

|   |   |
|---|---|
| e-Journal<br>Philosophie der<br>Psychologie | <b>PROBLEME DER INTROSPEKTION AN DER SCHNITTSTELLE<br/>ZWISCHEN ANALYTISCHER PHILOSOPHIE UND NEUROPHILOSOPHIE<br/>von Oliver Grimm (Mannheim)</b> |
|---|---|

### **Einleitung: Introspektion an der Schnittstelle von Neurowissenschaft und Philosophie**

Weshalb ist eine Betrachtung des Begriffs der "Introspektion" in den Neurowissenschaften interessant? Noch zu Anfang der Neuzeit wurde die Introspektion als selbstverständliches Werkzeug der Philosophie gesehen. So waren etwa Locke und Hume neben einer begrifflichen Fassung und Definition der Introspektion auch an einer psychologischen Theorie derselben interessiert. Das Konzept der Introspektion nahm innerhalb der Psychologie und später auch der Neurowissenschaften eine interessante Wende. Während der Begründer der modernen Psychologie, der (Proto-)Hirnforscher und Philosoph William James (James 1907), die Introspektion noch als selbstverständliche Fähigkeit des Geistes und sinnvolle Methode sah, geriet sie durch den Behaviorismus in Verruf. Erst mit dem Aufkommen der kognitiven Neurowissenschaften wurde im Zuge der Interpretation verschiedener empirischer Befunde der Begriff der Introspektion wieder "salonfähig". In den letzten 10 Jahren wird dieser vormals gemiedene Begriff in der Philosophie der Psychologie immer häufiger verwendet. Teils werden neuartige klinische und theoretische Konzepte damit erklärt, teils werden theoretische Konzepte auf den Begriff der Introspektion angewandt. Eine wichtige Funktion der analytischen Philosophie besteht in der Überprüfung der Definition eines Begriffs. Kann man "Introspektion" überhaupt sinnvoll gebrauchen? Eine interessante Studie, Bennett und Hackers Buch "Philosophical Foundations of Neuroscience", liefert eine wichtige Diskussionsgrundlage. Auf den Begriff der Introspektion gehen diese nur kurz ein, berühren jedoch in so vielen verwandten Bereichen wichtige Probleme der Begriffsdefinition, dass ich das Buch als Prototyp einer kompetenten analytischen Auseinandersetzung mit den Neurowissenschaften sehen möchte. Im Rahmen meines Aufsatzes soll deshalb ein kurzer Abriss der gängigen Introspektionskonzepte neben der Kritik Bennett und Hackers dargestellt werden werden. Insbesondere möchte ich die potentielle Leistungsfähigkeit alternativer Introspektionsbegriffe darlegen.

### **Ist die Introspektion eine Sonderform der so genannten 'mentalen Repräsentation'?**

Mentale Repräsentation ist ein weit verbreiteter Begriff in den Kognitionswissenschaften, sei es nun in einer mehr neurobiologischen, psychologischen oder auch philosophischen Ausrichtung. In einem einführenden Artikel in der "Stanford Encyclopedia of Philosophy" wird mentale Repräsentation wie folgt charakterisiert:

However, on the assumption that a representation is an object with semantic properties (content, reference, truth-conditions, truth-value, etc.), a mental representation may be more broadly construed as a mental object with semantic properties. As such, mental representations (and the states and processes that involve them) need not be understood only in computational terms. On this broader construal, mental representation is a philosophical topic with roots in antiquity and a rich history and literature predating the recent "cognitive revolution." Though most contemporary philosophers of mind acknowledge the relevance and importance of cognitive science, they vary in their degree of engagement with its literature, methods and results; and there remain, for many, issues concerning the representational properties of the mind that can be addressed independently of the computational hypothesis. (Pitt 2004)

***Introspektion und Mentale Repräsentation in klassischen Philosophien der Neuzeit***

Einige bedeutende Philosophen der Neuzeit beschäftigten sich intensiv mit der Introspektion. Zwar sind damalige Begriffe nicht ohne Vorbehalt in die gegenwärtige Debatte übertragbar, doch ist es faszinierend zu sehen, wie sich bereits dort Vorläufer der Konzepte und Fragestellungen der Kognitionswissenschaften an der Grenze zwischen Philosophie und Psychologie abzeichnen. Dies gilt auch für den Begriff der Introspektion.

Anhand einiger kurzer ausgewählter Beispiele möchte ich zeigen, wie Locke, Hume und Kant das Problem mentaler Repräsentate behandeln und wie sich diese Konzepte zu zeitgenössischen Fragestellungen verhalten.

Hume behandelt das Problem der mentalen Repräsentation vor allem in "An Enquiry concerning Human Understanding". Zentrale Konzepte für Hume sind die Wahrnehmung und die Ideen. Während alle Wahrnehmung bis auf die Leidenschaft ("passion") Dinge der Außenwelt repräsentieren können, können Ideen sowohl andere Ideen als auch Wahrnehmungen repräsentieren:

[N]o man, who reflects, ever doubted, that the existences, which we consider, when we say, this house and that tree, are nothing but perceptions in the mind, and fleeting copies or representations of other existences, which remain uniform and independent.<sup>1</sup>

Bei Lockes Betrachtung der Introspektion zeigt sich deutlich, dass er von der Metapher des "Nachinnen-Schauens" ausgeht. Offensichtlich scheint auch für ihn die Introspektion eine Form der inneren Wahrnehmung zu sein:

Zweitens ist die andere Quelle, aus der die Erfahrung den Verstand mit Vorstellungen versieht, die Wahrnehmung der Vorgänge in unserer eigenen Seele, wenn sie sich mit den erlangten Vorstellungen beschäftigt. Wenn die Seele auf diese Vorgänge blickt und sie betrachtet, so versehen sie den Verstand mit einer andern Art von Vorstellungen, die von Aussendungen nicht erlangt werden können; dahin gehören das Wahrnehmen, das Denken, Zweifeln, Glauben, Begründen, Wissen, Wollen und alle jene verschiedenen Thätigkeiten der eigenen Seele. Indem wir uns deren bewusst sind und sie in uns betrachten, so empfängt unser Verstand dadurch ebenso bestimmte Vorstellungen, wie von den unsere Sinne erregenden Körpern. Diese Quelle von Vorstellungen hat Jeder ganz in sich selbst, und obgleich hier von keinem Sinn gesprochen werden kann, da sie mit äusserlichen Gegenständen nichts zu thun hat, so ist sie doch den Sinnen sehr ähnlich und könnte ganz richtig innerer Sinn genannt werden. Allein da ich jene Quelle schon Sinneswahrnehmung nenne, so nenne ich diese: Selbstwahrnehmung, da die von ihr gebotenen Vorstellungen von der Seele nur durch Wahrnehmung ihres eigenen Thuns in ihr gewonnen werden können, unter Selbstwahrnehmung verstehe ich in dem Folgenden die Kenntniss, welche die Seele von ihrem eigenen Thun und seiner Weise nimmt, wodurch die Vorstellungen von diesen Thätigkeiten in dem Verstand entstehen. Diese beiden Dinge, d.h. die stofflichen, als die Gegenstände der Sinne, und die Vorgänge innerhalb unserer Seele als die Gegenstände der Selbstwahrnehmung sind für mich der alleinige Ursprung aller unserer Vorstellungen. Ich brauche hier das Wort: Vorgänge in einem weitem Sinne, wo es nicht blos die Thätigkeit der Seele in Bezug auf ihre Vorstellungen, sondern auch eine Art von Gefühlen

