

# Philosophie der Kognitionswissenschaft

## 1 Einleitung

### 1.1 Kognition, Kognitionswissenschaft und Philosophie der Kognition

Kognitionswissenschaft ist die Wissenschaft von Kognition. Was aber ist Kognition? Vor dem Hintergrund seines etymologischen Ursprungs in den lateinischen und griechischen Ausdrücken für erkennen, wahrnehmen oder wissen (*cognoscere*, *gignoskein*) wird der Ausdruck „Kognition“ zunächst einmal mit Problemlösen und Intelligenz assoziiert. Wir sind ständig mit Problemen unterschiedlichster Art konfrontiert: Wir müssen beim Schach den besten Zug finden, ein Theorem beweisen, herausfinden, warum der Rasenmäher Öl verliert, ein Tablett voller Gläser durch eine Party balancieren, Radfahrern ausweichen, den Hund zum Tierarzt bringen, beim Metzger Besorgungen machen und das Auto in Reparatur bringen, ohne dabei zu große Umwege zu fahren oder den Hund unnötig lange mit den Einkäufen allein im Wagen zu lassen usw. Als kognitive Leistung kann das gelten, was uns dies alles ermöglicht: Wir müssen etwa unsere Umgebung *wahrnehmen*, unsere *Aufmerksamkeit* auf etwas richten, uns an vergangene Lösungsversuche *erinnern*, aus gescheiterten *lernen*, aus unseren Erfahrungen *Schlüsse ziehen*, mit anderen *sprechen* oder nichtsprachlich mit ihnen *kommunizieren*, Lösungsstrategien *planen*, uns *entscheiden* usw. Die Kognitionswissenschaft lässt sich in diesem Sinne als ein integratives Forschungsprogramm auffassen, das eine empirisch wie begrifflich umfassende transdisziplinäre Untersuchung jener Leistungen anstrebt, die natürliche oder künstliche Systeme – zum Beispiel Menschen, andere Tiere, Computersimulationen oder Roboter – befähigen, durch intelligentes Verhalten Probleme effizient zu lösen (Stephan und Walter 2013). Neben dem Ideal einer „starken“ Künstliche-Intelligenz-Forschung, also der Erschaffung tatsächlich intelligenter künstlicher Systeme, und dem rein ingenieurwissenschaftlichen Interesse an künstlichen Systemen, die ein Verhalten zeigen, das zumindest beim Menschen Intelligenz erfordert (s. Kap. IV.15), verfolgt eine „schwache“ Künstliche-Intelligenz-Forschung dabei vor allem auch das Ziel, durch die Modellierung intelligenter Leistungen die Natur, Funktion und Organisationsprinzipien der kognitiven Leistungen des Menschen zu verstehen.

Mit dem Niedergang des radikalen Behaviorismus von John Watson und Burrhus Frederic Skinner Mitte des zwanzigsten Jahrhunderts setzte sich in der Psychologie (wieder) die Ein-

sicht durch, dass intelligentes Verhalten nur durch Rekurs auf interne mentale Prozesse zu erklären ist, die einen Akteur in die Lage versetzen, Probleme zu erkennen und in Form einer geeigneten Handlung einer Lösung zuzuführen. Im Zuge dieser „kognitiven Revolution“ (Gardner 1995; Sturm und Gundlach 2013) kam es zu einer wichtigen Abgrenzung, durch die sich Kognition und die Kognitionswissenschaft jedoch auch enger fassen lassen. In der ersten Hälfte des Jahrhunderts war die Psychologie nicht mehr uneingeschränkt die „Science of Mental Life“ gewesen, als die noch William James sie intendiert hatte (1890, 1), sondern zu einer „science of behavior“ geworden (Skinner 1953, Kap. 2). In diesem historischen Kontext betonte das Adjektiv „kognitiv“ den Unterschied zwischen einer rein behavioristischen Reiz-Reaktions-Psychologie einerseits und einer kognitiven Psychologie andererseits, die den Reizen und Reaktionen zur Erklärung intelligenten Verhaltens ausdrücklich informationsverarbeitende Strukturen im Gehirn – kognitive Prozesse – zwischenschaltete (z.B. Miller et al. 1960; Neisser 1967), weshalb man auch vom „Sandwichmodell“ von Kognition spricht (Hurley 1998): Das zentrale kognitive System wird von den peripheren sensorischen Input- und motorischen Outputsystemen eingerahmt wie der Belag eines Sandwichs von den Weißbrotscheiben (Fodor 1983). Da den zentralen informationsverarbeitenden Prozessen interne Repräsentationen zugrunde liegen sollten, lag es mit dem Aufkommen der Computerwissenschaft nahe, Kognition in Analogie zur Arbeitsweise von Computern zu verstehen, also im Sinne informationsverarbeitender Systeme, die durch syntaktisch spezifizierte algorithmische Berechnungsprozesse (*computations*) über interne Repräsentationen einen Input in einen Output überführen. Unter Kognition wurde in diesem Kontext daher die Verarbeitung von Information mittels regelgeleiteter Transformationen interner mentaler Repräsentationen durch geeignete Berechnungsprozesse verstanden. Entsprechend war die Kognitionswissenschaft zuständig für die Erforschung dieser Transformationsprozesse und umfasste insbesondere jene Disziplinen, die geistige Leistungen als Sache der funktionalen Struktur komplexer Systeme auffassen und diese sowohl auf der Ebene ihrer materiellen Implementation als auch auf abstrakteren computational-repräsentationalen Ebenen in den Blick nehmen können (Marr 1982), konkret die Anthropologie, Informatik, Linguistik, Neurowissenschaft, Philosophie und Psychologie (Miller 2003).

Leitmotiv der Philosophie der Kognition bzw. der Kognitionswissenschaft (Shapiro 2011; Walter 2014a) ist die Frage, wie Systeme strukturiert sein müssen, um kognitive Leistungen wie Wahrnehmen, Erinnern, Lernen, Schlussfolgern, Planen, Entscheiden usw. erbringen zu können. Zusätzlich zu den für die unterschiedlichen Leistungen jeweils spezifischen philosophischen Detailfragen stellt sich dabei eine Reihe von grundlegenden allgemeinen Fragen, insbesondere im Hinblick auf die zentralen Begriffe der *Berechnung* und der *Repräsentation*.

(1) Akzeptiert man die klassische, enge, Auffassung von Kognition, wonach kognitive Prozesse Berechnungsprozesse über interne mentale Repräsentationen sind, dann muss geklärt wer-

den, welche Arten von Repräsentationen bzw. Berechnungsprozessen maßgeblich sind: Ist Kognition im Sinne des traditionellen „Computermodells des Geistes“ eine Sache globaler Regeln folgender Berechnungsprozesse über Repräsentationen symbolischer Art, wie sie auch in digitalen Computern zu finden sind? Oder beruhen kognitive Leistungen stattdessen im Sinne des Konnektionismus auf Aktivierungsmustern in hochgradig vernetzten Verbänden einfacher Verarbeitungseinheiten, sogenannten neuronalen Netzen, die mit subsymbolischen Repräsentationen arbeiten und zwar keinen explizit vorgegebenen globalen Regeln folgen, aber dennoch Computer, das heißt „Berechner“, sind, die durch lokale Rechenoperationen an den einzelnen „Knoten“ einen Input in einen Output überführen (s. Abs. 2.1)?

