegen konstruierten nun eine „virtuelle Realitätskiste“, indem sie einen Spiegel vertikal in einen Pappkarton ohne Abdeckung einsetzten. Zwei Löcher in der Vorderseite des Kartons ermöglichten es dem Patienten, sowohl seinen echten als auch seinen Phantomarm hinein zu schieben. Ein Patient, der seit vielen Jahren unter einem paralysierten Phantomglied litt, wurde dann gebeten, das Bild seiner normalen Hand im Spiegel zu betrachten, um so – auf der Ebene des visuellen Inputs – die Illusion zu erzeugen, dass er zwei Hände sieht, obwohl er in Wirklichkeit nur das im Spiegel reflektierte Bild seiner intakten Hand sehen konnte. Die Fragestellung: Was geschieht mit dem Inhalt des phänomenalen Selbstmodells, wenn man jetzt die Versuchsperson bittet, auf beiden Seiten symmetrische Handbewegungen auszuführen? Ramachandran beschreibt ein typisches Resultat dieses Experiments:

¹⁴ Verwandte Überlegungen finden sich bei Grush 1997, 1998: 174; siehe auch Ramachandran & Rogers-Ramachandran 1996: 378.

Ich bat Philip, seine rechte Hand innerhalb der Kiste rechts vom Spiegel zu platzieren und sich vorzustellen, dass seine linke Hand (das Phantom) sich auf der linken Seite befindet. Dann gab ich die Instruktion: „Ich möchte, dass Sie gleichzeitig ihren rechten und ihren linken Arm bewegen“.

„Oh, das kann ich nicht“, sagte Philip. „Ich kann meinen rechten Arm bewegen, aber mein linker Arm ist eingefroren. Jeden Morgen beim Aufstehen versuche ich, mein Phantom zu bewegen, weil es sich immer in dieser seltsamen Stellung befindet, und weil ich das Gefühl habe, dass Bewegungen den Schmerz lindern könnten. Aber“ sagte er, während sein Blick abwärts an seinem unsichtbaren Arm entlang glitt, „ich war niemals in der Lage, auch nur den Funken einer Bewegung in ihm zu erzeugen.“

“Okay Philip – versuchen Sie es trotzdem.“

Philip drehte seinen Körper und bewegte seine Schulter in die richtige Stellung um sein lebloses Phantomglied in die Kiste „hinein zu schieben“. Dann hielt er seine rechte Hand neben die andere Seite des Spiegels und versuchte, synchrone Bewegungen zu machen. Als er in den Spiegel schaute, rang er plötzlich um Atem und rief dann aus: „Oh mein Gott! Oh mein Gott, Doktor! Das ist unglaublich. Ich glaube, ich werde verrückt!“ Er sprang auf und ab wie ein Kind. „Mein linker Arm ist wieder angeschlossen. Es ist, als ob ich in der Vergangenheit bin. Ganz viele Erinnerungen aus der Vergangenheit überfluten mein Bewusstsein. Ich kann meinen Arm wieder bewegen! Ich kann die Bewegung meines Ellenbogens spüren, auch die meines Handgelenks. Alles ist wieder beweglich.“

Nachdem er sich etwas beruhigt hatte, sagete ich: „Okay Philip – schließen Sie jetzt Ihre Augen.“

„Oh je,“ sagte er, und die Enttäuschung in seiner Stimme war deutlich zu hören, „es ist wieder eingefroren. Ich fühle wie meine rechte Hand sich bewegt, aber es gibt keinerlei Bewegungsempfindung im Phantom.“

„Öffnen Sie Ihre Augen.“

„Oh ja – jetzt bewegt es sich wieder.“¹⁵

¹⁵ Vgl. Ramachandran 1998: 47f (deutsche Übersetzung Thomas Metzinger). Die klinischen und experimentellen Details finden sich in Ramachandran & Rogers-Ramachandran 1996.



Ich hoffe, dass bereits deutlich geworden ist, wie solche neuen Daten den von mir eingeführten Begriff eines „Selbstmodells“ illustrieren: Was sich in diesem Experiment bewegt, *ist* das phänomenale Selbstmodell. Das plötzliche Auftreten von kinästhetischen Empfindungsqualitäten in der verlorenen Subregion des Selbstmodells wurde durch die Installation einer zweiten Quelle von „virtueller Information“ möglich gemacht. Sie machte den visuellen Modus der Selbstrepräsentation sozusagen wieder zugänglich und damit auch die betreffende Information wieder volitional verfügbar. Die willentliche Kontrolle wurde nun wieder möglich. Was das Experiment ebenfalls zeigt, ist wie phänomenale Eigenschaften durch komputationale und repräsentationale Eigenschaften determiniert werden. Körperliches Selbstbewusstsein hängt sehr eng mit Vorgängen im Gehirn zusammen.

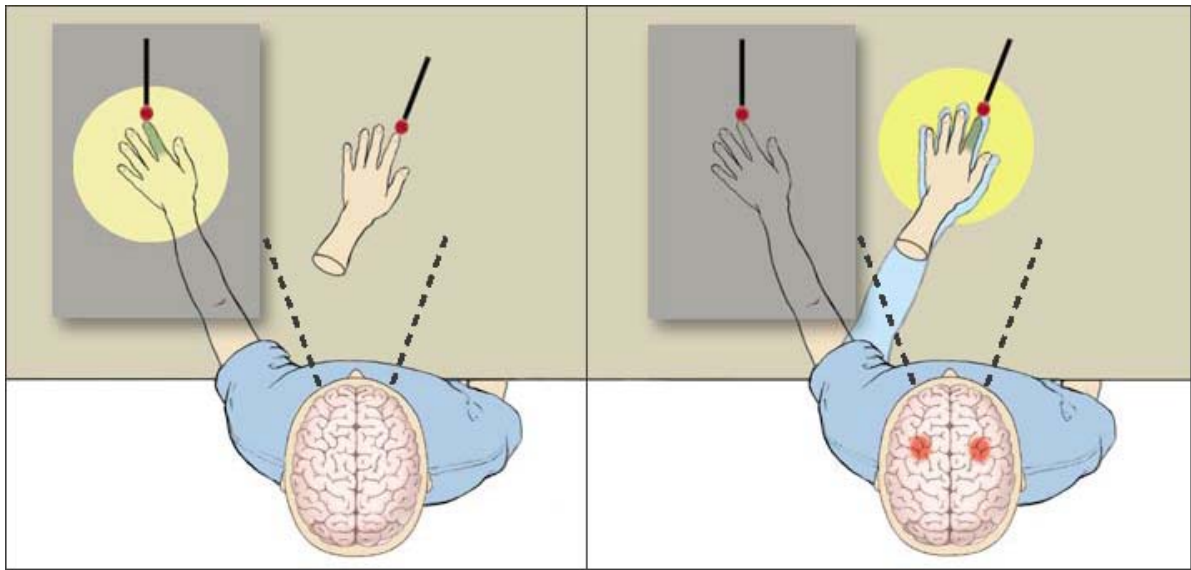
4. Der dritte Schritt: Eine repräsentationalistische Analyse der drei Zieleigenschaften