---

<sup>1</sup> David Hume, An Enquiry concerning Human Understanding, Abschnitt 1.2.3.15

umfasst, die mitunter aus ihnen entstehen, wie z.B. die Zufriedenheit oder Unzufriedenheit, welche aus einem Gedanken entspringt.<sup>2</sup>

Zwei der wichtigsten Proponenten der Idee, dass Introspektion eine Form der Wahrnehmung sei, Churchland und Armstrong, glauben, dass Kant ein Pionier dieser Ansicht war. Kant selbst geht tatsächlich auch von einer Wahrnehmungsmetapher aus. Er gebraucht zwar nicht den Begriff der Introspektion, jedoch spricht er vom "inneren Sinn" oder auch von der "inneren Anschauung":

Das Bewußtsein seiner selbst (Apperzeption) ist die einfache Vorstellung des Ich, und, wenn dadurch allein alles Mannigfaltige im Subjekt selbsttätig gegeben wäre, so würde die innere Anschauung intellektuell sein. Im Menschen erfordert dieses Bewußtsein innere Wahrnehmung von dem Mannigfaltigen, was im Subjekte vorher gegeben wird, und die Art, wie dieses ohne Spontaneität im Gemüte gegeben wird, muß, um dieses Unterschiedes willen, Sinnlichkeit heißen. Wenn das Vermögen, sich bewußt zu werden, das, was im Gemüte liegt, aufsuchen (apprehendieren) soll, so muß es dasselbe affizieren, und kann allein auf solche Art eine Anschauung seiner selbst hervorbringen, deren Form aber, die vorher im Gemüte zum Grunde liegt, die Art, wie das Mannigfaltige im Gemüte beisammen ist, in der Vorstellung der Zeit bestimmt; da es denn sich selbst anschauet, nicht wie es sich unmittelbar selbsttätig vorstellen würde, sondern nach der Art, wie es von innen affiziert wird, folglich wie es sich erscheint, nicht wie es ist.<sup>3</sup>

Ein bemerkenswerter Punkt ist, dass die Leistungsfähigkeit der modernen Begriffe von Introspektion und mentaler Repräsentation geeignet sind, die zitierten Philosophien in einen zeitgenössischen Kontext zu stellen. Zusammengefasst verwenden sowohl Locke als auch Kant Bewusstsein und Introspektion als wahrnehmungsartige Repräsentation zweiter Ordnung, die unsere psychologischen Zustände und Ereignisse repräsentiert (Lycan 1999).

### **Bedeutung der mentalen Repräsentation und der Introspektion in den kognitiven Neurowissenschaften**

Allgemein bezeichnet *Repräsentation* in der Philosophie eine Handlung, welche eigenen semiotischen Regeln gehorcht. Nach E. Cassirer ist zum einen der Richtungscharakter der Repräsentation wichtig, d.h. das Zeichen verweist auf etwas, zum anderen spielt in vielen Theorien, die von der Erkenntnistheorie bis hin zur Sprachphilosophie reichen, der Symbolcharakter der Repräsentation eine wichtige Rolle. Während der Begriff in der Philosophie des Mittelalters, z.B. bei Thomas von Aquin, noch allgemein verwendet wurde, erfuhr er in der Phänomenologie, dem Neukantianismus und später den Kognitionswissenschaften eine genauere Spezifizierung. Dort sind *mentale Repräsentationen* solche kognitiven Zustände, die auf ein Ziel gerichtet sind. Dies wird auch als *Intentionalität* bezeichnet. Für die Diskussion der Introspektion ist nun wichtig, dass für die mentalen Repräsentationen von einigen Philosophen und Kognitionswissenschaftlern gefordert wird, diese sollen *introspezierbar* sein. Gerade die Introspezierbarkeit kennzeichne mentale Repräsentationen als solche. Mentale Repräsentationen sind also ein Sammelbegriff für informationstheoretische Zeichensysteme, die neural im Gehirn verankert sind und deren

---

<sup>2</sup> Locke: *Versuch über den menschlichen Verstand. Philosophie von Platon bis Nietzsche*, S. 13857 (vgl. *Locke-Versuch Bd. 1*, S. 102)]

<sup>3</sup> Kant: *Kritik der reinen Vernunft. Philosophie von Platon bis Nietzsche*, S. 23749 (vgl. *Kant-W Bd. 3*, S. 93)