(2) Klärungsbedürftig ist auch das Verhältnis der beiden zentralen Begriffe der Berechnung und der Repräsentation. Auf der einen Seite binden manche im Zuge eines „semantischen“ Berechnungsbegriffs (Piccinini 2012) Berechnung von vornherein an die Verarbeitung von Repräsentationen – weshalb zum Beispiel Fodor (1975) argumentierte, jedes plausible Modell des menschlichen Geistes sei auf die Existenz interner Repräsentationen festgelegt. Auf der anderen Seite scheinen viele Kognitionswissenschaftler aber gerade mit einem Berechnungsbegriff zu arbeiten, der nicht voraussetzt, dass es Repräsentationen – zumindest nicht in so etwas wie Fodors Sinne – gibt (z.B. Milkowski 2013, Kap. 4; Piccinini 2008).

(3) Dies wiederum wirft unmittelbar die Frage möglicher Abstufungen im Repräsentationsbegriff auf: Ist einzig ein starker Repräsentationsbegriff im Sinne abstrakter, modalitätsunspezifischer und handlungsunabhängiger Repräsentationen maßgeblich oder gibt es auch schwächere Repräsentationsbegriffe (z.B. Gallagher 2008), die mit weniger voraussetzungsreichen, explizit handlungsbezogenen und mitunter modalitätsspezifischen sensomotorischen Repräsentationen arbeiten (z.B. Clark und Toribio 1994), und wenn ja, inwieweit sind diese explanatorisch fruchtbar (s. Abs. 2.3)?

(4) Ob Repräsentationen im Sinne eines schwächeren Repräsentationsbegriffs als Repräsentationen *sensu stricto* gelten können, ist auch im Zusammenhang mit der Frage relevant, ob man das klassische, enge, Bild von Kognition nicht womöglich aufgeben oder zumindest stark modifizieren sollte, weil Kognition mit Berechnungen und Repräsentationen wenig bis überhaupt nichts zu tun hat und im Rahmen eines mehr oder weniger radikalen Anticomputationalismus (z.B. van Gelder 1995) bzw. Antirepräsentationalismus (z.B. Hutto und Myin 2013) vielmehr eine Sache der dynamischen Gesamtorganisation des Gehirns, des übrigen Körpers und der aktiven Interaktion mit der Umwelt ist (Chemero 2009; s. Abs. 2.2).

(5) Schließlich wird vor dem Hintergrund der allgemeineren Auffassung von Kognition als demjenigen, was uns zu intelligentem Verhalten befähigt, unter dem Stichwort „sitierte Kognition“ (z.B. Robbins und Aydede 2009) seit geraumer Zeit diskutiert, ob sich die für die kognitiven

Leistungen eines Systems maßgeblichen Prozesse – ganz unabhängig davon, welche Kombination aus Computationalismus und Repräsentationalismus, Anticomputationalismus und Anti-repräsentationalismus für ihr Verständnis erforderlich ist – auf Strukturen in einer zentralen Verarbeitungseinheit, beim Menschen dem Gehirn, beschränken oder mitunter vielmehr den gesamten Körper umfassen oder sich sogar in die Umwelt hinein erstrecken (s. Abs. 2.2).

Der Philosophie der Kognition geht es demnach sowohl um das *Was* kognitiver Prozesse, also darum, ob und in welchem Sinne sie computational bzw. repräsentational sind, als auch um ihr *Wo*, das heißt darum, ob Kognition ausschließlich eine Sache von Prozessen in einer zentralen Verarbeitungseinheit wie dem Gehirn ist oder sich auch in den übrigen Körper oder die Umwelt eines Systems hinein erstreckt bzw. als relationales Phänomen überhaupt erst aus der wechselseitigen Interaktion mit der Umwelt hervorgeht.

## **1.2 Philosophie der Kognition, der Psychologie und der Neurowissenschaft**

So etwas wie eine Philosophie der Kognition konnte erst entstehen, als man Mitte des 20. Jahrhunderts in der Nachfolge der „kognitiven Revolution“ begann, aus interdisziplinärer Perspektive jenen Aspekt unserer *conditio humana* empirisch und begrifflich in den zu Blick nehmen, der unsere Fähigkeit zur intelligenten Problemlösung betrifft. Die Philosophie der Kognition bzw. der Kognitionswissenschaft schließt daher zwar in Teilen inhaltlich an das an, was als „Philosophie des Geistes“ bzw. „Philosophie der Psychologie“ bezeichnet wird (s. Kap. V.18), ist aber deutlich spezifischer. Die Probleme rund um die umfassenden metaphysischen Fragen nach dem Verhältnis von Gehirn und Geist, von Leib und Seele, nach Willensfreiheit, Qualia und intentionalem Gehalt etwa, die traditionell Gegenstand der Philosophie des Geistes sind, spielen in der Philosophie der Kognition in dem Maß keine oder allenfalls eine untergeordnete Rolle, wie sie für ihre zentrale Leitfrage, welche Teile eines Systems und seiner Umwelt auf welche Weise zu seinem intelligenten Verhalten beitragen, irrelevant sind (für ein liberaleres Verständnis, das derartige Fragen der Philosophie der Kognition zurechnet, vgl. Boden 2006).

Während die Philosophie der Kognition mit ihrem Fokus auf kognitiven Leistungen also enger ist als die klassische Philosophie des Geistes mit ihrem umfassenden Interesse an allen Aspekten geistigen Lebens, ist sie zugleich in zweierlei Hinsicht breiter als das, was oftmals als „Neurophilosophie“ oder „Philosophie der Neurowissenschaft“ bezeichnet wird (s. Kap. III.13).

Zum einen ist die Philosophie der Neurowissenschaft naturgemäß auf Systeme beschränkt, deren geistiges Leben eben auf die eine oder andere Weise in *neuronalen* Strukturen verankert ist, während es aus Sicht einer computational/repräsentationalen Auffassung von Kognition zunächst einmal gerade gleichgültig ist, wie die entsprechenden Berechnungsprozesse und Repräsentationen materiell implementiert sind. Unter anderem deshalb spielte die Erforschung

der unseren kognitiven Leistungen zugrunde liegenden neuronalen Mechanismen in der Kognitionswissenschaft zunächst eine eher untergeordnete Rolle. Heute sind die Neurowissenschaften zwar ein ebenso zentraler Bestandteil der Kognitionswissenschaft wie die kognitive Psychologie, die Künstliche-Intelligenz-Forschung oder die Informatik, können aber eben nur einen Ausschnitt des für die Kognitionswissenschaft Relevanten beleuchten. Marr (1982) betonte die Notwendigkeit des Zusammenspiels verschiedener Disziplinen im Rahmen seiner Unterscheidung von drei Analyseebenen, auf denen informationsverarbeitende Systeme beschrieben und ihre Funktions- und Arbeitsweise erklärt werden können. Auf der obersten Ebene, der Rechenebene (*computational level*), wird eine computationale Beschreibung der Aufgabe erstellt (etwa die Addition zweier Zahlen oder die Umwandlung der Retinastimulation in eine dreidimensionale Wahrnehmung). Dazu wird der für die Lösung der Aufgabe erforderliche Berechnungsprozess in Form einer mathematischen Funktion beschrieben und gezeigt, dass sich mit dieser Funktion die Aufgabe lösen lässt. Auf der mittleren, der algorithmischen, Ebene (*level of representation and algorithm*) wird das Repräsentationsformat von Input und Output spezifiziert und angegeben, mittels welcher Algorithmen das System die entsprechende Funktion berechnet. Auf der untersten Ebene, der Implementationsebene (*level of hardware implementation*), schließlich muss gezeigt werden, wie die verwendeten Repräsentationen und Algorithmen in einem konkreten System implementiert sind und wie die postulierten Mechanismen die ihnen zugeschriebene Funktion erfüllen können (etwa mittels Neuronen, Siliziumchips usw.). Je nach Problem mag dabei ein *top-down*- oder ein *bottom-up*-Ansatz angezeigt sein (s. Abs. 2.1), und in beiden Fällen kann die Neurowissenschaft zu einem besseren Verständnis beitragen, alleine kann sie aber eben nicht alle relevanten Aspekte beleuchten.

