

Empirische Perspektiven aus Sicht der Selbstmodell-Theorie der Subjektivität: Eine Kurzdarstellung mit Beispielen¹

Zusammenfassung: Dieser Beitrag bietet eine Kurzdarstellung der Selbstmodell-Theorie der Subjektivität (SMT; Metzinger, 2003a) für empirische Forscher. Anhand von empirischen Beispielen aus verschiedenen Wissenschaftsdisziplinen werden einige Mechanismen erörtert, durch die in einem physisch realisierten Informationsverarbeitungssystem wie dem menschlichen Gehirn Selbstbewußtsein entstehen könnte. Der Beitrag stellt zwei zentrale Begriffe vor: das »phänomenale Selbstmodell« (PSM) und das »phänomenale Modell der Intentionalitätsrelation« (PMIR). Dabei wird eine repräsentationalistische Analyse des bewußten Selbst und der Entstehung einer Erste-Person-Perspektive skizziert.

Stichworte: Bewußtsein; Selbstbewußtsein; Erste-Person-Perspektive; Meinigkeit; Agentivität; Selbstmodell; phänomenale Transparenz; Phantomglieder; Robotik; Gummihand-Illusion; Ganzkörper-Illusion; außerkörperliche Erfahrung; Evolution des Werkzeuggebrauchs

Das Ziel dieses Beitrags besteht darin, eine kurze Darstellung der »Selbstmodell-Theorie der Subjektivität« (SMT) anzubieten, die sich an naturwissenschaftlich orientierte Leser richtet, die selbst keine Berufsphilosophen sind, aber dennoch an philosophischen Theorien des Selbstbewußtseins interessiert sind.² Dazu werde ich eine Reihe empirischer Beispiele aus verschiedenen Disziplinen einsetzen, um einige Kernaussagen der Theorie zu veranschaulichen und die Reichweite des Erklärungspotentials sowie die Vorhersageleistung der SMT aufzuzeigen. Die SMT ist eine philosophisch-kognitionswissenschaftliche Theorie darüber, was es bedeutet, ein Selbst zu sein. Sie

¹ Dieser Text ist eine stark erweiterte, aktualisierte und überarbeitete Version eines Artikels, der 2008 in der Zeitschrift *Progress in Brain Research* erschienen ist. Ich danke Sarah Raphael für ihre Hilfe bei der Übersetzung einer ersten Rohfassung dieses Textes aus dem Englischen und Jennifer M. Windt für eine Vielzahl von kritischen Kommentaren und Verbesserungsvorschlägen. Ich bin auch Ingrid Götsch zu Dank verpflichtet, die mich in der Schlußphase auf zuverlässige und professionelle Weise unterstützt hat. Für ihre sorgfältige Hilfe bei der Schlußkorrektur danke ich Hannes Boelsen, Regina Fabry und Lisa Quadt.

² Eine populärwissenschaftliche Darstellung der wichtigsten Grundgedanken mit vielen Beispielen ist *Der Ego-Tunnel* (Berlin, 2009; ⁴2012), die umfassende englischsprachige Darstellung dieser Theorie ist T. Metzinger, *Being No One – The Self-Model Theory of Subjectivity*, (Cambridge, MA; 2003). Die kürzeste, kostenlos zugängliche Zusammenfassung findet sich in *Scholarpedia* 2 (10), S. 4174, eine sehr einfache deutsche Übersichtsdarstellung erschien 2005 unter dem Titel »Die Selbstmodell- Theorie der Subjektivität: Eine Kurzdarstellung in sechs Schritten« in C. S. Herrmann, M. Pauen, J. W. Rieger und S. Schickel (Hg.), *Bewußtsein. Philosophie, Neurowissenschaften, Ethik*, Stuttgart, S. 242–269. Eine etwas gehaltvollere, an den wichtigsten begrifflichen Werkzeugen orientierte Übersichtsdarstellung mit weiterführenden Literaturverweisen und ohne jeden Bezug auf empirische Daten ist »*Being No One – Eine sehr kurze deutsche Zusammenfassung*«, in T. Metzinger, *Grundkurs Philosophie des Geistes. Band 1: Phänomenales Bewußtsein*, Paderborn 2006, S. 424–475.

ist auch eine Theorie darüber, was es bedeutet, daß geistige Zustände »subjektive« Zustände sind und daß ein bestimmtes System eine »phänomenale Erste-Person-Perspektive« besitzt. Eine der ontologischen Hintergrundannahmen dieser Theorie ist, daß das Selbst keine »Substanz« im technischen philosophischen Sinne von etwas ist, das »sich selbst im Sein halten« kann, das also auch dann noch existieren könnte, wenn der Körper, das Gehirn oder der gesamte Rest des physikalischen Universums verschwänden. Es ist keine im ontologischen Sinne autonome, selbsttragende Entität, kein Einzelding oder ein mysteriöses *Etwas* im metaphysischen Sinne. Selbste in diesem Sinne gibt es in der Welt nicht: Selbste oder Subjekte gehören nicht zu den irreduziblen Grundbestandteilen der Wirklichkeit und sie dauern auch nicht über die Zeit hinweg an.³ Was existiert, ist natürlich das *Erleben* eines Selbst, ebenso wie die verschiedenen, ständig wechselnden Inhalte des Selbstbewußtseins. Das ist es auch, was Philosophen meinen, wenn sie über »das *phänomenale* Selbst« sprechen: die Art und Weise, in der man sich selbst *erscheint*, subjektiv und auf der Ebene des bewußten Erlebens. Der Begriff des *phänomenalen* Selbst muß dabei streng von dem des Selbst im Sinne einer *Substanz* unterschieden werden; denn Letzteres gibt es, wie wir eben gesehen haben, nicht. Im Folgenden wird es daher immer um das phänomenale Selbst gehen.

In der SMT wird die bewußte Erfahrung, ein Selbst zu sein, in erster Linie als Resultat von komplexen Informationsverarbeitungs- und Darstellungsprozessen im zentralen Nervensystem gesehen, als ein Ergebnis dynamischer Selbstorganisationsvorgänge im

³ Ein Überblick über verschiedene Weisen, in denen man auf philosophisch-begrifflicher Ebene die Existenz eines ontologisch autonomen »Selbst« bestreiten kann, findet sich in Metzinger 2011. Aus einer empirischen Perspektive ist das Folgende offensichtlich: Menschliche Wesen sind dynamische, sozial situierte *Systeme*; Selbstbewußtsein ist ein komplexer Vorgang, der zeitweise bestimmte Fähigkeiten erzeugt, die begrifflich am besten als globale Systemeigenschaften beschrieben werden (rationales Denken, selektive Aufmerksamkeitslenkung, flexible und kontextsensitive Handlungskontrolle, sprachliche Selbstbezugnahme usw.). Viele theoretische Probleme entstehen ganz einfach dadurch, daß diese Fähigkeiten und globalen Systemeigenschaften falsch beschrieben und dabei »verdinglicht« werden. Für Nicht-Philosophen wichtig ist hier deshalb vielleicht auch der Hinweis, daß das in der Alltagspsychologie, den Medien, aber auch in manchen akademischen Kontexten immer noch weit verbreitete Sprechen von »dem Ich« oder »meinem Selbst« einen schweren *logischen* Fehler darstellt. Das Personalpronomen der ersten Person Singular – der sprachliche Ausdruck »Ich« – verweist immer auf den Sprecher, der es aktuell verwendet. Seine logische Funktion ist nicht die eines Gattungsbegriffs oder die Bezugnahme auf ein konkretes Einzelding, sondern die der Selbstlokalisierung eines Sprechers in einem Äußerungskontext. In grammatischer und semantischer Hinsicht ist »Ich« also ein singulärer Term, der an einen bestimmten Äußerungskontext gebunden ist: Dieser Kontext besteht darin, daß der aktuelle Sprecher mit einem sprachlichen Werkzeug auf sich selbst zeigt. Trotzdem verwenden wir bei der sprachlichen Selbstbezugnahme den indexikalischen Ausdruck »Ich« sehr häufig so, als ob es sich dabei um einen Namen für ein inneres Ding oder eine Form von Objektreferenz handelt, also von Bezugnahme auf einen Gegenstand (Bennett & Hacker 2010: 12.4). Es gibt aber keine spezielle Gattung von Dingen (»Iche« oder »Selbste«), die man in sich tragen könnte wie ein Herz oder die man besitzen könnte wie ein Fahrrad oder einen Fußball. Und auch das in lebensweltlichen Kontexten allgegenwärtige Reden von unserem oder »meinem« Selbst ist logisch widersprüchlich, weil es dann ja schon jemanden geben müßte, der dieses Selbst »hat«, also ein Selbst hinter dem Selbst, das zu diesem in einer Besitzrelation steht. Das Selbst kann auch nichts »in mir« sein, weil dann ja das, mit dem ich identisch bin, nur ein konstituierender Teil von mir wäre.

Gehirn. Das phänomenale Selbst ist also kein Ding, sondern ein diskontinuierlicher Vorgang. Natürlich gibt es auch höherstufige, begrifflich vermittelte Formen des phänomenalen Selbstbewußtseins, die nicht nur neuronale, sondern auch *soziale* Korrelate besitzen.⁴ Die Theorie beginnt jedoch damit, sich auf die minimalen repräsentationalen und funktionalen Eigenschaften zu konzentrieren, die ein natürlich evolviertes informationsverarbeitendes System – wie der *Homo sapiens* – besitzen muß, um die Bedingungen der Möglichkeit für die Entstehung solcher höherstufigen Formen von Selbstbewußtsein und Intersubjektivität zu erfüllen. Die meisten Philosophen sind sich heutzutage darüber einig, daß das eigentliche Problem darin liegt, zunächst die einfachste und elementarste Form unseres Zielphänomens zu verstehen. Dies ist die nicht-begriffliche, präreflexive und vorsprachliche Schicht des Selbstbewußtseins. Deswegen ist die erste Frage, die wir zu beantworten haben, die folgende: Was sind (relativ zu einer konkreten Klasse von Systemen, z.B. *Homo sapiens* oder einem bestimmten zukünftigen Robotertyp) die *minimal hinreichenden Bedingungen* für das Entstehen eines phänomenalen Selbst? Später kann man dann auch fragen, was *notwendige* und hinreichende Bedingungen bei allen denkbaren Systemen sein könnten, die Beantwortung dieser Frage ist aber nicht das Ziel des vorliegenden Textes.

Die SMT nimmt an, daß die zur Debatte stehenden Eigenschaften repräsentationale und funktionale Eigenschaften des Gehirns sind. Mit anderen Worten: Phänomenales Selbstbewußtsein, diejenige psychologische Eigenschaft, die uns dazu befähigt, überhaupt eine Person zu werden, wird mit Hilfe von begrifflichen Werkzeugen der *subpersonalen* Beschreibungsebene analysiert. In der Philosophie des Geistes wird ein solches Verfahren manchmal auch als »Naturalisierungsstrategie« bezeichnet: Ein besonders komplexes und schwer verständliches Phänomen – zum Beispiel das Entstehen des phänomenalen Bewußtseins überhaupt, oder auch die mit ihm zusammen auftretende subjektive Innenperspektive – wird begrifflich auf eine Weise analysiert, die es empirisch behandelbar machen soll. Indem sie klassische Probleme ihrer eigenen Disziplin auf neue Weise formulieren, versuchen naturalistische Philosophen, die fraglichen Phänomenbereiche für interdisziplinäre Untersuchungen und naturwissenschaftliche Forschungsprogramme zu öffnen, beispielsweise für die Kognitions- oder Neurowissenschaften. Der amerikanische Philosoph Josh Weisberg hat den Begriff der »*method of interdisciplinary constraint satisfaction*« (MICS) geprägt (Weisberg 2006). Sie besteht darin, gleichzeitig auf einer Vielzahl verschiedener Beschreibungsebenen sowohl empiri-

⁴ In unserem eigenen Fall ist das Selbstmodell häufig auch ein *Personmodell*, also die mentale Repräsentation eines autonomen, rationalen Subjekts. Wir erleben uns dann nicht bloß als intelligente Organismen, sondern zum Beispiel als vernünftige, um ethische Integrität bemühte Personen. Wenn man solche hochstufigen Eigenschaften des Menschen – Rationalität, Moralität oder Personalität – wirklich ernst nehmen will, dann muß man die graduelle Entstehungsgeschichte des ganz spezifischen *subpersonalen* funktionalen Profils untersuchen, welches in unserem eigenen Fall die selbstorganisierende Dynamik der sozialen Anerkennungsbeziehungen überhaupt erst ermöglicht hat, durch die sich diese neuen Eigenschaften erstmals realisieren konnten. Die Beziehung zwischen begrifflichen und nicht-begrifflichen Inhalten des Selbstbewußtseins habe ich in Metzinger (2003a) genauer analysiert; Metzinger (2003b) ist eine frühere deutsche Version dieses Textes. Eine Hypothese zur Rolle des unbewußten Selbstmodells bei der Entstehung nicht begrifflich vermittelter Formen der sozialen Kognition findet sich in Metzinger und Gallese (2003), vgl. dazu auch Kapitel 6 in Metzinger (2012).

sche als auch begriffliche Auflagen zu erfüllen, etwa mit Blick auf eine umfassende Theorie des Selbstbewußtseins. Die Hoffnung besteht darin, ein komplexes Erkenntnisziel gewissermaßen zu «triangulieren», indem man verschiedene Methoden und Informationsquellen gleichzeitig nutzt, um zunächst heuristisch fruchtbare Arbeitsbegriffe zu konstruieren. Diese können dann verfeinert werden und zur Formulierung testbarer Hypothesen verwendet werden. Dabei ist es eine zentrale Aufgabe für die Philosophie der Kognitionswissenschaft, aus metatheoretischer Perspektive begriffliche Instrumente zu entwickeln, die eine *Integration* über die verschiedenen Ebenen der Analyse hinweg ermöglichen und im Idealfall einen formalen Rahmen bereitstellen, der dann verschiedene Datensätze und unterschiedliche theoretische Herangehensweisen zusammenführen kann. SMT ist ein Beispiel für einen solchen Versuch.

Eine allerletzte Vorbemerkung: Die MICS, der Naturalismus und die Suche nach einer reduktiven Erklärung für das phänomenale Selbst sind nicht der Ausdruck einer scientistischen Ideologie, sondern ganz einfach eine rationale Forschungsheuristik. Wenn es sich zum Beispiel zeigen sollte, daß es – wie viele glauben⁵ – etwas am menschlichen Selbstbewußtsein gibt, das sich dem naturwissenschaftlichen Zugriff *prinzipiell* entzieht, dann werden seriöse naturalistische Philosophen auch mit diesem Resultat zufrieden sein. Sie hätten nämlich das erreicht, was sie von Anfang an ihr Ziel war, und was Philosophen einen »epistemischen Fortschritt« nennen. Diese Art von Fortschritt könnte darin bestehen, in der Lage zu sein, jetzt auf wesentlich präzisere und gehaltvollere Weise beschreiben zu können, *warum genau* es auf bestimmte Fragen prinzipiell keine befriedigenden naturwissenschaftlichen Antworten geben kann. Daher sind die seriösesten philosophischen Anti-Naturalisten automatisch auch diejenigen, die das tiefgreifendste Interesse an neuen empirischen Forschungsergebnissen haben. Naturalismus und Reduktionismus sind keine Ideologien oder gar ein primitiver neuer Religionsersatz – im Gegenteil: Gerade Anti-Naturalisten und Anti-Reduktionisten, die an ein irreduzibles, essentiell subjektives Element im menschlichen Geist glauben, wird man daran erkennen, daß sie den stärksten Ehrgeiz besitzen, ihre philosophischen Argumente auch auf empirischer Ebene überzeugend darzustellen und auch dort den entscheidenden Punkt präzise zu isolieren.

Der erste Schritt: Was genau ist das Problem?

Was wir in alltagspsychologischen Kontexten fälschlicherweise »das Ich« nennen, ist, wie oben diskutiert, das phänomenale Selbst: derjenige Aspekt des Selbstbewußtseins, der unmittelbar im subjektiven Erleben gegeben ist, als Inhalt des phänomenalen Erlebens. Das phänomenale Selbst ist vielleicht die interessanteste Form von phänomenalem Inhalt überhaupt. Es verleiht unserem Bewußtsein zwei äußerst interessante *strukturelle* Merkmale: Zentriertheit und Perspektivität. Solange es ein phänomenales Selbst gibt, ist unser Bewußtsein ein zentriertes Bewußtsein und an das gebunden, was Philosophen eine »Erste-Person-Perspektive« (1PP) nennen. Zustände, die sich innerhalb dieses Bewußtseinszentrums befinden, sind dem Erleben nach *meine eigenen* Zustände, da sie mit einem Gefühl der »Meinigkeit« ausgestattet sind, das unabhängig

⁵ Vgl. beispielsweise Nagel, 1986, besonders Kapitel 4, vgl. dazu auch Metzinger 1995a und 2011.

von Sprache oder begrifflichem Denken ist. In all meinem bewußten Erleben und Handeln gehe ich außerdem ständig wechselnde Beziehungen zur Umwelt und zu meinen eigenen geistigen Zuständen ein. Ich erlebe mich selbst als *gerichtet* – auf Objekte der Wahrnehmung, auf andere Menschen oder auf die Inhalte meiner eigenen geistigen Zustände. Dieser Prozeß läßt eine subjektive Innenperspektive entstehen. Die Tatsache, daß ich eine solche Innenperspektive besitze, ist mir wiederum kognitiv verfügbar.⁶ Mit anderen Worten: Von den allermeisten anderen Tieren unterscheiden wir Menschen uns dadurch, daß wir nicht nur eine subjektiv erlebte Innenperspektive *haben*, sondern uns auch bewußt als *Wesen mit einer solchen Innenperspektive* konzeptualisieren können. Wir können uns diese Eigenschaft nämlich gedanklich und sprachlich selbst zuschreiben, zum Beispiel indem wir den Begriff eines »Subjekts« auf uns selbst anwenden.

Das erste Problem jedoch ist die Tatsache, daß wir eigentlich gar nicht so genau wissen, was wir meinen, wenn wir in dieser Weise über solche Fragen sprechen. Es geht nicht nur darum, daß wir nicht in der Lage sind, Begriffe wie »Selbst« oder »Subjekt« präzise zu definieren. Das eigentliche Problem ist, daß diese Begriffe sich oft nicht auf wahrnehmbare Gegenstände oder Bereiche der Wirklichkeit zu beziehen scheinen. Was wir deshalb zu allererst verstehen müssen, ist, wie die strukturellen Merkmale unseres inneren Erlebens unseren problematischen *Umgang* mit diesen Begriffen bestimmen. Um die Logik der Selbstzuschreibung psychologischer Eigenschaften zu analysieren und zu verstehen, worauf sich diese Begriffe tatsächlich beziehen, muß man zuerst die repräsentationale Tiefenstruktur des bewußten Erlebens selbst untersuchen. Drei höherstufige phänomenale Eigenschaften sind in diesem Zusammenhang besonders interessant:

- »*Meinigkeit*«: Dies ist eine höherstufige Eigenschaft einzelner Formen von phänomenalem Gehalt. Es handelt sich dabei um ein unmittelbar gegebenes, direktes und nicht-begriffliches Erleben des Besitzens, der subjektiven Aneignung. Hier sind einige umgangssprachliche Beispiele dafür, wie wir uns im alltagspsychologischen Diskurs auf diese phänomenalen Eigenschaften zu beziehen versuchen: »Ich erlebe *mein* Bein subjektiv immer schon als einen Teil von *mir selbst*«; »Ich erlebe *meine* Gedanken und Gefühle immer als Teil *meines eigenen* Bewußtseins«; »Meine Willensakte werden immer von *mir selbst* initiiert«.
- »*Selbstheit*«: Dieses erlebnismäßig nicht hintergehbare Gefühl, ein autonomes Selbst zu *sein*, ist das Herzstück, die phänomenale Kerneigenschaft, nach der wir suchen. Während es bei der *Meinigkeit* um erlebte Teil-Ganzes-Beziehungen geht, haben wir es hier mit einer *globalen* Form der Identifikation zu tun, nämlich mit dem Körper oder der Person als einer Ganzheit. Wieder können einige Beispiele veranschaulichen, wie wir mit sprachlichen Mitteln von außen auf dieses besonders interessante Merkmal unseres inneren Erlebens hinweisen: »Ich bin *jemand*«;

⁶ Baker (1998) enthält eine erste Einführung in die Problematik des kognitiven Selbstbezugs als eine potentielle Schwierigkeit für den philosophischen Naturalismus, Baker 2000 ist die gekürzte deutsche Übersetzung dieses Textes. Vgl. dazu auch Metzinger (2003a; Abschnitt 6.4.4) und insbesondere Metzinger (2003c). Eine kritische, interessante und übersichtliche Besprechung meiner eigenen Theorie der Erste-Person-Perspektive ist Baker (2009).

»Ich erlebe mich selbst als *identisch* über die Zeit hinweg«; »Die Inhalte meines Selbstbewußtseins bilden immer ein kohärentes *Ganzes*«; »Mit dem fundamentalen, nicht-begrifflichen Inhalt meines Selbstbewußtseins bin ich vor allen kognitiven oder reflexiven Operationen *immer schon* vertraut«.

- »*Perspektivität*«: In unserem Zusammenhang ist Perspektivität das dominante Strukturmerkmal des Bewußtseinsraums als Ganzem: Er wird sehr häufig durch ein handelndes und erlebendes Subjekt zentriert, durch ein aktives Selbst, das ständig wechselnde Beziehungen zu sich selbst und der Welt eingeht. Beispiele: »Meine Welt besitzt einen unverrückbaren Mittelpunkt und dieser Mittelpunkt bin immer *ich selbst*«; »Bewußtsein zu haben, bedeutet, eine *individuelle Erste-Person-Perspektive* zu besitzen«; »In meinem Erleben sowohl von Personen und Objekten in der Außenwelt als auch bei der Introspektion meiner eigenen geistigen Zustände bin ich stets an diese Innenperspektive gebunden – ich bin ihr Ausgangspunkt, ihr *Ursprung*«.

Der nächste Schritt besteht in einer repräsentationalen und funktionalen Analyse dieser Zieleigenschaften. Man muß fragen: Welche funktionalen und repräsentationalen Eigenschaften muß ein gegebenes informationsverarbeitendes System besitzen, um die fragliche *phänomenale* Eigenschaft zu instantiieren? Welche dieser Eigenschaften sind hinreichend und sind einige davon vielleicht sogar begrifflich *notwendig*? Was *genau* bedeutet es für ein solches System, die Welt ebenso wie seine eigenen geistigen Zustände aus einer Erste-Person-Perspektive zu erleben? 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 Feinheit des phänomenalen Erlebens Rechnung trägt. Ich werde jetzt versuchen, die Grundlinien eines solchen Begriffsrahmens in den übrigen fünf Schritten zu skizzieren.

Der zweite Schritt: Das Selbstmodell

Der zweite Schritt besteht darin, eine neue theoretische Entität einzuführen: das phänomenale Selbstmodell (PSM). Es bildet den wichtigsten Teil der repräsentationalen Basis für die Instantiierung der zu erklärenden phänomenalen Eigenschaften (Cummins, 1983). Was ist eine geistige »Repräsentation«? Ein repräsentationaler Zustand, beispielsweise im Gehirn, ist ein Zustand, der einen bestimmten *Inhalt* hat, weil er auf etwas in der Welt gerichtet ist. Der Gehirnzustand selbst ist der physikalische Träger; der Inhalt ist die Bedeutung dieses Zustandes für das System. Eine innere Repräsentation *handelt von* etwas, sie besitzt semantische Eigenschaften: Korrekte Repräsentation impliziert die erfolgreiche Bezugnahme auf etwas, eine *Referenz*. Ein repräsentationaler Zustand fungiert oft als ein Stellvertreter für etwas Externes, den Referenten; er repräsentiert, weil er etwas anderes »vertritt«. Jedoch kann dieses »Etwas« (Philosophen nennen es das »intentionale Objekt«) auch ein vergangenes Ereignis, ein potentiell zukünftiges Handlungsziel oder sogar eine reine Möglichkeit sein – in diesen Fällen spricht man von Repräsentationen als *Simulationen*. Diese simulieren lediglich *mögliche* Sachverhalte; sie repräsentieren eine Möglichkeit, keine Wirklichkeit. Die SMT ist in

erster Linie eine repräsentationalistische Theorie des Bewußtseins, da sie bewußte Zustände als repräsentationale Zustände und bewußte Inhalte als repräsentationale Inhalte analysiert.