Die Grundidee ist nun, dass Selbstbewusstsein in wesentlichen Aspekten eine *Integrationsleistung* ist: Alle repräsentationalen Zustände, die in das gegenwärtig aktive Selbstmodell eingebettet werden, gewinnen die höherstufige Eigenschaft

der phänomenalen Meinigkeit hinzu. Wenn dieser Einbettungsprozess gestört wird oder hypertrophiert, resultieren verschiedene neuropsychologische Syndrome oder veränderte Bewusstseinszustände. Werfen wir wieder einen Blick auf einige Beispiele, bei denen die phänomenale Meinigkeit verloren geht:

- Floride Schizophrenie: Bewusst erlebte Gedanken sind nicht mehr *meine* Gedanken
- Unilateraler Hemi-Neglekt: Mein Bein ist nicht mehr *mein* Bein.
- Depersonalisierung: Ich bin ein Roboter, verwandele mich in eine Marionette, volitionale Akte sind nicht mehr *meine* volitionalen Akte. (Was hier selektiv verloren geht, ist also das, was der Philosoph und Psychiater Karl Jasper „Vollzugsbewusstsein“ genannt hat.
- Manien: Ich bin die ganze Welt, alle Ereignisse in der Welt werden durch meine *eigenen* Willensakte kontrolliert.

Subjektiv erlebte „Meinigkeit“ ist also eine Eigenschaft einzelner Formen phänomenalen Gehalts, zum Beispiel der mentalen Repräsentation eines Beins, eines Gedankens oder eines Willensaktes. Diese Eigenschaft ist nicht notwendig mit ihnen verbunden, denn sie ist keine intrinsische, sondern eine *relationale* Eigenschaft. Ihre Verteilung über die Elemente eines bewussten Weltmodells kann variieren. Sie kann verloren gehen, und zwar genau dann, wenn dem System die Integration bestimmter einzelner repräsentationaler Inhalte ins Selbstmodell nicht mehr gelingt. Wenn das richtig ist, dann könnte man diese Eigenschaft zumindest prinzipiell operationalisieren, und zwar indem man nach einer empirisch überprüfbaren Metrik für die Kohärenz des Selbstmodells in den fraglichen Bereichen sucht. Man könnte auch empirisch untersuchen, *wie* und durch welche Teile des Gehirns ein bestimmter repräsentationaler Inhalt ins Selbstmodell eingebunden wird. Hier ist ein zweites konkretes Beispiel, diesmal allerdings für das, was ich mit "Meinigkeit" meine:



Die Gummihand-Illusion: Gesunde Versuchspersonen erleben ein künstliches Glied als einen Teil ihres eigenen Körpers. Die Versuchsperson beobachtet die Nachbildung einer menschlichen Hand, während eine ihrer eigenen Hände verdeckt ist (graues Rechteck). Sowohl die künstliche Gummihand wie auch die unsichtbare Hand werden wiederholt und gleichzeitig mit einem Stäbchen gestreichelt. Die grünen und gelben Bereiche stellen die jeweiligen taktilen und visuellen rezeptiven Felder für Neuronen im prämotorischen Cortex dar. Im rechten Bild sieht man die Illusion der Versuchsperson, bei der die gefühlte Berührung (in Grün) in Einklang mit der gesehenen Berührung gebracht wird (die phänomenal erlebte, illusorische Armstellung ist in Blau dargestellt). Die entsprechende Aktivierung von Neuronen im prämotorischen Cortex lässt sich experimentell sehr genau nachweisen.

Bei der Gummihand-Illusion werden eine gefühlte und eine gesehene Berührung vom Gehirn so mit einander verschmolzen, dass nicht nur eine propriozeptive Karte vorübergehend mit einer visuellen Karte in Übereinstimmung gebracht werden - also die Eigenwahrnehmung des Körpers mit dem, was man sieht - , sondern so, dass damit auch das Gefühl des „Besitzens“, die phänomenale „Meinigkeit“ auf die Gummihand übergeht. Die Versuchsperson erlebt die Gummihand als ihre *eigene* Hand und fühlt die Berührung *in* dieser Hand. Wenn man sie bittet, auf ihre verdeckte linke Hand zu zeigen, unterläuft ihr ein Fehler in Form einer Abweichung in Richtung auf die Gummihand.¹⁶ „Verletzt“ man einen Finger der Gummi-Hand durch eine Biegung nach hinten in eine physiologisch unmögliche Position, dann wird auch der phänomenal erlebte Finger als wesentlich weiter zurück gebogen erlebt, als er es in Wirklichkeit ist, zusätzlich zeigt sich eine deutliche messbare

¹⁶ Vgl. Botvinick & Cohen 1998: 756.

Hautwiderstandreaktion. Zwar berichteten nur 2 von 120 Versuchspersonen tatsächlich von einem echten Schmerzerlebnis, aber viele zogen ihre reale Hand zurück, rissen die Augen auf oder lachten nervös.¹⁷ Wenn man mit einem Hammer auf die Gummihand schlägt, zeigt die Person ebenfalls eine sehr deutliche Reaktion. Wieder zeigt sich, wie die phänomenale Zieleigenschaft direkt von repräsentationalen und funktionalen Vorgängen im Gehirn determiniert wird. Was wir als Teil unseres Selbst erleben, hängt also vom jeweiligen Kontext ab, und davon, welche Information vom Gehirn jeweils in unser Selbstmodell eingebettet wird.¹⁸

Wenden wir uns nun der zweiten Zieleigenschaft zu, dem bewusst erlebten Ichgefühl. Der höherstufigen phänomenalen Zieleigenschaft der „Ichhaftigkeit“ oder „präreflexiven Selbstvertrautheit“ dagegen entspricht die Existenz eines einzigen, kohärenten und zeitlich stabilen Selbstmodells, das den Mittelpunkt des repräsentationalen Gesamtzustands bildet. Wenn dieses repräsentationale Modul beschädigt ist, desintegriert oder wenn multiple Strukturen dieses Typs im System alternieren bzw. gleichzeitig aktiv sind, resultieren wiederum verschiedene neuropsychologische Störungsbilder oder veränderte Bewusstseinszustände:

- Anosognosien und Anosodiaphorien: Verlust höherstufiger Einsicht in bestehende Defizite, z.B. bei Verleugnung der eigenen Blindheit (*Antons Syndrom*).
- *Dissociative Identity Disorder* (DID¹⁹): Das System verwendet verschiedene und alternierende Selbstmodelle, um mit extrem traumatisierenden und sozial inkonsistenten Situationen umzugehen.
- „Ich-Störungen“: Eine große Klasse psychiatrischer Störungsbilder, die mit veränderten Formen des Erlebens der eigenen *Identität* einhergehen. Klassisches Beispiel: Schizophrenien.