Manipulation bestimmten Regeln folgt. Diese Regeln sind die Grundlage all unseres Denkens und wissenschaftlich beschreibbar.<sup>4</sup> Die Regeln der mentalen Repräsentationen folgen dabei Symbolsystemen, wie man sie beispielsweise in Programmiersprachen findet. Hier findet sich auch die begriffliche Nähe zu Diskussionen aus dem Bereich des Funktionalismus<sup>5</sup> innerhalb der *Philosophy of Mind*. Einen einheitlichen Repräsentationsbegriff gibt es aber nicht. Prinzipiell werden Vertreter eines symbolischen von Vertretern eines ikonischen Repräsentationsbegriffs unterschieden. Die Propositionalisten wie z. B. Fodor (Fodor et al. 1964) gehen von einer einheitlichen "language of thought" aus. Innerhalb der Neurowissenschaften gibt es zwar Diskussionen, ob und wie diese dem Geist zugrundeliegende Symbolstruktur implementiert ist, jedoch spielt das für das Repräsentationsargument gar keine Rolle: Fodor argumentiert, das "Wie" sei letztlich egal, es sei nur eine Frage der Zeit, wann die empirischen Einzelwissenschaften dessen Struktur klären. Geist und Denken seien letztlich aber mit keinem anderen Prinzip realisierbar. Auf philosophischer Ebene ist nun die Frage interessant, wie solche neural begründeten Repräsentationen ihren spezifischen Charakter erhalten. Gibt es zwischen Zeichen und Gegenstand eine Isomorphie? Wenn ja, wie ist diese beschaffen? Grundet sie evtl. in spezifischen neuralen Strukturen? Dies würde der Vereinbarkeit mit den empirischen Neurowissenschaften Möglichkeiten eröffnen. Gibt es eine kausale Abhängigkeit der inneren Zeichen von den äußeren Gegenständen? Ein Problem der erwähnten Computermetapher der Funktionalisten besteht in einer Grenze dieser Metapher, da Computerprogramme ihre Bedeutung durch eine Interpretation erhalten. Diese entsteht aber erst durch die sie bedienenden Menschen, die ihrerseits Intentionalität besitzen und Strategien zum Umgang mit ihrer (Computer-)Umgebung. Wie erhalten nun mentale Repräsentationen einen ontologischen und erfahrbaren Status, der sie von einem Toilettenmännchenpiktogramm unterscheidet? Letzteres stellt in seiner weit verbreiteten und leicht verständlich Bedeutung offensichtlich ein sehr leistungsfähiges Zeichen dar. Der Begriff der *Introspektion* erscheint dem Mainzer Neurophilosophen Thomas Metzinger nun geeignet, um eine bestimmte Klasse mentaler Zustände zu kennzeichnen. Damit erhält die Introspektion eine ganz neue Bedeutung, d.h. sie erscheint als zwangsläufig notwendiges mentales Konstrukt denkbar und nicht als unsichere (proto-)psychologische Erkenntnismethode:

Mental sind all jene Zustände, die prinzipiell introspezierbar sind, d.h. prinzipiell zu Gegenständen eines zumeist willkürlichen und zielgerichteten Prozesses innerer Aufmerksamkeit werden können. (...) Vielmehr ist Introspektion ein in manchen informationsverarbeitenden

<sup>4</sup> Diese Feststellung grenzt die kognitiven Neurowissenschaften vom Behaviorismus ab, der keinerlei interne kognitive Struktur als analysierbar annimmt. Während der Behaviorismus Begriffe wie Anschauung, Gedanke und Erinnerungsbild ignorierte und zurückdrängte, werden solche Begriffe in der gegenwärtigen Diskussion wieder aufgenommen. Insbesondere wird diskutiert, ob mentale Bilder, also bildhafte Vorstellungen, eine eigene Klasse von Symbolsystemen bilden oder ob die grundlegendere mentale Repräsentation nicht auf einer (vor-)sprachlichen Ebene mir sprachähnlicher propositionaler Struktur funktioniert.

<sup>5</sup> Der Funktionalismus innerhalb der Philosophie des Geistes propagiert vereinfacht gesagt eine Computeranalogie zur menschlichen Psyche. Die "Hardware" findet sich in der biologischen neuralen Gehirnstruktur, die "Software" ist das dem Denken zugrunde liegende Symbolsystem, das ähnlichen Regeln und Prinzipien gehorcht wie eine Programmiersprache. Dementsprechend wird in Gedankenexperimenten die prinzipielle Machbarkeit einer künstlichen Intelligenz erklärt, was beispielsweise J. Searle mit seinem Argument vom "chinesischen Zimmer" strikt leugnet.

Systemen ablaufender Vorgang, und zwar eine besondere Variante genau desselben Vorgangs (...): Introspektion ist eine interne Repräsentation mentaler Vorgänge. (Metzinger 1999)

Mentale Repräsentationen unterscheiden sich also von anderen Formen der Repräsentation, weil sie die Disposition besitzen, bewusst zu werden. Metzinger definiert *Mentalität* als eine dispositionale Eigenschaft einer bestimmten Klasse neuraler Strukturen. Ein Typus der Informationsverarbeitung innerhalb des Gehirns lässt sich als eine Datenklasse verstehen, die intentionalen Gehalt haben kann: Diese Datenklasse ist introspektiv erfahrbar.

In den Neurowissenschaften wurde der Begriff der Repräsentation ursprünglich unter Bezug auf ein simples Reiz-Reaktions-Schema von Neuronen verwendet. Diese reagieren auf einen bestimmten und definierten Stimulus in einer ebenso bestimmten Art und Weise. Eine klassische Vorstellung ist die des "Großmutter-Neurons", d.h. jedem Objekt der Außenwelt ist eine Nervenzelle zugeordnet. Solcherart repräsentierende Theorien erreichen inzwischen einen hohen Komplexitätsgrad. Anhand von einigen Einzelbeispielen versuchen sich Bennett und Hacker mit Theorien auseinanderzusetzen, die sich auf den mentalen Repräsentationsbegriff beziehen. Beispielhaft zitieren Bennett und Hacker Nicholas Humphrey, einen britischen Psychologen, der unter anderem mit Daniel Dennett zusammenarbeitete. Humphrey entwickelte in seinem Buch "The Inner Eye" eine Theorie des evolutionären Sinns von Introspektion. Humphrey gebraucht dabei das Konzept der "Theory of Mind", d.h. der kognitiven Fähigkeit, anderen Artgenossen intelligentes Sozialverhalten zuzuschreiben. Diese Fähigkeit kommt bereits bei vielen Säugetieren vor – ansonsten hätte wir sicherlich keine Freude an Haustieren – ist aber bei Menschenaffen und Menschen am ausgeprägtesten. Humphreys These ist, Introspektion ermögliche einem Wesen nicht nur den vorsprachlichen reflektierten Zugriff auf das eigene Denken und Fühlen, sondern darüber hinaus die Simulation des Verhaltens anderer (Humphrey 1993, c1986). In einem evolutionsbiologischen Kontext sei dies für soziale Tiere eine bedeutsame evolutionäre Neuerung. Es handelt sich also auch um eine repräsentationalistische Hypothese, d.h. Introspektion ist eine Metarepräsentation, eine Repräsentation unseres Denkens und Fühlens, wobei es sich ja auch um Repräsentationen handelt. Das Konzept des Geistes ist das einer Aussenweltsimulation, im Falle Humphreys, des Spezialfalles der sozialen Interaktion. Eine eigenständige Darstellung dieses Konzepts erfolgt in dem Abschnitt zur "Theory of Mind".