Eine unserer Schlüsselfragen war: Was ist die minimal hinreichende Menge an *repräsentationalen* Eigenschaften, die ein System entwickeln muß, um die relevanten Zieleigenschaften zu besitzen? Dies ist unsere erste, vorläufige Antwort: Das System muß eine kohärente und globale Selbstrepräsentation besitzen, ein konsistentes inneres Modell von sich selbst *als einer Ganzheit*. In unserem eigenen Fall ist das Selbstmodell eine episodisch aktive repräsentationale Einheit, deren Inhalt durch Eigenschaften des Systems selbst gebildet wird. Immer dann, wenn eine solche Selbstrepräsentation gebraucht wird, um die kausale Interaktion des Systems mit der äußeren oder inneren Umwelt zu regulieren, wird sie vorübergehend aktiviert – zum Beispiel, wenn Sie morgens aufwachen. Der SMT zufolge, passiert wenn Sie morgens aufwachen, wenn Sie *zu sich kommen*, das Folgende: Der biologische Organismus, der Sie sind, fährt sein PSM hoch, er aktiviert das bewußt erlebte Selbst. Dabei entsteht eine ganze Reihe von neuen funktionalen Eigenschaften. Erstens kann das System jetzt erstmals seine Aufmerksamkeit und auch andere seiner kognitiven Fähigkeiten auf sich selbst *als eine Ganzheit* richten. Damit werden dann auch globale Formen der Verhaltenskontrolle möglich, denn der erwachte Körper kann sich nun erstmals wieder *als eine Ganzheit* kausal selbst steuern. Zweitens aber *weiß* das System jetzt, daß es diese funktionalen Eigenschaften wiedererlangt hat, denn das PSM macht diese Information global verfügbar, stellt sie also auf der Ebene des bewußten Erlebens als *seine eigenen* Eigenschaften dar.

Mit anderen Worten: Was wir brauchen, ist eine umfassende Theorie des Selbstmodells von *Homo sapiens*.⁷ Ich gehe davon aus, daß eine solche Theorie in wesentlichen Teilen eine neurokomputationale Theorie sein wird (vgl. dazu beispielsweise Churchland, 1989; Howhy, 2013; Friston, 2010). Das bedeutet, daß es nicht nur eine wahre repräsentationale und funktionale Beschreibung des menschlichen Selbstmodells gibt, sondern auch eine wahre neurobiologische Beschreibung – zum Beispiel im Sinne eines weitgehend dezentralisierten, komplexen Aktivierungsmusters im Gehirn (ein klassisches Beispiel ist Damasio, 1999). Das PSM ist genau der Teil Ihres *mentalen* Selbstmodells, der gegenwärtig in die höchststufige integrierte Struktur, in Ihr globales Modell der Welt, eingebettet und deshalb auch für die Lenkung der introspektiven Aufmerksamkeit verfügbar ist (Yates, 1975; Baars, 1988; eine detailliertere Analyse möglicher Kriterien zur Unterscheidung von verschiedenen Stärken des Bewußtseins findet sich in Metzinger, 2003a, Kapitel 3). Ein wichtiger Aspekt dieses Gedankens ist, daß bestimmte Teile des Selbstmodells gleichzeitig unbewußt und über die kausalen Eigenschaften ihrer neuronalen Realisierung funktional aktiv sein können. Diese

⁷ Der methodologische Kern der akademischen Psychologie – wenn diese metatheoretische Randbemerkung eines philosophischen Beobachters erlaubt ist – kann jetzt auf neue und fruchtbare Weise analysiert werden. Psychologie ist *Selbstmodellforschung*. Sie ist eine wissenschaftliche Disziplin, die sich auf den repräsentationalen Gehalt, das funktionale Profil und die neurobiologische Realisierung des menschlichen Selbstmodells konzentriert, inklusive seiner ontogenetischen und evolutionären Geschichte und seiner notwendigen sozialen Korrelate.

Überlegung ist natürlich von zentraler Relevanz für die so genannte »psychosomatische« Medizin: »Psychosomatische Wechselwirkungen« sind in Wirklichkeit kausale Interaktionen zwischen bewußten und unbewußten Partitionen des Selbstmodells; ein traumatisierendes Erlebnis könnte zuerst auf der Ebene des PSM repräsentiert und dann funktional im unbewußten Selbstmodell weiterverarbeitet worden sein, wo es dann zum Beispiel einen direkten kausalen Beitrag zu einer Immunsuppression oder der Entwicklung somatoformer Störungen leistet. Die Enkodierung von traumatisierender Information im unbewußten Selbstmodell wäre also ein Vorgang, bei dem diese mit einer existierenden kognitiven Struktur (z.B. der autobiographischen oder emotionalen Selbstrepräsentation) verbunden und somit einer dauerhaften Speicherung zugeführt wird. Weil dieser Vorgang ein physikalischer Prozeß in einem komplexen System ist, kann er auch unerwartete kausale Effekte hervorrufen. Das PSM ist eine kohärente, multimodale Struktur, die unter anderem der Selbstlokalisierung dient und möglicherweise sogar auf einem teilweise angeborenen, »fest verdrahteten« Modell der räumlichen Eigenschaften des Systems beruht (mehr darüber im zweiten Beispiel⁸). In diesem Typ von Analyse wird der selbstbewußte Mensch als eine ganz bestimmte Art von informationsverarbeitendem System behandelt: Der subjektiv erlebte Gehalt des phänomenalen Selbst ist der repräsentationale Inhalt einer momentan aktiven, dynamischen Datenstruktur im zentralen Nervensystem dieses Systems.

Abgesehen von der repräsentationalen Beschreibungsebene kann man auch eine *funktionale* Analyse des Selbstmodells entwickeln. Während repräsentationale Zustände durch ihren Inhalt individuiert werden, ist ein funktionaler Zustand begrifflich charakterisiert durch seine *kausale Rolle*: die kausalen Beziehungen, in denen er zu Input-Ereignissen, Output-Ereignissen oder anderen inneren Vorgängen steht. Ein aktives Selbstmodell kann dann als ein subpersonaler funktionaler Zustand beschrieben werden: Ein PSM ist eine zusammenhängende (und möglicherweise komplexe) Menge von Kausalbeziehungen, die jeweils zu einem bestimmten Zeitpunkt realisiert sein können oder auch nicht. Da dieser funktionale Zustand durch einen konkreten neurobiologischen Zustand realisiert ist, spielt er automatisch eine bestimmte kausale Rolle im System. Beispielsweise kann er ein Element in einer intelligenten Form von Informationsverarbeitung sein. Aus der Perspektive der klassischen Kognitionswissenschaft kann man diesen Gedanken vielleicht so verdeutlichen: Das Selbstmodell ist ein *transientes computationales Modul*, das vom System vorübergehend aktiviert wird, um seine Interaktion mit der Umwelt zu regulieren. Mit anderen Worten: Wenn Sie morgens aufwachen, d.h. wenn das System, das Sie sind, »zu sich kommt«, wird dieses *transiente computationale Modul* aktiviert – der Moment des »Erwachens« ist genau der Moment, in dem dieses neue Instrument der intelligenten Informationsverarbeitung in Ihrem Gehirn verfügbar wird. Dies geschieht, weil Sie jetzt ein bewußtes Selbstmodell brauchen, um eine flexible und kontextsensitive Form der sensomotorischen Integration zu entwickeln, um komplexes und adaptives Verhalten hervorzubringen und um Ihre Aufmerksamkeit auf den Körper als Ganzes zu richten und ihn zu kontrollieren. Das

⁸ Vgl. auch den fünften Abschnitt in O'Shaughnessy, 1995 und seinen Gebrauch des Begriffs eines »*Long-term body image*«; Metzinger, 1993, 1996, 1997; Damasio, 1994, 1999. Ein guter Einstieg in die empirische Literatur ist Tsakiris 2011.

bewußte Selbstmodell hat auch eine metarepräsentationale Schicht, denn es bietet dem Organismus eine explizite Repräsentation seiner eigenen *Fähigkeiten* an.

Die Entwicklung von immer effizienteren Selbstmodellen als einer neuen Art von »virtuellen Organen« ist – und diesen Punkt darf man nicht übersehen – auch eine Voraussetzung für die Bildung von komplexen Gesellschaften. Plastische und immer komplexere Selbstmodelle erlaubten nicht nur die fortlaufende Optimierung somato-motorischer, perzeptiver und kognitiver Funktionen, sondern machten auch die Entwicklung von sozialer Kognition und kooperativem Verhalten möglich. Das aktuell bekannteste Beispiel ist natürlich das menschliche System der Spiegelneuronen, ein Teil unseres unbewußten Selbstmodells, in dem die Körperbewegungen und Selbstmodelle anderer Agenten in der Umwelt eine permanente funktionale Resonanz finden, was durch einen komplexen Prozeß der motorischen Emulation – einer »*embodied simulation*«, wie Vittorio Gallese (2005) es treffend beschreibt – geschieht, z.B. immer dann, wenn wir zielgerichtetes Verhalten in unserer Umgebung beobachten. Solche wechselseitig gekoppelten Selbstmodelle sind wiederum die fundamentalen repräsentationalen Ressourcen dafür, sich in die Perspektive eines anderen hineinzuversetzen, Empathie und ein Gefühl der Verantwortung zu empfinden, aber ebenso für metakognitive Leistungen wie die Entwicklung eines Selbst*konzepts* und einer *theory of mind* (vgl. hierzu Bischof-Köhler, 1996, 1989; bezüglich möglicher neurobiologischer Korrelate solcher basalen sozialen Fähigkeiten, die gut in den hier skizzierten Rahmen passen, vgl. Gallese und Goldman, 1998; Metzinger und Gallese, 2003).

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 das einführt, was in der Philosophie des Geistes oft eine *teleofunktionalistische Hintergrundannahme* genannt wird (vgl. etwa Millikan, 1984, 1993; Bieri, 1987; Dennett, 1987; Dretske, 1988, 1998; Lycan, 1996⁹). Die Entwicklung und Aktivierung dieses komputationalen Moduls spielt eine Rolle *für* das System: Das funktionale Selbstmodell besitzt dann auch eine wahre evolutionsbiologische Beschreibung, d.h. es war eine Waffe, die in einem »kognitiven Wettrüsten« erfunden und kontinuierlich optimiert wurde (Clark, 1989, S. 61). Die funktionale Instantiierungsbasis für die phänomenal erlebte Erste-Person-Perspektive kann als spezifische kognitive Errungenschaft angesehen werden: als die Fähigkeit, egozentrische repräsentationale Räume zu nutzen und sich dabei selbst als einen epistemischen Agenten zu modellieren (vgl. hierzu den letzten Abschnitt dieses Beitrags). Mit anderen Worten: Phänomenale Subjektivität (das Entstehen einer subsymbolischen, nicht-begrifflichen Erste-Person-Perspektive) ist eine Eigenschaft, die nur instantiiert wird, wenn das betreffende System ein kohärentes Selbstmodell aktiviert und dieses in sein globales Weltmodell integriert.

Die Existenz eines stabilen Selbstmodells erlaubt die Entstehung dessen, was Philosophen manchmal als die »Perspektivität des Bewußtseins« bezeichnen: das Vorhandensein eines einzelnen, kohärenten und zeitlich stabilen Realitätsmodells, das repräsentati-

⁹ Die wichtigsten Texte finden sich in deutscher Übersetzung (zusammen mit weiterführenden Literaturhinweisen, die auch für Nicht-Philosophen geeignet sind) in Band 2 und 3 von T. Metzinger, *Grundkurs Philosophie des Geistes*, Module L-15 und I-9 bis I-11.

tional auf ein einzelnes, kohärentes und zeitlich stabiles phänomenales Subjekt zentriert ist, also auf ein Echtzeit-Modell des Systems *im Akt des Erlebens* oder *im Akt des Wissens* (diesen Ausdruck hat zuerst Antonio Damasio verwendet, vgl. 1999, S. 168ff). Dieses strukturelle Merkmal des globalen Darstellungsraums kann dann zur episodischen Instantiierung einer zeitlich ausgedehnten und nicht-begrifflichen Erste-Person-Perspektive führen. Wenn diese globale repräsentationale Eigenschaft verlohrenght, verändert sich auch die Phänomenologie, was zu verschiedenen neuropsychologischen Störungen oder veränderten Bewußtseinszuständen führt. Manche meiner Leser könnten jetzt den Eindruck gewinnen, daß all dies doch sehr abstrakt sei. Ein Selbstmodell ist jedoch nichts Abstraktes – es ist absolut konkret, sowohl im subjektiven Erleben als auch auf der Ebene der neuronalen Dynamik. Ein erstes, mittlerweile klassisches Beispiel soll vor Augen führen, was ich – unter vielem anderen – mit dem Begriff »Selbstmodell« meine.

Die Plastizität des PSM hat der indische Neuropsychologe Vilayanur Ramachandran in einer Reihe faszinierender Experimente demonstriert, in denen er mit Hilfe von Spiegeln Synästhesien und Bewegungsillusionen in Phantomgliedern auslöste (vgl. Ramachandran und Rogers-Ramachandran, 1996; eine populäre Darstellung findet sich in Ramachandran und Blakeslee, 1998, 46ff. Ich bin Ramachandran für die Überlassung der Abbildung im Text zu Dank verpflichtet). Phantomglieder sind subjektiv erlebte Gliedmaßen, die typischerweise nach dem Verlust eines Armes oder einer Hand oder nach chirurgischer Amputation auftreten. In manchen Fällen, zum Beispiel nach einer nicht-traumatischen Amputation durch einen Chirurgen, haben Patienten den subjektiven Eindruck, daß sie ihr Phantomglied willentlich kontrollieren und bewegen können. Das neurofunktionale Korrelat dieser phänomenalen Konfiguration könnte darin bestehen, daß Motorbefehle, die im motorischen Kortex entstehen, immer noch durch Teile des Parietallappens überwacht werden und – da es keine widersprechende Rückmeldung des amputierten Gliedes gibt – anschließend in den Teil des Selbstmodells integriert werden, der als *Motoremulator* dient (verwandte Überlegungen finden sich bei Grush 1997, 1998, S. 174; vgl. auch Ramachandran und Rogers-Ramachandran, 1996, S. 378). In anderen Fällen geht die subjektiv erlebte Beweglichkeit und Kontrolle der Phantomglieder verloren. Diese alternativen Konfigurationen könnten etwa durch eine präamputationale Lähmung als Folge peripherer Nervenschädigungen oder durch das längere Fehlen einer propriozeptiven und kinästhetischen »Rückmeldung«, die eine Bewegung bestätigen könnte, entstehen. Das Resultat auf der phänomenologischen Beschreibungsebene wäre dann ein gelähmtes Phantomglied.

Ramachandran und seine Kollegen konstruierten eine »Virtuelle-Realität-Kiste«, indem sie einen Spiegel vertikal in einen Pappkarton ohne Deckel einsetzten (Abb. 1 verdeutlicht das Grundprinzip). Der Patient, der seit vielen Jahren unter einem gelähmten Phantomglied litt, wurde dann gebeten, sowohl seinen echten Arm als auch seinen Phantomarm in zwei Löcher in der Vorderseite der Box zu schieben. Als nächstes wurde der Patient gebeten, seine gesunde Hand im Spiegel zu betrachten. Auf der Ebene des visuellen Inputs schaffte dies die Illusion, beide Hände zu sehen, obwohl er tatsächlich nur das Spiegelbild seiner intakten Hand sah. Was geschah nun mit dem Inhalt des PSM, als man die Versuchsperson bat, auf beiden Seiten symmetrische

Abb. 1: Spiegelbild-induzierte Synästhesie. Der vormals gelähmte Teil eines halluzinierten Arms wird wieder für die bewußte Handlungskontrolle verfügbar gemacht, indem man eine virtuelle Quelle visuellen Feedbacks bereitstellt. Abb. mit freundlicher Genehmigung von Vilanayur Ramachandran.



Handbewegungen auszuführen? Ramachandran beschreibt ein typisches Ergebnis dieses Experiments wie folgt:

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. »Ich möchte, daß sie gleichzeitig ihren rechten und ihren linken Arm bewegen«, wies ich ihn an.

»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 »hineinzuschieben«. 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 bewegt sich wieder.«

Nachdem er sich etwas beruhigt hatte, sagte ich: »Okay, Philip, schließen Sie jetzt ihre Augen.«

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

»Öffnen Sie ihre Augen.«¹⁰

¹⁰ Vgl. Ramachandran 1998, 47f. Die klinischen und experimentellen Details finden sich in Ramachandran und Rogers-Ramachandran, 1996.

Es sollte mittlerweile deutlich geworden sein, wie diese experimentellen Resultate den von mir eingeführten Begriff des »Selbstmodells« illustrieren: was sich in diesem Experiment bewegt, *ist* das PSM. Das plötzliche Auftreten von kinästhetischen Bewegungsempfindungen in der verlorenen Subregion des Selbstmodells wurde durch die Installation einer zweiten Feedback-Quelle ermöglicht, durch die Rückmeldung von »virtueller Information«. Dadurch wurde umgehend eine neue funktionale Eigenschaft erzeugt, nennen wir sie »Verfügbarkeit für selektive motorische Kontrolle«. Indem hier der visuelle Modus der Selbstsimulation zugänglich wurde, wurde die entsprechende Information auch volitional verfügbar. Jetzt war die willentliche Kontrolle über das gelähmte Körperteil wieder möglich. Dieses Experiment zeigt auch, wie phänomenale Eigenschaften sehr direkt durch computationale und repräsentationale Eigenschaften determiniert werden. Körperliches Selbstbewußtsein hängt sehr eng mit Vorgängen im Gehirn zusammen.

Auch im nächsten Beispiel geht es um die Phänomenologie von Phantomgliedern. Wie »unwirklich« oder »geisterhaft« sind Phantomglieder? Können wir die »subjektiv erlebte Wirklichkeit« des bewußten Selbst messen? Eine neuere Fallstudie, die von dem bekannten Schweizer Neuropsychologen Peter Brugger und seinen Kollegen an der Universitätsklinik Zürich durchgeführt wurde, verwendete eine Sieben-Punkte-Skala, um die subjektiv wahrgenommene Lebendigkeit und »Tatsächlichkeit« von Phantomgliedern genauer zu bewerten. Über mehrere Sitzungen hinweg zeigten sich hochgradig konsistente Aussagen für die Versuchsperson »AZ«, eine damals 44-jährige Akademikerin, die ohne Unterarme und Beine geboren worden war. Seit sie sich erinnern konnte, hatte sie das bewußte Erlebnis ihrer nicht existierenden Unterarme (einschließlich der Finger) und ihrer Beine (einschließlich der FüÙe sowie des großen und des kleinen Zehs).

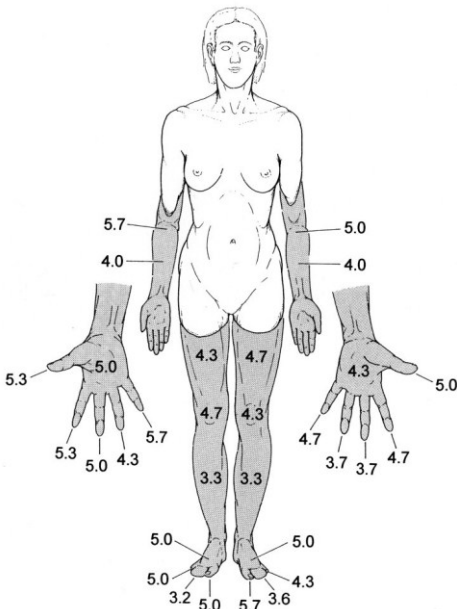


Abb. 2: Hinweise auf eine angeborene Komponente des körperlichen Selbstmodells? Phantomglieder (graue Flächen) bei einer Person, die ohne Gliedmaßen geboren wurde. Die Zahlen sind Intensitätsbewertungen für die gefühlte Präsenz verschiedener Phantom-Körperteile auf einer Sieben-Punkte-Skala von 0 (keinerlei bewußtes Erleben) bis 6 (lebhaftester Eindruck, also genauso realistisch wie das Erleben der tatsächlich existierenden Körperteile). (Mit freundlicher Genehmigung von Peter Brugger, Zürich.)

Die Phantome waren, wie die Abbildung oben zeigt, nicht so realistisch wie der Inhalt ihres nicht-halluzinatorischen Körpermodells. Eine funktionelle Magnetresonanztomographie (fMRI) während von ihr vorgestellten bzw. durchgeführten Bewegungen mit den Phantomhänden zeigte keine Aktivierung der primären Areale für Somatosensorik und Motorik, dafür jedoch eine beidseitige Aktivität im prämotorischen und parietalen Kortex. Eine direkte Stimulation der primären Areale für Somatosensorik und Motorik mit einem die Schädeldecke durchdringenden Magnetfeld rief auf zuverlässige und wiederholbare Weise bewußte Empfindungen in den Phantomfingern und der Phantomhand hervor, und zwar jeweils auf der Seite, die der Stimulationsstelle gegenüberlag. Eine Reizung in prämotorischen und parietalen Bereichen löste ähnliche Phantomempfindungen aus, dann jedoch ohne durch motorische Befehle hervorgerufene Potenziale im Stumpf.

Diese Daten zeigen deutlich, daß Körperteile, die nie physisch entwickelt waren, in sensorischen und motorkortikalen Arealen phänomenal simuliert werden können. Sind sie Bestandteile eines angeborenen Körpermodells? Oder könnte es vielleicht eine parsimonischere Erklärungsstrategie geben, die auf die besondere Motivationslage eines schwerbehinderten Kindes und die Existenz von Spiegelneuronen in Hirnregionen wie etwa BA 44 hinweist? Könnte die Komplettierung des Körpermodells der Patientin vielleicht über die frühkindliche, visuelle Betrachtung von sich bewegenden anderen Menschen erfolgt sein, bei der Arme und Beine sozusagen über eine unbewußte Form der motorischen Perspektivenübernahme in ihr Selbstmodell »hineingespiegelt« wurden? Da ich Philosoph und nicht Neuropsychologe bin, werde ich mich hier weiterer laienhafter Spekulationen enthalten.

Neuere Theorien aus der Erforschung des Schmerzerlebens in Phantomgliedern weisen jedoch auf die mögliche Existenz einer genetisch determinierten Neuromatrix hin, deren Aktivierungsmuster die Grundlage dieser stabilen unbeweglichen Teile des Selbstmodells und den eher unveränderlicheren Hintergrund der körperlichen Selbstempfindung bilden könnten (die so genannte »Phylomatrix des Körperschemas«¹¹). Ein weiterer interessanter empirischer Befund ist die Tatsache, daß über 20% der Kinder, die ohne einen Arm oder ein Bein geboren wurden, später das realistische, bewußte Erleben eines Phantomgliedes entwickeln. Im Zusammenhang mit phänomenaler »Echtheit« und bezüglich der Integration des körperlichen Selbstmodells als Ganzes in das bewußte Realitätsmodell ist es ebenso interessant zu sehen, daß in diesem Fall »die Wahrnehmung ihrer Phantomglieder nur dann vorübergehend unterbrochen wird, wenn ein Objekt oder eine Person in ihre gefühlte Position eindringt oder wenn sie sich selbst in einem Spiegel sieht«.¹²

Was haben die Phänomenologien von Ramachandrans und Bruggers Versuchspersonen gemeinsam? Der Übergang von Stumpf zu Phantomglied ist mit Blick auf die

¹¹ Vgl. dazu Melzack, 1989, 2005; zum Begriff einer »*neurosignature*« des körperlichen Selbstbewußtseins vgl. Melzack, 1992, S. 93; eine wichtige Studie zu Phantomgliedern nach Aplasie und früher Amputation ist Melzack et al., 1997. Siehe dazu auch Hilit et al., 2012.

¹² Vgl. Brugger et al., 2000, S. 6168. Weitere Details bezüglich des phänomenologischen Profils finden sich ebenfalls an dieser Stelle; eine interessante experimentelle Folgestudie, die die Intaktheit des phänomenalen Modells der kinästhetischen und posturalen Gliedeigenschaften demonstriert, findet sich in Brugger et al., 2001.

phänomenale Qualität der »Meinigkeit« nahtlos; subjektiv sind sie beide Teil desselben körperlichen Selbst, da das Gefühl des Besitzens sich auf beide gleichermaßen erstreckt. Es gibt keine Lücke oder einen plötzlichen Sprung im Gefühl des Besitzens. Die Entstehung des körperlichen Selbstmodells basiert auf einem subpersonalen, automatischen Prozeß der Merkmalsbindung, auf einem Vorgang, der Kohärenz erzeugt und kausal beeinflußbar ist. Aber was genau wird erlebt? Was ist der *Inhalt* des Erlebens? In *De Anima* sagt Aristoteles, die Seele sei die Form des physischen Körpers, die allerdings auch mit ihm zusammen im Tode vergeht (*Über die Seele*, II: 412a, 412b–413a). Nach Spinoza ist die Seele die Vorstellung, die der Körper von sich selbst entwickelt (*Die Ethik*, II: 12 und 13). In aktueller Terminologie könnte man sagen, daß eine »Vorstellung« schlicht eine mentale Repräsentation ist – genauer gesagt eine *Selbstrepräsentation* – und daß der Gehalt des Selbstbewußtseins der introspektiv zugängliche Teil dieser Selbstrepräsentation ist, nämlich das PSM, das durch SMT postuliert wird. Gestalteigenschaften – wie die Körperform – sind *globale* Eigenschaften eines Wahrnehmungsobjektes. Könnte Selbstmodellierung dann nicht ein neuronaler Mechanismus sein, um genau solche globalen Eigenschaften für den Organismus selbst intern darzustellen? Könnte sie nicht ein inneres Werkzeug sein, um neue Erkenntnisse über den Organismus *als Ganzes* zu gewinnen? Plato entwickelte in seinem Dialog *Meno* als erster den philosophischen Gedanken, daß manche unserer Ideen angeboren sein könnten. Und dies ist nach wie vor eine interessante Frage in der heutigen Neurowissenschaft des Selbstbewußtseins: Besitzt das PSM eine angeborene Komponente, ist es eine »eingeborene Idee des Körperselbst«? Ist das bewußte Körperbild möglicherweise eine Art von »fixer Idee«, verankert in einem angeborenen und genetisch festgelegten Kernzustand?