Durch das Vorhandensein eines stabilen Selbstmodells entsteht fast immer auch die „Perspektivität des Bewusstseins“ in Form von vorübergehenden Subjekt-Objekt-Beziehungen (vgl. Abschnitt 6)²⁰. Dieses strukturelle Merkmal des globalen Darstellungsraums führt episodisch zur Instantiierung einer zeitlich

¹⁷ Vgl. Armel & Ramachandran 2003: 1503.

¹⁸ Vgl. insbesondere Botvinick & Cohen, 1998, und die Neurimaging-Studie von Ehrsson, Spence & Passingham 2004, Botvinick 2004.

¹⁹ Zu den aktuellen diagnostischen Kriterien vgl. DSM-IV: 300.14. Warum Ichstörungen für Philosophen interessant sind untersucht Metzinger 2004a.

²⁰ Vgl. Nagel 1992, Metzinger 1993, 1995a, 2000b, insbesondere 2005a.

ausgedehnten und nicht-begrifflichen Erste-Person-Perspektive, und es kann ebenfalls verloren gehen.

- Vollständige Depersonalisierung: Verlust der phänomenalen Erste-Person-Perspektive, begleitet von dysphorischen Zuständen und funktionalen Defiziten („Angstvolle Ich-Auflösung“²¹).
- Mystische Erfahrungen: Selbstlose und nicht-zentrierte Globalzustände, die als nicht-pathologisch und nicht-bedrohlich erlebt bzw. beschrieben werden. („Ozeanische Selbstentgrenzung“; „Der große Blick von nirgendwo“)

Wenn man dem Reichtum und der Vielfalt menschlicher Erlebnisformen theoretisch gerecht werden will, dann muss man anerkennen, dass es auch *aperspektivische* Formen des bewussten Erlebens gibt. Die Selbstmodelltheorie stellt die begrifflichen Mittel zur Verfügung, um dies zu tun.²²

5. Der vierte Schritt: Die funktionale Zentrierung des phänomenalen Raums durch leibliche Verankerung

Ich habe eben zwischen einer repräsentationalen und einer funktionalen Analyse der Erste-Person-Perspektive unterschieden. Das zentrale theoretische Problem auf der funktionalen Beschreibungsebene entsteht nun durch folgende Frage: Auf welche Weise unterscheidet sich das phänomenale Selbstmodell überhaupt von den anderen phänomenalen Modellen, die gegenwärtig aktiv sind? Durch welche funktionale Eigenschaft wird es als Mittelpunkt ausgezeichnet, wodurch genau wird es zum stabilen *Zentrum* des phänomenalen Darstellungsraums?

Hier ist meine vorläufige Antwort: Das Selbstmodell ist die einzige repräsentationale Struktur, die im Gehirn durch eine *kontinuierliche Quelle intern generierten Inputs* verankert ist. Immer dann, wenn es überhaupt zu bewusstem Erleben kommt (also zur Aktivierung eines stabilen, integrierten Modells der Wirklichkeit), existiert auch diese kontinuierliche Quelle internen, propriozeptiven Inputs. Es gibt im Wesentlichen vier Typen von intern erzeugter Information, die ein persistierendes funktionales Bindeglied zwischen dem phänomenalen Selbstmodell und seiner körperlichen Basis im Gehirn

²¹ Vgl. Dittrich 1985.

²² Weitere neurophänomenologische Fallstudien finden sich in Metzinger 2003a, Kapitel 4 und 7.

erzeugen:

- Input aus dem Vestibulärorgan: Der Gleichgewichtssinn.
- Input aus dem invarianten Teil des Köperschemas: Das „Hintergrundgefühl“ im räumlichen Modell des Körpers.
- Input aus Sensoren in den Eingeweiden, aber auch aus den Blutgefäßen, z.B. aus den kardiovaskulären Mechanosensoren: „Bauchgefühle“ und somatoviszzerale Formen der Selbstpräsentation.
- Input aus bestimmten Teilen des oberen Hirnstamms und des Hypothalamus: Hintergrundemotionen und „Gestimmtheiten“, verankert in der kontinuierlichen homöostatischen Selbstregulation des „internen Milieus“, also der biochemischen Landschaft in unserem Blut.

Entscheidend sind hier nicht die neurobiologischen Details, sondern die sehr plausible Annahme, dass es einen Teil des menschlichen Selbstmodells gibt, der hochgradig stimuluskorreliert ist und ausschließlich auf *intern* erzeugter Information beruht. Die konstante Aktivität derjenigen Regionen des Köperselbstes, der unabhängig von externem Input ist, wird – das ist meine These - zum funktionalen *Mittelpunkt* des phänomenalen Darstellungsraums. Marcel Kinsbourne hat in diesem Zusammenhang von einem „*background 'buzz' of somatosensory input*“ gesprochen²³, Antonio Damasio von einem *core self*²⁴ ich selbst habe an anderer Stelle den Begriff eines „phänomenalen Selbstpräsentats“ eingeführt.²⁵

Als ein erstes Beispiel dafür, was man unter einem Selbstmodell verstehen kann, hatte ich Ramachandrans Experiment zur Mobilisierung eines gelähmten Phantomglieds vorgestellt. Ein Selbstpräsentat ist derjenige Teil des Phantomglieds, der auch ohne jede Bewegung ständig im Bewusstsein bleibt. Wenn man *diesen* Teil verliert, verliert man im subjektiven Erleben auch die leibliche Präsenz – man wird zu einem „körperlosen Wesen“.²⁶ Neuere

²³ Kinsbourne 1995: 217. Umfangreiche theoretische Überlegungen und zahlreiche neuere empirische Befunde zur leiblichen Verankerung des bewussten Erlebens finden sich in Damasio 1999.