Ein anderer Protagonist repräsentationalistischer Theorien ist P. N. Johnson-Laird, der mit Hilfe von Computermodellen das Konzept der Repräsentation nachstellt (Johnson-Laird 1983, 1996). Ohne auf dessen Thesen im Detail einzugehen, ist damit die Nähe der repräsentationalistischen Theorien zu den Computerwissenschaften offensichtlich.

Was ist der gemeinsame Nenner der Repräsentationalen Theorie des Mentalen, kurz *RTM*? Eine traditionelle Einteilung der *RTMs* unterscheidet Kognitionen, die keine phänomenale Qualität<sup>6</sup> haben (und damit keine Qualia) von Wahrnehmungen, die phänomenale Qualität haben, aber keinen konzeptuellen Inhalt. Vereinfacht ausgedrückt, lässt diese Taxonomie zwei Arten der Repräsentation durch mentale Zustände zu: sprachliche und nicht-sprachliche wie z. B. Vorstellungen in Form von inneren Bildern.

---

<sup>6</sup> *Phänomenale Qualität* meint hierbei die ganz spezifische Erlebnisqualität des Menschen. Nicht das beobachtete Verhalten interessiert dabei, sondern das jeweilige Fühlen und Erleben eines Menschen, das seinen jeweiligen Seinszustand ausmacht. Dieses Konzept wird in der Philosophie des Geistes häufig in der Frage gebraucht, ob ein "neuraler Zombie", ein Wesen ohne dieses spezifische subjektive Erleben, aber mit exakt denselben neuralen Vorgängen, sich in irgendeiner Form von einem "echten" Menschen unterscheidet.

In einigen historischen Debatten (Locke, Hume und Aristoteles) wurde angenommen, die nichtkonzeptuellen Repräsentationen ("Ideas" bei Hume) seien die einzigen Repräsentationen. Der Verstand stelle die Welt mittels solcher mentaler Konstrukte dar, die den ursprünglich wahrgenommenen Dingen ähneln. Dies bedeutete, dass alle repräsentationalen Zustände eine spezifische phänomenale Qualität haben, sozusagen eine Simulation der Außenwelt. Viele Philosophen lehnten jedoch eine solche Betrachtungsweise ab. Solche mentalen Zustände seien nicht generalisierbar, zeigten eine hohe Ambiguität<sup>7</sup> und könnten deshalb nicht als Grundlage logischen oder mathematischen Denkens funktionieren. Aus diesem Grund sei eine Theorie des Mentalen, die sich nur auf mentale Zustände verlasse, die in mehr oder weniger komplexer Form den Hume'schen "Ideas" ähnelten, nur unzureichend.

In zeitgenössischen Meinungsverschiedenheiten geht es meist um die Existenz und Natur phänomenaler Eigenheiten und die Rolle, die sie bei der Bestimmung des Inhaltes einer Wahrnehmung spielen. Braucht es diese spezifische Phänomenale Erfahrung, die Qualia, um einer Wahrnehmung Inhalt und Referenz zu verleihen? Während D. Dennett dies verneint, betonen andere Philosophen, der Sinn und Nutzen der *Qualia* sei als spezifische Referenz des Repräsentierten in seiner Beziehung zur Außenwelt zu suchen (Rey 1998). Das besondere Erleben "markiert" sozusagen Gedachtes und Erlebtes. In Bezug auf den Begriff der Introspektion bedeutet dies, die Introspektion hat eine Markierungsfunktion, d.h. besonders bedeutsame mentale Ergebnisse werden mit ihrer Hilfe aus dem alltäglichen "stream of consciousness" hervorgehoben.

#### **Die Ablehnung des Repräsentationsbegriffs durch Bennett und Hacker**

Die sprachanalytische Argumentation beider Autoren wehrt sich gegen die Idee, im Gehirn gebe es eine Repräsentation der Aussenwelt. Das "Feuern" der Neuronen im visuellen okzipitalen Kortex, der Sehrinde, sei keine symbolische Beschreibung eines Objekts im visuellen Feld. Der Zusammenhang zwischen neuraler Struktur, einem möglichen Korrelat des Repräsentats, und dem Repräsentandum sei nur vergleichbar mit dem Zusammenhang zwischen Rauch und Feuer. Sage man "Rauch bedeutet Feuer", decke dies den Bedeutungsgehalt aller neurokognitiven Aussagen ab. Mit Repräsentation habe dies aber nichts zu tun, da Repräsentationen einen semantischen Gehalt besäßen, der aber dem "Rauch" in seiner Symbolisierung des "Feuers" völlig abginge.

Hierbei sind zwei Punkte kritisch anzumerken. Erstens ist der Begriff der Repräsentation innerhalb der Kognitionswissenschaften fruchtbar und blickt auf eine jahrzehntelange Tradition zurück. Die Ablehnung eines solchen Konzeptes würde die Kognitionswissenschaften um Jahrzehnte in der Diskussion zurückwerfen bzw. wäre es unklar, in welche Begrifflichkeiten man die Theoriegebäude überführen müsste, um einen solchen Rückschritt zu vermeiden. Zweitens ist die von Bennett und Hacker gegebene Definition und Sprechweise von "Repräsentation" zu einseitig. Dies mag mit dem reduzierten *Gehirn*-Begriff der Autoren zusammenhängen. Begreift man das Gehirn von seinem ontologischen Status her als lediglich eine "graue Masse aus Nervenzellen", so ist die Rede vom *repräsentierenden* Gehirn sinnlos. Begreift man das Gehirn als einen dynamischen Systemzustand, der weniger eine umschriebene Substanz ist als vielmehr ein Prozess, so kann es sinnvoll sein, von Repräsentationen innerhalb dieses Prozesses zu sprechen.

---

<sup>7</sup> *Ambiguität* meint in diesem Zusammenhang, die Erlebniseindrücke des Denkens seien zu unklar, um in einem präzisen Begriff fassbar zu sein.

## Alternativkonzepte der Introspektion in den Kognitionswissenschaften

### *Introspektion als eine Variante der Theory of Mind*

Innerhalb der kognitiven Neurowissenschaften ist das Konzept der *Theory of Mind* zu großer Bedeutung aufgestiegen. Vereinfacht gesagt, meint dieser Begriff die Fähigkeit, unser Gegenüber mental zu simulieren. Diese Simulation eines sozial Anderen ermöglicht diesem Gründe für sein Handeln zuzuschreiben. Innerhalb der Neuropsychiatrie und Neuropsychologie ist dieser Begriff insbesondere bei schweren Störungen zwischenmenschlicher Interaktion, z.B. dem Autismus, erfolgreich angewandt worden. Innerhalb der Philosophie des Geistes kann er herangezogen werden, um das Phänomen der Introspektion zu konzeptualisieren. Hierbei ist die Introspektion eine Sonderform der mentalen Simulation unseres Gegenübers, eben die Simulation unseres eigenen Denken und Fühlens. Aber erst diese Fähigkeit erlaubt es, den Anderen als nach menschlichen Beweggründen handelndes Wesen wahrzunehmen, da die Simulation des Anderen ein Spezialfall unserer Eigenwahrnehmung, der Introspektion, ist.