Ich will, wie gesagt, an dieser Stelle nicht spekulieren, jedoch auf einen weiteren Punkt hinweisen, der häufig übersehen wird und der nicht die Phänomenologie, sondern die *semantischen Eigenschaften* und den epistemischen Status des Körpermodells in unserem Gehirn betrifft. Ich mache dazu zunächst ohne jedes Argument die ontologische Hintergrundannahme, daß eine bewußtseinsunabhängige Außenwelt und der physische Körper tatsächlich existieren. Das bewußt erlebte Körpermodell ist dann genau diejenige Struktur, die von Anfang an eine intrinsische Semantik besitzt: Es ist sozusagen von Anbeginn geerdet und funktional in seinem Bezugsgegenstand verankert. Sein Referent existiert immer, wenn auch das körperliche Selbstmodell existiert, und zwar ganz einfach deshalb, weil das Modell selbst ein physikalischer *Teil* des Körpers ist – zum Beispiel ein spezifisches dynamisches Aktivierungsmuster, eine verteilte subsymbolische Repräsentation im Gehirn. Weil zumindest der Träger des *phänomenalen* Körpermodells immer intern realisiert ist, gibt es genau hier eine garantierte Referenz, und damit auch ein Element von Gewißheit.¹³ Natürlich könnten sich prinzipiell alle Inhal-

¹³ Deshalb sind all jene Fälle theoretisch interessant, in denen dieses Element von Gewißheit scheinbar verlorengeht. Das könnte genau dann der Fall sein, wenn man sich selbst – im Traum, in asomatischen OBEs – als nicht körperliches, rein mentales Selbst erlebt. Wichtig ist es jedoch, hier immer genau zwischen der Phänomenologie der Gewißheit und dem erkenntnistheoretischen Aspekt zu unterscheiden: Es gibt einen phänomenalen Kernaspekt der raumzeitlichen Selbstlokalisierung, welcher möglicherweise die einfachste Form des Selbstbewußtseins darstellt (Blanke & Metzinger 2009; Windt 2014). Das Traumbeispiel

te dieses Modells als Fehlrepräsentationen entpuppen, aber die grundlegende Existenzannahme – daß es also überhaupt eine physische Basis des erlebten Vorgangs gibt – ist gegeben. Dieser Zusammenhang ist dann wichtig, wenn man unter naturalistischen Hintergrundannahmen die Selbstorganisation bedeutungstragender Zustände verstehen will, etwa die evolutionäre Entstehung intentionaler und semantischer Eigenschaften in Biosystemen: Wer solche Eigenschaften explanatorisch auf funktionale Eigenschaften reduzieren will, der muß verständlich machen, wie kognitive Funktionen letztlich aus der Wechselwirkung zwischen Wahrnehmung und körperlicher Bewegung entstanden sind, aus einer langen Geschichte dynamischer Umweltinteraktionen. Intentionalität ist dann notwendig ein verkörpertes Phänomen: Der semantische Inhalt unseres Weltmodells entfaltet sich schrittweise aus seiner leiblichen Verankerung heraus. Motorische Primitive wurden zu semantischen Primitiven, durch die zielgerichtete Interaktion mit der Umwelt haben wir diese sozusagen schrittweise »mit Bedeutung infiziert«. ¹⁴ Kommen wir nun zum dritten Beispiel. Es stammt aus einer völlig anderen wissenschaftlichen Disziplin, nämlich aus dem faszinierenden neuen Feld der evolutionären Robotik. Es zeigt eine Reihe weiterer Aspekte auf, die der begriffliche Rahmen der SMT als theoretische Möglichkeiten voraussagt und für empirische Erklärungsversuche zu öffnen versucht. Erstens: Ein Selbstmodell kann vollkommen unbewußt sein und es kann häufig als das Produkt eines automatischen »*bottom-up*«-Prozesses der *dynamischen Selbstorganisation* angesehen werden. Zweitens: Es ist kein »Ding« (oder das Modell eines Dings), sondern basiert auf einem kontinuierlichen, fortlaufenden Modellierungsprozeß. Drittens: Es besitzt einen hohen Grad an *Plastizität* (d.h. es kann durch Lernvorgänge verändert werden). Und viertens: Es basiert in seinem Ursprung nicht auf Sprache oder reflexivem, begrifflichem Denken, sondern sehr wahrscheinlich auf dem Versuch, motorisches Verhalten auf erfolgreiche Weise zu organisieren. Es ist ein komputationales Werkzeug, um globale Handlungskontrolle und die Strukturierung des Wahrnehmungsraumes zu ermöglichen. Genauer gesagt hat ein Körpermodell die Funktion, sensorische Eindrücke mit motorischem Output auf noch intelligentere, kontextsensitivere und flexiblere Art zu integrieren. Bereits die unbewußten Vorläufer des PSM brachten eine neue Art von Intelligenz und Widerstandsfähigkeit mit sich.

Bongard et al. (2006) haben einen künstlichen »Seestern« entworfen, der stufenweise ein explizites inneres Selbstmodell entwickelt. Diese vierbeinige Maschine benutzt Beziehungen zwischen Motorik und Sensorik (*sactuation-sensation relationships*), also zwischen selbsterzeugten Körperbewegungen und Rückmeldungen aus der durch diese aktiv veränderten Sinneswahrnehmung, um indirekt Rückschlüsse auf ihre eigene Struktur zu ziehen, und sie nutzt dann dieses Selbstmodell, um Vorwärtsbewegungen und verschiedene intelligente Gangarten zu erzeugen. Wenn ein Teil eines ihrer Beine entfernt wird, paßt sie ihr Selbstmodell an die neue Körperform an und erzeugt alter-

zeigt jedoch, daß auch dieser Aspekt auf einer Fehlrepräsentation beruhen könnte, die für das Subjekt nicht als solche zu erkennen ist.

¹⁴ Dieser Zusammenhang ermöglicht es dann auch, in *Gruppen* von Biosystemen über dynamisch miteinander gekoppelte Selbstmodelle eine geteilte Semantik für öffentliche Repräsentationssysteme zu evolvieren – etwa für von allen Gruppenmitgliedern verständliche Gesten, Lautäußerungen und sprachliche Symbole. Siehe dazu Steels & Hild, 2012.

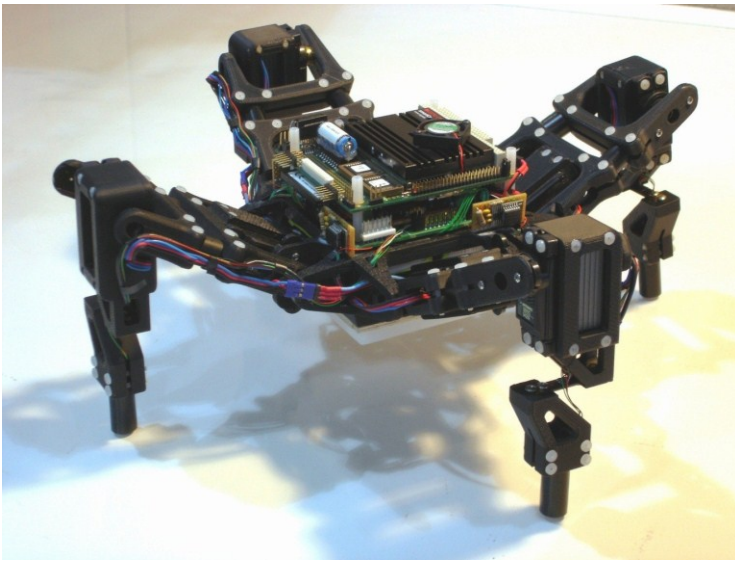


Abb. 3: Der Seestern, ein vierbeiniger Roboter, der sich mit Hilfe eines internen Selbstmodells fortbewegt, das sich in ihm entwickelt hat und welches er ständig verbessert. Wenn er ein Bein verliert, kann er sein internes Selbstmodell und seine Bewegungsabläufe anpassen. (Foto: Josh Bongard.)

native Schrittmuster und Formen des Gehens – sie lernt zu hinken. Mit anderen Worten: Anders als die Phantomgliedpatienten in den Beispielen 1 und 2 (und wie die meisten gewöhnlichen Patienten) ist die Maschine in der Lage, ihre Körperrepräsentation nach dem Verlust eines Gliedes zu *restrukturieren*. Sie kann lernen. Dieses Konzept könnte nicht nur zur Entwicklung von robusteren Maschinen nützlich sein und Aufschluß über Selbstmodellierung bei Tieren geben. Es ist auch theoretisch interessant, weil es zum ersten Mal zeigt, daß ein physisches System in der Lage ist, »autonom und mit wenig Vorwissen seine eigene Topologie selbstständig wiederzuentdecken« (S. 1120), wie es die Autoren beschreiben, und zwar durch eine fortlaufende Optimierung der Parameter seines eigenen Selbstmodells. Der Seestern entwickelt also nicht nur ein inneres Selbstmodell, sondern benutzt dieses Selbstmodell kontinuierlich dazu, intelligentes Verhalten zu erzeugen. Abbildung 4 gibt einen Überblick über diesen Prozeß. Wie man sieht, führt der Roboter anfangs eine Zufallsbewegung aus und registriert die daraus resultierenden Sensordaten (man kann zum Beispiel auch an die Strampelbewegungen, das »motorische Plappern« eines menschlichen Säuglings denken). Die Modellsynthese-Komponente erstellt dann ein Set von 15 möglichen Selbstmodellen, indem es stochastische Optimierungsstrategien einsetzt, um die beobachtete Beziehung zwischen Motorik und Sensorik zu erklären. Der Roboter erzeugt nun im nächsten Schritt ein exploratives motorisches Verhalten in der wirklichen Welt, mit dessen Hilfe er die Nichtübereinstimmung zwischen den Vorhersagen maximiert, die durch diese virtuellen, miteinander konkurrierenden Selbstmodelle geliefert werden. Diese Handlung wird also physisch ausgeführt und anschließend werden die fünfzehn Selbstmodell-Versionen mit Hilfe der neuen Daten verbessert. Wenn sich die Modelle schließlich angleichen, wird das beste Modell von der Verhaltenssynthese-Komponente benutzt, um ein gewünschtes Verhalten zu generieren, das dann vom Roboter ausgeführt werden kann. Wenn der Roboter unerwartete sensomotorische Muster oder ein externes Signal infolge unvorhergesehener morphologischer Veränderungen entdeckt, initiiert er erneut den wechselnden Zyklus des Modellierens und des

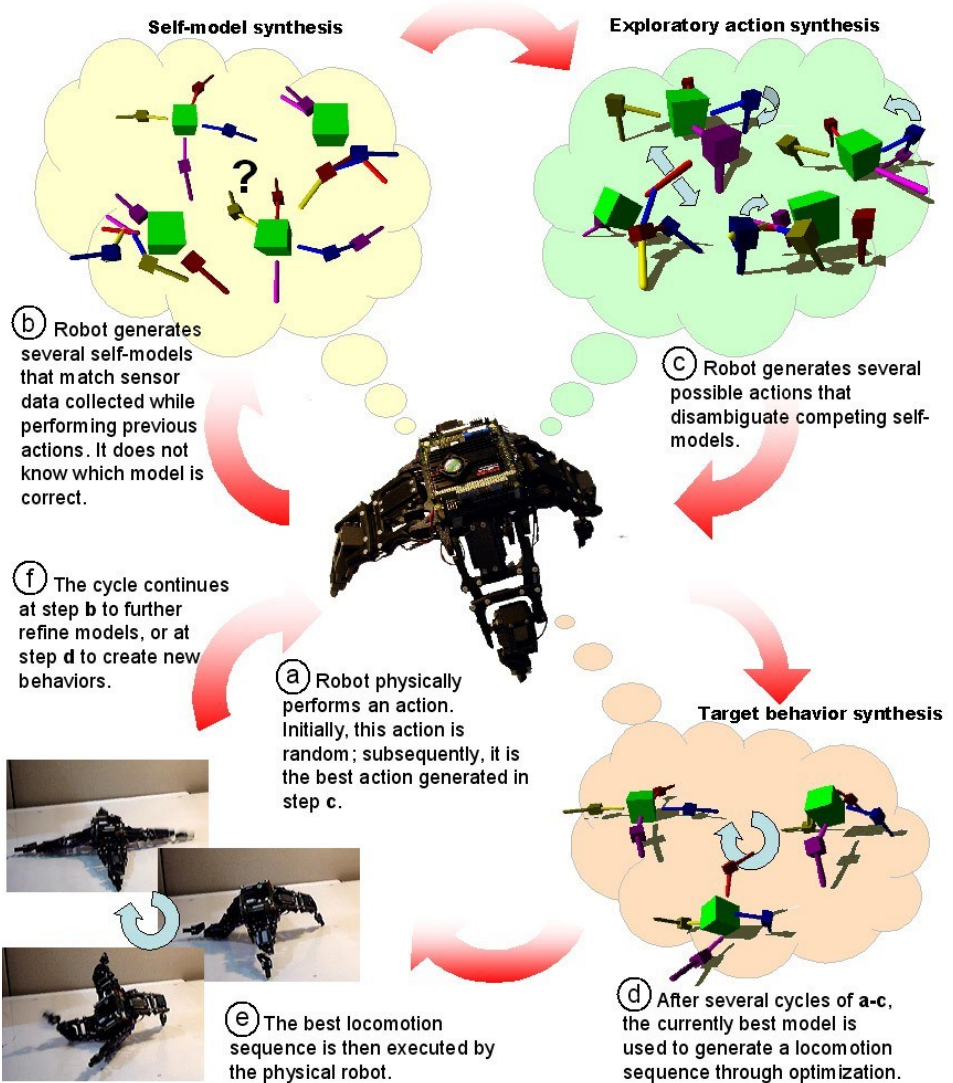


Abb. 4: In den Schritten (A) und (B) geschieht die Synthese des Selbstmodells. Der Roboter führt zunächst eine physische Handlung durch (A). Anfänglich ist diese Handlung rein zufällig; später ist es die beste Handlung, die in Schritt (C) entdeckt wurde. Dann erzeugt der Roboter eine Reihe von Selbstmodellen, die zu den Sinnesdaten passen sollen, welche während vorangegangener Handlungen gesammelt wurden. Er weiß nicht, welches Modell korrekt ist (B). Die Synthese von Probandlungen erfolgt in Schritt (C): Der Roboter erzeugt verschiedene mögliche Körperbewegungen, die zu einer eindeutigen Entscheidung zwischen den miteinander im Wettstreit stehenden Selbstmodellen führen. Darauf folgt die Synthese des Zielverhaltens in Schritt (D): Nach einer Reihe von Durchläufen durch die Stufen (A) bis (C) wird das aktuell beste Modell dazu benutzt, durch weitere Optimierung eine Abfolge von Fortbewegungsmustern zu erzeugen. Die beste Abfolge von Fortbewegungsmustern wird dann durch das physikalische System ausgeführt (E). (Grafik mit freundlicher Genehmigung von Josh Bongard.)

explorativen Handelns, um neue Selbstmodelle zu produzieren, die auch diese Veränderungen berücksichtigen. Das Beste dieser neuen Modelle wird dann benutzt, um ein kompensatorisches Verhalten zu erzeugen und die Funktionalität wiederherzustellen. Technische Details beiseite: Was sind die philosophischen Konsequenzen aus dem dritten Beispiel? Erstens: Man muß kein lebendiges Wesen sein, um ein Selbstmodell zu besitzen. Nicht-biologische SMT-Systeme sind möglich. Zweitens: Ein SMT-System kann vollkommen unbewußt sein, d.h. es muß kein PSM, kein *phänomenales* Selbstmodell besitzen. Bewußtsein ist offensichtlich ein zweiter Schritt.¹⁵ Drittens: Das Selbstmodell unterstützt Planungsvorgänge und schnelle Lernprozesse auf verschiedene Art und Weise. Es macht ein System eindeutig intelligenter und anpassungsfähiger, sein repräsentationaler Inhalt ist eine *Vorhersage* und es entsteht in einem systematischen Wechselspiel zwischen virtuellem und wirklichem Verhalten. Viertens: Es ist das, was ich oben ein virtuelles Modell oder ein »virtuelles Organ« genannt habe, und eine seiner Hauptfunktionen besteht in der Aneignung eines Körpers mit Hilfe eines globalen morphologischen Modells, um diesen als Ganzes zu kontrollieren. An anderer Stelle habe ich für diese Art von Selbstkontrolle den Begriff des »*second-order embodiment*« eingeführt (»*third-order embodiment*« bezieht sich dann auf die wesentlich kontextsensitivere, bewußte Verarbeitungsebene).¹⁶ Wenn mir der Gebrauch einer Metapher erlaubt ist: Eine der theoretischen Intuitionen ist hier, daß die Entstehung eines aktiven Selbstmodells es einem physischen System erstmals erlaubt, seine eigene niedrigstufige Mikrodynamik mit Hilfe eines einzelnen, funktional integrierten, inneren Modells von sich selbst als einer Ganzheit zu »versklaven« und sich dadurch erstmals *global* zu kontrollieren, sich also sozusagen die eigene Hardware auf kausaler Ebene »anzueignen«. Diese Aneignung erleben wir subjektiv als »Meinigkeit«, und sie ist gleichzeitig der entscheidende erste Schritt auf dem Weg dazu, ein autonomer Agent zu werden.

Der dritte Schritt: Eine repräsentationalistische Analyse der drei Zieleigenschaften

Der Grundgedanke ist nun, daß Selbstbewußtsein in wesentlichen Aspekten eine *Integrationsleistung* ist: Indem repräsentationale Zustände in das gegenwärtig aktive Selbstmodell eingebettet werden, gewinnen sie die höherstufige Eigenschaft der phänomenalen Meinigkeit hinzu. Wenn dieser Einbettungsprozeß gestört wird, hat dies verschiedene neuropsychologische Syndrome oder veränderte Bewußtseinszustände zur Folge (Fallstudien hierzu finden sich in Kapitel 7 in Metzinger, 2003a). Zum Beispiel kann man Somatoparaphrenien (Vallar & Ronchi, 2009), Xenomelien (Giummarra et al., 2011, Hilti et al., 2012) oder bestimmte Positivsymptome der Schizophrenie wie die Gedankeneingebung als funktionale Konfigurationen analysieren, in denen das System existierende Repräsentationen von Körperteilen oder der eigenen kognitiven Vorgänge nicht mehr in das PSM integrieren kann. Werfen wir einen Blick auf solche

¹⁵ Vgl. Metzinger, 1995b, 2000a als ersten Überblick; eine zusätzliche Liste von zehn Auflagen, die für bewußtes Erleben erfüllt sein müssen, findet sich in Metzinger, 2003a, Abschnitt 3.2.

¹⁶ Vgl. dazu Metzinger, 2006b, 2013b. Eine wichtige neuere Veröffentlichung, die über das *Starfish*-Modell hinausgeht und die das verdeutlicht, was ich auf philosophisch-begrifflicher Ebene mit »Verkörperung zweiter Stufe« meine, ist Schilling & Cruse, 2012.

konkreten Beispiele, in denen die phänomenale Meinigkeit, das subjektive Gefühl des Besitzens, selektiv verloren geht.

- Floride Schizophrenie: »Bewußt erlebte Gedanken sind nicht mehr *meine* Gedanken.«
- Somatoparaphrenie, Xenomelie (*body identity integrity disorder*, BIID; Giummarra, 2011, Hilti et al., 2013), unilateraler Hemi-Neglekt: »Mein Bein ist nicht *mein* Bein.«
- Depersonalisierung (DP; Michal & Beutel, 2009) und Beeinflussungswahn: »Mein Körper als Ganzer wird mir fremd und erscheint als unwirklich; ich bin ein Roboter, bewege mich wie ein Automat, verwandele mich in eine von außen gesteuerte Marionette, die Gedanken werden mir eingegeben oder entzogen, volitionale Akte sind nicht mehr *meine* volitionalen Akte.«¹⁷
- Manien: »Ich bin die ganze Welt, alle Ereignisse in der Welt werden durch *meine eigenen* Willensakte kontrolliert.«

Subjektiv erlebte »Meinigkeit« ist eine Eigenschaft einzelner Formen phänomenalen Gehalts, zum Beispiel die der mentalen Repräsentation eines Beins, eines Gedankens oder eines Willensaktes. Diese Eigenschaft, das Gefühl des Besitzens, ist nicht notwendigerweise mit ihnen verbunden, d.h. sie ist keine intrinsische, sondern eine *relationale* Eigenschaft. Daß man einen Gedanken oder ein Körperteil bewußt als seinen eigenen erlebt ist keine essentielle, strikt notwendige Eigenschaft des bewußten Erlebens dieses Gedankens oder Körperteils. Es hätte auch anders sein können: In anderen phänomenologischen Zusammenhängen verschwindet die Meinigkeit. Ihre Verteilung über die Elemente eines bewußten Weltmodells kann variieren. Wenn dem System die Integration bestimmter einzelner repräsentationaler Inhalte ins Selbstmodell nicht mehr gelingt, geht sie verloren. Wenn dies zutrifft, sollte es zumindest prinzipiell möglich sein, diese Eigenschaft zu operationalisieren, indem man nach einer empirisch überprüfbaren Metrik für die Kohärenz des Selbstmodells in den betreffenden Bereichen sucht. Man könnte auch empirisch untersuchen, wie und durch welche Teile des Gehirns eine bestimmte Art von repräsentationalem Inhalt ins Selbstmodell eingebunden wird. Lokale Körperillusionen wie die Gummihand-Illusion (Botvinick & Cohen, 1998), manche Störungen der Willkürmotorik (Sebanz & Prinz, 2006), aber auch das Phänomen der halluzinierten Agentivität (Wegner & Wheatley, 1999) erscheinen als Fehlrepräsentationen, in denen bereits im Gehirn aktive repräsentationale Inhalte in das Selbstmodell eingebettet und dadurch automatisch mit der phänomenale Eigenschaft der »Meinigkeit« versehen werden: Was immer vom Gehirn funktional in das gegenwärtig aktive PSM eingebettet wird, wird von der betreffenden Person unhintergebar als *eigener* Zustand erlebt. Hier ist ein klassisches Beispiel für das, was ich »Meinigkeit« nenne.

¹⁷ Was hier in manchen Fällen selektiv verloren geht, ist also das, was der Philosoph und Psychiater Karl Jaspers »Vollzugsbewußtsein« genannt hat; vgl. Jaspers [1946] 1973, S. 102. Siehe auch Simeon & Abugel, 2006, Ratcliffe, 2008.

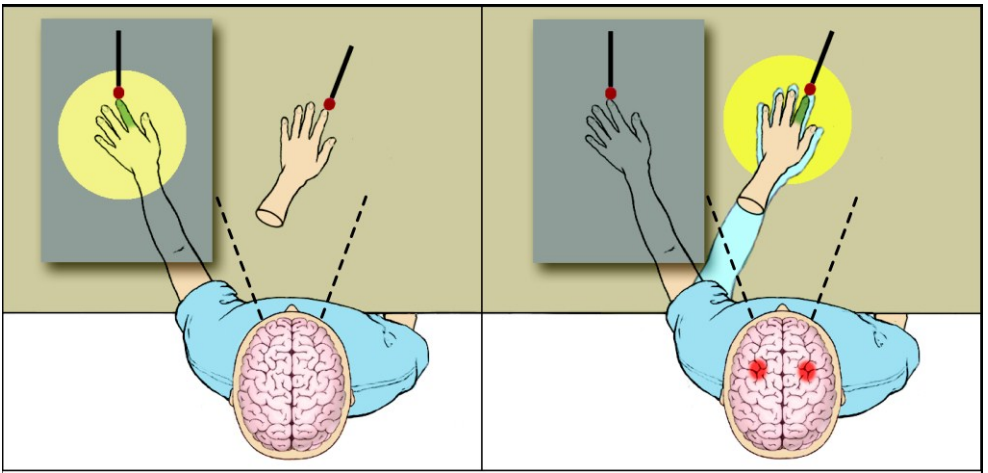


Abb. 5: Die *Gummihand-Illusion*. Eine gesunde Versuchsperson erlebt eine künstliche Gummihand als Teil ihres eigenen Körpers. Der Proband betrachtet eine Nachbildung einer menschlichen Hand, während seine eigene Hand verdeckt ist (dunkelgraues Feld). Sowohl die künstliche Gummihand als auch die unsichtbare Hand werden wiederholt und genau gleichzeitig mit einem Stäbchen gestreichelt. Die hellen Bereiche um die Hand und die schwarzen Bereiche um den Zeigefinger stellen die jeweiligen taktilen und visuellen rezeptiven Felder für Neuronen im prämotorischen Kortex dar. In der rechten Abbildung sieht man die Illusion der Versuchsperson, bei der die gefühlten Berührungen in Einklang gebracht werden mit den gesehenen Berührungen durch das Stäbchen (die dunklen Felder stellen Gebiete erhöhter Hirnaktivität dar; die phänomenal erlebte, illusorische Armstellung wird durch die helle Umrißlinie dargestellt). Die mit dem Auftreten der Illusion einhergehende Aktivierung von Neuronen im prämotorischen Kortex läßt sich experimentell nachweisen. (Abb. Litwak Illustrations Studio, 2004.)