²⁴ Vgl. Damasio 1999.

²⁵ Vgl. Metzinger 1993: 156ff, und insbesondere 2003a, Abschnitt 5.4.

²⁶ Natürlich existieren auch hier wieder die entsprechenden phänomenologischen Zustandsklassen. In Metzinger 1993 und 1997 habe ich in diesem Zusammenhang auf Oliver Sacks' Beispiel der „körperlosen Frau“ hingewiesen, siehe auch Metzinger 2003a.

Forschungsergebnisse aus der Erforschung des Schmerzerlebens in Phantomgliedern deuten auf die Existenz einer genetisch determinierten Neuromatrix hin, deren Aktivitätsmuster Grundlage dieser starren Teile des Körperbilds und des invarianten Hintergrunds der Körperempfindung sein könnte („Phylomatrix des Körperschemas“²⁷). Ein Beispiel: Über 20% der Kinder, die ohne Arme oder Beine geboren wurden, haben später ein realistisches und bewusst erlebtes Phantomglied.

6. Der fünfte Schritt: Autoepistemische Geschlossenheit - Die Transparenzannahme und das naiv-realistische Selbstmissverständnis

Das zentrale theoretische Problem auf der *repräsentationalen* Beschreibungsebene dagegen ergibt sich daraus, dass man mir leicht vorwerfen könnte, ich würde mit der Einführung des Begriffs „Selbstmodell“ einen Etikettenschwindel betreiben. Es scheint keine notwendige Verbindung von den funktionalen und repräsentationalen Basiseigenschaften zu den *phänomenalen* Zieleigenschaften der „Meinigkeit“, „präreflexiven Selbstvertrautheit“ und „Perspektivität“ zu geben. All das könnte sich durchaus ereignen, ohne dass es zur Entstehung eines echten phänomenalen Selbst oder einer subjektiven Innenperspektive kommt: Man kann sich vorstellen, dass biologische Informationsverarbeitungssysteme durch ein Selbstmodell zentrierte Darstellungsräume entwickeln und benutzen, *ohne* dass Selbstbewusstsein entsteht. Ein „Selbstmodell“ ist noch lange kein Selbst, sondern nur eine Repräsentation des Systems als Ganzen - eben bloß ein *Systemmodell*. Damit aus der funktionalen Eigenschaft der Zentriertheit aber die phänomenale Eigenschaft der Perspektivität werden kann, muss aus dem Modell des Systems ein phänomenales Selbst werden. Die philosophischen Kernfrage lautet deshalb: Wie entsteht in einem bereits funktional zentrierten Repräsentationsraum ein bewusstes Selbst und das, was wir als die phänomenale Erste-Person-Perspektive zu bezeichnen gewohnt sind? Oder: Wie wird aus dem Selbstmodell ein *Selbstmodell*?

Ein genuines bewusstes Selbst - so lautet meine Antwort - entsteht immer

²⁷ Vgl. Melzack 1989, zum Begriff einer „Neurosignatur“ 1992: 93, eine wichtige Studie zu Phantomgliedern nach Aplasien und Frühamputationen ist Melzack *et al.* 1997; siehe dazu Metzinger 2003a, Kapitel 7 und wieder Damasio 1994, 1999.

genau dann, wenn das System das von ihm selbst aktivierte Selbstmodell auf der Ebene des bewussten Erlebens nicht mehr *als* Modell erkennt. Wie also kommt man von der funktionalen Eigenschaft der „Zentriertheit“ und der repräsentationalen Eigenschaft der „Selbstmodellierung“ zu der phänomenalen Eigenschaft der „präreflexiven Selbstvertrautheit“? Die Lösung liegt in dem, was Philosophen manchmal „phänomenale Transparenz“ nennen.²⁸ Die vom System eingesetzten repräsentationalen Zustände sind *transparent*, d.h. sie stellen die Tatsache, *dass* sie Modelle sind, nicht mehr auf der Ebene ihres Gehalts dar. Deshalb schaut das System durch seine eigenen repräsentationalen Strukturen „hindurch“, als ob es sich in direktem und unmittelbarem Kontakt mit ihrem Inhalt befände. Transparenz in dem hier definierten Sinne ist ausschließlich eine Eigenschaft *bewusster* Zustände, unbewusste Repräsentationen sind also weder transparent noch opak. Ich vertrete zwei kausale Hypothesen bezüglich der Entstehungsgeschichte transparenter phänomenaler Zustände. Erstens: Die fraglichen Datenstrukturen werden so schnell und zuverlässig aktiviert, dass das System sie nicht mehr als solche erkennen kann, z.B. wegen des mangelnden zeitlichen Auflösungsvermögens *metarepräsentationaler* Funktionen. Außerdem hat es – zweitens – allem Anschein nach keinen evolutionären Selektionsdruck auf die entsprechenden Teile der funktionalen Architektur gegeben: Der naive Realismus ist für biologische Systeme wie uns selbst eine funktional adäquate Hintergrundannahme gewesen.

Transparenz ist eine besondere Form der Dunkelheit. In der Phänomenologie des visuellen Erlebens bedeutet Transparenz, dass wir etwas nicht sehen können, weil es durchsichtig ist. Phänomenale Transparenz *im Allgemeinen* dagegen bedeutet, dass etwas Bestimmtes dem subjektiven Erleben nicht zugänglich ist, nämlich der Repräsentationscharakter der Inhalte des bewussten Erlebens. Diese Analyse bezieht sich auf alle Sinnesmodalitäten und insbesondere auf das integrierte phänomenale Modell der Welt als ganzer. Das *Mittel* der Darstellung kann selbst nicht noch einmal als solches dargestellt werden und darum wird das erlebende System notwendigerweise in einen naiven Realismus verstrickt, weil es sich selbst als in direktem Kontakt mit dem Inhalt seines Bewusstseins erleben muss. Was es nicht erleben kann, ist die Tatsache, dass sein Erleben immer in einem *Medium* stattfindet. Eine vollständig

²⁸ Eine kurze deutsche Erläuterung des Begriffs "phänomenale Transparenz" findet sich in Metzinger 2003b.

transparente Repräsentation zeichnet sich dadurch aus, dass die Mechanismen, die zu ihrer Aktivierung geführt haben, und die Tatsache, dass es einen konkreten inneren Zustand gibt, der ihren Gehalt trägt, introspektiv nicht mehr erkannt werden können. Die Phänomenologie der Transparenz ist die Phänomenologie des naiven Realismus.