Zum anderen laufen „situiertere“ Ansätze im Zuge einer liberaleren Auffassung von Kognition darauf hinaus, dass kognitive Prozesse gerade nicht auf Vorgänge im Gehirn beschränkt sind, sondern mitunter auf die eine oder andere näher zu spezifizierende Weise den gesamten Körper und sogar die natürliche, technische oder soziale Umwelt umfassen (s. Abs. 2.2). Auch in diesem Fall kann die Neurowissenschaft zwar wichtige Einsichten bereitstellen, aber eben nicht die komplette „kognitive Maschinerie“ erfassen.

Vor dem Hintergrund der in Abschnitt 1 skizzierten Überlegungen geht Abschnitt 2 auf einige wichtige primär ontologische Fragen ein, die sich ergeben, wenn man ein „gehirnzentriertes“ Bild von Kognition zugunsten der Vorstellung aufgibt, dass sich Kognition mitunter über das Gehirn hinaus in den übrigen Körper oder die Umwelt hinein erstreckt.

## **2 Situiertere Kognition: die Rolle von Körper und Umwelt**

Spätestens gegen Ende der 1980er Jahre geriet die unter anderem durch das klassische Sandwichmodell verkörperte Auffassung, dass Kognition auf die Vorgänge in einer zentralen Verarbeitungseinheit wie dem Gehirn beschränkt ist, nachhaltig in die Kritik. Angestoßen durch konvergierende Erkenntnisse in verschiedenen Bereichen der Kognitionswissenschaft – unter anderem der Robotik (z.B. Brooks 1991), der Linguistik (z.B. Lakoff/Johnson 1980, 1999) und der Entwicklungspsychologie (z.B. Thelen/Smith 1994) – begann sich zu dieser Zeit die Auffassung durchzusetzen, dass weder die auf Computation und Repräsentation setzenden traditionellen Positionen (s. Abs. 2.1) noch dezidiert anticomputationalistische und antirepräsentationalistische Alternativen kognitive Leistungen umfassend erklären können, solange sie nicht der Tatsache Rechnung tragen, dass Kognition ganz wesentlich auch vom übrigen Körper eines Systems sowie seiner aktiven Einbettung in seine natürliche, technische und soziale Umwelt abhängt (s. Abs. 2.2): Intelligente Akteure sind keine isolierten Denker, deren Verhalten ausschließlich auf der wiederholten Abarbeitung starrer Wahrnehmen-Denken-Handeln-Zyklen beruht, und Kognition beschränkt sich nicht auf das Gehirn, sondern ist im gesamten Körper und der Umwelt *situiert* (Robbins und Aydede 2009; Shapiro 2011; Walter 2014a). Diese Situiertheitsthese gilt inzwischen als Wesensmerkmal einer „Kognitionswissenschaft zweiter Generation“, deren ausdrücklich handlungsbezogener Ansatz das fleischlose Credo des „Ich denke“ der ersten Generation ersetzt hat. Es handelt sich dabei bislang jedoch nicht um ein wohldefiniertes Forschungsprogramm (s. Abs. 3.1), sondern allenfalls um einen losen Verbund philosophischer Erwägungen, empirischer Studien, psychologischer Modelle und kognitionswissenschaftlicher Anwendungen, die einzig das Ideal eines ganzheitlichen Ansatzes von Gehirn, Körper und Umwelt eint, ohne dass dabei immer klar würde, wie die diversen Vorstellungen, denen oft nachgesagt wird, als „Konterrevolution“ der kognitiven Revolution der Anfangsjahre mit dem Versuch der Etablierung eines „neues Paradigmas“ (Stewart et al. 2010) begegnet zu sein, im Detail auszubuchstabieren sind (s. Abs. 2.3 und 2.4) und ob es sich dabei tatsächlich um eine grundlegende Revolution oder doch bestenfalls um Perspektivenverschiebungen und graduelle Revisionen handelt (z.B. Villalobos 2013).

## **2.1 Die Kognitionswissenschaft erster Generation: das Computermodell des Geistes und der Konnektionismus**

Ihren historisch einflussreichsten Ausdruck fand die klassische Konzeption von Kognition in der Vorstellung, dass Kognition ein Prozess der formalen Regeln folgenden Transformation symbolischer Strukturen, das heißt eine Form von Informationsverarbeitung – genauer gesagt: Symbolverarbeitung – in einem physisch realisierten Symbolsystem (*physical symbol system*; Newell und Simon 1976) ist. Kognitive Prozesse sind demnach syntaktischen Regeln folgende sequenzielle Berechnungsprozesse über symbolische mentale Repräsentationen, die in Analogie zu Computern einen Input in einen Output überführen. Der Geist ist entsprechend so et-

was wie ein Programm, das heißt die durch computational-repräsentationale Prozesse in der Hardware des Gehirns implementierte Software. Unter dem Einfluss dieses Computermodells des Geistes war die Kognitionswissenschaft methodologisch zunächst von einem *top-down*-Ansatz geprägt, der mit computationalen Modellen begann und dann über die algorithmische Zergliederung kognitiver Leistungen zur Ebene ihrer materiellen Implementierung voranschritt. Anfänglich lag das Augenmerk daher vor allem auf sogenannten „deterministischen Problemen“, für die sich eine entsprechende computationale Analyse anbietet, weil Anfangs- und Zielzustände klar definiert und die Problemlöseoperatoren überschaubar sind, zum Beispiel deduktives Schließen, Theorembeweisen oder die Lösung von Aufgaben wie dem Turm-von-Hanoi-Problem. Da Menschen bei der Anwendung logischer Schlussregeln und systematischen Suchen in großen Problemräumen schnell an ihre Grenzen stoßen, waren digitale Computer dem Menschen auf diesen Gebieten bald überlegen. Umgekehrt jedoch kamen Computer sehr viel schlechter mit unvollständigem, mehrdeutigem und vagem Input zurecht und hatten Schwierigkeiten bei Generalisierungen (Induktion) und der Hypothesenbildung (Abduktion), sodass sie zum Beispiel im Hinblick auf Leistungen wie Wahrnehmung, Kategorisierung oder die Verarbeitung natürlicher Sprache dem Gehirn deutlich unterlegen waren.