Der von Bennett und Hacker angeführte Nicholas Humphrey sieht die Entwicklung der Introspektion als evolutionäre Strategie, die es sozialen Tieren ermöglichte, die Handlungsmotive anderer Sozialpartner zu erfassen:

The trick with which nature came up with was introspection ... [the] examination of the contents of consciousness. (Humphrey 1984)

Das grundlegende Problem einer solchen Vorstellung, dass Introspektion und die Simulation der mentalen Zustände Anderer Hand in Hand gehen, ist der *Homunculus-Vorwurf*. Angenommen, es gäbe ein Modul innerhalb der Geist-Gehirn-Struktur, welches innere Zustände beobachtet bzw. erzeugt und diese Fertigkeit zur Simulation der mentalen Zustände Anderer gebraucht, dann muss dieses Modul eine Art beobachtender Homunkulus sein. Und was beobachtet bzw. introspeziert innerhalb dieses Homunkulus? Die Folge wäre ein Regress ad infinitum oder, wie der Neurophilosoph Daniel D. Dennett argumentiert:

Efforts to identify the self – a mortal and material soul, you might say – with a particular subsystem in the brain run into snags at every turn. I call this the fallacy of the cartesian Theater, the place where it all comes together for conscious appreciation and decision (...) All the work done by the imagined homunculus in the Cartesian Theater must be distributed around to lesser agencies in the brain. (Dennett c1991)

Dennett schlägt also vor, auf die Idee des Kartesianischen Theaters zu verzichten. Man könne sehr wohl ein Netzwerk aus mentalen Modulen vorstellen, das die Fähigkeit zur Introspektion entwickelt, ohne dass man auf den Homunculus rekurren müsse. Der grundlegende sprachliche Fehler ist dabei nicht die implizite Vorstellung des Kartesianischen Theaters, sondern die Tatsache, dass in der gegenwärtigen Fachliteratur die Verwendung des Begriffs *Bewusstsein* beinahe automatisch das Konzept *Selbst* impliziert. Aber verwendet man das Konzept *Selbst*, so gerät man automatisch in eine Zuschreibungsfalle, die in der Frage mündet, wer denn da wahrnehme, handle oder introspeziere?

### ***Introspektion als eine Form der inneren Wahrnehmung***

Die älteste Vorstellung der Introspektion greift auf die Vorstellung zurück, es handle sich um eine Art der inneren Vorstellung. Eine Variante davon betrachtet die Introspektion des Bewusstseins als eine Art von "neural monitoring", also eine Art von neuralem Überwachungs- und Beobachtungsmechanismus. Erst diese Beobachtung brächte uns mentales Geschehen zu Bewusstsein. Nur eine bestimmte Kategorie mentalen Geschehens ist aber auch introspezierbar. Das Introspektionsargument ist also auch hier, ähnlich wie bei Th. Metzinger, eine Prädikatsklasse mentalen Geschehens. Im Mittelpunkt dieser Betrachtung steht jedoch die Metapher vom "Auge des Bewusstseins", die zentral von Bennett und Hacker als falsch angesehen wird.

Der amerikanische Philosoph W. Lycan (Lycan et al. 1995) begründet auch, was ein Vorteil der "Inner Sense" Theorie ist. Diese unterscheidet die Introspektion von der Aufmerksamkeit der (Wahrnehmungs-)Psychologie und Neurowissenschaft, außerdem unterscheidet die Introspektion bewusste Zustände von reiner "mentation"<sup>8</sup>

### **Die Probleme einer sprachanalytischen Kritik könnten in unreflektierten Vorannahmen liegen**

Bennett und Hacker stellen in ihrem Buch "Philosophical Foundations of Neuroscience" (Bennett et al. 2003) einen langen Katalog an Fakten über den korrekten Sprachgebrauch hinsichtlich kognitiver Vorgänge auf. Nicht immer ist einsichtig, wie sie auf die jeweiligen Fakten kommen. Teilweise orientieren sie sich sehr am herkömmlichen Sprachgebrauch. Scheint die Neurowissenschaft zu sehr von diesem abzuweichen, so begeht sie einen sprachlichen Fehler. Streckenweise scheint diese Auslegung zu puristisch, hemmt sie doch die Möglichkeit des Wissenschaftlers, selbstständig neue Begriffe zu schaffen oder alten Begriffen in einem neuen Kontext neues Leben einzuhauchen. So gehen die Argumente, die sie zur Abwehr des Konzeptes der Introspektion als innerer Wahrnehmung wählen, von Prämissen aus, die oft nicht ganz einsichtig sind.

But introspection involves no perceptual organ (...). It involves no observation – one does not observe one's thought, or descry one's desires or intentions.

Zwar scheint auf den ersten Blick offensichtlich, dass die Introspektion (oder das Gehirn) keine Wahrnehmungsorgane sind, aber auf den zweiten Blick fragt man sich nach einer plausiblen und ausreichend genauen Begründung. Eine Definition des "wahrnehmenden Organs" wird nicht geleistet, vielmehr lehnen sich Bennett und Hacker immer wieder am herkömmlichen Sprachgebrauch an. Hierbei besteht die Gefahr, dass sie Sprache über mentale Vorgänge in einer derart naiven Form verwenden, dass man von einer "folk psychology", einer naiven, umgangssprachlichen Populärpsychologie sprechen muss. Die Gefahr einer solchen *folk psychology* besteht in der unreflektierten Übernahme von impliziten Vorstellungen, wie mentales Geschehen in Menschen abläuft.

Derselbe Vorwurf lässt sich auch bei der Behandlung der Wahrnehmung durch Bennett und Hacker machen. Hier wird erstmals eine Begriffsdefinition der Wahrnehmung gemacht, die sich jedoch auf die Aufzählung des herkömmlichen Sprachgebrauchs bezieht. So gebe es fünf Sinne, wobei der

---

<sup>8</sup> Mentation (engl.): =abstraktes Denken; in der deutschen Fachterminologie findet sich kein Begriff, der das abstrakte, serielle Denken in meist propositionalen linguistischen Strukturen bezeichnet.