Phänomenale Repräsentationen sind transparent, weil ihr Inhalt und vor allem dessen Existenz in allen möglichen Kontexten festzustehen scheint: Das Buch, das Sie jetzt in Händen halten, wird dem subjektiven Erleben nach immer nur dieses Buch bleiben - egal wie sich die äußere Wahrnehmungssituation ändert. Was Sie erleben, ist nicht ein „aktiver Objektemulator“, der gerade in ihr globales Realitätsmodell integriert worden ist, sondern einfach nur der *Gehalt* des zugrunde liegenden Repräsentationsvorgangs, eben: dieses *Buch* als Ihnen selbst hier und jetzt anstrengungslos gegebenes. Die beste Art und Weise, sich den Begriff der Transparenz klarzumachen, besteht nämlich darin, zwischen dem Vehikel und dem Gehalt einer Repräsentation zu unterscheiden, zwischen repräsentationalem Träger und repräsentationalem Inhalt.²⁹

Der repräsentationale Träger ihres Erlebnisses ist ein bestimmter Vorgang im Gehirn. Diesen Vorgang - der in keiner konkreten Weise etwas „Buchhaftes“ an sich hat - erleben Sie nicht bewusst, er ist transparent in dem Sinne, dass Sie durch ihn hindurch schauen. *Worauf* Sie schauen ist sein repräsentationaler Inhalt, eben die sensorisch gegebene Existenz eines Buchs, hier und jetzt. Der Inhalt ist also eine abstrakte Eigenschaft des konkreten repräsentationalen Zustands in ihrem Kopf. Wenn der repräsentationale Träger ein gut und zuverlässig funktionierendes Instrument zur Wissensgewinnung ist, dann erlaubt er Ihnen dank seiner Transparenz „durch ihn hindurch“ direkt auf die Welt, auf das Buch zu schauen. Er macht die von ihm getragene Information global verfügbar, ohne dass Sie sich darum kümmern müssen, *wie* das geschieht. Das Besondere an der phänomenalen Variante der Repräsentation ist nun, dass Sie diesen Inhalt auch dann, wenn Sie halluzinieren und es das Buch gar nicht gibt, immer noch als maximal *konkret*, als absolut eindeutig, als direkt und unmittelbar gegeben erleben. Phänomenale Repräsentationen sind solche, für die wir die Unterscheidung zwischen repräsentationalem Gehalt und repräsentationalem Träger im subjektiven Erleben nicht machen können.

²⁹ Vgl. dazu auch Dretske 1998: 45ff .

Es gibt natürlich Gegenbeispiele, und sie sind hilfreich, um den Begriff der „Transparenz“ noch besser zu verstehen. *Opake* phänomenale Repräsentationen entstehen zum Beispiel dann, wenn die Information, *dass* es sich bei ihrem Inhalt um das Resultat eines inneren Darstellungsvorgangs handelt, plötzlich global verfügbar wird. Wenn Sie entdecken, dass es das Buch in Wirklichkeit gar nicht gibt, dann wird die Halluzination zur Pseudohalluzination: Auch auf der Ebene des Erlebens selbst ist jetzt die Information verfügbar, dass Sie nicht auf die Welt schauen, sondern auf einen aktiven repräsentationalen Zustand, der im Moment allem Anschein nach kein gutes Instrument zur Wissensgewinnung ist. Der phänomenale Buchzustand wird undurchsichtig. Was Sie verlieren, ist die *sensorische* Transparenz. Ihnen wird die Tatsache bewusst, dass Wahrnehmungen durch Sinnesorgane erzeugt werden und dass diese Organe nicht in allen Situationen absolut zuverlässig funktionieren.

Nehmen wir weiter an, dass Sie jetzt sogar plötzlich entdecken, dass sich nicht nur die Buchwahrnehmung, sondern auch Ihr gesamtes philosophisches Nachdenken über das Problem des Bewusstseins in einem Traum ereignet. Dann wird dieser Traum zum Klartraum.³⁰ Die Tatsache, dass Sie momentan nicht in einer Welt leben, sondern nur in einem *Weltmodell* wird nun global verfügbar: Sie können diese Information zur Handlungskontrolle, im weiteren Nachdenken oder für die Aufmerksamkeitslenkung einsetzen. Was Sie verlieren, ist *globale* Transparenz. Interessanterweise ist kognitive Verfügbarkeit allein nicht hinreichend, um den naiven Realismus des phänomenalen Erlebens zu durchbrechen. Man kann sich nicht einfach aus dem phänomenalen Modell der Wirklichkeit „hinausdenken“, indem man seine Meinungen über dieses Modell ändert: Die Transparenz phänomenaler Repräsentationen ist kognitiv nicht penetrabel, phänomenales Wissen ist nicht dasselbe wie begrifflich-propositionales Wissen.

Diesen Gedanken muss man nun im letzten Schritt wieder auf das Selbstmodell anwenden. Meine Hauptthese ist: Wir selbst sind Systeme, die erlebnismäßig nicht in der Lage sind, ihr eigenes subsymbolisches Selbstmodell *als* Selbstmodell zu erkennen. Deshalb operieren wir unter den Bedingungen eines „naiv-realistischen Selbstmissverständnisses“: Wir erleben uns selbst, als wären wir in direktem und unmittelbarem epistemischen Kontakt mit uns selbst. Und

³⁰ Warum Klarträume eine philosophisch relevante Klasse von Bewusstseinszuständen sind, erläutere ich in Metzinger 2003a, Abschnitt 7.2.5.

auf diese Weise entsteht – das ist der Kern der Selbstmodelltheorie - erstmals ein basales „Ichgefühl“, ein für das betreffende System unhintergehbare phänomenales Selbst. Sehr poetisch ausgedrückt hat diesen Zusammenhang Antonio Damasio: „Das Selbst ist die Antwort auf eine Frage, die nie gestellt wurde.“³¹

6. Der sechste Schritt: Das phänomenale Modell der Intentionalitätsrelation

Aus einem transparenten Modell der Welt entsteht eine Wirklichkeit. Aus einem transparenten Modell des Systems entsteht ein in diese Wirklichkeit eingebettetes Selbst. Wenn nun noch eine transparente Darstellung der wechselnden *Beziehungen* entsteht, die dieses Selbst im Wahrnehmen und im Handeln vorübergehend zu Gegenständen und anderen Personen in dieser Wirklichkeit aufbaut, dann tritt das hervor, was ich zu Beginn die „phänomenale Erste-Person-Perspektive“ genannt habe. Eine genuine Innenperspektive entsteht genau dann, wenn das System sich *für sich selbst* noch einmal als mit der Welt interagierend darstellt, diese Darstellung aber wieder nicht *als* Darstellung erkennt. Es besitzt dann ein bewusstes Modell der Intentionalitätsrelation. Sein Bewusstseinsraum ist ein perspektivischer Raum und seine Erlebnisse sind jetzt *subjektive* Erlebnisse.