In den 1980er Jahren begann vor diesem Hintergrund ein alternativer *bottom-up*-Ansatz (wieder) Erfolge zu feiern, der das klassische Computermodell für physiologisch unrealistisch hielt, stattdessen davon ausging, dass das Gehirn einige kognitive Leistungen gerade deshalb erbringen kann, weil es eben *keine* sequenzielle Symbolverarbeitung betreibt und genau auf jenen Gebieten erfolgreich war, auf denen digitale Computer und der *top-down*-Ansatz des Computermodells ihre Schwächen hatten: Statt aus computationalen Modellen gewonnene Regeln und Repräsentationen explizit in seriellen Systemen symbolisch zu kodieren, setzte man unter dem Stichwort *Parallel Distributed Processing* (Rumelhart et al. 1986) in Anlehnung an informationsverarbeitende Strukturen im Gehirn auf parallel arbeitende künstliche neuronale Netze, die sich die erforderlichen Regeln und Repräsentationen aufgrund ihrer Verbindungsstruktur implizit selbst aneignen. Dieser sogenannte Konnektionismus (*connection*: Verbindung) modifizierte die Grundannahmen des Computermodells, gab sie aber nicht vollends auf: Neuronale Netze arbeiten zwar nicht seriell, sind aber dennoch informationsverarbeitende Systeme, die Eingabe- in Ausgabevektoren überführen, und die Aktivierungsmuster in solchen Netzen gehen zwar nicht auf explizite globale Regeln zurück, erfordern lokal aber immer noch regelgeleitete, parallel an den einzelnen Knoten eines Netzes stattfindende Operationen und fungieren auf diese Weise als „verteilte“ bzw. „subsymbolische“ Repräsentationen (Hinton et al. 1986; Smolensky 1988).

Das klassische Computermodell und der Konnektionismus waren sich auch in einer anderen entscheidenden Hinsicht einig: Als Kognition galt beiden dasjenige, was in einer zentralen Ver-

arbeitungseinheit wie dem Gehirn einen Input in einen Output überführt, während die Prozesse jenseits davon, im übrigen Körper oder der Umwelt, bestenfalls als Quelle des Inputs bzw. als Schauplatz des Outputs dienen, aber kein Teil der eigentlichen „kognitiven Maschinerie“ waren. Zum Bruch mit dieser traditionellen Vorstellung kam es erst mit dem Dynamizismus und situierten Ansätzen, die heutzutage der „Kognitionswissenschaft zweiter Generation“ zugeordnet werden.

## **2.2 Die Kognitionswissenschaft zweiter Generation: situierte Kognition**

Der traditionelle Fokus auf Berechnung und Repräsentation wurde in den 1990er Jahren zunächst vom Dynamizismus kritisiert (z.B. Kelso 1995; Thelen und Smith 1994; van Gelder 1995, 1998), der sowohl den digitalen Computer als auch neuronale Netze als Modell des Geistes zurückwies und Kognition stattdessen als Prozess in dynamischen Systemen verstand, für den Berechnungen und Repräsentationen – ganz gleich, ob global oder lokal, symbolisch oder subsymbolisch – nebensächlich bzw. irrelevant sind. Kognition ist für den Dynamizismus vielmehr etwas, das im Rahmen der reziproken Echtzeitinteraktion eines körperlich auf charakteristische Weise verfassten Systems mit einer entsprechend strukturierten Umwelt entsteht. Als derart dynamischer Prozess lässt sich Kognition überhaupt nicht in einzelne, prinzipiell separierbare Berechnungsschritte aufteilen und daher statt mit computational-repräsentationalen Modellen sehr viel besser mit Differenzialgleichungen beschreiben und mit den Methoden der Theorie dynamischer Systeme erklären (vgl. Walter 2014a, Kap. 4).

Die für das klassische Sandwichmodell von Kognition charakteristische Dreiteilung in sensorischen Input, motorischen Output und die „eigentliche kognitive Maschinerie“ dazwischen sowie die damit einhergehende Degradierung von Körper und Umwelt geriet in etwa zur selben Zeit aber auch ganz unabhängig vom Dynamizismus in die Kritik, als sich in weiten Teilen der Kognitionswissenschaft die Auffassung durchzusetzen begann, dass weder die auf Computation und Repräsentation setzenden traditionellen Positionen noch der Anticomputationalismus und Antirepräsentationalismus des Dynamizismus erfolgreich sein werden, solange sie ignorieren, dass Kognition in dem Sinne *situiert* ist, dass kognitive Prozesse wesentlich auch vom gesamten Körper eines Systems sowie von seiner aktiven Einbettung in seine natürliche, technische und soziale Umwelt abhängen. Im Kern geht es Anhängern der unterschiedlichen Spielarten dieses neuen Ansatzes darum, dass Kognition eben keine rein neuronale Angelegenheit ist, die losgelöst von Körper und Umwelt, gewissermaßen „offline“, zwischen sensorischen Eingangs- und motorischen Ausgangssignalen stattfindet, sondern vor allem auch „online“ im Zuge der reziproken Interaktion mit der Umwelt entsteht. Insbesondere die folgenden fünf Positionen haben sich dabei im Laufe der Debatte herauskristallisiert (Walter 2014a,b).



(1) *Die Rolle des Körpers: embodied cognition.* Oft wird betont, dass Kognition auch in der spezifischen körperlichen Verfasstheit eines Systems gründet, das heißt verkörperlicht (*embodied*) ist. Diese Verkörperlichungsthese soll sich nicht in einem vergleichsweise trivialen Einfluss des Körperlichen dergestalt erschöpfen, dass unser geistiges Leben anders aussähe, wenn wir uns zum Beispiel mittels Echoortung orientierten oder sich unsere Ohren statt seitlich am Kopf nebeneinander auf dem Rücken befänden. Dem Körper soll vielmehr eine wesentliche Rolle in unserem geistigen Leben zukommen. Insbesondere soll er nicht länger als bloßes Outputvehikel fungieren, das als kontingente „Behausung“ des eigentlichen kognitiven Systems lediglich als Instrument zur Überführung interner psychischer Vorgänge in Handlungen dient. Der Körper soll vielmehr aufgrund seiner spezifischen materiellen Beschaffenheit selbst aktiv zur energie- und berechnungseffizienten Lösung von Problemen beitragen, etwa indem die biologischen, physiologischen oder morphologischen Details eines Systems Funktionen übernehmen, die traditionell internen Repräsentationen und zentralen Kontrollprozessen zukamen (*organic* oder *morphological computing*; z.B. Pfeifer und Bongard 2007), oder indem abstrakte, vermeintlich amodale kognitive Prozesse – der „Belag“ des klassischen Sandwichs – modalitätsspezifische sensomotorische Repräsentationen aus den peripheren Systemen „wiederverwerten“ (*shared circuits* oder *neural reuse*; z.B. Goldman 2012; Hurley 2008).

(2) *Die Rolle der Umwelt (I): embedded cognition.* Anhängern situierter Kognition geht es jedoch nicht nur um den Körper, sondern auch darum, dass kognitive Systeme mithilfe ihres Körpers ihre Umwelt so ausnutzen können, dass diese nicht mehr nur passiver Schauplatz eines intern bereits in allen Details durchgeplanten behavioralen Outputs ist, sondern wiederum selbst aktiv zur Genese intelligenten Verhaltens beiträgt. Kognition ist also mitunter in dem Sinne von unserer Umwelt abhängig oder situativ eingebettet (*embedded*), dass ein Akteur sie als Gerüst (*scaffold*) nutzt, um seinen internen kognitiven Aufwand möglichst gering zu halten. Professionelle Barmixer etwa assoziieren verschiedene Drinks mit bestimmten Glasformen, Dekorationen usw. und können so durch geeignete Anordnung umfangreiche Bestellungen in der richtigen Reihenfolge abarbeiten – eine Fähigkeit, die sie einbüßen, sobald sie einheitliches Zubehör verwenden und die Abfolge in Ermangelung eines externen *scaffolds* „im Kopf“ behalten müssen (Clark 2001, 141; vgl. auch Beach 1993). Die durch die körperliche Verfasstheit eines Systems ermöglichte und strukturierte Echtzeitinteraktion mit der Umwelt macht dabei die exakte und umfassende Repräsentation relevanter Weltausschnitte, der traditionell eine so zentrale Rolle zukam, zumindest zum Teil verzichtbar, indem sie eine Situation schafft, in der wir wie etwa der Barmixer nicht mehr „im Kopf“ mit internen Repräsentationen jonglieren müssen, sondern die relevanten Informationen bei Bedarf der Welt selbst entnehmen können. Kognition soll dabei jedoch nur von der situativen Einbettung eines Akteurs in seine Umwelt *abhängen*, die Erinnerungen des Barmixers etwa werden aber offenbar nicht durch externe Faktoren wie die Anordnung der Gläser mit *konstituiert*.