Sinn der taktilen Wahrnehmung über kein spezifisches Wahrnehmungsorgan verfüge. Aus neurobiologischer Sicht sind solche Definitionen haarsträubend unpraktisch. Ungeachtet der Tatsache, dass man die faktischen Aufzählungen Bennett und Hackers bezweifeln kann, sind die verwendeten Begriffsdefinitionen für einen Neurowissenschaftler unbrauchbar. Legen seine Ergebnisse nahe, dass er Wahrnehmung nicht mit den Wahrnehmungsbegriffen der Umgangssprache fassen kann, z.B. weil Teile des Gehirns Eigenschaften zeigen, die man vormals dem wahrnehmenden Organ zuschrieb, kann man ihm nicht vorwerfen, sich nicht an diese Sprachkonvention zu halten, wenn ihn die Faktizität seiner Experimente zu neuen Konzepten zwingt. Der Satz "Ich fühle das Messer in meiner Hand" kennzeichnet ja nicht die Hand als das primär wahrnehmende Organ. Ebenso kann man von einem Hirnforscher nicht verlangen, er müsse das Auge als das primär wahrnehmende visuelle Organ anerkennen, wenn er versucht, das Verhalten zu beschreiben, das ein Patient zeigt, dessen Augen intakt sind, der jedoch nach Zerstörung (okzipitaler) Gehirnteile neural erblindet ist (Solms et al. 2004). Obwohl dessen Auge intakt ist, sprechen wir von Blindheit. Aber was ist erblindet? Das Gehirn? Nein, das Gehirn sei kein wahrnehmendes Organ, also könne man so nicht sinnvoll sprechen, argumentieren Bennett und Hacker und erschweren hier eher die Transportation wissenschaftlicher Erkenntnis.

Der Begriff der Repräsentation und der Introspektion als Systemeigenschaften eines Gehirn-Geist-Komplexes ist außerordentlich leistungsfähig, gestattet er doch die Verbindung klassischer Philosophie (vgl. Locke, Kant u.a.), Psychologie, Neurowissenschaften und Philosophie. Ganz klar ändert sich dabei der Begriff in einer Weise, die von Bennett und Hacker als grundlegend falsch angesehen wird. Ihr sprach-puristischer Ansatz läuft jedoch Gefahr, innovationsfeindlich zu sein und dazu noch die sprachimplizite "folk psychology", eine möglicherweise falsche Vorstellung von Geist und Bewusstsein, zu transportieren. Auch an dem von Bennett und Hacker diskutierten Beispiel *Schmerz* möchte ich kurz prinzipielle Probleme ihrer Argumentkette zeigen. In der Neurophysiologie werden die neuronalen Verbindungen zwischen den Wahrnehmungsrezeptoren der Haut (oder anderer Organstrukturen) und dem kortikalen Gehirn als Schmerzbahnen oder auch Schmerzzentren bezeichnet. Im Sinne einer Hacker'schen Sprachkritik müsste man nun sagen, dies sei grundlegend falsch, da Schmerz ein Prozess sei, niemals könne ein Gehirnbereich Schmerz "haben", "empfinden" oder gar "darstellen". In ähnlicher Weise, argumentiert er, befände sich der Zahnschmerz nicht im Gehirn, sondern im Zahn. Wobei es weniger um den ontologischen Status der phänomenalen Qualität geht, sondern um die Definition der Sprachverwendung. Zahnschmerz impliziert, der Schmerz sei im Zahn. Genauso wird mit dem Thema "Karten im Gehirn" oder auch mentale Repräsentation verfahren. Da der Begriff des Schmerzrezeptors oder der Schmerzbahn in der Neurophysiologie völlig etabliert ist, scheint er nicht nur empirisch ausreichend verifiziert worden zu sein, auch auf einer wissenschaftstheoretisch abstrakteren Ebene hat er seine Leistungsfähigkeit gezeigt, lassen sich doch viele verschiedene Befunde, Hypothesen und Theorien sinnvoll mit seiner Hilfe diskutieren. Wie ein Verzicht auf den Begriff im Sinne des Hacker'schen Sprachpurismus aussehen kann, wird von seiner Seite nicht klar.

Ein Schmerzbegriff, der sowohl das Gehirn als auch die Schmerzzentren und Schmerzbahnen Schmerz erfahren lässt, gestattet in Kombination mit dem oben dargestellten "Inner Sense" Introspektionskonzept ein zwangloses, aber nichtsdestotrotz sinnvolles Reden über Schmerzphänomene:

Fourth, the Inner Sense View sorts out a long standing issue about sensations and feelings: consider pain. A minor pain may go unfelt, or so we sometimes say. Even quite a bad pain may

not be felt if attention is distracted by sufficiently pressing concerns (...). Sometimes the word 'pain' is used just to mean the first-order representation of damage or disorder. But sometimes 'pain' means a conscious feeling or mode of awareness, and on that usage the phrase 'unfelt pain' is [a] first-order representation and the second-order scanning all together.<sup>9</sup>

Zusammengefasst lässt sich meine Kritik an Bennett und Hacker an zwei Punkten festmachen. Erstens zeigt sich bei Austausch des Gehirnbegriffs gegen äquivalente Positionen innerhalb der Kognitionswissenschaften (einschließlich der Philosophie), dass sich eine Vielzahl von Begriffen über Geist, mentale Repräsentation und Introspektion in Luft auflösen. Ein sinnvolles Reden über mentale Phänomene ist bei einer solchen Reductio ad absurdum nicht mehr möglich. Zweitens ist der Gehirnbegriff selbst zu eng gefasst. Die ontologische Struktur des Gehirnbegriffs wird von Bennett und Hacker als ganz selbstverständlich vorausgesetzt, das Argument des mereologischen Fehlschlusses (dem Gehirn psychologische Prädikate zuschreiben) verliert seine Brisanz, wenn man den starren Gehirnbegriff ändert. Sehr wohl werden innerhalb der Neurowissenschaften Konzepte präsentiert, die das Gehirn als einen dynamischen Prozess, als einen Geist-Gehirn-Komplex verstehen. Der Vorwurf des mereologischen Fehlers findet sich in besser ausgeführter Variante bei Daniel Dennett, der die Idee eines physikalischen Substrates kritisiert, das alle neurobiologischen Informationsbegriffe zusammenführt und diese (evtl. via Introspektion) dem Bewusstsein zuführt. Ein solches physikalisches Substrat, das man leicht mit einem naiven Gehirnbegriff identifizieren kann, bezeichnet er als "kartesisches Theater". Vor diesem Hintergrund sind aus einer sprachanalytischen Sicht "naive" Gehirn-Geist-Konzepte abzulehnen. Das Phänomen der Introspektion dient hierbei als Beispiel, um die Unvereinbarkeit mit einem solch naiv-reduktionistischen Konzept darzustellen, wobei der Sprachreduktionismus Bennett und Hackers keine Lösung darstellt.