Die Intentionalitätsrelation ist in der Hauptsache die Wissensbeziehung zwischen Subjekt und Objekt: Ein mentaler Zustand wird dadurch zu einem Träger von Wissen, dass er über sich selbst hinaus verweist – gewissermaßen wie ein Pfeil, der aus dem Geist eines Menschen auf einen Gegenstand in der wirklichen oder sogar in einer möglichen Welt zeigt. Philosophen sagen dann, dass dieser Zustand einen *intentionalen Inhalt* besitzt. Der Inhalt ist das, worauf der Pfeil zeigt. Dieser Inhalt kann ein Bild, eine Aussage oder auch ein Handlungsziel sein. Wenn viele solcher Pfeile im Bewusstsein verfügbar sind, dann entsteht eine zeitlich ausgedehnte Erste-Person-Perspektive.³² Es gibt dann nicht mehr nur ein neurobiologisch verankertes Kernselbst, ein Selbstpräsentat, sondern auch eine dynamische, phänomenale Simulation des Selbst als eines über ständig wechselnde Wissens- und Handlungsbeziehungen

³¹ Vgl. Damasio 1999: 316.

³² Eine wesentlich ausführlichere Darstellung dieses Punkts findet sich, mit Blick auf bewusste Willensakte, in Metzinger 2005a.

in die Welt eingebundenen Subjekts. Der bewusste Mensch ist ein System, das bei einzelnen repräsentationalen Akten die Repräsentationsbeziehung selbst noch einmal *ko-repräsentieren* kann. Der Inhalt höherstufiger Formen des Selbstbewusstseins ist immer eine Relation: Das Selbst *im Moment des Erkennens*,³³ das Selbst *im Akt des Handelns*.

Natürlich ist die Art und Weise, in der wir diese Relation subjektiv erleben, eine stark vereinfachte Version der realen Prozesse – gewissermaßen eine funktional adäquate Konfabulation. Die Evolution hat auch in diesem Fall wieder eine einfache, eine elegante Lösung favorisiert. Das virtuelle Selbst, das sich in der phänomenalen Welt bewegt, besitzt kein Gehirn, kein Motorsystem und keine Sinnesorgane: Teile der Umgebung erscheinen direkt in seinem Geist, der Wahrnehmungsprozess ist anstrengungslos und unmittelbar. Auch Körperbewegungen werden scheinbar „direkt“ ausgelöst. Solche Effekte sind typisch für unsere Form des subjektiven Erlebens, und sie sind – als neurokomputationale Strategie betrachtet – die Vorteile einer benutzerfreundlichen Oberfläche. Das, was wir eben als „Transparenz“ kennen gelernt haben, ist eine Art, die *Geschlossenheit* dieser multimodalen, hochdimensionalen Oberfläche zu beschreiben. Das phänomenale Selbst ist der Teil dieser Oberfläche, den das System benutzt, um sich selbst als ein Ganzes zu fühlen, um sich für sich selbst als erkennendes Ich darzustellen und um sich selbst als Agenten zu begreifen. Dieser virtuelle Agent „sieht mit den Augen“ und „handelt mit den Händen“. Er weiß nicht, dass er einen visuellen Cortex besitzt, und genau so wenig, dass er einen motorischen Cortex besitzt. Er ist das Interface, welches das System benutzt, um sich seine eigene Hardware funktional anzueignen, um *autonom* zu werden. Die intentionalen Pfeile, die diesen Agenten mit Gegenständen und anderen Selbsten innerhalb des gerade aktiven Wirklichkeitsmodells verbinden, sind phänomenale Repräsentationen von vorübergehend auftretenden Subjekt-Objekt-Beziehungen – und auch sie können nicht *als* Repräsentationsprozesse erkannt werden.

All dies spielt sich innerhalb eines phänomenalen Gegenwartsfensters ab. Die Inhalte des phänomenalen Erlebens erzeugen nämlich nicht nur eine Welt, sondern auch eine *Gegenwart*.³⁴ Wenn man so will, dann ist phänomenales Bewusstsein sogar in seinem Kern genau dies: Die Erzeugung einer

³³ Vgl. Damasio 1999: 168ff.

³⁴ Vgl. Metzinger 2003a, Abschnitt 3.2.2.

Gegenwartsinsel im physikalischen Fluss der Zeit.³⁵ Erleben heißt „Gegenwärtigsein“. Es bedeutet, Information in einer sehr speziellen Weise zu verarbeiten. Sie besteht darin, bereits repräsentierte Einzelereignisse immer wieder und kontinuierlich zu zeitlichen Gestalten zusammenzufassen, zu einem Moment im Bewusstsein. Viele empirische Daten zeigen heute, dass die bewusst erlebte Gegenwart in einem bestimmten Sinne eine *erinnerte* Gegenwart ist.³⁶ Auch das phänomenale Jetzt ist in diesem Sinne selbst ein repräsentationales Konstrukt, es ist eine *virtuelle* Gegenwart, und an diesem Punkt kann man sich erstmals klarmachen, was es überhaupt bedeutet, zu sagen, dass der phänomenale Raum ein virtueller Raum ist: Sein Inhalt ist eine *mögliche* Realität.³⁷ Der Realismus des phänomenalen Erlebens entsteht dadurch, dass in ihm eine Möglichkeit – die beste Hypothese, die es im Moment gibt – unhintergebar als eine Wirklichkeit – eine *Aktualität* – dargestellt wird. Auch diesen Punkt muss man nun am Ende wieder auf den Sonderfall der phänomenalen Selbstmodellierung anwenden: Weil die Virtualität des Selbstmodells und die Virtualität des Gegenwartsfensters nicht auf der Ebene des subjektiven Erlebens verfügbar sind, wird das in ihnen dargestellte System zu einem *jetzt anwesenden Subjekt*.