(2) *Die Rolle der Umwelt (II): extended cognition.* Wenn der Umwelt im Hinblick auf unsere kognitiven Leistungen aber tatsächlich eine so wichtige Rolle zukommt, mit welchem Recht betrachten wir das Gehirn dann noch als alleinigen Konstituenten von Kognition? Was qualifiziert Prozesse im Gehirn (oder Körper) als kognitive Prozesse *sensu stricto*, während die Umwelt Kognition nur unterstützt und ermöglicht, selbst aber nicht als kognitiv gilt? Sind kognitive Prozesse stattdessen nicht in dem Sinne erweitert (*extended*), dass sie sich unter bestimmten Bedingungen auch über die organismischen Grenzen eines Akteurs hinaus in die Umwelt hinein erstrecken, indem sie nicht nur durch neuronale, sondern auch durch Prozesse jenseits der organismischen Grenzen konstituiert werden? Wenn zum Beispiel schriftliche Notizen das Verhalten eines Alzheimerpatienten auf dieselbe Weise leiten, wie neuronal abgespeicherte Gedächtnisinhalte unser Tun, wäre es dann nicht ein Zeichen „bio-chauvinistischer Voreingenommenheit“ (Clark 2008,77), weigerten wir uns, sie ebenso als genuinen Teil der materiellen Konstituenten seiner Erinnerungen, Überzeugungen usw. anzuerkennen, wie wir es bei neuronalen Prozessen in unserem Fall ganz selbstverständlich tun (Clark und Chalmers 1998)?

(4) *Die Rolle des Sozialen: distributed cognition.* In eine ähnliche Richtung geht die Vorstellung, dass Kognition mitunter in dem Sinne verteilt (*distributed*) ist, dass sich kognitive Prozesse über die Grenzen von Individuen hinaus in soziale Komplexe aus interagierenden Akteuren und ihren technischen Ressourcen erstrecken. Die Navigation eines Schiffes etwa konnte vor dem GPS-Zeitalter von einem Einzelnen überhaupt nicht bewältigt werden, sondern war insofern „verteilt“ als der Navigationsprozess die koordinierte Interaktion mehrerer Akteure und ihrer technischen Hilfsmittel erforderte und dabei insbesondere auch von den Eigenheiten der sozialen Hierarchie der Beteiligten abhing (Hutchins 1995). Auch solche „verteilten“ kognitive Prozesse sollen wie „erweiterte“ weder auf das Gehirn noch den übrigen Körper beschränkt sein, sondern sich in die Umwelt eines Individuums hinein erstrecken (s. Abs. 2.4).

(5) *Die Rolle der Interaktion: enactivism.* Der sogenannte Enaktivismus schließlich versteht Kognition als ein Merkmal lebendiger Organismen, das erst in der aktiven Interaktion autonomer und adaptiver Systeme mit ihrer Umgebung hervorgebracht (*enacted*) wird. Höherstufige kognitive Leistungen sollen eine Weiterentwicklung jenes grundlegenden Prozesses einer Sinnstiftung (*sense-making*) sein, durch den einfachste Lebewesen die sie umgebende Welt perspektivisch in einen Ort von subjektiver Bedeutung und damit in eine *Umwelt* im eigentlichen Sinne transformieren (Thompson und Stapleton 2009). Der Enaktivismus gibt das Computermodell des Geistes auf (Varela et al. 1991) und betont die sogenannte „Kontinuitätsthese von Leben und Geist“ (*continuity thesis of life and mind*), wonach lebendige Systeme kognitive Systeme sind und der Prozess des Lebens ein Prozess der Kognition ist (Maturana und Varela 1980). Hinter dieser zunächst eigenwillig erscheinenden Gleichsetzung steht die Überlegung, dass sich Kognition als Prozess der Sinnstiftung aus der Interaktion autonomer und adaptiver

Systeme mit ihrer Umgebung ergibt und lebendige Organismen immer schon im entsprechenden Sinne autonom und adaptiv sind. Leben und Kognition sind daher insofern eins, als höhere kognitive Leistungen Prinzipien folgen, unter die auch schon einfachste Lebensformen fallen – auch wenn Kognition in Menschen natürlich anders ausgeprägt ist als etwa in einzelligen oder Amöben. Kognition ist für den Enaktivismus also eine ganzheitliche Sache einer der verkörperlichten Interaktion eines Individuums mit seiner Umwelt und hat ihren Platz mithin nicht im Gehirn, ist aber weder erweitert noch verteilt, sondern ein wesentlich relationales Phänomen (s. Abs. 2.4).

Einige der von Anhängern situierter Kognition hervorgehobenen Punkte wurden auf die eine oder andere Weise bereits im Zusammenhang mit dem Dynamizismus aufgeworfen. Allerdings laufen situierte Ansätze insofern quer zu der Debatte zwischen Computationalisten und Dynamizisten, als situierte Ansätze oft sowohl computationalistisch als auch dynamizistisch motiviert sind: Während die Vorstellung verteilter kognitiver Prozesse in der Regel vor dem Hintergrund einer computationalistischen Konzeption von Kognition entwickelt wird (Hutchins 1995) und der Enaktivismus ausdrücklich einen Antirepräsentationalismus vertritt (Varela et al. 1991, Kap. 7), finden sich die Vorstellungen, dass kognitive Prozesse verkörperlicht, situativ eingebettet oder erweitert sind, sowohl im Kontext computationalistisch/repräsentationalistischer Ansätze als auch im Zuge anticomputationalistisch/antirepräsentationalistischer Überlegungen. In der Philosophie der Kognition geht es aktuell mithin um zwei verschiedene und zumindest grundsätzlich unabhängige Fragen, die bislang allzu oft nicht klar genug unterschieden werden. Auf der einen Seite sind sich Computationalisten und Dynamizisten uneins über das *Was* kognitiver Prozesse, also darüber, wie kognitive Systeme am besten zu beschreiben sind: Hat Kognition wesentlich etwas mit Berechnung und Repräsentation zu tun, was auch viele Anhänger der zweiten Generation akzeptieren, oder sind die Begriffe der Berechnung und Repräsentation nachrangig oder gar völlig verzichtbar, wie etwa der Dynamizismus oder der Enaktivismus behaupten? Auf der anderen Seite geht es um das *Wo* von Kognition, also darum, wo in der Welt kognitive Prozesse – ganz gleich, ob sie klassisch oder dynamizistisch beschrieben werden müssen – zu finden sind und in welchem Verhältnis der „kognitive Kern“ eines Systems zu „kognitiven Ressourcen“ in seinem übrigen Körper und der Umwelt steht: Sind kognitive Prozesse ausschließlich eine neuronale Angelegenheit oder können sie sich auch in den übrigen Körper und die Umwelt eines Systems hinein erstrecken bzw. als wesentlich relationales Phänomen erst aus der wechselseitigen Interaktion mit der Umwelt hervorgehen? In beiden Bereichen gibt es trotz unbestreitbarer Erkenntnisfortschritte noch stets eine Reihe offener grundlegender Fragen, insbesondere mit Blick auf die vermeintlich trennscharfen, in der Praxis aber erstaunlich diffusen Unterscheidungen zwischen den verschiedenen Lagern in den Debatten um das *Was* (s. Abs. 2.3) und das *Wo* (s. Abs. 2.4) kognitiver Prozesse.