Bei der *Gummihand-Illusion* wird das Gefühl, mit einem Stäbchen gestreichelt zu werden, so mit der korrespondierenden visuellen Wahrnehmung verschmolzen, daß das Gehirn vorübergehend die propriozeptive Karte (der Körpereigenwahrnehmung der Versuchsperson) in Übereinstimmung bringt mit einer visuellen Karte (derjenigen Körperteile, welche die Versuchsperson in jenem Moment sieht). Gleichzeitig wird das Gefühl des Besitzens oder die phänomenale »Meinigkeit« auf die Gummihand übertragen. 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, wird ihre Armbewegung automatisch in Richtung der Gummihand abweichen (Botvinick und Cohen, 1998, S. 756). Wenn einer der Finger der Gummihand durch ein Verbiegen nach hinten in eine physiologisch unmögliche Position »verletzt« wird, erlebt die Versuchsperson ihren echten phänomenalen Finger ebenso als viel weiter zurückgebogen als er in Wirklichkeit ist. Zusätzlich zeigt sich eine deutlich meßbare Hautwiderstandsreaktion. Zwar berichteten in einer Studie nur 2 von 120 Versuchspersonen von einem echten Schmerz Erlebnis, aber viele zogen ihre reale Hand zurück, rissen ihre Augen auf oder lachten nervös (Armel und Ramachandran, 2003, S. 1503). Versuchspersonen zeigten ebenfalls eine deutliche Reaktion, wenn man mit einem Hammer auf die Gummihand schlug. 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 vom jeweiligen Kontext

ab, und davon, welche Information unser Gehirn in das gegenwärtig aktive Selbstmodell integriert.¹⁸ Die faszinierende Frage ist natürlich diese: Könnten auch *Ganzkörper-Illusionen* existieren? Die Antwort lautet »ja« und ich werde in Beispiel 5 auf diesen Punkt zurückkommen.

Zuvor wende ich mich aber der zweiten Zieleigenschaft zu, dem bewußt erlebten »Ich-Gefühl«. Intuitiv könnte man dieses Gefühl als »globale Meinigkeit« konzeptualisieren: das subjektive Gefühl, den Körper *als Ganzes* zu besitzen und sich phänomenal mit ihm zu *identifizieren*. Diese Beschreibung würde uns aber in begriffliche Probleme führen, da sie ein unsichtbares Selbst »hinter« dem Körper einführt, das diesen besitzt. Worum es vielmehr geht, ist, daß der Körper auf einer hohen Ebene zwei Tatsachen für sich selbst darstellt: seine eigene Existenz als eine Ganzheit und die Fähigkeit, sich selbst als eine solche Ganzheit kausal zu kontrollieren. Es geht also um die Repräsentation von Existenz und Autonomie. Methodologisch ist es dabei wichtig, zunächst die *einfachste* Form der Zieleigenschaft zu isolieren (Blanke & Metzinger, 2009). Das phänomenale »Ich-Gefühl« entspricht der Existenz eines einzelnen, kohärenten und zeitlich stabilen Selbstmodells, das den Mittelpunkt des repräsentationalen Gesamtzustands bildet. Wenn dieses repräsentationale Modul beschädigt ist oder desintegriert, oder wenn multiple Strukturen dieses Typs im System alternieren oder sogar gleichzeitig aktiv sind, hat das wiederum verschiedene neuropsychologische Störungen oder veränderte Bewußtseinszustände zur Folge:

- Anosognosie und Anosodiaphorien: Verlust höherstufiger Einsicht in bestehende Defizite, z.B. bei blinden Patienten, die ihre eigene Blindheit verleugnen (*Anton's Syndrom*).
- *Dissociative Identity Disorder* (DID): Das System verwendet verschiedene und alternierende Selbstmodelle, um mit extrem traumatisierenden und sozial inkonsistenten Situationen umzugehen (zu den aktuellen diagnostischen Kriterien vgl. ICD-10-CM F44.81).
- »Ich-Störungen« oder Identitätsstörungen: Eine große Klasse psychiatrischer Störungen, die mit veränderten Formen des Erlebens der eigenen *Identität* einhergehen. Schizophrenien sind ebenso ein klassisches Beispiel wie das Cotard-Syndrom, reduplikative Paramnesie oder wahnhaftes Fehlidentifikation (eine Erörterung der Frage, warum Identitätsstörungen aus philosophischer Perspektive interessant sind, findet sich in Metzinger, 2004a).

Durch das Vorhandensein eines stabilen Selbstmodells entsteht fast immer auch die »Perspektivität des Bewußtseins« in Form von vorübergehenden Subjekt-Objekt-Beziehungen.¹⁹ Dieses strukturelle Merkmal des globalen Darstellungsraums führt episodisch zur Instantiierung einer zeitlich ausgedehnten und nicht-begrifflichen Erster-Person-Perspektive. Es kann ebenfalls verloren gehen.

¹⁸ Vgl. vor allem Botvinick & Cohen, 1998, und die Neuroimaging-Studie von Botvinick, 2004; Ehrsson et al., 2004.

¹⁹ Vgl. Abschnitt 6; dazu auch Nagel, 1986; Metzinger, 2011; zur Frage nach Beispielen für bewußte Realitätsmodelle, die (wie vielleicht der akinetische Mutismus) egozentrisch, aberaperspektivisch sind, siehe insbesondere Metzinger, 2006a).

- Zerfall des Selbstbewußtseins: Verlust der phänomenalen Erste-Person-Perspektive, begleitet von dysphorischen Zuständen und funktionalen Defiziten (»Angstvolle Ich-Auflösung«; vgl. Dittrich, 1985, 2006; Dittrich, Lamparter & Maurer, 2006; Studerus, Gamma & Vollenweider, 2010).
- 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«; vgl. Dittrich, 1985, 2006; Dittrich, Lamparter & Maurer, 2006; Studerus, Gamma & Vollenweider, 2010)

Um dem Reichtum und der Vielfalt menschlicher Erlebnisformen gerecht zu werden, muß man das Vorhandensein bestimmter aperspektivischer und selbstloser Formen des bewußten Erlebens anerkennen. Phänomenologisch ist ein *nicht-subjektives* Bewußtsein – phänomenales Erleben, das nicht an ein Selbst oder eine individuelle Erste-Person-Perspektive gebunden ist – nicht nur eine Möglichkeit, sondern Realität, auch wenn wir diesen Gedanken unvorstellbar²⁰ finden mögen. Philosophisch interessant sind dabei insbesondere all jene Zustandsklassen, in denen bewußte Personen spontan den Gebrauch des Erste-Person-Pronomens »Ich« einstellen, zum Beispiel beim Cotard-Syndrom oder im Falle länger andauernder spiritueller Erfahrungen. Die SMT stellt die begrifflichen Mittel zur Verfügung, um diese speziellen Fälle zu erfassen (weitere neurophänomenologische Fallstudien finden sich in Metzinger, 2003a, Kapitel 4 und 7).

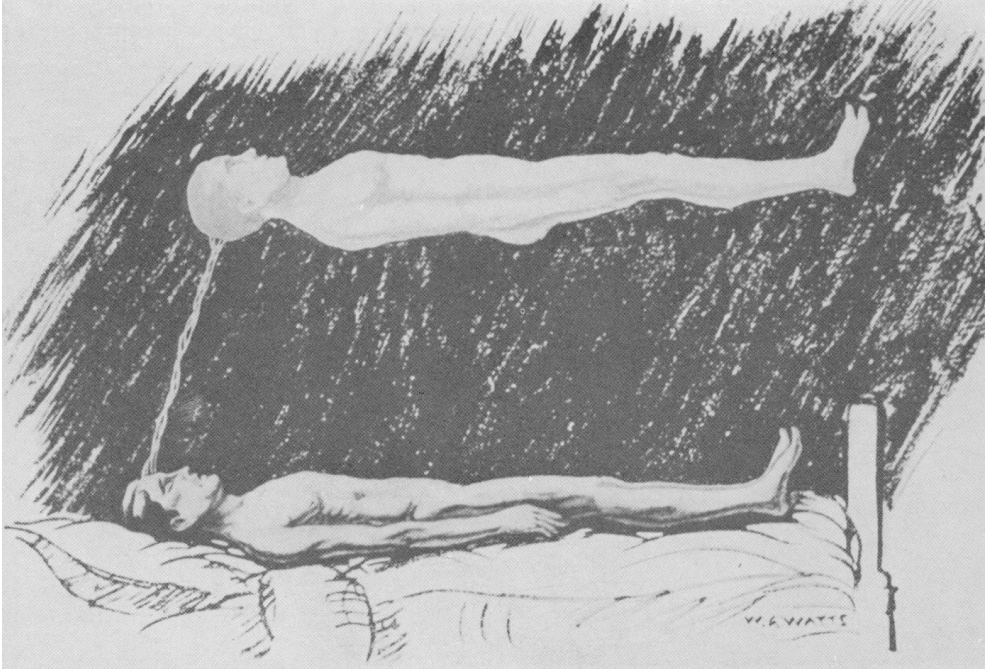
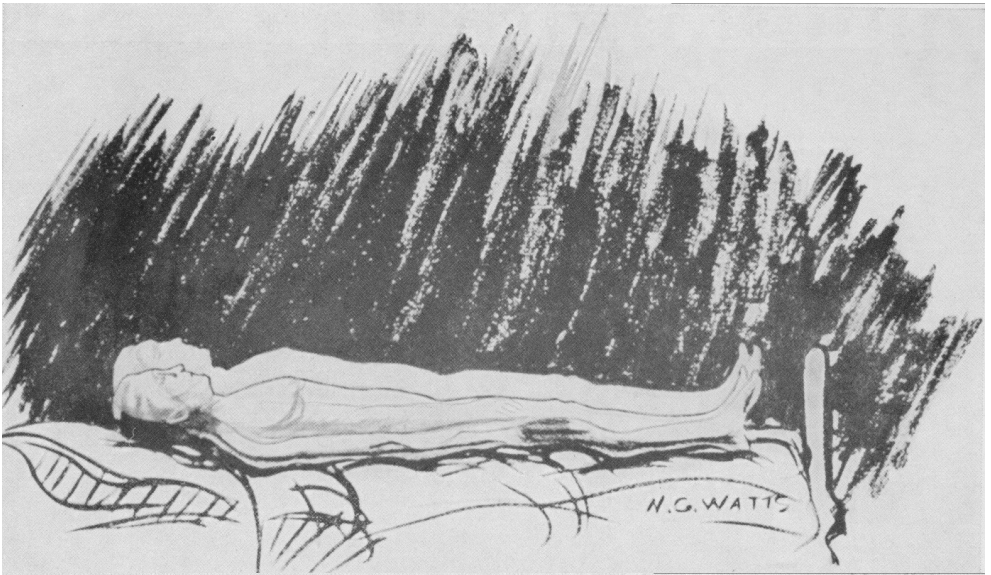
Das fünfte Beispiel soll dieses Prinzip in einem anderen Gegenstandsbereich verdeutlichen. Wenn man die notwendigen begrifflichen Instrumente hat, kann man nämlich nicht nur die Feinheiten und die Variabilität des menschlichen Erlebens ernstnehmen. Man kann auch neue interdisziplinäre Forschungsprogramme entwickeln, die in »Tabu-Zonen« eindringen und Aufschluß geben über Phänomene, die bisher nur Projektionsfläche für esoterische Folklore und metaphysische Ideologien waren. Könnte es eine integrierte Art von körperlichem Selbstbewußtsein geben – sei es nun das eines beweglichen Köpers, der für willentliche Kontrolle voll verfügbar ist, oder das eines gelähmten Körpers – das in seiner Gesamtheit eine phänomenologische Konfabulation ist? Kurz gesagt: Könnte es ein gleichzeitig *halluziniertes* und *körperliches* Selbst geben? Ist es denkbar, daß bei einer menschlichen Versuchsperson so etwas wie eine Ganzkörper-Analogie der Gummihand-Illusion oder ein »globalisiertes Phantomgliederlebnis« – das Erleben eines *Phantomkörpers* – entsteht? Die Antwort lautet »Ja«. Es gibt eine bekannte und gut dokumentierte Klasse von phänomenalen Zuständen, in denen die erlebende Person die unhintergehbare und höchst realistische bewußte Erfahrung macht, den eigenen physischen Körper zu verlassen und sich – häufig in der Form eines ätherischen Doppelgängers – außerhalb dieses Körpers zu bewegen. Mit anderen Worten: Es gibt eine Klasse (oder zumindest ein stabiles Bündel) von eng

²⁰ »Unvorstellbar« oder »kontraintuitiv« finden menschliche Wesen all jene Bewußtseinszustände, in denen es kein phänomenales Selbst gibt, genau deshalb, weil man sie nicht aktiv simulieren kann (denn dies würde notwendig die Phänomenologie des inneren Handelns erzeugen, also »kognitive Agentivität«) und weil solche Zustände, wenn es sie bereits gegeben hat, schlecht in das autobiographische Selbstmodell integriert werden können (denn sie sind ja kein Teil der *eigenen* phänomenalen Biographie). Vgl. dazu Metzinger, 2011.

verwandten phänomenalen Realitätsmodellen, die typischerweise durch zwei Komponenten charakterisiert und definiert werden: zum einen durch eine *visuelle Repräsentation* des eigenen Körpers in einer perzeptuell unmöglichen, externalisierten Dritte-Person-Perspektive (z.B. als sich selbst von oben sehend, etwa auf dem Bett, einem Operationstisch oder der Straße liegend); zum anderen durch eine *zweite Repräsentation* des eigenen Körpers, die sich typischerweise (aber nicht in allen Fällen) frei im Raum bewegt. Dieses zweite Körpermodell ist der Ort, an dem sich das phänomenale Selbst befindet. Es bildet nicht nur den »wahren« Fokus des eigenen phänomenalen Erlebens, sondern fungiert auch als eine integrierte Repräsentation aller kinästhetischen Qualia und aller nicht-visuellen Arten von Propriozeption. Diese Klasse phänomenaler Zustände wird »außerkörperliche Erfahrung« (AKE; häufig wird auch die Abkürzung »OBE« verwendet, für *out-of-body experience*) genannt (eine kurze Zusammenfassung wissenschaftlicher Studien findet sich in Metzinger, 2009, 135ff). Berichte über solche Erlebnisse finden wir in allen Kulturen und aus allen Zeiten. An anderer Stelle (Metzinger, 2005b, weitere Verweise finden sich in Lenggenhager et al., 2007) habe ich dafür argumentiert, daß ein wichtiger Ursprung unseres traditionellen, alltagsphänomenologischen Begriffs der »Seele« in ehrlichen und ernstzunehmenden Erste-Person-Berichten über subjektive Erlebnisse dieser spezifischen neurophänomenologischen Klasse von Zuständen liegen könnten.

AKEs treten häufig ganz von selbst beim Einschlafen, aber auch nach schweren Unfällen oder während Operationen auf. Derzeit ist es nicht klar, ob der Begriff einer AKE eine klar umrissene Menge von notwendigen und hinreichenden Bedingungen besitzt. Stattdessen könnte der Begriff einer AKE ein Cluster-Begriff sein, der durch ein lockeres Bündel von Kriterien und durch unterschiedliche (und sich möglicherweise überschneidende) Teilmengen von phänomenologischen Auflagen konstituiert wird, die jeweils eine Menge von hinreichenden, aber nicht notwendigen Bedingungen bilden. Andererseits ist die AKE eindeutig so etwas wie ein phänomenologischer *Prototyp*. Daß das Phänomen nicht selten ist und einen allgemeinen Kern besitzt, zeigt sich an der einfachen Tatsache, daß viele Leser in irgendeiner Form bereits von dieser Art Erfahrung gehört haben werden.

Man kann eine repräsentationalistische Analyse von AKEs entwickeln, indem man sie als eine Klasse von abweichenden Selbstmodellierungsprozessen beschreibt. Phänomenologische Berichte über »Seelenreisen« wären dann zum Beispiel Berichte über den repräsentationalen Inhalt des PSM während eines solchen devianten Bewußtseinszustands. Eine zentrale Eigenschaft dieser Klasse von abweichenden PSM besteht in der Koexistenz von *zwei* Modellen: (a) einer mehr oder weniger wahrheitsgetreuen Repräsentation des körperlichen Selbst, häufig als aus einer externen, visuellen Perspektive gesehen, welches aber *nicht* als das Zentrum des globalen Realitätsmodells fungiert, und (b) einem zweiten Selbstmodell, das dem subjektiven Erleben nach in erster Linie propriozeptive, taktile und kinästhetische Wahrnehmungen integriert – wobei Gewichts-, Temperatur- und Schmerzempfindungen interessanterweise sehr selten sind – und das spezielle Form- und Gestalteeigenschaften besitzt, die nicht unbedingt wahrheitsgetreu sein müssen. Obwohl nur eines dieser Modelle das System als Subjekt der Erfahrung, also *als erlebend* darstellt, sind beide Modelle in denselben räumlichen Bezugsrahmen eingebettet (deshalb sind solche Zustände *außerkörperliche Er-*



fahrungen). Dieser Bezugsrahmen ist ein egozentrischer Bezugsrahmen. Es handelt sich hierbei um den am häufigsten beschriebenen Bewußtseinszustand, in dem zwei Systemmodelle zum gleichen Zeitpunkt aktiv sind. Selbstverständlich ist nur eines von ihnen der »Ort der Identität«, der Ort, wo sich der Phänomenologie nach auch das befindet, was man in der Philosophie das Handlungssubjekt nennen würde. Das andere Selbstmodell – das des physischen Körpers, der etwa unten auf dem Bett oder dem Operationstisch liegt – ist streng genommen gar kein Selbstmodell, weil es nicht als

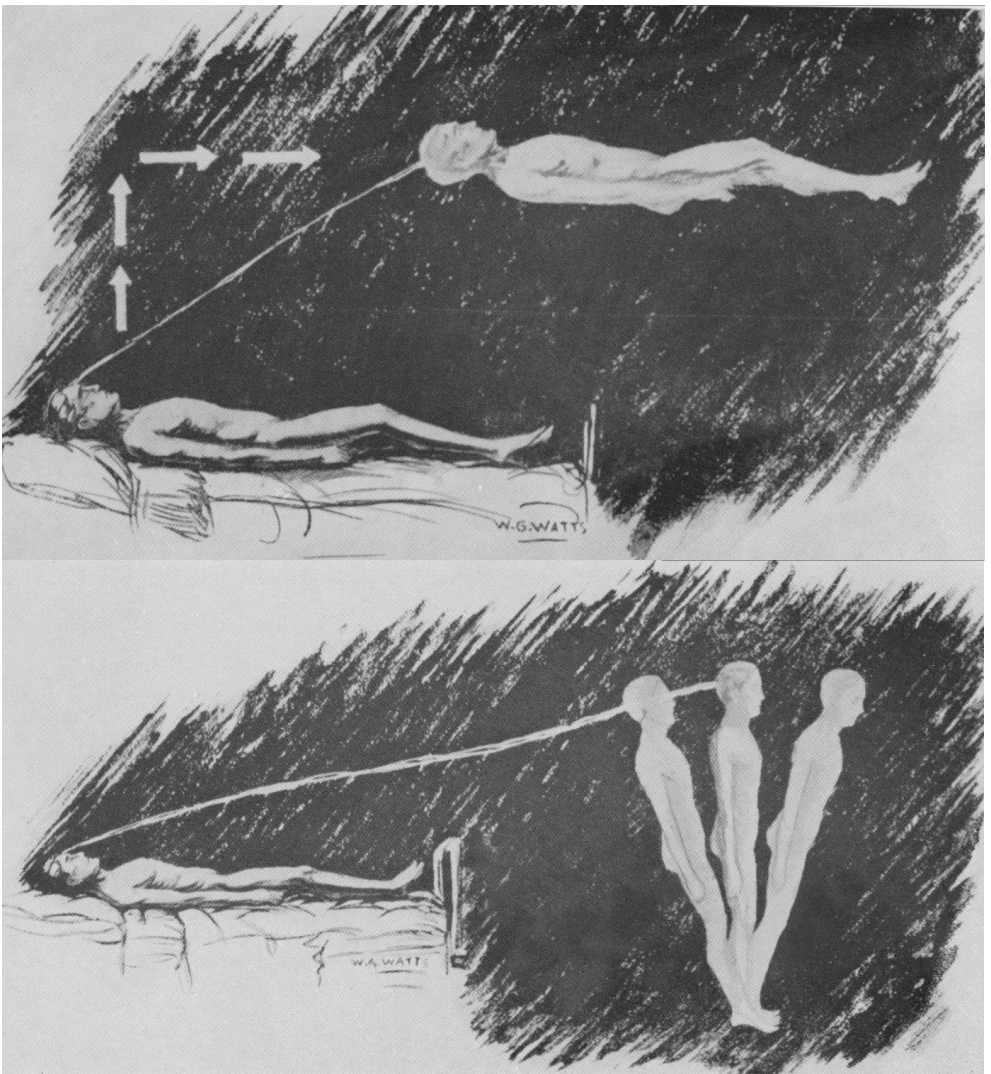


Abb. 6a: Kinematik des phänomenalen Körperbils zu Beginn einer außerkörperlichen Erfahrung: Das »klassische« Bewegungsmuster nach S. Muldoon und H. Carrington, *The Projection of the Astral Body* (1929). Deutsch: *Die Aussendung des Astralkörpers* (2001).

Ursprung der 1PP fungiert. Dieses zweite Selbstmodell ist kein *Subjektmodell*. Es ist zum Beispiel nicht der Ort im Raum, von dem aus Sie Ihre Aufmerksamkeit willentlich auf die Welt richten. Auf der anderen Seite ist es immer noch Ihr eigener Körper, den Sie da unten von außen sehen. Sie erkennen ihn visuell als Ihren eigenen, aber nun ist es nicht mehr der Körper *als Subjekt*, als Ort des Wissens, des Handelns und des bewußten Erlebens, der Teil der Wirklichkeit, mit dem Sie sich *identifizieren*. All das aber zeichnet das Ego aus. Solche phänomenalen Konfigurationen sind deshalb aufschlußreich, weil sie uns gestatten, verschiedene funktionale Schichten im bewußten Selbst des Menschen zu unterscheiden. Sehen wir uns jetzt zwei klassi-

sche phänomenologische Beschreibungen der AKE an, die spontan in einem alltäglichen, nicht-pathologischen Kontext auftauchten.

Ich erwachte nachts – es dürfte gegen drei Uhr morgens gewesen sein – und stellte fest, daß ich mich überhaupt nicht bewegen konnte. Ich wußte ganz genau, daß ich nicht träumte, denn ich war voll bei Bewußtsein. Voller Angst über diesen Zustand hatte ich nur ein Ziel, nämlich mich wieder bewegen zu können. Ich nahm alle meine Willenskraft zusammen und versuchte, auf die Seite zu rollen: Etwas rollte, aber nicht mein Körper – etwas, was ich war, mein ganzes Bewußtsein mit allen seinen Empfindungen. Ich rollte auf den Boden neben dem Bett. Dabei hatte ich das Gefühl, als sei ich nicht körperlos, sondern mein Leib bestehe aus einer Art Substanz, die eine Mischung zwischen gasförmig und flüssig sei. Das Erstaunen, gemischt mit Verblüffung, die mich packten, als ich wohl spürte, daß ich auf den Boden fiel, aber der erwartete harte Aufprall ausblieb, habe ich bis heute nicht vergessen. Eigentlich hätte ich meinen Kopf an der Kante meines Nachttischchens aufschlagen müssen, wenn die Bewegung mit meinem normalen Körper abgelaufen wäre. Am Boden liegend, befahlen mich eine unheimliche Angst und Panik. Ich wußte, daß ich einen Körper besaß, und ich hatte nur einen großen Wunsch, diesen wieder zu beherrschen. Mit einem jähen Ruck gelang es mir, diesen wieder zu bewegen, ohne daß ich wußte, wie ich zu ihm zurückgelangt war.