Zum Schluss biete ich meinen Lesern noch eine Metapher an, die den eben skizzierten Gedankengang noch einmal illustrieren und verdeutlichen soll. Die Metapher ist besonders interessant, denn sie ist nicht nur anschaulich, sondern sie enthält auch einen logischen Fehler. Es ist die „Verwechslungsmetapher“: *De facto* sind wir selbst Systeme, die sich selbst ständig mit dem Inhalt des von ihnen selbst erzeugten subsymbolischen Selbstmodells „verwechseln“. Indem wir dies tun, generieren wir eine stabile und kohärente Ich-Illusion, die wir auf der Ebene des bewussten Erlebens nicht transzendieren können. Und genau das ist es, was es bedeutet, eine nicht-begriffliche Erste-Person-Perspektive zu besitzen, einen präreflexiven, phänomenalen Standpunkt, der allen späteren Formen begrifflich vermittelten und reflexiven Selbstbewusstseins zugrunde liegt, allen späteren Formen von sozialer Kognition und Ich-Du-Beziehungen: Die Tatsache, dass wir unser subsymbolisches Selbstmodell nicht *als* Modell

³⁵ Vgl. Ruhnau 1995 und die dort gegebenen Literaturhinweise, insbesondere auf die Arbeiten von Ernst Pöppel.

³⁶ Vgl. z.B. Edelmann 1989.

³⁷ Meine Vorstellungen konvergieren in diesem Punkt sehr stark mit denen von Antti Revonsuo: *Virtual reality* ist die beste technische Metapher für phänomenales Bewusstsein, die wir gegenwärtig besitzen. Vgl. Revonsuo 1995, 2000.

erleben können. Der Kern der Subjektivität des Mentalen liegt also in diesem Akt der „Selbstverwechslung“, in der Transparenz des Selbstmodells. Aber Vorsicht! - verfangen Sie sich nicht in meiner didaktischen Metapher, in der Idee der Selbstverwechslung. Bei näherem Hinsehen enthalten nämlich der allzu populäre Begriff der „Ich-Illusion“ und die von mir eben angebotene Metapher des „Sich-mit-seinem-eigenen-inneren-Bild-von-sich-selbst-Verwechselns“ einen logischen Fehler: Etwas, das noch kein epistemisches Subjekt in einem starken Sinne begrifflich-propositionalen Wissens ist, *kann* sich überhaupt noch nicht mit irgendetwas verwechseln. Wahrheit und Falschheit, Realität und Illusion gibt es für ein biologisches Informationsverarbeitungssystem auf der fraglichen Entwicklungsstufe noch gar nicht. Es gibt zudem niemand *im* System, der sich täuschen könnte, denn der Homunkulus existiert nicht. Hüten Sie sich deshalb vor allem auch auf *theoretischer* Ebene immer davor, sich mit ihrem eigenen Bild von sich selbst zu verwechseln. Hüten Sie sich vor dem kleinen Männchen im Kopf!

- Armstrong, K. C. & Ramachandran, V. S. (2003). Projecting sensations to external objects: Evidence from skin conductance response. *Proceedings of the Royal Society of London: Biological*, **270**, 1499-1506.
- Baars, B.J. (1988). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.
- Baker, L.R. (1998). The first-person perspective: A test for naturalism. *American Philosophical Quarterly*, **35**, 327-46.
- Baker, L.R. (2000). Die Perspektive der ersten Person: Ein Test für den Naturalismus. In G. Keil und H. Schnädelbach (Hrsg.), *Naturalismus - Philosophische Beiträge*. Frankfurt am Main: Suhrkamp.
- Bermúdez, J.L., Marcel, A. & Eilan, N. (1995)[eds]. *The Body and the Self*. Cambridge, MA: MIT Press.
- Cummins, R. (1983). *The Nature of Psychological Explanation*. Cambridge, MA: MIT Press.
- Baker, L.R. (1998). The first-person perspective: A test for naturalism. *American Philosophical Quarterly*, **35**, 327-46.
- Bieri, P. (1987). *Evolution, Erkenntnis und Kognition*. In Lüttersfelds, W. (Hrsg.), *Transzendente oder Evolutionäre Erkenntnistheorie?* Darmstadt: Wissenschaftliche Buchgesellschaft
- Bischof-Köhler, D. (1989). *Spiegelbild und Empathie*. Bern: Huber. Nachdruck 1993.
- Bischof-Köhler, D. (1996). Ichbewusstsein und Zeitvergegenwärtigung. Zur Phylogenese spezifisch menschlicher Erkenntnisformen. In Barkhaus, A., Mayer, M., Roughley, N. & Thürnau, D. (Hrsg.), *Identität, Leiblichkeit, Normativität. Neue Horizonte anthropologischen Denkens*. Frankfurt am Main: Suhrkamp.
- Botvinick, M. & Cohen, J. (1998). Rubber hand “feel” touch that eyes see. *Nature*, **391**, 756.
- Botvinick, M. (2004). Probing the neural basis of body ownership. *Science*, **305**, 782-3.
- Churchland, P.M. (1989). *A Neurocomputational Perspective*. Cambridge, MA / London: MIT Press.
- Damasio, A. (1994). *Descartes' Error*. New York: Putnam/Grosset.