## 2.3 Antirepräsentationalismus versus Repräsentationalismus

Der Dynamizismus hat sich in der Anfangszeit vor allem durch seine Repräsentationsskepsis als vermeintlich radikaler Gegenpol zu den klassischen Ansätzen etabliert. Ein berühmtes Beispiel ist der Fliehkraftregler, der sogenannte Watt-Governor, ein Regelkreis mit negativer Rückkoppelung, mit dessen Hilfe das Schwungrad einer Dampfmaschine trotz schwankender Dampfmenge konstant angetrieben werden kann. Am Schwungrad wird ein vertikales Gestänge mit zwei beweglichen Armen angebracht, die sich aufgrund ihres Gewichts und der Zentrifugalkraft heben und senken, wenn das Schwungrad schneller bzw. langsamer wird. Die Arme sind über einen Hebelmechanismus so mit einem Ventil verbunden, dass der Dampfzufluss abnimmt, wenn sie sich heben, und zunimmt, wenn sie sich senken. Dadurch pendelt sich das Schwungrad auf der gewünschten Drehzahl ein: Wird die Maschine zu langsam, senken sich die Arme, weshalb der Dampf zunimmt und die Drehzahl steigt; mit steigender Drehzahl aber heben sich die Arme wieder, sodass der Dampf abnimmt und die Maschine langsamer wird usw. Laut van Gelder (1995, 351–354) kann man nicht sagen, dass der Watt'sche Fliehkraftregler die erforderliche Anpassung des Ventils in irgendeinem substantziellen Sinne *berechnet*, da es unter anderem gar keine *Repräsentationen* von Größen wie Drehzahl, Dampfmenge usw. gibt, die als Grundlage eines Berechnungsprozesses dienen könnten. Dagegen hat zum Beispiel Bechtel (1998) eingewendet, dass der Winkel der Arme systematisch mit der Geschwindigkeit variiert, daher das Ventil über die aktuelle Geschwindigkeit informieren kann und diese mithin sehr wohl *repräsentiert*, während van Gelder seine Repräsentationsskepsis mit dem Hinweis darauf verteidigt, dass erstens Korrelation nicht Repräsentation ist, zweitens Winkel und Geschwindigkeit nur annähernd korrelieren, weil die Arme nur verzögert auf Geschwindigkeitsänderungen reagieren, und drittens die Geschwindigkeit ebenso vom Winkel abhängt wie umgekehrt, aber nicht beides das jeweils andere repräsentieren kann (1995, 352–353). Wie auch immer dieser Disput konkret zu entscheiden sein mag, er deutet auf ein grundsätzliches Problem hin: Offenbar lässt sich immer ein entsprechend schwacher Repräsentationsbegriff formulieren, der auch auf dynamische Systeme und andere vermeintlich repräsentationsfreie kognitive Systeme – etwa die „Mrobots“ von Brooks (1991) – anwendbar ist (z.B. Markman und Dietrich 2000). Das zeigt jedoch nur, dass die *metaphysische* These des Dynamizismus, dass kognitive Systeme dynamische Systeme *sind* (van Gelder 1998), uninteressant ist: So wie grundsätzlich jedes, *eo ipso* also auch jedes kognitive System als dynamisches System beschreibbar ist, so lässt sich mit einem entsprechend schwachen Repräsentationsbegriff jedes kognitive System als repräsentationales System beschreiben. In diesem Sinne „sind“ kognitive Systeme also ebenso dynamisch wie repräsentational. Viel entscheidender ist die *explanatorische* These, dass sich kognitive Systeme mit dynamizistischen Begriffen und Methoden so beschreiben lassen, dass repräsentationale Beschreibungen explanatorisch überflüssig werden: Eine repräsentationale Beschreibung des Fliehkraftreglers zum Beispiel

hat anscheinend keinen zusätzlichen Erklärungswert, da sich sein Zustand ja für jeden beliebigen Zeitpunkt mittels entsprechender Entwicklungsgleichungen exakt vorhersagen lässt.

Aus der bloßen Tatsache, dass computational/repräsentationale Modellierungsansätze beim Fliehkraftregler keinen explanatorischen Mehrwert haben, folgt allerdings natürlich nicht, dass der Dynamizismus generell richtig und der Fliehkraftregler, wie van Gelder (1995, 381) behauptet, als Paradigma für die Arbeitsweise kognitiver Systeme einschlägiger ist als ein Computer, denn warum sollte das, was für den Fliehkraftregler gilt, auch für (alle) kognitive(n) Systeme gelten? Mit Blick auf eine Erklärung kognitiver Leistungen scheint der Dynamizismus bestenfalls jene einfachen, nur in einem sehr rudimentären Sinne „kognitiven“ Leistungen modellieren zu können, die sich in dynamischen Echtzeitinteraktionen mit der Umwelt vollziehen, während er höherstufige kognitive Leistungen offenbar nicht einfangen kann: Wie soll ein System zum Beispiel allein durch die Interaktion mit seiner aktuellen Umwelt und ohne Repräsentationen und Planungsprozesse Probleme lösen, die eine Einschätzung kontrafaktischer Sachverhalte erfordern, also etwa beim Schach drei Züge vorausdenken, Ziele für den kommenden Urlaub abwägen oder „im Geiste“ geometrische Formen rotieren (z.B. Edelman 2003)? Der Dynamizismus muss auch solche klassischen kognitiven Fähigkeiten erklären können, möchte er den Vorwurf vermeiden, sich mit seiner Hinwendung zu zweibeinigem Laufen (z.B. Thelen und Smith 1994) oder rhythmischen Fingerbewegungen (z.B. Haken et al. 1985) als Paradebeispielen gar nicht länger mit Kognition *sensu stricto* zu beschäftigen und damit auch keine Alternative zu traditionellen Ansätzen darzustellen, sondern schlicht das Thema zu wechseln und statt Kognition wieder ausschließlich die einfachen Reiz-Reaktions-Schemata des Behaviorismus in den Blick zu nehmen. Die reaktiven Agenten der verhaltensbasierten Robotik zum Beispiel, welche die potenzielle Schlichtheit der Grundlagen intelligenten Verhaltens und die Möglichkeit von „*Intelligence without representation*“ (Brooks 1991) demonstrieren sollen, sind zwar computational und repräsentational extrem sparsam, kommen jedoch keinesfalls völlig ohne klassische Elemente aus. Der Nachweis, dass intelligentes Verhalten ganz allgemein ohne jede Art von Berechnung und Repräsentation zuwege gebracht werden kann, steht mithin trotz erheblicher und zweifellos eindrucksvoller Fortschritte in Teilbereichen nach wie vor aus.