Die Häufigkeit des Auftretens von AKEs reicht von 8% in der gesamten Bevölkerung bis zu 25% bei Studenten, mit extrem häufigem Vorkommen in bestimmten Bevölkerungsgruppen wie, um nur ein Beispiel zu nennen, 42% bei Schizophrenen (Blackmore, 1986; vgl. auch Blackmore, 1982; ein Überblick und weitere Verweise finden sich in Alvarado, 1986, 2000, S. 18; Irwin, 1985, S. 174, Metzinger 2009, S. 135ff). Es wäre jedoch falsch, anzunehmen, daß AKEs typischerweise bei Menschen mit schweren psychiatrischen Störungen oder neurologischen Defiziten vorkommen. Ganz im Gegenteil: Der Großteil der Berichte über außerkörperliche Erfahrungen stammt von gewöhnlichen Menschen in alltäglichen Situationen. Bleiben wir deshalb bei nicht-pathologischen Situationen und sehen wir uns ein weiteres paradigmatisches Beispiel an, das wieder vom schweizerischen Biochemiker Ernst Waelti berichtet wird:

Benommen ging ich gegen elf Uhr nachts ins Bett und versuchte einzuschlafen. Ich war unruhig und wälzte mich häufig, was meine Frau veranlaßte, kurz zu schimpfen. Ich zwang mich nun, still dazuliegen. Eine Weile dämmerte ich so dahin, als ich das Bedürfnis spürte, meine Hände, die auf der Bettdecke lagen, hinaufzuziehen und in eine bequemere Lage zu bringen. Im gleichen Moment merkte ich, daß ich mich überhaupt nicht mehr bewegen konnte und mein Körper in einer Art Starre lag. Gleichzeitig konnte ich die Hände aus meinen körperlichen Händen herausziehen, als wären die letzteren ein Paar steife Handschuhe. Die Loslösung begann deutlich spürbar an den Fingerspitzen, fast mit einem fühlbaren Geräusch, einer Art Knistern. Es war dies genau die Bewegung, die ich eigentlich mit meinen körperlichen Händen hatte ausführen wollen. Mit dieser Bewegung löste ich mich von meinem Körper und schwebte kopfseitig aus ihm heraus. Ich richtete mich auf, als hätte ich fast kein Gewicht mehr. Trotzdem hatte ich einen Körper mit richtigen Gliedern. Sie haben sicher einmal gesehen, wie elegant sich eine Qualle im Wasser bewegt. Mit der gleichen Leichtigkeit konnte ich mich bewegen.

Ich legte mich waagrecht in der Luft hin und schwebte in zwei Meter Höhe wie ein Schwimmer, der sich vom Rand eines Schwimmbeckens abgestoßen hatte, über

das Bett hinweg. Ein beglückendes Gefühl der Befreiung stieg in mir auf. Doch bald packte mich wieder die uralte, allen Lebewesen innewohnende Angst, meinen physischen Körper zu verlieren. Sie genügte, um mich zurück in meinen Leib zu treiben.²¹

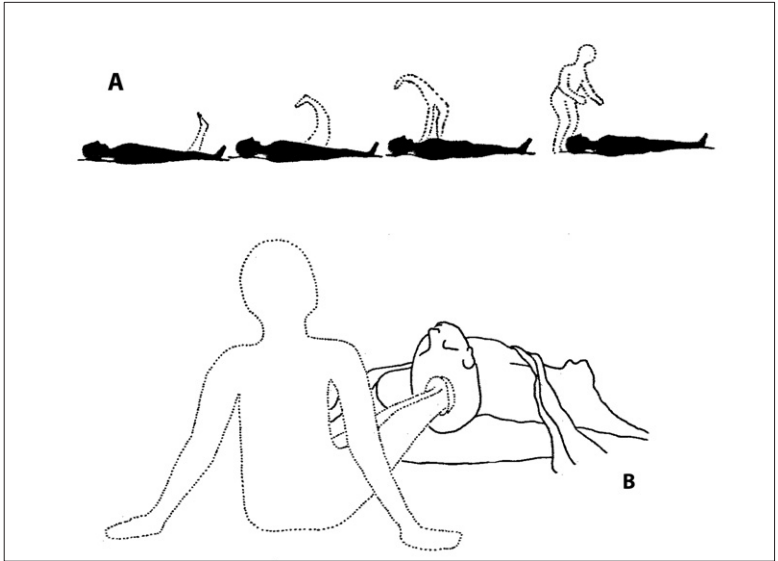


Abb. 6b: Kinematik des bewußten Körperbildes zu Beginn einer außerkörperlichen Erfahrung. Zwei alternative, aber gleichermaßen charakteristische Bewegungsmuster, wie sie der Schweizer Biochemiker Ernst Waelti im Jahr 1983 beschrieb.

Eine Schlaflähmung, wie in den beiden oben zitierten Beispielen, ist allerdings keine notwendige Bedingung für das Auftreten von AKEs. Sie kommen häufig während Extremsportarten vor, zum Beispiel beim Klettern in großen Höhen oder bei Marathonrennen.

Eine schottische Frau schrieb, daß sie eine AKE hatte, als sie 32 Jahre alt war und für einen Marathon trainierte. »Nachdem ich ungefähr 12-13 Meilen gelaufen war ..., hatte ich das Gefühl, nicht durch meine eigenen Augen zu schauen, sondern von woanders aus. ... Ich fühlte mich, als ob etwas meinen Körper verlassen würde, und obwohl ich immer noch weiterlief und mir die Landschaft ansah, sah ich auch mich selbst laufen. Meine Seele, oder was auch immer es war, schwebte irgendwo über meinem Körper, hoch genug, um die Baumspitzen und die kleinen Hügel zu sehen.« (Alvarado, 2000, S. 184; Übersetzung des Autors)

Die klassische außerkörperliche Erfahrung enthält zwei Selbstmodelle: eines, das visuell aus einer externen Perspektive dargestellt wird, und eines, das das Zentrum der phänomenalen Welt bildet, aus dem die Erste-Person-Perspektive hervorgeht. Was die repräsentationalistische und funktionalistische Analyse von AKEs schwierig macht, ist die Tatsache, daß viele *verwandte* Phänomene existieren, z.B. autoskopische Phänomene während epileptischer Anfälle, bei denen nur das erste Kriterium erfüllt ist (eine neurologische Kategorisierung findet sich in Brugger et al., 1997; eine ausgezeichnete neuere Sammlung von Fallstudien ist Heydrich et al., 2010). Devinsky et al. (1989, S. 1080) unterscheiden zwischen Autoskopie in Form einer komplexen, halluzinatori-

²¹ Die Zitate stammen aus E. Waelti, *Der dritte Kreis des Wissens*, Seiten 18 und 25.

schen (visuellen) Wahrnehmung des eigenen Körpers als extern, bei der »das Bewußtsein der Person ... meist als in ihrem Körper befindlich erlebt wird«, und einer zweiten Form, der klassischen AKE, die das Gefühl, den Körper zu verlassen und ihn aus einem anderen Blickwinkel zu sehen, beinhaltet. Autoskopische Anfälle kommen wahrscheinlich häufiger vor als früher angenommen. Die Autoren fanden ein Vorkommen von 6,3% unter ihren Patienten (Devinsky et al., 1989, S. 1085). Anfälle, die keine motorischen Symptome oder den Verlust des Bewußtseins umfaßten und vom Patienten möglicherweise nicht als solche erkannt werden, können tatsächlich häufiger vorkommen als allgemein angenommen (eine Fallstudie eines Patienten, der zunächst einige Jahre lang AKEs hatte und erst später unter generalisierten epileptischen Anfällen litt, findet sich in Vuilleumier et al., 1997, S. 116).

Welche Funktion könnte diese Art Erfahrung *für* den Organismus als Ganzes haben? Hier ist ein spekulativer Vorschlag von Devinsky und seinen Kollegen:

Es gibt mehrere mögliche Vorteile, die sich aus dissoziativen Phänomenen wie der Autoskopie ergeben können. Wenn zum Beispiel ein Beutetier mit hoher Wahrscheinlichkeit von einem Raubtier gefaßt wird, kann ein fingierter Tod von Bedeutung für das Überleben sein. Außerdem legen Berichte von Überlebenden von Nahtoderfahrungen im Gefecht oder beim Bergsteigen nahe, daß die mentale Klarheit, die mit einer Dissoziation einhergeht, es Menschen ermöglichen kann, bemerkenswerte Rettungsmanöver auszuführen, die sonst vielleicht nicht möglich wären. Daher könnte die Dissoziation ein neuronaler Mechanismus sein, der es ermöglicht, inmitten eines Nahtodtraumas die Ruhe zu bewahren. (Devinsky et al., 1989, S. 1088; Übersetzung des Autors)

Begrifflich ist es durchaus denkbar, daß ein Informationsverarbeitungssystem unter physischem oder emotionalem Streß gezwungen sein kann, eine »repräsentationale Arbeitsteilung« einzuführen, indem es verschiedene repräsentationale Funktionen auf zwei oder mehrere unterschiedliche Selbstmodelle aufteilt (wie zum Beispiel bei dem Phänomen, das früher »multiple Persönlichkeitsstörung« genannt wurde, vgl. Metzinger, 2003a, Abschnitt 7.2.4). Die AKE könnte im Zusammenhang mit traumatisierenden Situationen ein Fall von vorübergehender funktionaler Modularisierung sein, ein Fall einer durchaus »zweckdienlichen«, d.h. funktional adäquaten Trennung von repräsentationalen Inhaltsebenen im PSM. Wenn das System beispielsweise von somatosensorischem Input abgeschnitten ist oder überflutet wird mit Streßsignalen und Informationen, die die allgemeine Integrität des Selbstmodells selbst bedrohen, könnte es von Vorteil sein, die aktuell laufende bewußte Repräsentation höherer kognitiver Funktionen - also Aufmerksamkeit, begriffliches Denken und willentliche Selektionsprozesse- in ein *separates* Modell des Selbst einzubetten. Dies könnte einen hohen Grad an lokaler Kohärenz und integrierter Verarbeitung, d.h. von »mentaler Klarheit«, ermöglichen, indem verschiedene Funktionen wie Propriozeption, Aufmerksamkeit und Kognition funktional abgeschottet und damit *modularisiert* werden, um zumindest einige dieser Funktionen in einer lebensbedrohlichen Situation aufrecht zu erhalten. Fast alle notwendige systembezogene Information ist nach wie vor global verfügbar und höherstufige Prozesse wie Aufmerksamkeit und Kognition können diese Informationen nach wie vor handhaben, da sie weiterhin auf integrierte Weise präsentiert

sind. Allerdings verändert sich ihre Verteilung auf spezifische Subregionen des phänomenalen Raumes als Ganzes erheblich. Nur eines der beiden Selbstmodelle ist wirklich in der multimodalen Gesamtszene »situert«; nur eines der beiden ist unmittelbar verkörpert und virtuell selbstpräsent in dem Sinne, daß es tatsächlich als autonomer Agent in einen innerlich simulierten Verhaltensraum integriert ist.

Es ist seit langem bekannt, daß AKEs nicht nur bei gesunden Personen vorkommen, sondern auch in bestimmten klinischen Populationen (z.B. Epileptikern). In einer jüngeren Studie konnten Olaf Blanke und seine Kollegen die betreffende Gehirnläsion oder Funktionsstörung des temporoparietalen Übergangs bei fünf von sechs Patienten lokalisieren. Ebenfalls war es ihnen erstmals möglich, einen AKE-ähnlichen Zustand mit Hilfe von direkter elektrischer Stimulation hervorzurufen. Die generelle Arbeitshypothese ist, daß zwei verschiedene Fehlfunktionen oder Randbedingungen notwendig sind, um eine AKE hervorzurufen. Der eine ist die Desintegration verschiedener Schichten des Selbstmodells oder des »personalen Raums« (hervorgerufen durch das Scheitern der Integration propriozeptiver, taktiler und visueller Informationen bezüglich des eigenen Körpers); zusätzlich ist eine zweite Desintegration zwischen externem, »extrapersonalem«, visuellem Raum und dem inneren Bezugsrahmen erforderlich, der durch vestibuläre Information erzeugt wird. Die Erfahrung, den eigenen Körper in einer Position zu sehen, die nicht mit seiner gefühlten Position übereinstimmt, könnte daher von einer zerebralen Funktionsstörung im temporoparietalen Übergang ausgelöst werden, die beide Formen der funktionalen Desintegration verursacht und so zu der oben beschriebenen repräsentationalen Konfiguration führt.

Durch das Mapping evozierter Potentiale haben Blanke und seine Kollegen auch nachgewiesen, daß eine selektive Aktivierung des temporoparietalen Übergangs (nach 330-400 ms) stattfindet, wenn gesunde Freiwillige sich geistig vorstellen, sie seien in genau der Position und nähmen die visuelle Perspektive ein, die für eine AKE charakteristisch sind. Gleichzeitig ist es möglich, diese mentale Transformation des körperlichen Selbstmodells zu beeinträchtigen, indem man die Informationsverarbeitung an diesem spezifischen Ort mit transkranieller Magnetstimulation blockiert. Bei einer epileptischen Patientin mit AKEs, die durch eine Verletzung des temporoparietalen Übergangs ausgelöst wurden, konnte durch aktive innere Imagination des PSM der AKE (d.h. durch die mentale Simulation einer AKE, wie sie sie bereits vorher erlebt hatte) gezeigt werden, daß eine partielle Aktivierung des Krampfherdes stattgefunden hatte (Blanke et al., 2005, eine wichtige neuere Studie ist Ionta et al., 2011). Demnach existiert eine anatomische Überschneidung zwischen den neuronalen Korrelaten dieser drei sehr ähnlichen Typen von phänomenalem Inhalt.

Was momentan am meisten gebraucht wird, ist ein experimentelles Design, das AKEs unter Laborbedingungen kontrollierbar und zu einem reproduzierbaren Phänomen bei gesunden Versuchspersonen macht. Dieses Zwischenziel zu erreichen wäre nicht nur vom empirischen, sondern auch vom philosophischen Standpunkt aus von großer Bedeutung. Wenn man die funktionale Feinstruktur des vorsprachlichen körperlichen Selbstbewußtseins besser verstünde, indem man durch eine überzeugende repräsentationalistische Analyse der phänomenalen *Entkörperlichung* schrittweise ein besseres Verständnis der ihm zu Grunde liegenden kausalen Architektur entwickelte, dann würden dadurch sicher die klassischen Probleme des nicht-begrifflichen Selbstbe-

wußtseins und des Ursprungs einer Erste-Person-Perspektive in einem neuen Licht erscheinen. Insbesondere wäre es von großer theoretischer Bedeutung, wenn ein Minimalmodell des phänomenalen Selbstbewußtseins erstellt werden könnte, zum Beispiel indem man die Möglichkeit eines minimalen Ichgefühls *ohne Agentivitäts-Komponente* empirisch nachweisen könnte (Blanke & Metzinger, 2009). Daher möchte ich ein kurzes Beispiel meiner eigenen neueren Forschungen vorstellen. Beispiel 5 ist eine mittlerweile klassische Studie, die sich aus einer interdisziplinären Kooperation zwischen Neurowissenschaft und Philosophie des Geistes ergeben hat und ein experimentelles Design zugrunde legt, das ursprünglich aus philosophischen Überlegungen entwickelt wurde (Details finden sich in Lenggenhager et al., 2007; ein populärwissenschaftlicher Überblick ist Metzinger, 2009, ein ausgezeichnete Überblick über die weitere Entwicklung dieses Forschungsgebiets ist Blanke, 2012).

Die klassische Gummihand-Illusion (Beispiel 4) sagt lediglich etwas über die lokale phänomenale Zieleigenschaft des »Besitzens« (für Körperteile) aus, aber nichts über die globale phänomenale Qualität »Selbstheit« (sozusagen die Eigenschaft des Besitzens für den *ganzen* Körper). Um die Selbstlokalisierung und die subjektive Identifikation mit dem *ganzen* Körper nicht-invasiv zu manipulieren und das globale phänomenale Ichgefühl *per se* zu untersuchen, wurde ein Experiment entwickelt, das auf klinischen Daten von neurologischen Patienten mit AKEs basiert. Aus diesen Daten geht hervor, daß die räumliche Einheit zwischen Selbst und Körper gestört werden kann, was in manchen Fällen zu der bemerkenswerten Erfahrung führt, daß das bewußte Selbst sich in einer außerkörperlichen Position befindet. Das Ziel des Experiments war es deshalb, AKEs bei gesunden Versuchspersonen auszulösen, um die phänomenalen Zieleigenschaften der Selbst-Identifikation und Selbst-Lokalisation zu untersuchen. Die Hypothese war zunächst, daß Teilnehmer der Studie unter den richtigen experimentellen Bedingungen einen visuell präsentierten Körper als ihren eigenen erleben würden, was eine Verschiebung des subjektiv erlebten körperlichen Selbst in eine Position außerhalb der eigenen körperlichen Grenzen hervorrufen würde. Kann man eine Ganzkörperanalogie der Gummihand-Illusion schaffen, eine Illusion, in der gesunde Teilnehmer einen virtuellen Körper als ihren eigenen erleben und ihr Selbst außerhalb ihrer körperlichen Grenzen, in einer anderen Position im Raum, lokalisieren? Bigna Lenggenhager und Tej Tadi setzten Techniken der virtuellen Realität ein, um die mögliche Herbeiführung von AKEs durch einen multisensorischen Konflikt zu untersuchen. Im ersten Experiment sahen die Versuchspersonen ihren Körper von hinten, der aus einer Distanz von zwei Metern gefilmt und auf ein *head-mounted display* (siehe Abb. 7) projiziert wurde. Eine Minute lang wurde entweder synchron oder asynchron zum virtuell gesehenen Körper der Rücken der Versuchspersonen taktile stimuliert. Die globale Selbstzuschreibung, also die erlebte Identifikation mit der virtuellen Figur, wurde dann mit Hilfe eines Fragebogens gemessen, der von der Gummihand-Illusion übernommen und adaptiert wurde. Die Verschiebung der Selbstlokalisierung wurde gemessen, indem die mit einer Augenbinde versehenen Versuchspersonen direkt nach dem Auftreten der Illusion desorientiert und dann gebeten wurden, wieder in ihre Ausgangsposition zurückzukehren (Abb. 7).

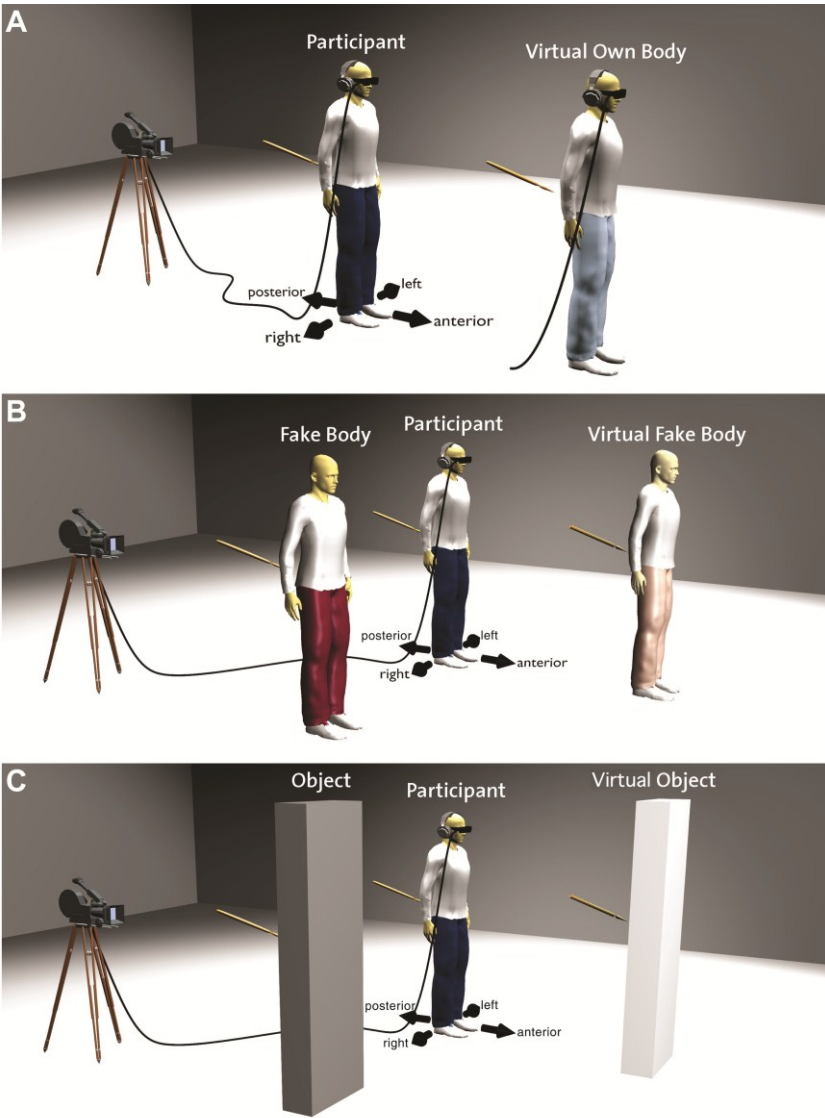


Abb. 7: Die Erzeugung einer Ganzkörper-Variante der Gummiband-Illusion. (A) Der Versuchsteilnehmer (dunkle Hose) sieht durch ein HMD (ein *head-mounted display* ist ein auf dem Kopf getragenes visuelles Ausgabegerät, das am Computer erzeugte Bilder auf einem augennahen Bildschirm darstellt oder sogar direkt auf die Netzhaut projiziert) seinen eigenen virtuellen Körper (hellere Hose) in einer dreidimensionalen Darstellung zwei Meter vor sich stehen und synchron oder asynchron am Rücken des Teilnehmers gestreichelt werden. Bei anderen Versuchsbedingungen sieht der Teilnehmer entweder (B) einen virtuellen künstlichen Körper (nämlich den Rücken einer Schaufensterpuppe; helle Hose) oder (C) ein virtuelles nichtbiologisches Objekt, das keinerlei Ähnlichkeit mit einem menschlichen Körper hat, welches aber ebenfalls synchron oder asynchron an der Rückseite gestreichelt wird. Dunkle Farben zeigen den tatsächlichen Standort des physischen Körpers oder Objekts an, während helle Farben den virtuellen Körper bzw. das virtuelle Objekt anzeigen, wie es im HMD gesehen wird. (Illustration von M. Boyer.)

Während die Versuchspersonen gestreichelt wurden, wurde ihnen entweder ihr eigener Rücken (»Eigenkörper-Bedingung«), der Rücken einer Schaufensterpuppe (»Kunstkörper-Bedingung«) oder ein unbelebtes Objekt (»Objekt-Bedingung«) gezeigt, die gestreichelt wurden und deren Bild entweder direkt (synchron) oder mit zeitlicher Verzögerung (asynchron) auf das *head-mounted display* projiziert wurde. Nach der Stimulierungsphase wurden die Versuchspersonen passiv im Raum versetzt und dann gebeten, in ihre Ausgangsposition zurückzukehren und einen modifizierten »Gummihand-Fragebogen« auszufüllen. Die Ergebnisse der Fragebogenauswertung zeigten, daß die Versuchspersonen bei den synchronen »Eigenkörper- und »Kunstkörper«-Situationen oft das Gefühl hatten, die betrachtete virtuelle Figur sei ihr eigener Körper. Dieser Effekt war in der »Objekt-Bedingung« und unter allen asynchronen Stimulationsbedingungen weniger wahrscheinlich. Die synchronen Bedingungen zeigten auch eine signifikant größere Verschiebung in Richtung des projizierten echten oder künstlichen Körpers als die asynchronen und die Kontrollbedingungen. Diese Daten legen nahe, daß Selbstlokalisierung – auf Grund widersprüchlichen visuell-somatosensorischen Inputs – genauso anfällig für Fehlidentifizierung und Fehllokalisierung ist wie es bereits von anderen Körperteilen berichtet wurde, beispielsweise bei der Gummihand-Illusion.

Illusorische Selbstlokalisierung in einer Position außerhalb des eigenen Körpers macht deutlich, daß das körperliche Selbstbewußtsein und die Phänomenologie der Identifikation prinzipiell von einer korrekten Repräsentation der Position des eigenen physischen Körpers im Raum getrennt werden können. Insofern unterscheidet sich dieses Phänomen von der Gummihand-Illusion, in der das »Ichgefühl« konstant blieb und lediglich die Selbstattribution und Lokalisierung der stimulierten Hand manipuliert wurde. Bedeutet die illusorische Selbstlokalisierung in einer Position außerhalb des eigenen Körpers jetzt, daß wir experimentell vollständige AKEs hervorgerufen haben? Nein, dies war nur ein erster Schritt und der Effekt ist auch wesentlich schwächer als die Gummihand-Illusion, er ähnelt für viele Versuchspersonen eher der Phänomenologie der Heautoskopie. Allerdings ist es recht deutlich, was die nächsten Schritte sein müssen. Außerkörperliche Erfahrungen sind gekennzeichnet durch eine Verlagerung des phänomenal erlebten Selbst hin zu einem außerkörperlichen Ort, einer außerkörperlichen, typischerweise visuell-räumlichen Perspektive (obwohl es auch rein audiotrische AKEs gibt), sowie durch die Tatsache, daß man den eigenen Körper dann von diesem außerkörperlichen Ort (des Selbst) aus sieht. Da die vorliegende Illusion weder mit einem expliziten *disembodiment* – also mit dem Gefühl, den eigenen Körper verlassen zu haben oder das Körpergefühl gänzlich verloren zu haben – noch mit einer Verschiebung der visuell-räumlichen Perspektive einherging, haben wir nur einige Aspekte einer AKE hervorgerufen. Die Erfahrung ähnelte damit eher der mit AKEs verwandten Heautoskopie, die ebenso bei neurologischen Patienten beobachtet wurde (weitere Literaturhinweise finden sich in den Originalpublikationen, Metzinger, 2009 und Blanke, 2012). Diese sind, wie bei der Autoskopie, durch eine illusorische visuelle Repräsentation des Körpers – einen visuellen Doppelgänger –, sowie durch einen raschen Wechsel der Selbstlokalisierung zwischen dem illusorischen und dem echten Körper gekennzeichnet.