- Damasio, A. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. Harcourt Brace & Company.
- Dennett, D.C. (1987b). *The Intentional Stance*. Cambridge, MA und London: MIT Press.
- Dittrich, A. (1985). *Ätiologie-unabhängige Strukturen veränderter Wachbewusstseinszustände*. Stuttgart: Enke.
- Dretske, F. (1988). *Explaining Behavior - Reasons in a World of Causes*. Cambridge, MA: MIT Press.
- Dretske, F. (1998). *Die Naturalisierung des Geistes*. Paderborn: mentis.
- Edelman, G.M. (1989). *The Remembered Present: A Biological Theory of Consciousness*. New York: Basic Books.
- Ehrsson, H.H., Spence, C. & Passingham, R.E. (2004). That's my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science*, **305**, 875-7.
- Gallese, V. & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, **2**, 493-501.
- Grush, R. (1997). The architecture of representation. *Philosophical Psychology*, **10**, 5-25.
- Grush, R. (1998). Wahrnehmung, Vorstellung, und die sensomotorische Schleife. In Heckmann, H.-D. & Esken, F. (Hrsg.), *Bewusstsein und Repräsentation*. Paderborn: mentis.
- Kinsbourne, M. (1995). Awareness of one's own body: An attentional theory of its nature, development, and brain basis. In Bermúdez et al. 1995.
- Lycan, W.G. (1996). *Consciousness and Experience*. Cambridge, MA: MIT Press.
- Melzack, R. (1989). Phantom limbs, the self and the brain: The D.O. Hebb memorial lecture. *Canadian Psychology*, **30**, 1-16.
- Melzack, R. (1992). Phantom limbs. *Scientific American*, **266**, 90-6.
- Melzack, R., Israel, R., Lacroix, R. & Schultz, G. (1997). Phantom limbs in people with congenital limb deficiency or amputation in early childhood. *Brain*, **120** (Pt 9), 1603-20.
- Metzinger, T. (1993;1999). *Subjekt und Selbstmodell. Die Perspektivität phänomenalen Bewusstseins vor dem Hintergrund einer naturalistischen Theorie mentaler Repräsentation*. Paderborn: mentis.
- Metzinger, T. (1995a). Perspektivische Fakten? Die Naturalisierung des „Blick von nirgendwo“. In G. Meggle & J. Nida-Rümelin (1997)[Hrsg.], *ANALYOMEN 2 - Perspektiven der Analytischen Philosophie*. Berlin und New York: de Gruyter. S. 103-10.
- Metzinger, T. (1995b; 5., erweiterte Auflage 2005)[Hrsg.]. *Bewusstsein - Beiträge aus der Gegenwartsphilosophie*. Paderborn: mentis.
- Metzinger, T. (1996). „Niemand sein“. In S. Krämer (Hrsg.), *Bewusstsein - Philosophische Positionen*. Frankfurt am Main: Suhrkamp.
- Metzinger, T. (1997). Ich-Störungen als pathologische Formen mentaler Selbstmodellierung. In Northoff, G. [Hrsg.], *Neuropsychiatrie und Neurophilosophie*. Paderborn: mentis.
- Metzinger, T. (2000a). *Die Selbstmodell-Theorie der Subjektivität: Eine Kurzdarstellung für Nicht-Philosophen in fünf Schritten*. In W. Greve (Hrsg.), *Psychologie des Selbst*. Weinheim: BELTZ / Psychologie Verlags Union.
- Metzinger, T. (2000b). The *subjectivity* of subjective experience: A representationalist analysis of the first-person perspective. In T. Metzinger (ed), *Neural Correlates of Consciousness - Empirical and Conceptual Questions*. Cambridge, MA: MIT Press. Überarbeitete Version (2004): *Networks*, **3-4**, 33-64.
- Metzinger, T. (2000c). *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. Cambridge, MA: MIT Press.
- Metzinger, T. (2003a;2004). *Being No One. The Self-Model Theory of Subjectivity*. Cambridge, MA: MIT Press.

- Metzinger, T. (2003b). *Phänomenale Transparenz und kognitive Selbstbezugnahme*. In U. Haas-Spohn (Hrsg.), *Intentionalität zwischen Subjektivität und Weltbezug*. Paderborn: mentis. Seite 411-459.
- Metzinger, T. (2003c). *Der Begriff einer Bewusstseinskultur*. In G. Kaiser (Hrsg.), *Jahrbuch 2002/2003 des Wissenschaftszentrums Nordrhein-Westfalen*. Düsseldorf: Wissenschaftszentrum Nordrhein-Westfalen. S. 150-171.
- Metzinger, T. (2003d). Phenomenal transparency and cognitive self-reference. *Phenomenology and the Cognitive Sciences*, **2**, 353-393. doi:10.1023/B:PHEN.0000007366.42918.eb
- Metzinger, T. (2004a). Why are identity-disorders interesting for philosophers? In Thomas Schramme und Johannes Thome (Hrsg.), *Philosophy and Psychiatry*. Berlin: de Gruyter.
- Metzinger, T. (2004b). Appearance is not knowledge: The incoherent strawman, content-content confusions and mindless conscious subjects. Eingeladener Kommentar zu Alva Noë und Evan Thompson: „Are there neural correlates of consciousness?“, in einer Sondernummer von *Journal of Consciousness Studies*, **11 (1)**, 67-71.
- Metzinger, T. (2005a). Conscious volition and mental representation: Towards a more fine-grained analysis. In N. Sebanz und W. Prinz (Hrsg.), *Disorders of Volition*. Cambridge, MA: MIT Press.
- Metzinger, T. (2005b). Précis of „Being No One“. In *PSYCHE - An Interdisciplinary Journal of Research on Consciousness*. URL: <http://psyche.cs.monash.edu.au/>
- Metzinger, T. (2005c). The pre-scientific concept of a “soul”: A neurophenomenological hypothesis about its origin. In Markus F. Peschl (Hrsg.), *Die Rolle der Seele in der Kognitionswissenschaft und der Neurowissenschaft - Auf der Suche nach dem Konzept/Substrat der Seele*. Würzburg: Königshausen und Neumann.
- Metzinger, T. & Gallese, V. (2003). The emergence of a shared action ontology: building blocks for a theory. In G. Knoblich, B. Elsner, G. von Aschersleben, und T. Metzinger (eds), *Self and Action*. Special issue of *Consciousness & Cognition* (12:4), December 2003, 549-571.
- Millikan, R.G. (1984). *Language, Thought, and other Biological Categories*. Cambridge/MA: MIT-Press.
- Millikan, R.G. (1993). *White Queen Psychology and Other Essays for Alice*. Cambridge/MA: MIT-Press.
- Nagel, T. (1986). *Der Blick von nirgendwo*. Frankfurt am Main: Suhrkamp.
- Newen, A. & Vogeley, K. (2000)[Hrsg.]. *Das Selbst und seine neurobiologischen Grundlagen*. Paderborn: mentis.
- O'Shaughnessy, B. (1995). Proprioception and the body image. In Bermúdez *et al.* 1995.
- Ramachandran, V.S. & Blakeslee, S. (1998). *Phantoms in the Brain*. New York: William Morrow and Company, Inc.
- Ramachandran, V.S. & Rogers-Ramachandran, D. (1996). Synaesthesia in phantom limbs induced with mirrors. *Proceedings of the Royal Society London*, **B**, 377-86.
- Revonsuo, A. (1995). Consciousness, dreams, and virtual realities. *Philosophical Psychology*, **8**, 35-58.
- Revonsuo, A. (2000). Prospects for a scientific research program on consciousness. In Thomas Metzinger (ed), *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. Cambridge, MA: MIT Press.
- Ruhnau, E. (1995). Zeit-Gestalt und der Beobachter. In Metzinger 1995.
- Yates, J. (1985). The content of awareness is a model of the world. *Psychological Review*, **92**, 249-84.