Die Frage, warum etwa zweibeinig Laufen oder rhythmische Fingerbewegungen als kognitive Leistungen im eigentlichen Sinne anzusehen sein sollten, ist für den Dynamizismus umso dringlicher, als nicht alle dynamischen Systeme kognitive Systeme sind und der Dynamizismus daher anscheinend kaum einfangen kann, was an kognitiven Systemen eigentlich das charakteristisch Kognitive ist (z.B. Clark 2008, 26). Während computationalistische Ansätze eine klare Vorstellung davon haben, was Kognition ist, das heißt durch welche Art von Mechanismen kognitive Leistungen implementiert sind, und so neue Phänomene vorhersagen und diese Hypothesen überprüfen können, bleibt beim Dynamizismus insbesondere unklar, wie er in Er-

mangelung einer Theorie von Kognition, die über die bloße Beschreibung der zeitlichen Entwicklung kognitiver Systeme mittels Differenzialgleichungen hinausgeht, neue Phänomene vorhersagen und diese Vorhersagen überprüfen kann (vgl. aber Chemero 2009, Kap. 4.4).

Im Hinblick auf die Frage des *Was* liegt die Bringschuld derzeit also noch immer bei anticomputationalistischen und antirepräsentationalistischen Ansätzen. Der an sich berechtigte Vorwurf, die traditionellen Positionen hätten einem von den peripheren sensomotorischen Systemen isolierten zentralen kognitiven System zu viel Gewicht beigemessen, sollte nicht in eine generelle Repräsentationsskepsis umschlagen. Die Vorstellung, dass intelligentes Verhalten maßgeblich durch neuronal realisierte Berechnungsprozesse über interne Repräsentationen erklärt werden kann, gehört auch nach der für die zweite Generation charakteristischen Erweiterung des Forschungsfeldes zu den erfolgreichsten Modellierungsansätzen der Kognitionswissenschaft überhaupt. Die These, dass kognitive Prozesse von wesentlich handlungsbezogenen Repräsentationen Gebrauch machen, verspricht dabei vieles von dem, worauf die zweite Generation zu Recht hingewiesen hat, einzufangen, ohne dass man dafür den begrifflichen und methodologischen Rahmen der ersten Generation vollständig über Bord werfen müsste. Statt eine strikte Opposition von Antirepräsentationalismus und Repräsentationalismus zu propagieren, scheint es daher fruchtbarer zu klären, wie sich die lohnenden Einsichten vermeintlich antirepräsentationalistischer Positionen mit einem handlungsbezogenen Repräsentationsbegriff einfangen und umsetzen lassen, der anspruchsloser als die begrifflichen Repräsentationen des klassischen Computermodells, aber eben immer noch ein Repräsentationsbegriff ist, sodass die zunächst vermeintlich alternativlose Gegenüberstellung der entsprechenden Positionen mit Blick auf das *Was* kognitiver Prozesse sowohl begrifflich als auch empirisch zusehends aufgelöst wird. Ganz ähnliche Abgrenzungsprobleme zwischen anscheinend unverträglichen Positionen gibt es auch für die Debatte um das *Wo* kognitiver Prozesse.

## **2.4 Wie entscheiden wir, ob kognitive Prozesse eingebettet, erweitert, verteilt oder schlicht „nirgendwo“ sind?**

Die Debatte um situierte Kognition bereichert die Kognitionswissenschaft zweifellos: Klassische Probleme wie Planen, Sprachverstehen oder Schlussfolgern sind in den Hintergrund getreten und die Aufmerksamkeit richtet sich stattdessen viel liberaler auf Herausforderungen, welche die aktive Echtzeitinteraktion verkörperter Akteure mit ihrer Umwelt erfordern, etwa die Navigation in belebten Umgebungen oder energieeffiziente und dezentral gesteuerte Fortbewegungsmethoden. Diese Liberalisierung hat bislang jedoch noch nicht zu mehr Klarheit geführt. Zum einen ist die Terminologie nach wie vor ungeklärt. Während einige von situiertem Kognition sprechen (z.B. Robbins und Aydede 2009), verwenden andere den Ausdruck „*situated*“ synonym zu „*embedded*“ (z.B. Shapiro 2010) oder „*extended*“ (z.B. Wilson und Clark

2009) oder „*embodied*“ als Oberbegriff (z.B. Fingerhut et al. 2013; Shapiro 2011). Zudem bleibt oft unklar, was mit Schlagworten wie „*embedded*“ oder „*enacted*“ im Detail gemeint ist und wie sich die Positionen inhaltlich genau zueinander verhalten (Walter 2014b). Zwar spricht gegenwärtig einiges dafür, dass Kognition entgegen der traditionellen Vorstellung nicht ausschließlich „im Kopf“ ist. Um jedoch die weitergehende Frage beantworten zu können, wo genau kognitive Prozesse ihren Platz haben, muss man die entsprechenden Positionen und die dahinterstehenden Überlegungen zum einen auf theoretischer Ebene voneinander abgrenzen, indem man die entscheidenden Unterschiede herausarbeitet. Um zu verhindern, dass es dabei um bloße Fragen philosophischer Nomenklatur ohne ein *fundamentum in re* geht, sollte sich diese Begriffsklärung zum anderen aber immer auch in einem praktischen Unterschied niederschlagen, das heißt, es sollte im Hinblick auf konkrete extraorganismische Ressourcen – etwa den Gebrauch von sogenannten *Brain-Computer-Interfaces* zur Kommunikation mit Locked-In-Patienten (Kyselo 2011; Walter 2010) – deutlich werden, wie grundsätzlich wohlbegründet zu entscheiden wäre, welche der verschiedenen theoretischen Beschreibungen (wenn überhaupt) zutrifft. Während dabei auf begrifflicher Seite durchaus Fortschritte erzielt wurden, ist deren praktische Relevanz derzeit zumeist bestenfalls unklar (Walter 2014c, Abs. 3).

Die Abgrenzung der Erweiterungsthese zu der These, dass kognitive Prozesse verteilt sind, erscheint intuitiv klar: Offenbar umfassen verteilte kognitive Prozesse nicht nur Artefakte, sondern auch andere Akteure, und erweiterte Prozesse integrieren extraorganismische Ressourcen in die kognitive Binnenarchitektur eines Individuums, während es einen solchen „kognitiven Kern“ bei verteilten Prozessen gerade nicht gibt – es ist der Alzheimerpatient Otto, der sich mithilfe seines Notizbuchs erinnert (Clark und Chalmers 1998), aber es ist nicht ein einzelner Seemann, der mithilfe der übrigen hierarchisch organisierten Crew und diversen Instrumenten navigiert (Hutchins 1995), sondern das entsprechende Kollektiv. Allerdings wäre in diesem Fall Otto das Subjekt *erweiterter* mentaler Zustände, wenn er sich auf sein Notizbuch verlässt, während dann, wenn er sich stattdessen in ansonsten analoger Weise auf seine Frau, seine Sekretärin und deren Notizen verlässt, das entsprechende *Kollektiv* das Subjekt *verteilter* mentaler Zustände wäre. Man könnte diese Spannung erstens dadurch aufzulösen versuchen, dass man auch erweiterte Erinnerungen nicht Otto, sondern dem Otto-plus-Notizbuch-Komplex zuschreibt. Zum einen widerspräche dies jedoch unserer etablierten Zuschreibungspraxis mentaler Zustände, und zum anderen hätte dieses System nach wie vor ein Individuum als „kognitiven Kern“, den es bei verteilten kognitiven Prozessen so nicht gibt. Zweitens könnte man auch dann Otto die Erinnerungen zuschreiben, wenn er sich auf seine Mitmenschen verlässt, und von verteilten Prozessen nur bei Kollektiven ohne Individuen als „kognitive Kerne“ sprechen. Unklar bliebe dann allerdings, wie und wo die Grenze gezogen werden soll zwischen hybriden Prozessen, die einem Individuum zugeschrieben werden, und jenen, die Kollektiven zukommen. Drittens könnte man verteilte Prozesse immer Individuen zuschreiben und