Um nur ein Beispiel für zukünftige Forschungsziele zu nennen: Ich bin der Meinung, daß eine vorübergehende Episode visuell-vestibulärer Desintegration eine weitere notwendige Bedingung für die Erzeugung voll entwickelter AKEs und den vollständigen Transfer des »Ichgefühls«, der phänomenalen Eigenschaften von Selbstlokalisierung und Identifikation, auf den illusorischen Körper darstellen könnte. Es müssen mindestens zwei räumliche Bezugsrahmen im Gehirn schrittweise funktional dissoziiert werden, um nicht nur eine »Teleportations-AKE« zu erhalten, sondern eine realistische Austrittsphänomenologie, also eine graduelle Bewegungsbahn durch den phänomenalen Raum (siehe Abb. 6a). Wenn es nicht um plötzliche AKEs wie nach Unfällen oder direkter elektrischer Stimulation des Gehirns geht, dann sollte die Austrittsphänomenologie am Beginn der Erfahrung diese Entkopplung verschiedener räumlicher Bezugsrahmen auf der Ebene des Bewußtseins widerspiegeln. Warum ist diese Überlegung aus theoretischer Perspektive relevant und warum ist es so schwer, ihn experimentell zu untersuchen? In Standardsituationen und im Gegensatz zu allen anderen bewußten Modellen von Teilen der jeweils als real erlebten Situation ist das menschliche Selbstmodell durch einen kontinuierlichen Fluß von selbstgeneriertem Input im Gehirn verankert (diesen Punkt habe ich in Metzinger, 2003 näher ausgeführt). Es existiert eine permanente kausale Verbindung zum physischen Körper selbst: Es sind der interozeptive Input aus Eingeweiden und Blutgefäßen, der ständige Informationsfluß aus dem Gleichgewichtsorgan und auch der emotionale Hintergrundzustand, die das PSM des Menschen fest in seiner neuronalen Basis verankern und es schwer machen, es sozusagen aus dem Körper »herauszukopieren« oder auf ein anderes Trägersystem zu portieren. Um die SMT besser verstehen zu können, müssen wir uns nun diesem Punkt zuwenden – er erklärt, warum unser bewußtes Realitätsmodell ein *zentriertes* Realitätsmodell ist.

Der vierte Schritt: Das körperliche Selbst als eine funktionale Verankerung des phänomenalen Raums

Ich habe oben die Unterscheidung zwischen einer repräsentationalen und einer funktionalen Analyse der Erste-Person-Perspektive gemacht. Das zentrale theoretische Problem auf der funktionalen Beschreibungsebene kann mit folgender Frage zusammengefaßt werden: Was genau ist der spezifische Unterschied zwischen dem PSM und anderen phänomenalen Modellen, die gegenwärtig im System aktiv sind? Gibt es eine charakteristische kausale Signatur des PSM? Durch welche *funktionale Eigenschaft* wird es zum stabilen Zentrum des phänomenalen repräsentationalen Raums?

Hier ist meine vorläufige Antwort: Das Selbstmodell ist die einzige repräsentationale Struktur, die durch eine *kontinuierliche Quelle intern generierten Inputs* im Gehirn verankert ist. Nennen wir dies die »Hypothese der permanenten kausalen Verbindung«. Immer dann, wenn bewußtes Erleben entsteht (also wenn ein stabiles, integriertes Modell der Wirklichkeit aktiviert wird), existiert auch diese kontinuierliche Quelle interozeptiven Inputs.²² Das menschliche Selbstmodell besitzt eine dauerhafte und maximal direkte

²² Die Bedeutung von Interozeption für das Selbstbewußtsein wird in einer ganzen Reihe von neueren Studien von Manos Tsakiris und Kollegen untersucht; siehe etwa Tsakiris et al.,

kausale Verknüpfung im Gehirn, und zwar durch die Eigenwahrnehmung des Körpers. Es hat Teile, die durch *permanente* Formen der Verarbeitung von *permanenten* Formen selbstgenerierten Inputs, zum Beispiel auch aus niedrigstufiger, homöostatischer Autoregulation, realisiert werden. Um diesen grundlegenden Gedanken anders auszudrücken: Der Körper ist das einzige Wahrnehmungsobjekt, vor dem das Gehirn nie weglaufen kann. Es gibt eine Reihe offensichtlicher Kandidaten für Quellen hoher Invarianz, zum Beispiel die folgenden vier Typen von intern erzeugter Information, die während Episoden des bewußten Erlebens ein dauerhaftes funktionales Bindeglied zwischen dem PSM und seiner körperlichen Basis im Gehirn erzeugen:

- Input aus dem Vestibulärorgan: der Gleichgewichtssinn.
- Input aus dem autonom aktiven, invarianten Teil des Körperschemas: Das andauernde »Hintergrundgefühl« im räumlichen Körpermodell, das unabhängig ist von externem Input, z.B. durch Bewegungswahrnehmung.
- Input aus Sensoren der Eingeweide, aber auch aus den Blutgefäßen, zum Beispiel aus den kardiovaskulären Mechanosensoren: »Bauchgefühle« und somatoviszzerale Formen der Selbstrepräsentation.
- Input aus bestimmten Teilen des oberen Hirnstamms und des Hypothalamus: Hintergrundemotionen und -stimmungen, die in der kontinuierlichen homöostatischen Selbstregulation des »internen Milieus« verankert sind, also der biochemischen Landschaft in unserem Blut.

Philosophisch gesehen sind es nicht so sehr die neurobiologischen Details, die entscheidend sind, sondern vielmehr die plausible Annahme, daß es einen bestimmten Teil des menschlichen Selbstmodells gibt, der sich durch einen hohen Grad an Stimuluskorrelation auszeichnet und ausschließlich auf *intern* generierter Information beruht. Diese Ebene des PSM ist direkt und permanent in Stimuli aus dem Inneren des Körpers verankert. Diese Tatsache ist phänomenologisch relevant, weil sie den entscheidenden Beitrag zur quasi-cartesianischen Phänomenologie der Substantialität (»Ich bin eine ontologisch autonome Entität, die sich selbst im Sein halten kann«) und der Selbstgewißheit des Subjekts (»Ich weiß, daß ich *weiß*, daß ich selbst existiere«) leistet. Sie ist aber auch erkenntnistheoretisch relevant, weil sie auf der funktionalen Ebene die garantierte Referenz erzeugt, von der oben die Rede war. Erinnern Sie sich noch an die Patientin AZ aus dem zweiten Beispiel? Der schwächere Grad an phänomenaler »Lebhaftigkeit« oder »Echtheit« in ihren Phantomgliedern könnte eben gerade die Abwesenheit von permanenter *bottom-up*-Stimulation widerspiegeln, die unter normalen Umständen von existierenden physischen Gliedern verursacht wird. Marcel Kinsbourne hat in diesem Zusammenhang von einem »*background 'buzz' of somatosensory input*« (Kinsbourne, 1995, S. 217) gesprochen. Um die Phänomenologie, die auf der repräsentationalen Beschreibungsebene an diesem »puren Körperlichkeitsgefühl« beteiligt ist, zu erfassen, möchte ich zwischen Selbstrepräsentation und Selbststrepräsentation

2011, Ainley et al., 2012. Ein interessantes komputationales Modell ist Seth, Suzuki & Critchley, 2012.

unterscheiden.²³ Phänomenologisch gesehen bezieht sich der erste Begriff auf das rein sensorische Gefühl der körperlichen Präsenz, welches interessanterweise einhergeht mit einer subjektiven Wahrnehmung zeitlicher Unmittelbarkeit und der auf Erfahrung beruhenden Sicherheit, eine direkte, nicht-inferentielle Selbstkenntnis zu besitzen. Was genau ist diese tiefste Schicht des phänomenalen Selbst? Warum ist es der Ursprung der Erste-Person-Perspektive? Meine Hypothese besagt, daß nach der raumzeitlichen Selbstlokalisierung die konstante Selbstorganisationsaktivität derjenigen Regionen des körperlichen Selbst, die von externem Input unabhängig sind, das funktionale *Zentrum* des phänomenalen Repräsentationsraums bildet.

Als ein erstes Beispiel dafür, was man unter einem Selbstmodell verstehen kann, hatte ich Ramachandrans Experiment zur Mobilisierung eines gelähmten Phantomglieds vorgestellt. Eine Selbst*präsentation* ist genau der Teil des Phantomglieds, der unabhängig von Bewegungen immer bewußt bleibt. Wenn man *diesen* Teil verliert, verliert man auch das subjektive Erleben körperlicher Präsenz – man wird zu einem »körperlosen Wesen«.²⁴ Jedoch könnte es sogar weitere, allgemeinere empirische Perspektiven geben, aus denen das Selbstmodell zwangsläufig mit der Basisaktivität des Gehirns, wie man sie im Ruhezustand beobachten kann, verbunden ist (vgl. Raichle et al., 2001; Gusnard, 2005).

Der fünfte Schritt: Autoepistemische Geschlossenheit – Die Transparenzannahme und das naive-realistische Selbstmißverständnis

Das zentrale theoretische Problem auf der *repräsentationalen* Beschreibungsebene besteht darin, daß man mir leicht vorwerfen könnte, ich würde mit der Einführung des Begriffs »Selbstmodell« das eigentliche Problem terminologisch umgehen. Die zentrale Eigenschaft des Selbstbewußtseins – die Tatsache, daß wir uns immer als Selbst erleben, aber nicht als Modell – scheint durch die SMT somit nicht erklärt zu werden. Aber zum einen ist ein Selbstmodell natürlich nicht ein statischer repräsentationaler Träger, das Modell eines mysteriösen Einzeldings, das wir dann *das Selbst* nennen. Es ist ein kontinuierlicher und auf das betreffende System selbst gerichteter Prozeß, der globale Eigenschaften des Organismus in ihrer zeitlichen Entwicklung verfolgt und überwacht. Zum anderen scheint es keine notwendige Verbindung zwischen den die-

²³ Zu einer ausführlichen theoretischen Behandlung dieses Themas und zahlreichen jüngeren empirischen Ergebnissen zum Körper als Anker des bewußten Erlebens vgl. Damasio (1999). Antonio Damasio verwendet den Begriff eines Kernselbst (*core self*) und an anderer Stelle (Metzinger, 1993, S. 156ff; Metzinger, 2003a, Abschnitt 5.4) habe ich den technischen Begriff der »phänomenalen Selbstrepräsentation« (gegenüber der Selbst*repräsentation*) eingeführt. Auf der Ebene der Körperrepräsentation ist die Selbst*präsentation* das, was AZ in ihren Phantomgliedern fehlt, während Selbst*repräsentation* das ist, was sie tatsächlich besitzt – obwohl dies offensichtlich auch eine Form der *Felb*repräsentation ist, da der Referent dieser Repräsentation nie existiert hat.

²⁴ Auch hierzu existieren die entsprechenden phänomenologischen Zustandsklassen. In Metzinger (1993) und Metzinger (1997) habe ich Oliver Sacks' Beispiel der »körperlosen Lady« besprochen. Vgl. in diesem Zusammenhang auch den berühmten Fall von Ian Waterman, der in Metzinger (2003a) diskutiert wird. Interessant ist jedoch auch die neuere Traumforschung, die zeigt, wie es möglich ist, daß das phänomenale Subjekt stabil sein kann, sich aber trotzdem lediglich als ausdehnungsloser Punkt in einem räumlichen Bezugsrahmen selbst lokalisiert, Windt, 2010, 2014.

sen Vorgang charakterisierenden funktionalen und repräsentationalen Eigenschaften einerseits und den *phänomenalen* Zieleigenschaften der »Meinigkeit«, »präreflexiven Selbstheit« und »Perspektivität« andererseits zu geben – zumindest gewissen modalen Intuitionen zufolge. Manche denken: All das könnte sich durchaus ereignen, ohne daß es zu Bewußtsein, zur Entstehung eines echten phänomenalen Selbst oder einer subjektiven Innenperspektive kommt. Solange der Begriff »Bewußtsein« noch so extrem inhaltsleer ist, wie dies in der gegenwärtigen Debatte der Fall ist, scheinen Zombies zumindest logisch möglich zu bleiben: Es ist denkbar, daß biologische Informationsverarbeitungssysteme einen Darstellungsraum entwickeln und erfolgreich anwenden können, in dessen Zentrum ein Selbstmodell steht, *ohne* dabei auch phänomenales Selbstbewußtsein zu entwickeln. Etwas interessanter ist der mögliche Einwand, daß die Entstehung der spezifischen phänomenalen Qualität der *Selbstheit* – sogar dann, wenn man *die phänomenologische Ebene als gegeben voraussetzt*, d.h. sogar in einem System, das schon bewußt ist – nicht offensichtlich oder selbstverständlich ist. Was genau würde, mit begrifflicher Notwendigkeit, ein echtes Ego erzeugen? Ein »Selbstmodell« ist noch lange kein Selbst, sondern nur eine Repräsentation des Systems als einer Ganzheit – es ist nicht mehr als ein integriertes *Systemmodell*. Wenn die funktionale Eigenschaft der Zentriertheit und die repräsentationale Eigenschaft der Selbstmodellierung zu der phänomenalen Zieleigenschaft eines echten, bewußt erlebten »Ichgefühls« führen sollen, dann muß das bewußte Systemmodell zu einem phänomenalen Selbst werden. Die entscheidende philosophische Frage ist deshalb die folgende: Wie führt die Existenz eines funktional zentrierten Darstellungsraums notwendigerweise zu der Entstehung eines bewußten Selbst und dann zu dem, was wir gemeinhin eine phänomenale Erste-Person-Perspektive nennen? Mit anderen Worten: Wie wird das Systemmodell zu einem *Selbstmodell*?

Ein genuines, bewußtes Selbst – so lautet meine Antwort – entsteht immer genau in dem Moment, in dem das System das Selbstmodell, das es gerade auf der Ebene des bewußten Erlebens generiert, nicht mehr *als* Modell erleben kann. Wie kommt man also von der funktionalen Eigenschaft der »Zentriertheit« und der repräsentationalen Eigenschaft der »Selbstmodellierung« zu der phänomenalen Eigenschaft der »präreflexiven Selbstvertrautheit«? Die Lösung hängt mit dem zusammen, was Philosophen »phänomenale Transparenz« nennen (eine kurze Erläuterung des Begriffs findet sich in Metzinger, 2003c; Metzinger, 2003b ist der deutsche Vorläufer). Die vom System generierten repräsentationalen Zustände sind *transparent*, d.h. sie stellen die Tatsache, daß sie Modelle sind, nicht mehr auf der Ebene ihres Inhalts dar. Folglich – und dies ist zunächst nur eine visuell-phänomenologische Metapher, die vor über hundert Jahren von dem britischen Philosophen George Edward Moore eingeführt wurde – schaut das System durch seine eigenen repräsentationalen Strukturen »hindurch«, als ob es sich in direktem und unmittelbarem Kontakt mit ihrem Inhalt befände. Dies ist wohlbemerkt lediglich eine Aussage über die *Phänomenologie* des Systems. Es ist keine erkenntnistheoretische Aussage über den Besitz von Wissen: Man kann vollkommen verblendet sein und nichts oder sehr wenig über die Realität wissen (oder auch über den eigenen Geist) und sich dennoch der Phänomenologie der Gewißheit oder des *Wissens, daß man weiß*, erfreuen. Phänomenale Transparenz ist also nicht *epistemische* Transparenz, oder Descartes' klassische – und mittlerweile empirisch widerlegte –

Vorstellung, daß wir uns über den Inhalt unseres eigenen Geistes nicht täuschen können. Transparenz ist in diesem Zusammenhang ausschließlich als eine Eigenschaft des *bewußten* Zustands definiert. Unbewußte Zustände sind weder transparent noch opak. Phänomenale Transparenz steht ebenso wenig in direkter Verbindung mit dem dritten technischen Begriff in der Philosophie, nämlich mit »referentieller Transparenz«. Wesen, die weder eine Sprache noch das begriffliche Denken beherrschen, könnten deshalb durchaus phänomenal transparente Zustände haben, ebenso wie Systeme, die einer fundamentalen Selbsttäuschung unterliegen. Der naive Realismus ist hier keine Überzeugung, kein spezifischer semantischer Kontext oder eine kognitive, epistemische Einstellung, sondern das, was diese überhaupt erst ermöglicht: eine Eigenschaft des phänomenalen Erlebens selbst.

Ich vertrete zwei Hypothesen bezüglich der mikrofunktionalen Basis und der evolutionären Entstehungsgeschichte transparenter phänomenaler Zustände. Erstens: In einem sehr kleinen Zeitfenster werden die fraglichen neuronalen Datenstrukturen so schnell und zuverlässig aktiviert, daß das System sie nicht mehr als solche erkennen kann, z.B. wegen der vergleichsweise langsamen zeitlichen Auflösung *metarepräsentationaler* Funktionen oder weil sie zu einem stabilen, Bayes-optimalen Modell gehören, für das der Vorhersage-Irrtum bereits stark minimiert wurde (vgl. Howhy 2013, Friston 2010). Introspektiv ist der Konstruktionsprozeß unsichtbar. Zweitens: Durch ein viel größeres Zeitfenster betrachtet gab es anscheinend keinen evolutionären Selektionsdruck auf die entsprechenden Teile unserer funktionalen Architektur. Der naive Realismus war für biologische Systeme wie uns eine funktional adäquate Hintergrundannahme. Wir mußten wissen: »Vorsicht! Es ist ein Wolf in der Nähe!«, aber nicht: »Eine Wolf-Repräsentation ist momentan in meinem Gehirn aktiv!«

Phänomenale Transparenz hat also doch etwas mit Wissen zu tun (und nicht nur mit dem subjektiven Erleben *per se*): Es ist eine besondere Form der Dunkelheit. Es ist ein Mangel an Wissen. Epistemologisch gesehen ist sie ein impliziter Mangel an Selbstkenntnis, an Wissen über die funktionale Tiefenstruktur des eigenen Geistes, der aber selbst wiederum nicht wieder explizit repräsentiert wird. Wie Franz Brentano ([1874] 1973, S. 165f) und heute Daniel Dennett (1991, S. 359) deutlich gemacht haben, ist die Repräsentation von Abwesenheit nicht dasselbe wie Abwesenheit von Repräsentation. In transparenten Zuständen existiert keine Repräsentation vorhergehender Verarbeitungsstufen oder dynamisch instabiler Zustände. In der Phänomenologie des visuellen Erlebens bedeutet Transparenz, daß wir etwas nicht sehen können. Phänomenale Transparenz *im Allgemeinen* dagegen bedeutet, daß der Repräsentationscharakter des Gehalts des bewußten Erlebens selbst für das subjektive Erleben nicht zugänglich ist: Wir können nicht mehr sehen, daß es sich hier um eine *Vorhersage* handelt, um ein sehr gutes inneres Modell, dessen Funktion darin besteht, Überraschungen in der Zukunft zu vermeiden. Diese Analyse bezieht sich auf alle Sinnesmodalitäten, insbesondere auf das integrierte phänomenale Modell der Welt als Ganzer. Weil das *Mittel* der Darstellung als solches nicht dargestellt werden kann, wird das erlebende System notwendigerweise in einen naiven Realismus verstrickt; es erlebt sich selbst in direktem Kontakt mit dem Inhalt seines eigenen bewußten Erlebens. Was es nicht erleben kann, ist die Tatsache, daß sein Erleben immer in einem *Medium* stattfindet – und eben dies ist gemeint, wenn wir von der »Unmittelbarkeit« des phänomenalen Bewußtseins sprechen.

In einer vollständig transparenten Repräsentation können die Mechanismen, die zu ihrer Aktivierung führen, und die Tatsache, daß es einen konkreten inneren Zustand gibt, der seinen Gehalt trägt, introspektiv nicht mehr erkannt werden. Im Jargon der traditionellen analytischen Philosophie des Geistes gesprochen: »Nur Inhaltseigenschaften sind introspektiv verfügbar, Eigenschaften des repräsentationalen Trägers sind unzugänglich.« Die Phänomenologie der Transparenz ist daher die Phänomenologie des naiven Realismus.

Viele phänomenale Repräsentationen sind transparent, weil ihr Inhalt und vor allem dessen Existenz in allen möglichen Kontexten festzustehen scheint: Das Buch, das Sie jetzt in den Händen halten, wird dem subjektiven Erleben nach immer nur dieses Buch bleiben – unabhängig davon, wie sich die äußerliche Wahrnehmungssituation verändert. Sie erleben niemals, daß ein »aktiver Objektemulator« in Ihrem Gehirn gerade in Ihr globales Realitätsmodell integriert wird. Sie erleben lediglich den *Inhalt* des zugrunde liegenden Repräsentationsvorgangs: das *Buch* als Ihnen jetzt und hier anstrengungslos gegebenes Buch. Die beste Art und Weise, sich den Begriff der Transparenz klarzumachen, besteht darin, zunächst ganz traditionell zwischen dem Vehikel und dem Gehalt einer Repräsentation zu unterscheiden, zwischen repräsentationalem Träger und repräsentationalem Inhalt (vgl. dazu auch Dretske, 1998, S. 45ff).

Der repräsentationale Träger Ihres bewußten Erlebens ist ein bestimmter Prozeß im Gehirn. Dieser Prozeß – der in keiner Weise etwas »Buchhaftes« an sich hat – wird nicht bewußt erlebt; er ist transparent in dem Sinne, daß Sie, phänomenologisch gesehen, durch ihn hindurch schauen. Worauf Sie schauen ist sein repräsentationaler Inhalt, die durch die eigene Sinneswahrnehmung vermittelte Existenz eines Buches, hier und jetzt. Aus der begrifflichen Außenperspektive betrachtet ist der Inhalt eine abstrakte Eigenschaft des konkreten repräsentationalen Zustands im Gehirn. Wenn der repräsentationale Träger ein gut und zuverlässig funktionierendes Instrument zur Wissensgewinnung ist, erlaubt er Ihnen dank seiner Transparenz, »durch ihn hindurch« direkt auf die Welt, auf das Buch in Ihrer Hand zu schauen. Er macht die von ihm getragene Information global verfügbar, ohne daß Sie – die Person als Ganzes – sich darum kümmern müssen, *wie* das geschieht. Das Besondere an den meisten phänomenalen Repräsentationen ist, daß Sie diesen Inhalt aus der Innenperspektive auch dann als maximal *konkret* und unzweideutig, als direkt und unmittelbar gegeben erleben, wenn das fragliche Objekt – das Buch in Ihrer Hand – gar nicht existiert, sondern nur eine Halluzination ist. Phänomenale Repräsentationen scheinen genau die Art von Repräsentationen zu sein, bei denen wir repräsentationalen Gehalt und repräsentationalen Träger auf der Ebene des subjektiven Erlebens nicht unterscheiden können.

Es gibt natürlich Gegenbeispiele und diese sind hilfreich, um den Begriff der »Transparenz« noch besser zu verdeutlichen. *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 plötzlich entdecken, daß es das Buch in Ihrer Hand in Wirklichkeit nicht gibt, wird die Halluzination zu einer Pseudohalluzination. Die Information, daß Sie nicht auf die Welt, sondern »auf« einen aktiven repräsentationalen Zustand schauen, der im Moment allem Anschein nach nicht als ein zuverlässiges Instrument zur Wissensgewinnung fungiert, ist jetzt auch auf der Ebene des subjektiven Erlebens selbst

verfügbar. Der phänomenale Buchzustand wird undurchsichtig. Die *sensorische* Transparenz geht verloren. Ihnen wird die Tatsache bewußt, daß Ihre Wahrnehmungen durch Sinnesorgane erzeugt werden und daß diese Organe nicht in allen Situationen absolut zuverlässig funktionieren. Sie erleben das Buch jetzt nicht nur plötzlich als Repräsentation, sondern erleben es als *Fehlre*präsentation.