In ihrer Ablehnung der Introspektion beziehen sich Bennett und Hacker auch auf das Wittgenstein'sche Privatsprachenargument. Wittgenstein argumentiert in den Philosophischen Untersuchungen, es könne keine Sprache geben, die sich nur auf das innere Erleben eines Sprechers bezieht (Wittgenstein 2003). Sprache ist nur sinnvoll in einem Kontext, wo Sprachregeln zwischen Sprechern definiert werden. Eine Privatsprache kann aber gar nicht über gemeinsam definierte Regeln verfügen, also existiert sie nicht. Bennett und Hacker wenden dieses Argument auf verschiedene Bereiche der Kognitionswissenschaften an. Die implizite Unterscheidung zwischen einem mentalen Innenraum und einem objektiven Außenraum erscheint ihnen vor diesem Hintergrund als nicht sinnvoll. Es gibt keinen Innenraum, der sprachlich exklusiv für den Sprecher existiert, sondern nur gemeinsame Sprachregeln. Wenn jemand meint, von einem exklusiven inneren Erleben berichtet zu können, so ist dies aber sinnlos. Beispielsweise gilt der Schmerz als vermeintlich exklusive Erlebnisqualität, jedoch verweisen Bennett und Hacker darauf, dass das Schmerzverhalten erst in einem definierten Kontext Sinn erhält. In unserem herkömmlichen Sprachgebrauch schiene es sinnlos, wenn jemand lache, dabei aber erkläre, er leide starke Schmerzen.<sup>10</sup> Mit dem Verweis auf die Kontextabhängigkeit unserer alltäglichen Sprachhandlungen

---

<sup>9</sup> Lycan, Internal Monitoring, Seite 4

<sup>10</sup> Interessanterweise begründen Bennett und Hacker dies nicht, sondern schreiben lediglich, es sei "an intelligible possibility that pain might systematically be correlated with smiling and laughing" (S.82). Da viel der argumentativen Struktur des Buches von diesem Argument abhängt, überrascht, wie wenig ausführlich diese Begründung ist.

wird die Introspektion als Möglichkeit eines Zugriffs auf ein exklusives und privates inneres Erleben abgelehnt.

### **Spracherwerb nach Bennett und Hacker – Eine Kritik des Privatsprachenarguments**

In ihrer Ablehnung der Neurowissenschaften als für die Philosophie relevante Disziplin verweisen Bennett und Hacker darauf, dass Sprache in einem Kontext erworben wird, wo ein Kind eben lernt, einem bestimmten Verhalten – z.B. Schmerz – eine Bedeutung zuzuweisen. Die empirischen Befunde der Neurowissenschaft sind dabei gleichgültig, philosophisch geht es um eine präzise Begriffsdefinition. Wittgensteins Privatsprachenargument wird gebraucht, um die prinzipielle Unterscheidung von (mentalem) Innen und Außen, direktem und indirektem Zugriff (auf mentales Geschehen), privilegiertem Zugriff und Introspektion sowie Subjektivität und Privatheit als sprachliche Verwirrung abzulehnen. Bennett und Hacker fassen dabei die Positionen der Gegenseite wie folgt zusammen:

In short, psychological vocabulary is thus conceived to consist of names of subjective experiences, mental states or processes to which each person has private, indirect and privileged access. And the words are conceived to be given their meanings by being linked to the experience, states or processes. (S. 97).

Um die Sichtweise repräsentationalistischer Theorien des Geistes – wie eingangs skizziert – aufzugreifen, könnte man auch sagen, die Wortbedeutung besteht in der Repräsentation von Erfahrung, mentalen Zuständen und Prozessen. Wittgenstein stellte nun die Frage, ob wir die Bedeutung von Wörtern wie "Denken", "Fühlen" oder auch "Schmerz" erlernen, indem wir sie mit einer Erfahrung dieser mentalen Modalität assoziieren. Tatsächlich kann aber eine subjektive Erfahrung keine operative Regel für den Sprachgebrauch liefern, da wir nie wissen könnten, ob andere Personen dieselbe Erfahrung mit derselben Wortbedeutung verbinden. Wäre die Erfahrung Schmerz aber nur privat und introspektiv zugänglich, so wüsste niemand mit letzter Sicherheit, ob wir alle dieselbe Erfahrung auch teilten. Eine problematische Wende in der an Wittgenstein angelehnten Argumentation ergibt sich jedoch in folgendem Absatz:

We do not teach a child to use the word 'pain' by getting him to associate the word with a sensation he feels, and to use the memory of that sensation as a standard of correctness for the subsequent use of the word. (...)The primitive or elementary pain-utterance is learnt as an extension of natural expressive behavior. (...) [He] no more needs ground, either 'introspective' or behavioral, for crying out 'It hurts' (...) than he needs ground for screaming in pain. (...) Obviously, a child does not learn to identify by means of introspection an inner phenomenon to which it has 'access' and then to name it (...). (S.100 - 102).

Bennett und Hacker liefern in einem längeren Absatz also eine Darstellung des frühkindlichen Spracherwerbs, die stark an behavioristische Konditionierungs-Schemata erinnert.

Bei ihrer abgewandelten Verwendung des Privatsprachenarguments schleicht sich bei Bennett und Hacker damit jedoch selbst eine empirische Begründung ein, die zweifeln lässt, ob die rigide Trennung zwischen begriffsklärender Philosophie und empirischer Neurowissenschaft immer sinnvoll ist. Sie präsentieren eine Theorie des Spracherwerbs, die sich ganz klar empirisch überprüfen ließe. Mit diesem Argument steht und fällt jedoch ihre Verwendung des

Privatsprachenarguments. In einer Kritik des Buches in "Mind" werden die Position und die Kritik folgendermaßen zusammengefasst:

They go so far as to say that, 'the supposition that scientific evidence may contravene a philosophical analysis is...risible' (p.404). This seems somewhat problematic given that, as part of their account of the meanings of psychological predicates, Bennett and Hacker offer an empirical story as to how psychological vocabulary is learnt. They claim that self-ascriptions are learnt, not by associating names with sensations, but as replacements of natural expressive behaviour, and they offer an account of how this might be achieved via parent-child interaction (pp.100-103). But this is precisely the sort of account that one would expect to be supported (or otherwise) by an empirical account of language acquisition.