Nehmen wir weiter an, daß Sie jetzt sogar plötzlich entdecken, daß nicht nur die Buchwahrnehmung, sondern auch Ihr gesamtes philosophisches Nachdenken über das Problem des Bewußtseins in einem Traum stattfindet. Dann wird dieser Traum zu einem Klartraum (warum Klarträume eine philosophisch relevante Klasse von Bewußtseinszuständen sind, erläutere ich in Metzinger, 2003a, Abschnitt 7.2.5; mehr zu diesem Thema ist zu finden in Windt und Metzinger, 2007; Noreika et al., 2010; insbesondere in Windt, 2014). Die Tatsache, daß Sie momentan nicht eine Welt, sondern nur ein *Weltmodell* erleben, würde global verfügbar: Sie könnten diese Information deshalb jetzt auch zur Kontrolle Ihrer Handlungen und Ihrer Gedanken oder für die Aufmerksamkeitslenkung nutzen. Sie würden die *globale* Transparenz verlieren. Interessanterweise ist kognitive Verfügbarkeit allein aber nicht hinreichend, um den naiven Realismus des phänomenalen Erlebens zu durchbrechen. Man kann sich nicht einfach aus dem phänomenalen Modell der Wirklichkeit »herausdenken«, indem man seine Meinungen über dieses Modell ändert: Die Transparenz phänomenaler Repräsentationen ist kognitiv nicht penetrel, 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 sind Systeme, die erlebnismäßig nicht in der Lage sind, ihr eigenes subsymbolisches Selbstmodell *als* Selbstmodell zu erkennen. Aus diesem Grund operieren wir, phänomenologisch gesehen, unter den Bedingungen eines »naiv-realistischen Selbstmißverständnisses«: Wir erleben uns selbst, als wären wir in direktem und unmittelbarem epistemischen Kontakt zu uns selbst. Denn – das ist der begriffliche Punkt – ein phänomenal transparentes Selbstmodell erzeugt mit logischer Notwendigkeit das bewußte Erlebnis, *sich selbst unendlich nah zu sein*. Die Kernaussage der SMT ist, daß auf diese Weise ein basales »Ich-Gefühl«²⁵ entsteht, ein für das betreffende System unhintergebares phänomenales Selbst. Der Inhalt des nicht-begrifflichen Selbstbewußtseins ist der Inhalt eines transparenten PSM. Dies legt mich auch auf eine spezifische empirische Voraussage fest: Wenn das PSM seine Transparenz verlöre und opak würde, wenn der Organismus als Ganzer in der Lage wäre, sein momentanes Selbstmodell erlebnismäßig *als* ein Modell zu erkennen, dann würde die Eigenschaft der Selbstheit verschwinden, das »Ich-Gefühl« würde sich auflösen. In phänomenologischen Standardkonfigurationen aber ist die Entität, die auf das Buch in

²⁵ Auf begrifflicher und empirischer Ebene die konstitutiven bzw. (relativ zu einer spezifischen Klasse von Systemen) minimalen Bedingungen für die Entstehung eines phänomenalen Selbst zu isolieren, ist eines der wichtigsten Erkenntnisziele für die Selbstmodell-Theorie der Subjektivität. Dieses Forschungsprogramm ist mit dem Begriff der *minimal phenomenal selfhood* (MPS; Blanke & Metzinger, 2009) verknüpft. Ein guter Kandidat ist die phänomenal transparente Selbstlokalisierung in einem räumlichen und zeitlichen Bezugsrahmen, wobei Agentivität und das Vorhandensein eines expliziten Körpermodells keine notwendigen Bedingungen zu sein scheinen (vgl. Metzinger, 2013; Windt, 2010, 2014).

ihren Händen schaut, selbst eine Form von transparentem phänomenalem Inhalt. Und dies trifft ebenso auf die »Gerichtetheit«, die mit diesem Akt der visuellen Aufmerksamkeit einhergeht, zu – also auf die scheinbar direkte Beziehung, die Subjekt und Objekt zu verbinden scheint.

Der sechste Schritt: Das phänomenale Modell der Intentionalitätsrelation

Gehen wir noch einen Schritt weiter, bevor wir zum Schluß kommen. Das phänomenale Erleben des »Ein-Selbst-Seins« ist aufs Engste nicht nur mit dem Gefühl des Besitzens verbunden, sondern auch mit dem Erleben von Agentivität. Es geht nicht nur darum, ein transparentes Selbstmodell zu besitzen, sondern auch um die Repräsentation von Gerichtetheit, d.h. darum, mit Zielobjekten und Zielzuständen dynamisch verbunden zu sein. Dazu nun zwei weitere empirische Beispiele, und zwar aus den experimentellen Neurowissenschaften, die Makaken als Versuchsobjekte benutzen.

In der klassischen Neurologie wurde die Hypothese eines »Körperschemas« aufgestellt, einer unbewußten, aber kontinuierlich aktualisierten Karte der Körperform und -haltung im Gehirn.²⁶ Neuere Studien zeigen, wie japanische Makaken darauf trainiert werden können, Werkzeuge zu benutzen, obwohl sie in ihrem natürlichen Umfeld kaum Gebrauch von Werkzeugen machen (vgl. Maravita und Iriki, 2004). Während des erfolgreichen Gebrauchs von Werkzeugen finden Veränderungen in spezifischen neuronalen Netzen in ihren Gehirnen statt. Aus dieser Erkenntnis geht hervor, daß die Werkzeuge vorübergehend in ihr Körperschema integriert werden. Wenn ein Futtermügelchen außerhalb der Reichweite ihrer Hände ausgegeben wird und sie erfolgreich eine Harke benutzen, um es näher heranzuziehen, kann man eine Veränderung ihres körperlichen Selbstmodells im Gehirn beobachten. Es sieht tatsächlich so aus, als sei das bewußte Modell ihrer Hand bis zur Spitze des Werkzeugs erweitert worden. Eine genauere Art zu beschreiben, was hier geschieht, ist die folgende: Auf der Ebene des bewußten Realitätsmodells des Affen werden Eigenschaften der Hand auf die weiter entfernte Spitze des Werkzeugs übertragen. Wir kennen denselben Effekt bei Menschen: In unserem Fall kann wiederholte Übung die Spitze eines Werkzeuges zu einem Teil unserer eigenen Hand machen, zu einem Teil, der genauso »führend« und geschickt wie unsere eigenen Finger benutzt werden kann.

Mit anderen Worten: Neuere neurowissenschaftliche Daten stützen eindeutig die Ansicht, daß wir durch Werkzeuge nicht nur in der Lage sind, unsere Reichweite zu erweitern. Sie zeigen, daß jede erfolgreiche Erweiterung des Handlungsraums auch im neuronalen Träger des Körperbildes im Gehirn wiedergespiegelt wird. Das Gehirn konstruiert ein »internalisiertes« Bild des Werkzeugs, indem es rasch in das bereits existierende Körperbild als Ganzes integriert wird – es wird funktional »einverleibt«. Natürlich können wir nicht wirklich sagen, ob Affen tatsächlich bewußt das Gefühl des Besitzens erleben oder ob es sich hier nur um unbewußte Mechanismen handelt.

²⁶ Die Terminologie war nie ganz klar, aber sie unterscheidet oft zwischen einem unbewußten »Körperschema« und einem bewußten »Körperbild«. Eine philosophische Sicht auf die begriffliche Unklarheit, die beide Begriffe umgibt, findet sich in Gallagher (2005); einen guten Einstieg in die empirische Literatur bietet Maravita (2006). Eine ausgezeichnete neuere Überblicksdarstellung ist Iriki und Sakura (2008).

Wir wissen aber um mehrere Ähnlichkeiten zwischen Makaken und Menschen, die diese Annahme plausibel machen. Dies könnte auch einer der Anfänge der mentalen *Simulation* einer momentanen Gerichtetheit auf ein Zielobjekt oder einen Zielzustand sein. Und dieser Schritt führt uns zum zweiten Hauptaspekt der Selbstheit oder des »Ich-Gefühls«: Neben globalem *Besitzen* müssen wir *Agentivität* verstehen, das Phänomen der flexiblen globalen Handlungskontrolle.

Ein spannender Aspekt dieser neuen Daten ist, daß sie Aufschluß geben über die Evolution des Werkzeuggebrauchs. Eine notwendige Voraussetzung für die Erweiterung des eigenen Handlungsraums und der eigenen Befähigungen durch Werkzeuggebrauch scheint nämlich ganz eindeutig die Fähigkeit zu sein, Werkzeuge in ein bereits gegebenes Selbstmodell zu integrieren. Man kann nur dann zielgerichtet und intelligent Gebrauch von Werkzeugen machen, wenn das Gehirn sie zeitweise als Teil des eigenen Selbst repräsentiert. Intelligenter Werkzeuggebrauch war eine entscheidende Errungenschaft in der menschlichen Evolution. Man kann annehmen, daß eine Art elementarer funktionaler Baustein des menschlichen Werkzeuggebrauchs schon in den Gehirnen unserer Vorfahren existiert hat. Durch evolutionären Druck und noch nicht vollständig verstandene kulturelle Mechanismen hat sich dieser Grundbaustein dann schnell zu dem entwickelt, was wir heute bei Menschen beobachten.²⁷ Ich denke, einer der interessantesten Aspekte der neueren soziokulturellen Entwicklung besteht genau darin, daß Menschen sich mit Hilfe ihres neuartigen Selbstmodells in einem bis dahin ungekannten Ausmaß *gegenseitig* zu instrumentalisieren versuchen: Menschen versuchen nämlich sehr häufig, ihren eigenen Handlungsraum zu vergrößern, indem sie *andere* Menschen kontrollieren und als Werkzeuge einsetzen. Meine spekulative These ist, daß diese Tatsache die Evolution von Großgesellschaften auf entscheidende Weise angetrieben hat.

Es gibt heute ein schnell wachsendes Forschungsfeld, in dem Ingenieure und Neurowissenschaftler zusammenarbeiten: *Brain-Machine Interfaces* (ein kurzer Überblick findet sich in Lebedev & Nicolelis, 2006). Eine weitere Anwendung des oben skizzierten Grundgedankens besteht darin, künstliche Glieder oder Roboter direkt durch bestimmte Verbände von kortikalen Neuronen anzusteuern und zu kontrollieren, also eine Maschine in die Lage zu versetzen, motorische Befehle auszuführen, die in einem biologischen Gehirn erzeugt werden. Mittlerweile können Affen bereits über das Internet in Echtzeit die Schrittmuster von humanoiden Robotern auf der anderen Seite der Welt fernsteuern, von der *Duke University* in Amerika zum *Computational Brain Project* bei der *Japan Science and Technology Agency* in Japan, und zwar ausschließlich durch die Aufzeichnung ihrer Gehirnaktivität. Miguel Nicolelis schreibt:

Die verblüffendste Entdeckung war, daß der Affe, als wir das Laufband abschalteten und er aufhörte, seine Beine zu bewegen, in der Lage war, die Fortbewegung des Roboters für einige Minuten aufrechtzuerhalten – nur durch seine Gedanken und ausschließlich mit Hilfe des visuellen Feedbacks vom Roboter in Japan.

²⁷ Vgl. Iriki et al., 1996; Maravita and Iriki, 2004.

In unserem Zusammenhang ist die vielleicht interessanteste Beobachtung in diesem Experiment (zu den Details vgl. Carmena et al., 2003), daß der Affe allmählich seinen eigenen Arm vernachlässigt, obwohl er immer noch Teil seines biologischen Körpers ist. Das heißt, daß in seiner Selbstrepräsentation im Gehirn bestimmte Veränderungen abzulaufen scheinen, weil er jetzt versucht, das Feedback in einer neuen Art von motorischer Aufgabe und mit einem anderen Zielzustand zu kontrollieren. Dabei optimiert er ein neues Set von motorischen Parametern, indem er versucht, einen realen Roboterarm oder sogar einen virtuellen Arm zu kontrollieren, den er auf dem Bildschirm vor sich sieht. Die Abstimmung der Neuronen, ihre *tuning properties*, verändern sich also. Lebedev und Nicolelis (2006, S. 542) beschreiben diesen Effekt wie folgt:

Es ist bemerkenswert, daß die Gliedmaßen der Tiere, nachdem diese angefangen hatten den Aktuator durch ihre neuronalen Aktivitäten direkt zu kontrollieren, aufhörten, sich zu bewegen, während die Tiere weiterhin den Aktuator durch das Generieren geeigneter Modulationen ihrer kortikalen Neuronen kontrollierten. Die sparsamste Interpretation dieser Ergebnisse ist die, daß das Gehirn in der Lage war, eine graduelle Assimilation des Aktuators in die Karten, die den Körper repräsentierten, durchzuführen.

Aus der Perspektive der SMT ist die plausibelste Interpretation die, daß die Repräsentationen der alten Körperteile des Affen bestimmte funktionale Eigenschaften verlieren, sobald das Tier eine innere Repräsentation dieses neuen Aktuators erfolgreich in sein bewußtes Selbstmodell einbettet. Die alten Körperteile werden vorübergehend immer weniger zugänglich für die Aufmerksamkeitsverarbeitung und treten nach und nach aus dem bewußten Erleben zurück. Diese Beispiele führen uns zwei weitere wichtige Einsichten vor Augen. Zum einen ist es offensichtlich, daß das phänomenale Selbstmodell oft ein wichtiger Teil einer *Kontrollhierarchie* ist: Es ist ein abstraktes Werkzeug, ein Mittel zur Überwachung bestimmter kritischer Aspekte desjenigen Prozesses, in dem der Organismus flexible, adaptive Verhaltensmuster erzeugt. Zum anderen ist es hochgradig plastisch, weil mehrere Repräsentationen von Objekten *außerhalb* des Körpers vorübergehend in das Selbstmodell integriert werden können. Dies trifft nicht nur auf Gummihände zu, sondern ist umso mehr auf Werkzeuge im Allgemeinen anwendbar – Erweiterungen körperlicher Organe, die erfolgreich kontrolliert werden müssen, um intelligentes und zielgerichtetes Verhalten zu erzeugen. Das Selbstmodell ist das funktionale Fenster, durch das das Gehirn mit dem Körper *als einer Ganzheit* interagieren kann, und umgekehrt. Wenn der Körper durch Stöcke, Steine, Haken oder Roboterarme erweitert wird, muß auch das Selbstmodell erweitert werden. Erst wenn eine integrierte Repräsentation des Körpers-plus-Werkzeug existiert, kann das erweiterte System des Körpers-plus-Werkzeug ein Teil der Kontrollhierarchie des Gehirns werden. Anders gefragt: Wie sollte man überhaupt lernen, ein Werkzeug intelligent – also: auf flexible und kontextsensitive Weise – zu benutzen, *ohne* es in das bewußte Selbst zu integrieren? Das bewußte Selbstmodell ist ein virtuelles Organ, das uns erlaubt, Feedbackschleifen zu *besitzen* und Kontrollprozesse zu initiieren, aufrechtzuerhalten und flexibel anzupassen. Einige Elemente des Kontrollkreislaufs sind physisch (wie das Gehirn und Werkzeuge), andere sind virtuell (wie das Selbstmodell und die Zielzustandssimulation).

Natürlich ist es unter SMT denkbar, auch komplette Roboter oder virtuelle Körper (Avatare) in das PSM einzubetten und dadurch kausal zu kontrollieren. Eine empirische Vorhersage unter den begrifflichen Annahmen von SMT ist nämlich, daß es prinzipiell möglich sein muß, das menschliche Selbstmodell unter Umgehung des nicht-neuronalen, biologischen Körpers in einer kausal direkteren Weise an künstliche Effektoren und Sensoren zu koppeln und dadurch funktional in einer technologisch erzeugten Umwelt zu situieren. In einer ehrgeizigen Pilotstudie haben Ori Cohen, Doron Friedman und ihre Kollegen demonstriert, daß es mit funktionaler Echtzeit-Magnetresonanztomographie möglich ist, die motorischen Intentionen einer Versuchsperson auszulesen und direkt als hochstufige Motorbefehle an einen humanoïden Roboter zu übertragen, der sie in körperliche Handlungen transformiert, während die Versuchsperson gleichzeitig das gesamte Experiment durch die Augen des Roboters visuell miterleben kann (Cohen et al., 2012). Der Vorgang basiert auf bildhaften Bewegungsvorstellungen, die es den Versuchspersonen erlauben, »direkt mit Ihrem PSM zu handeln«,²⁸ indem sie aus einem Scanner in Israel einen humanoïden HO-AP3-Roboter in Frankreich kontrollieren.

Oben habe ich bereits drauf hingewiesen, daß Menschen (und auch manche andere Tiere) oft das Verhalten oder die Befindlichkeit einer *anderen* Person kontrollieren wollen. Wir »instrumentalisieren« und »beschlagnahmen« einander, manchmal machen wir uns sogar zu »Leibeigenen«. Menschen erweitern sich ständig – nicht nur mit Stöcken, Steinen, Harken oder Roboterarmen, sondern auch mit den Gehirnen und Körpern *anderer* Menschen (Metzinger & Gallese, 2003). Der Übergang von biologischer zu kultureller Evolution ist ohne Frage unmittelbar mit der Realisierung neuer und spezifischer funktionaler Eigenschaften durch das PSM der Primaten verbunden. Dies ist eine der interessantesten Fragen für die Zukunft: Was genau war die Veränderung im PSM des *Homo sapiens*, gegenüber dem PSM des Schimpansen, die zur Explosion von Kultur und der Entwicklung komplexer Gesellschaften führte? Meine zweite spekulative Arbeitshypothese wäre hier die, daß es nicht allein komplexer Werkzeuggebrauch in der sozialen Dimension *per se* war, sondern auch die Fähigkeit, einen viel größeren

²⁸ An andere Stelle habe ich die ethischen Implikationen dieser historisch neuartigen Handlungsformen untersucht und den Begriff einer »PSM-action« eingeführt, einer Handlung, bei der ein menschlicher Akteur Teile seines Selbstmodells zu *offline*-Simulationen benutzt, die dann unter Umgehung des proximalen, nicht-neuronalen *embodiments* zu Konsequenzen in der Welt außerhalb des biologischen Körpers führt. Es ist aus verschiedenen Gründen denkbar, daß die Handlungsautonomie – und damit auch die ethische Verantwortlichkeit – des Akteurs im Sinne von Impulskontrolle und Abbruchmöglichkeiten in solchen Situationen eingeschränkt ist. Vgl. Metzinger, 2013a.



VISUAL FEEDBACK

MOTOR IMAGERY

MOTION
DATABASE



Abb. 8: Direktes Handeln mit dem PSM durch *robotic re-embodiment*: Das Ziel des Experiments bestand darin, es einer Versuchsperson in Israel durch »direkte Gedankenkontrolle« über das Internet zu ermöglichen, einen Roboter in Frankreich zu kontrollieren. Eine Video-Demonstration findet sich unter <http://www.youtube.com/user/TheAVL2011>. Abbildung mit freundlicher Genehmigung von Doron Friedman.

Teil der Kontrollhierarchie *offline* laufen zu lassen, sie in Simulationen zu benutzen und gleichzeitig ein opakes (d.h. ein nicht-transparentes) PSM zu erzeugen.²⁹ Es war die Fähigkeit, sich selbst bewußt *als* repräsentierend, *als* auf einen Zielzustand gerichtet zu repräsentieren. Es war der Unterschied zwischen dem bloßem Besitzen einer Erster-Person-Perspektive und der mentalen Fähigkeit, eben diese Tatsache noch einmal explizit zu repräsentieren.

²⁹ Die SMT sagt, daß der Inhalt eines transparenten Zustandes als gegeben erlebt wird, der eines opaken aber als selbst konstruiert (Metzinger, 2003a, c; 2009). Ein interessanter Gedanke von Jennifer Windt beruht auf der Beobachtung, daß die frühen Werkzeuge unserer Vorfahren nicht nur benutzt, sondern – vor der Einbettung ins Selbstmodell – selbst erst hergestellt werden mußten. Affen dagegen stellen ihre eigenen Werkzeuge kaum oder nur sehr selten her. Die ersten Werkzeuge unserer Vorfahren wurden anscheinend in wenigen Schritten (4 bis 5 Steinschläge) hergestellt, komplexere Werkzeuge, die man dann für verschiedene Tätigkeiten nutzen und auch über längere Zeit behalten konnte (eine Voraussetzung für die geographische Ausbreitung), haben dann mehrere durchdachte Arbeitsschritte (und verschiedene andere Werkzeuge) erfordert – also wahrscheinlich auch mentale *offline*-Simulationen. *Trial and error* hätte wohl kaum zu dieser Komplexität und auch zu der Effizienz der Herstellung geführt, außerdem wurde bei der Benutzung dann tatsächlich etwas ins Selbstmodell eingebettet, das vorher durch eigenes körperliches Handeln und ganz konkret und sinnlich wahrnehmbar *selbst* konstruiert worden war.

Werfen wir deshalb einen Blick auf die repräsentationale Architektur, die dem subjektiven Erleben von Gerichtetheit im Allgemeinen zugrunde liegt. Phänomenologisch gesehen entsteht aus einem transparenten Weltmodell eine Wirklichkeit. Aus einem transparenten Modell des Systems entsteht ein Selbst, das in diese Wirklichkeit integriert ist. Wenn nun noch ein transparentes Modell der vorübergehenden und sich kontinuierlich verändernden Beziehungen zwischen dem wahrnehmenden und handelnden Selbst und den Objekten und Personen in dieser Wirklichkeit entsteht, resultiert dies in dem, was ich oben eine »phänomenale Erste-Person-Perspektive« genannt habe. Eine genuine Innenperspektive entsteht genau dann, wenn sich das System *für sich selbst* als momentan mit der Welt interagierend darstellt und wenn es diese Darstellung wiederum nicht als solche erkennt. Es besitzt dann ein bewußtes Modell der Intentionalitätsrelation.³⁰ Es stellt sich als auf bestimmte Aspekte der Welt gerichtet dar. Sein phänomenaler Raum ist ein *perspektivischer* Raum und seine Erlebnisse sind jetzt *subjektive* Erlebnisse.

Die Intentionalitätsrelation ist in der Hauptsache eine Wissensbeziehung zwischen Subjekt und Objekt: Ein mentaler Zustand wird dadurch zu einem Träger von Wissen, daß er über sich selbst hinaus verweist – wie ein Pfeil, der aus dem Geist eines Menschen auf ein Objekt in der realen oder sogar nur in einer möglichen Welt zeigt. Philosophen sagen, daß diese Art von mentalem Zustand einen *intentionalen Inhalt* besitzt. Der Inhalt ist das, worauf der Pfeil zeigt. Dieser Inhalt kann ein Bild, eine Aussage oder sogar ein Handlungsziel sein. Wie Philosophen sagen, gibt es »praktische Intentionalität« in dem Sinne, daß man auf bestimmte »Erfüllungsbedingungen« (wie z.B. ein Handlungsziel) gerichtet ist, und es gibt eine »theoretische Intentionalität« in dem Sinne, daß man auf »Wahrheitsbedingungen« (z.B. eines Satzes, also ein Erkenntnisziel) gerichtet ist. Wenn viele solcher Pfeile im Bewußtsein verfügbar sind, d.h. vom Gehirn auf der funktionalen Ebene der globalen Verfügbarkeit dargestellt werden, dann entsteht eine zeitlich ausgedehnte Erste-Person-Perspektive. Kurz gesagt: Es ist eine Sache, ein biologischer Organismus zu sein, der die Welt darstellt, es ist aber etwas ganz Anderes, sich selbst auf der Ebene des bewußten Erlebens als *darstellend* darzustellen, und zwar in Echtzeit und während der dynamische Vorgang der Repräsentation sich tatsächlich ereignet. Die SMT beschäftigt sich mit Letzterem. Es gibt dann nicht mehr nur ein neuronal verankertes Kernselbst, eine Selbstrepräsentation, sondern auch eine dynamische, phänomenale Simulation des *Selbst als eines Subjekts*, das durch ständig wechselnde Wissens- und Handlungsbeziehungen in die Welt eingebettet ist. Natürlich wäre zum zentralen Konzept eines »phänomenalen Modells der Intentionalitätsrelation« (PMIR) sehr viel mehr zu sagen.³¹ Der Kerngedanke ist aber der folgende: Der bewußte Mensch ist ein System, das bei einzelnen repräsentationalen Akten die Repräsentationsbeziehung selbst noch einmal dynamisch *ko-repräsentieren* kann. Das Instrument, das es für diesen Zweck benutzt, ist das PMIR. Das PMIR ist lediglich ein weiteres virtuelles Organ, genau wie das PSM. Der Inhalt höherstufiger

³⁰ Vgl. dazu insbesondere den zweiten Teil von Metzinger 2006c.

³¹ Die PMIR-Theorie ist natürlich komplexer, als ich es in diesem kurzen Überblicksartikel darstellen kann. Neben Metzinger (2003a) empfehle ich Lesern, die sich für die Idee interessieren, Metzinger (2006c). Eine detailliertere Besprechung, die sich besonders auf die repräsentationale Struktur bewußter Willensakte bezieht, ist zu finden in Metzinger (2006a).

Formen des Selbstbewußtseins ist nämlich immer eine Relation: Das Selbst *im Moment des Erkennens* (Damasio, 1999, S. 168ff), das Selbst *im Akt des Handelns*. Ein Subjekt zu sein bedeutet, die Fähigkeit zu besitzen, diese intentionale Beziehung selbst zu korrepräsentieren und sie gleichzeitig durch Interaktion mit einer Welt aktiv zu konstruieren.

Natürlich ist die Art und Weise, in der wir diese Subjekt-Objekt-Beziehungen subjektiv erleben, eine vereinfachte Version der eigentlichen Prozesse – sie ist gewissermaßen eine funktional adäquate Konfabulation. Die Evolution hat auch in diesem Fall wieder eine einfache, aber elegante Lösung favorisiert. Das virtuelle Selbst, das sich in der phänomenalen Welt bewegt, besitzt kein Gehirn, kein Motorsystem und keine Sinnesorgane: Bestimmte Teile der Umgebung erscheinen direkt in seinem Geist, der Wahrnehmungsprozeß wird als anstrengungslos und unmittelbar erlebt. 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. Was oben als »Transparenz« definiert wurde, ist eine Art, die *Geschlossenheit* dieser multimodalen, hochdimensionalen Benutzeroberfläche des Gehirns zu beschreiben. Das phänomenale Selbst ist der Teil dieser Oberfläche, den das System benutzt, um *sich selbst* als ein Ganzes zu repräsentieren und sein eigenes Verhalten vorherzusagen, um sich für sich selbst als ein erkennendes »Ich« darzustellen und 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 oder einen motorischen Kortex besitzt. Das PSM ist das interne Interface, welches das System benutzt, um sich seine eigene Hardware funktional anzueignen, um seine eigenen niedrigstufigen Dynamiken zu kontrollieren, um zeitliche entfernte Zielzustände als *seine eigenen* darzustellen und um *autonom* zu werden. Die intentionalen Pfeile, die diesen Agenten mit Objekten und anderen Selbsten innerhalb des gerade aktiven Wirklichkeitsmodells verbinden, sind phänomenale Repräsentationen von vorübergehend auftretenden Subjekt-Objekt-Beziehungen – und oft können auch sie nicht als Repräsentationsprozesse erkannt werden. In Standardsituationen ist die bewußt erlebte Erste-Person-Perspektive der Inhalt eines transparenten PMIR.

All dies spielt sich innerhalb eines phänomenalen Gegenwartsfensters ab. Die Inhalte des phänomenalen Erlebens erzeugen nicht nur eine Welt, sondern auch eine *Gegenwart* (vgl. Metzinger, 2003a, Abschnitt 3.2.2). Wenn man so will, ist der Kern des phänomenalen Bewußtseins die Erzeugung einer Gegenwartsinsel im physikalischen Fluß der Zeit (vgl. Ruhnau, 1995 und die dort gegebenen Literaturhinweise, insbesondere auf die Arbeiten von Ernst Pöppel). Minimales Selbstbewußtsein ist Selbstlokalisierung innerhalb eines räumlichen und eines zeitlichen Bezugsrahmens (Blanke & Metzinger, 2009): Subjektives Erleben heißt »hier sein« (Immersion in einer Szene) und immer auch »jetzt sein« (Präsenz). Bewußtsein ist Gegenwärtigsein. Es bedeutet, Informationen in einer sehr speziellen Weise zu verarbeiten. Diese besteht darin, bereits repräsentierte Einzelereignisse immer wieder und kontinuierlich zu zeitlichen Gestalten zusammenzufassen, zu einem bewußt erlebten Moment. Viele empirische Daten zeigen heute, daß die bewußt erlebte Gegenwart in einem bestimmten Sinne eine *erinnerte* Gegenwart ist. So gesehen ist sogar das phänomenale »Jetzt« ein repräsentationales Kon-

strukt, eine *virtuelle* Gegenwart. An diesem Punkt wird erstmals deutlich, was es bedeutet, wenn man sagt, daß der phänomenale Raum ein virtueller Raum ist: Sein Inhalt ist eine *mögliche* Realität, eine Vorhersage.³² Der Realismus des phänomenalen Erlebens entsteht dadurch, daß in ihm eine Möglichkeit – die beste Hypothese, die es im Moment gibt – unhintergebar als Wirklichkeit oder als eine *Aktualität* dargestellt wird. Mit anderen Worten: Die Mechanismen, die zeitliches Erleben und unser subjektives Gegenwertsgefühl herstellen, sind ebenfalls transparent. Auch diesen Punkt muß man nun am Ende auf den Sonderfall der phänomenalen Selbstmodellierung anwenden: Weil die Virtualität des Selbstmodells *und* die Virtualität des Gegenwartsfensters nicht auf der Ebene des subjektiven Erlebens selbst verfügbar sind, wird das in ihnen dargestellte System zu einem *jetzt anwesenden Subjekt*.

SMT löst das Homunkulus-Problem, weil wir jetzt sehen können, daß kein »kleiner Mann im Kopf« gebraucht wird, um den Inhalt mentaler Repräsentationen zu interpretieren und »auszulesen«. Gleichzeitig ist sie ein höchst sparsames Erklärungsmodell, da sie es uns erlaubt, die Entstehung von Selbstbewußtsein zu erklären, ohne anzunehmen, daß ein substantielles Selbst existiert und mit dem PSM gleichzeitig einen für zukünftige Entwicklungen offenen Nachfolgebegriff anbietet.³³ Soll all dies bedeuten, daß das Selbst – ontologisch gesehen - nur eine Illusion ist? Auf den zweiten Blick enthalten der allzu populäre Begriff der »Ich-Illusion« und die Metapher des »Sich-mit-seinem-eigenen-inneren-Bild-von-sich-selbst-Verwechselns« einen logischen Fehler: *Wessen* Illusion sollte dies sein? Von einer Illusion zu sprechen setzt die Annahme voraus, daß es jemanden gibt, der sie *hat*. Etwas, das kein epistemisches Subjekt in einem starken Sinne begrifflich-propositionalen Wissens ist, *kann* sich überhaupt nicht mit irgendetwas anderem verwechseln. Wahrheit und Falschheit, Realität und Illusion gibt es für ein biologisches Informationsverarbeitungssystem auf der fraglichen Entwicklungsstufe noch gar nicht. Bisher haben wir ja nur eine kognitionswissenschaftlich fundierte Theorie der Phänomenologie der Selbstheit, des »Ich-Gefühls«, aber noch keine Theorie der Selbsterkenntnis. Ich habe hier lediglich kurz ausgeführt, wie eine *phänomenale* Erste-Person-Perspektive das Produkt der natürlichen Evolution sein könnte. Subjektivität im *epistemischen* Sinne, eine epistemische Erste-Person-Perspektive, wäre ein weiterer Schritt. Natürlich ist die Phänomenologie des »Ich-Gefühls«, des nicht-begrifflichen Selbstbewußtseins, die wichtigste Voraussetzung für diesen Schritt, weil es die zentrale notwendige Bedingung für die Konstitution eines genuin *reflexiven*, begrifflichen Selbstbewußtseins ist. In gewisser Hinsicht besteht darin der ganze Sinn der Theorie: Wenn wir die höheren Formen von Subjektivität und Intersubjektivität ernst nehmen wollen, müssen wir zunächst bescheiden und vorsichtig sein, uns auf die fundierenden Ursprünge auf der Ebene nicht-begrifflicher Inhalte und der selbstorganisierenden neuronalen Dynamik im Gehirn konzentrieren. Meine Leser werden nicht überrascht sein, wenn der Autor dieses Aufsatzes der Überzeugung ist, daß subjektives *Wissen* am Ende dann eben genau das Wissen ist, das mit ei-

³² Meine eigenen Vorstellungen ähneln in diesem Punkt denen von Antti Revonsuo: *Virtuelle Realität* ist die beste technische Metapher für phänomenales Bewußtsein, die wir gegenwärtig besitzen. Vgl. Revonsuo (1995, 2000), und besonders Revonsuo (2006).

³³ Zu den verschiedenen Weisen, auf die man ein Anti-Realist bezüglich »des« Selbst sein kann vgl. Metzinger, 2011.

nem spezifischen inneren Modus der Präsentation zusammenhängt, nämlich als Wissen *unter einem phänomenalen Modell der Intentionalitätsrelation*. Was wir heute die »Erste-Person-Perspektive« nennen, ist eine hochspezifische, neuronal realisierte Weise des Gegebenseins, ein abstraktes Datenformat, welches im Laufe der biologischen Evolution schrittweise entstanden ist. Subjektivität im erkenntnistheoretischen Sinne kann deshalb zumindest prinzipiell ebenfalls naturalisiert werden. Dies ist aber nur dann möglich, wenn wir eine überzeugende evolutionäre und neurowissenschaftliche Erklärung dafür finden, wie sich genau diese repräsentationale Struktur, der dynamische, indexikalische innere Modus des Gegebenseins, den wir heute »die subjektive Innenperspektive« nennen, in einem selbstorganisierenden physikalischen Universum überhaupt erst entwickeln konnte. Worum es geht, ist, daß ein natürlich evolviertes System beginnt, sich intern selbst *als epistemischen Agenten* zu modellieren, und dabei gleichzeitig mit Hilfe des PSM und des PMIR eine neue Modalität des Erkennens erzeugt, eine neue Form der Darstellung von Wissen über die Welt. Letztlich ist es offensichtlich, daß jedes einzelne Vorkommnis eines PSM bzw. PMIR identisch ist mit einem spezifischen Zeitfenster in der kontinuierlichen, dynamischen Selbstorganisation kohärenter Aktivität, die lokal in einem einzelnen biologischen Gehirn stattfindet. In diesem kontinuierlich fortlaufenden Prozeß auf der subpersonalen Ebene gibt es aber keinen kleinen Agenten und auch keinen bösen Dämon, der als *Urheber* einer Illusion angesehen werden könnte. Und es gibt auch keine starre subpersonale Entität, die als das *Subjekt* der Illusion angesehen werden könnte. Es gibt niemanden *im* System, der sich täuschen oder mit etwas verwechseln könnte, der Homunkulus existiert nicht. Selbstbewußtsein ist eine Eigenschaft des Systems als Ganzem.

Literatur:

- Ainley, V., Tajadura-Jiménez, A., Fotopoulou, A. & Tsakiris, M. (2012). Looking into myself: Changes in interoceptive sensitivity during mirror self-observation. *Psychophysiology* 49(11): 1672-6.
- Alvarado, C.S. (2000). Out-of-Body Experiences, in E. Cardena et al. (eds.), *Varieties of Anomalous Experience: Examining the Scientific Evidence*. Washington, DC: American Psychological Association.
- Armel, 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.
- Baker, L.R. (2008). Naturalism and the first-person perspective. In G. Gasser (ed.), *How successful is Naturalism? Publications of the Austrian Ludwig Wittgenstein Society*. Frankfurt am Main: Ontos Verlag.

- Bennett, M.R. & Hacker, P.S.M. (2010). *Die philosophischen Grundlagen der Neurowissenschaften*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Bermúdez, J.L., Marcel, A. & Eilan, N. (1995)[eds]. *The Body and the Self*. Cambridge, MA: MIT Press.
- Bieri, P. (1987). Evolution, Erkenntnis und Kognition. In W. Lütterfelds (Hrsg), *Transzendente oder Evolutionäre Erkenntnistheorie?* Darmstadt: Wissenschaftliche Buchgesellschaft.
- Bischof-Köhler, D. (1989, 21993). *Spiegelbild und Empathie*. Bern: Huber.
- Bischof-Köhler, D. (1996). Ichbewußtsein und Zeitvergegenwärtigung. Zur Phylogenese spezifisch menschlicher Erkenntnisformen. In A. Barkhaus, M. Mayer, N. Roughley & D. Thürnauf (Hrsg.), *Identität, Leiblichkeit, Normativität. Neue Horizonte anthropologischen Denkens*. Frankfurt am Main: Suhrkamp.
- Blanke, O. (2012). Multisensory mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience* 13(8): 556-571.
- Bongard, J., Zykov, V. & Lipson, H. (2006). Resilient machines through continuous self-modeling. *Science* 314: 1118.
- Botvinick, M. & Cohen, J. (1998). Rubber hands »feel« touch that eyes see. *Nature* 391: 756.
- Botvinick, M. (2004). Probing the neural basis of body ownership. *Science* 305: 782-3.
- Brugger, P., Regard, M., and Landis, T. (1997). Illusory reduplication of one's own body: phenomenology and classification of autoscopic phenomena. *Cognitive Neuropsychiatry* 2: 19-38.
- Brugger, P., Kollias, S.K., Müri, R.M., Crelier, G., Hepp-Reymond, M.-C., and Regard, M. (2000). Beyond re-membering: Phantoms sensations of congenitally absent limbs. *Proceedings of the National Academy of Science USA* 97: 6167-72.
- Brugger, P., Regard, M., and Shiffar, M. (2001). Hand movement observation in a person born without hands: Is body scheme innate? *Journal of Neurology, Neurosurgery, and Psychiatry* 70: 276.
- Carmena, J. M., Lebedev, M. A., Crist, R. E., O'Doherty, J. E., Santucci, D. M., Dimitrov, D. F., Patil, P. G., Henriquez, C. S., and Nicolelis, M. A. L. (2003). Learning to Control a Brain-Machine Interface for Reaching and Grasping by Primates. *PLoS Biology* 1:193-208.
- Churchland, P.M. (1989). *A Neurocomputational Perspective*. Cambridge, MA / London: MIT Press.
- Clark, A. (1989). *Microcognition - Philosophy, Cognitive Science, and Parallel Distributed Processing*. Cambridge, MA: MIT Press.
- Cohen, O., Druon, S., Lengagne, S., Mendelsohn, A., Malach, R., Kheddar, A. &Friedman, D. (2012). MRI-based robotic embodiment: A pilot study. *IEEE International Conference on Biomedical Robotics and Biomechatronics*, June 24-27, Roma, Italy.
- Cummins, R. (1983). *The Nature of Psychological Explanation*. Cambridge, MA: 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. (1987). *The Intentional Stance*. Cambridge, MA und London: MIT Press.
- Devinsky, O., Feldmann, E., Burrowes, K., and Bromfield, E. (1989). Autoscopic phenomena with seizures. *Archives of Neurology* 46: 1080-8.

- Dittrich, A. (1985). *Ätiologie-unabhängige Strukturen veränderter Wachbewusstseinszustände*. Stuttgart: Enke.
- Dittrich, A. (1996). *Ätiologie-unabhängige Strukturen veränderter Wachbewusstseinszustände. Ergebnisse empirischer Untersuchungen über Halluzinogene I. und II. Ordnung, sensorische Deprivation, hypnagogie Zustände, hypnotische Verfahren sowie Reizüberflutung*. Berlin: VWB.
- Dittrich, A., Lamparter, D., Maurer & M. (2006). *5D-ABZ: Fragebogen zur Erfassung Außergewöhnlicher Bewusstseinszustände*. Zürich: PSIN PLUS.
- 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.
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience* 11: 127-38.
- Gallagher, S. (2005). *How the Body shapes the Mind*. Oxford: Oxford University Press.
- Gallese, V. & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2, 493-501.
- Gallese, V. (2005). Embodied simulation: from neurons to phenomenal experience. *Phenomenology and the Cognitive Sciences* 4: 23-38.
- Gallup, G.G., Anderson, J.R., Platek, S.M. (2011). Self-recognition. In S. Gallagher (ed.), *The Oxford Handbook of the Self*. Oxford: Oxford University Press.
- Gallup, G.G., Jr. (1997). On the rise and fall of self-conception in primates. In J.G. Snodgrass and R.L. Thompson (eds.), *The Self across Psychology – Self-Recognition, Self-Awareness, and the Self Concept*. New York, NY: *Annals of the New York Academy of Sciences* 818: 4-17.
- Grush, R. (1997). The architecture of representation. *Philosophical Psychology* 10: 5-25.
- Grush, R. (1998). Wahrnehmung, Vorstellung, und die sensomotorische Schleife. In H.-D. Heckmann F. & Esken (Hrsg.), *Bewusstsein und Repräsentation*. Paderborn: mentis.
- Guimera, M.J., et al. (2011). Body integrity identity disorder: deranged body processing, right fronto-parietal dysfunction, and phenomenological experience of body incongruity. *Neuropsychological Review* 21: 320–333.
- Hilti, L. M., Hanggi, J., Vitacco, D. A., Kraemer, B., Palla, A., Luechinger, R., ... (2012). The desire for healthy limb amputation: structural brain correlates and clinical features of xenomelia. *Brain* 136(1): 318-329.
- Howhy, J. (2013). *Mind in Prediction*. Oxford: Oxford University Press.
- Iriki, A., Tanaka, M., and Iwamura, Y. (1996). Coding of modified body schema during tool-use by macaque post-central neurons. *Neuroreport* 7: 2325-2330.
- Iriki, A. & Sakura, O. (2008). The neuroscience of primate intellectual evolution: natural selection and passive and intentional niche construction, *Phil. Trans. R. Soc. B* 363: 2229–2241.
- Jaspers, K. (1946). *Allgemeine Psychopathologie*. Berlin und Heidelberg.
- Kinsbourne, M. (1995). Awareness of one's own body: An attentional theory of its nature, development, and brain basis. In Bermúdez et al. 1995.
- Lenggenhager, B., Tadi, T., Metzinger, T. & Blanke, O. (2007). Video Ergo Sum: Manipulating bodily self-consciousness. *Science* 317: 1096-1099.

- Lycan, W.G. (1996). *Consciousness and Experience*. Cambridge, MA: MIT Press.
- Maravita, A., and Iriki, A. (2004). Tools for the body (schema). *Trends in Cognitive Sciences* 8: 79-86.
- Maravita, A. (2006). From »body in the brain« to »body in space«: Sensory and intentional components of body representation. In G. Knoblich, I. Thornton, M. Grosjean & M. Shiffrar (eds.), *Body Perception from the Inside Out*. New York: Oxford University 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.
- Melzack R. (2005). Evolution of the neuromatrix theory of pain. The Privithi Raj Lecture: presented at the Third World Congress of World Institute of Pain, Barcelona 2004. *Pain Pract* 5: 85-94.
- 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)[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 Neuropsychologie*. Paderborn: mentis.
- Metzinger, T. (2000a). 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. (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) (ed.). *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. Cambridge, MA: MIT Press.
- Metzinger, T. (2003). Phänomenale Transparenz und kognitive Selbstbezugnahme. In U. Haas-Spohn (Hrsg.), *Intentionalität zwischen Subjektivität und Weltbezug*. Paderborn: mentis.
- 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 & T. Metzinger (eds.), *Self and Action*. Special issue of *Consciousness & Cognition* (12:4), December 2003, 549-571.
- 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). Phenomenal transparency and cognitive self-reference. *Phenomenology and the Cognitive Sciences* 2: 353-393.
- Metzinger, T. (2004a). Why are identity-disorders interesting for philosophers? In T. Schramme und J. Thome (eds.), *Philosophy and Psychiatry*. Berlin: de Gruyter.
- Metzinger, T. (2004b). Appearance is not knowledge: The incoherent strawman, content-content confusions and mindless conscious subjects. Invited commentary for Alva Noë und Evan Thompson: »Are there neural correlates of consciousness?«, in a special issue of *Journal of Consciousness Studies* 11(1): 67-71.
- Metzinger, T. (2005b). [Out-of-body experiences as the origin of the concept of a »soul«](#). *Mind and Matter* 3(1): 57-84.
- Metzinger, T. (2005c). Die Selbstmodell-Theorie der Subjektivität: Eine Kurzdarstellung in sechs Schritten. In C. S. Herrmann, M. Pauen, J. W. Rieger & S. Schicktanz (Hrsg.), *Bewusstseins: Philosophie, Neurowissenschaften, Ethik*. Stuttgart: UTB/Fink .
- Metzinger, T. (2006a). Conscious volition and mental representation: Towards a more fine-grained analysis. In N. Sebanz und W. Prinz (ed.), *Disorders of Volition*. Cambridge, MA: MIT Press.
- Metzinger, T. (2006b). [Reply to Gallagher: Different conceptions of embodiment](#). In *PSYCHE – An Interdisciplinary Journal of Research on Consciousness*, 12 (4).
- Metzinger, T. (2006c). *Being No One – Eine sehr kurze deutsche Zusammenfassung*. In T. Metzinger, *Grundkurs Philosophie des Geistes, Band 1: Phänomenales Bewusstsein*. S. 424-475.
- Metzinger, T. (2008). Empirical perspectives from the self-model theory of subjectivity: A brief summary with examples. In R. Banerjee & B.K. Chakrabarti (eds.), *Progress in Brain Research* 168: 215-246. Amsterdam: Elsevier.
- Metzinger, T. (2009). *Der Ego Tunnel. Eine neue Philosophie des Selbst: Von der Hirnforschung zur Bewusstseinsethik*. Berlin: Bloomsbury Verlag.
- Metzinger, T. (2011). *The No-Self-Alternative* (Chapter 11). In S. Gallagher (ed.), *The Oxford Handbook of the Self*. Oxford, UK: Oxford University Press. S. 279-296.
- Metzinger, T. (2013a). Two principles for robot ethics. In E. Hilgendorf & J.-P. Günther (Hrsg.), *Robotik und Gesetzgebung*. Baden-Baden: Nomos.
- Metzinger, T. (2013b). First-order embodiment, second-order embodiment, third-order embodiment: From spatiotemporal self-location to minimal phenomenal selfhood. In Lawrence Shapiro (ed.), *The Routledge Handbook of Embodied Cognition*. London: Routledge.
- Michal, M. & Beutel, M.E. (2009). Weiterbildung CME: Depersonalisation/Derealisation – Krankheitsbild, Diagnostik und Therapie. *Z Psychosom Med Psychother* 55: 113-140.
- 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.

- Noreika, V., Windt, J.M., Lenggenhager, B., & Karim, A.A. (2010). New perspectives for the study of lucid dreaming: from brain stimulation to philosophical theories of self-consciousness. Commentary on »The neurobiology of consciousness: Lucid dreaming wakes up« by J. Allan Hobson. *International Journal of Dream Research* 3(1) 36-46.
- 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.
- Ratcliffe, M. (2008). *Feelings of Being: Phenomenology, Psychiatry and the Sense of Reality*. Oxford & New York: Oxford University Press.
- 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 T. Metzinger (ed.), *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. Cambridge, MA: MIT Press.
- Revonsuo, A. (2006). *Inner Presence*. Cambridge, MA: MIT Press.
- Ruhnau, E. (1995). Time-Gestalt and the observer. In Metzinger 1995.
- Schilling, M. & Cruse, H. (2012). What's next: recruitment of a grounded predictive body model for planning a robot's actions. *Frontiers in Psychology* 3: 383. doi: 10.3389/fpsyg.2012.00383
- Seth, A.K., Suzuki, K. & Critchley, H.D. (2012). An interoceptive predictive coding model of conscious presence. *Frontiers in Psychology* 2(395). doi: 10.3389/fpsyg.2011.00395
- Simeon, D., & Abugel, J. (2006). *Feeling Unreal: Depersonalization Disorder and the Loss of the Self*. Oxford/New York: Oxford University Press.
- Steels, L., & Hild, M. (2012). *Language Grounding in Robots*. Boston, MA: Springer US.
- Studerus, E., Gamma, A. & Vollenweider, F.X. (2010). Psychometric evaluation of the altered states of consciousness rating scale (OAV). *PLoS ONE* 5(8): e12412.
- Trivers, R. (2012). *The Folly of Fools: The Logic of Deceit and Self-Deception in Human Life*. New York: Basic Books. Deutsch: *Betrug und Selbstbetrug: Wie wir uns selbst und andere erfolgreich belügen* (2013; Berlin: Ullstein).
- Tsakiris, M. (2010). My body in the brain: a neurocognitive model of body-ownership. *Neuropsychologia* 48(3): 703-12.
- Tsakiris, M. (2011). The sense of body ownership. In S. Gallagher (ed.), *The Oxford Handbook of the Self*. Oxford: Oxford University Press.
- Tsakiris, M., Jiménez, A.T. & Costantini, M. (2011). Just a heartbeat away from one's body: interoceptive sensitivity predicts malleability of body-representations. *Proceedings of the Royal Society B: Biological Sciences* 278(1717): 2470-2476.
- Tsakiris, M., Longo, M.R. & Haggard, P. (2010). Having a body versus moving your body: neural signatures of agency and body-ownership. *Neuropsychologia* 48(9): 2740-9.
- Vallar, G. & Ronchi, R. (2009). Somatoparaphrenia: a body delusion. A review of the neuropsychological literature. *Experimental Brain Research* 192(3): 533-551.
- Von Hippel, B., Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences* 34: 1-56.

- Vuilleumier, P., Despland, P.A., Assal, G., and Regli, F. (1997). Héautoscopie, exta-se et hallucinations expérientielles d'origine épileptique. *Revue Neurologique* 153: 115–119.
- Waelti, E. (1983) *Der dritte Kreis des Wissens*. Interlaken: Ansata.
- Windt, J.M. (2010). The immersive spatiotemporal hallucination model of dreaming. *Phenomenology and the Cognitive Sciences* 9: 295-316.
- Windt, J.M. & Metzinger, T. (2007). The philosophy of dreaming and self-consciousness: What happens to the experiential subject during the dream state? In D. Barrett & P. McNamara (eds.), *The New Science of Dreaming*. Estport, CT: Praeger Imprint/Greenwood Publishers.
- Windt, J.M. (2014). *Dreaming*. Cambridge, MA: MIT Press.
- Yates, J. (1985). The content of awareness is a model of the world. *Psychological Review* 92: 249-84.