Ich möchte mich nicht mit den Details einer entwicklungspsychologischen Kritik einer solchen Vorstellung von Spracherwerb auseinandersetzen, sondern beispielhaft darauf hinweisen, dass die strikte Trennung zwischen Empirie und philosophischer Begriffserklärung auch Bennett und Hacker nicht gelingt.

Bezogen auf die Kritik der Introspektion lässt sich sagen, dass Bennett und Hacker hauptsächlich einen naiven Introspektionsbegriff kritisieren. Wie eingangs gezeigt, gibt es jedoch Introspektions-Modi, die für eine sinnvolle wissenschaftliche Diskussion fruchtbar sind und hinreichend definiert erscheinen. Bennett und Hacker muss recht gegeben werden, wenn sie eine naive "Retrospektion" ablehnen, d.h. das Reden über ein inneres Erleben. Introspektion kann jedoch auch als vorsprachliches Erleben, als Zugriff auf mentale Repräsentationen verstanden werden. Die Kritik Bennett und Hackers ist in vielem deckungsgleich mit dem Introspektionsbegriff der behavioristischen Psychologie. Diese lehnte – zu Recht – das Sprechen über mentales Geschehen als unzuverlässig ab.

### **Zusammenfassung**

Aufgrund neuer Diskussionen in der Psychologie und Neurobiologie ist es spannend, den Begriff der Introspektion neu zu beleben. Introspektion wurde implizit und explizit in verschiedenen Philosophien der Neuzeit verwendet (Hume, Locke und Kant). Dies nimmt in vielem die gegenwärtige Diskussion in den Kognitionswissenschaften vorweg. Diese charakterisiert heute die Introspektion als eine Sonderform mentaler Repräsentation. Die Vorstellung, der menschliche Geist "repräsentiere" Erlebnisse und Sinneseindrücke, d.h. Denken symbolisiere tatsächliches und vorgestelltes Geschehen, ist nicht neu, gewinnt jedoch aufgrund vieler empirischer Erkenntnisse an Schärfe. Viele Philosophen und Gehirnforscher verstehen Introspektion deshalb als eine Art Metarepräsentation, d.h. die Repräsentation von Repräsentiertem im Geist. Anders ausgedrückt das Denken über Denken ist ein Prozess zweiter Ordnung. Bennett und Hacker lehnen jedoch den Repräsentationsbegriff deutlich ab. Ihre Argumente führen auch zu einer Ablehnung eines solchen Introspektionsbegriffs. Ebenso lehnen sie die Vorstellung von der Introspektion als innerer Wahrnehmung oder als "Theory of Mind" ab. Ein zentrales Problem ihrer sprachanalytischen Kritik ist hierbei die ungewollte Verwendung impliziter empirischer Fakten.

Bei genauer Betrachtung der sprachanalytischen Untersuchung schleicht sich der Verdacht ein, dass die sprachliche Analyse-Ebene Teil des Missverständnisses ist. Vielleicht ist ein sprachlicher Introspektionsbegriff mit den von Bennett und Hacker vorgebrachten Argumenten deutlich abzulehnen, aber wie eingangs charakterisiert, bezieht sich der Gebrauch des Introspektionsbegriffs in der Neurophilosophie und der kognitiven Neurowissenschaften auf einen

vorsprachlichen Prozess. Bennett und Hacker legen zwar überzeugend dar, weshalb eine sprachliche Introspektion – hauptsächlich bedingt durch das Privatsprachenargument – sinnlos ist. Eine vorsprachliche repräsentationalistische Verwendung des Introspektionsbegriffs, also ein Zugriff auf Bedeutung durch Repräsentation sprachlichen Geschehens, ist jedoch sinnvoll. Wird Denken als vorsprachliche Struktur verstanden, greift keinesfalls das Privatsprachenargument. Ein privater Satz von mentalen Strukturen und Regeln ist dabei durchaus sinnvoll. Daneben ist eine deutliche Schwäche des Privatsprachenarguments, wie es von Bennett und Hacker dargestellt wird, dass es implizite Annahmen über den Spracherwerb macht, die eindeutig einer empirischen Verifizierung zugänglich sind. Dass dieses Modell des menschlichen Spracherwerbs von Theoretikern wie Chomsky oder Pinker völlig abgelehnt würde, da diese einen bedeutenden Anteil einer angeborenen vorsprachlichen Komponente sehen, ist nur noch eine Nebensache.

### Literatur- und Quellenverzeichnis

- Bennett, Max R. and Hacker, Peter M. S. *Philosophical foundations of neuroscience*. Berlin [u.a.]: Blackwell, 2003.
- Dennett, Daniel Clement. *Consciousness explained*. Boston: Little, Brown and Co, c1991.
- Fodor, Jerry A. and Katz, Jerrold J. *The structure of language*. Englewood Cliffs, N.J.: Prentice- Hall, 1964.
- Humphrey, Nicholas. *Consciousness regained*. Oxford: Oxford University Press, 1984.
- Humphrey, Nicholas. *The inner eye*. London: Vintage, 1993, c1986.
- James, William. *The principles of psychology*. London: McMillan, 1907.
- Johnson-Laird, P. N. *Mental models*. Cambridge, Mass: Harvard University Press, 1983.
- Johnson-Laird, Philip Nicholas. *Der Computer im Kopf: Formen und Verfahren der Erkenntnis*. München: Dt. Taschenbuch-Verl., 1996.
- Lycan, William. *Consciousness as Internal Monitoring. Philosophical Perspectives*, 1995, 9.
- Metzinger, Thomas. *Subjekt und Selbstmodell: die Perspektivität phänomenalen Bewusstseins vor dem Hintergrund einer naturalistischen Theorie mentaler Repräsentation*. Paderborn: Mentis, 1999.
- Pitt, David. *Mental Representation*. , 2004.
- Rey, G. "A Narrow Representationalist Account of Qualitative Experience," in J.E. Tomberlin, eds., *Language, Mind, and Ontology*. Atascadero, USA: Ridgeview Publishing, 1998,.
- Solms, Mark and Turnbull, Oliver. *Das Gehirn und seine innere Welt: Neurowissenschaft und Psychoanalyse*. Zürich: Walter, 2004.
- Wittgenstein, Ludwig, Schulte, Joachim. *Philosophische Untersuchungen*. Frankfurt am Main: Suhrkamp, 2003.