andere Akteure schlicht als extraorganismische Ressourcen *inter pares* betrachten. Allerdings wird die Identifikation eines „kognitiven Kerns“ in dem Maße problematisch, wie ein Kollektiv aus gleichrangigen Akteuren besteht: Es war eben nicht ein bestimmter Abgeordneter, der 1955 unter Zuhilfenahme seiner Kollegen die Pariser Verträge ratifizierte, sondern der Deutsche Bundestag, und es ist genau dieser Unterschied, der dagegen spricht, erweiterte und verteilte Kognition gleichzusetzen: Navigations- und Ratifizierungsprozesse erweitern sich nicht auf dieselbe Weise *von* einem Seemann oder Abgeordneten *in* ein soziales Kollektiv, wie sich Ottos Erinnerungen *von* Otto *in* sein Notizbuch erweitern sollen – im ersten Fall gibt es schlicht keinen „kognitiven Kern“, aus dem heraus sich der Geist „ausdehnen“ könnte (vgl. auch Hutchins 2014). Diese theoretischen Unklarheiten führen zu Abgrenzungsproblemen auf praktischer Ebene: Solange es unklare Grenzfälle gibt, ist nicht ausgemacht, dass wir immer nichtarbiträr entscheiden können, ob ein konkreter Anwendungsfall eher Otto als Paradebeispiel „erweiterter Kognition“ oder eher einem Bundestagsbeschluss und der Navigation eines Schiffes als Paradebeispiel „verteilter Kognition“ gleicht, sodass die Frage „Erweitert oder verteilt?“ womöglich nicht immer wohl begründet zu beantworten ist.

Etwas ganz Ähnliches gilt insbesondere auch für die Abgrenzung gegenüber Ansätzen, die Kognition als situativ eingebettet, aber nicht erweitert betrachten (Rupert 2004, 2009). Auch hier gibt es intuitiv klare Beispiele: Prozessor und Arbeitsspeicher sind Konstituenten eines Computers, während die Stromproduktion im Elektrizitätswerk nur eine kausale Vorbedingung seines Funktionierens ist; ebenso scheint die Retina konstitutiv für visuelle Wahrnehmung zu sein, während geeignete Lichtverhältnisse offenbar nur eine kausale Bedingung, nicht aber konstitutiver Teil des Wahrnehmungsprozesses sind. Andere Fälle hingegen sind intuitiv weniger klar: Ob eine externe Festplatte konstitutiver Teil eines Computers oder kausale Bedingung seines Funktionierens ist, mag in der Praxis ebenso wenig grundsätzlich zu entscheiden sein, wie etwa die Frage, ob Brillen, Hörgeräte oder Ottos Notizbuch Konstituenten entsprechender kognitiver Prozesse sind oder lediglich als extraorganismische Ressourcen kausal zu deren Funktionieren beitragen. Für solche Grenzfälle scheint sich die für die Debatte um die Erweiterungsthese zentral gewordene Unterscheidung zwischen Abhängigkeit und Konstitution, selbst wenn sie theoretisch völlig klar sein mag, kaum praktisch niederschlagen zu können (Walter 2014b). Die Rivalität zwischen der Einbettungsthese und der Erweiterungsthese verkäme damit zu einem abstrakten metaphysischen Disput, in dem man sich über das Phänomen an sich einig ist und nur über seine korrekte Beschreibung streitet, obwohl diese Wortklauberei folgenlos bleibt: Für die Kognitionswissenschaft, die Medizin, die Pflegewissenschaft oder Ethikkommissionen scheint es nämlich schlicht unerheblich zu sein, ob etwa die Wiederherstellung der kommunikativen Fähigkeiten von Locked-In-Patienten durch *Brain-Computer-Interfaces* lediglich kausal hervorgebracht oder aber zum Teil mit konstituiert wird, solange zum Beispiel beides dieselben kontrafaktischen Konditionale impliziert.



Auch der Verweis auf eine enge Kopplung eines Akteurs mit extraorganismischen Ressourcen reicht nicht aus, um die Erweiterungsthese gegenüber der Einbettungsthese auszuzeichnen. In Erwiderung auf den Einwand von Adams und Aizawa (2008), dass der Schluss von einer bloßen Abhängigkeit auf eine Konstitutionsthese einem „*coupling/constitution fallacy*“ gleichkommt, haben Anhänger der Erweiterungsthese darauf hingewiesen, dass extraorganismische Faktoren unter bestimmten Bedingungen so intensiv mit der kognitiven Binnenarchitektur eines Akteurs interagieren, dass man statt von einer Kopplung zwischen separaten Entitäten nur noch von einem einzigen kognitiven Prozess sprechen sollte. Allerdings ist es bislang nicht gelungen, zufriedenstellende Kriterien dafür anzugeben, wann eine Kopplung zu einer „bloßen“ Integration führt, in deren Zuge extraorganismische Ressourcen im Sinne der Einbettungsthese als Hilfsmittel furchtbar gemacht werden, und wann eine Kopplung im Sinne der Erweiterungsthese eine echte ontologische „Verschmelzung“ extraorganismischer Ressourcen mit einem kognitiven Kernsystem zur Folge hat. Jeder Versuch, extraorganismische Ressourcen über eine entsprechend enge Kopplung als konstitutiven Teil kognitiver Prozesse auszuzeichnen, muss insbesondere eine sogenannte „kognitive Inflation“ (*cognitive bloat*) vermeiden, das heißt verhindern, dass jeder an einer kognitiven Leistung beteiligte extraorganismische Faktor *eo ipso* zu einem konstitutiven Teil aufgewertet wird (Sprevak 2010) – gute Lichtverhältnisse tragen wesentlich zu visuellen Wahrnehmungsleistungen bei, sollten deshalb aber nicht automatisch als Konstituenten menschlichen Sehens gelten. Eine entsprechende Abgrenzung der Erweiterungsthese gegenüber der ontologisch konservativeren Einbettungsthese lässt sich aber eben gerade solange nicht überzeugend bewerkstelligen, wie entsprechende Kriterien für eine „substanzielle“ im Gegensatz zu einer „ontologisch harmlosen“ Kopplung fehlen.

Für die Abgrenzung gegenüber dem Enaktivismus schließlich gilt ebenfalls etwas ganz Ähnliches. Auf theoretischer Ebene hat sich mittlerweile herauskristallisiert, dass sich der Enaktivismus von allen anderen situierten Ansätzen dadurch unterscheidet, dass Letztere eine These darüber aufstellen, wo in der Welt kognitive Prozesse zu finden sind, während diese Frage für den Enaktivismus streng genommen unsinnig ist (Walter 2014a, Kap. 10), weil Kognition als wesentlich relationalem Phänomen überhaupt kein Ort zukommt (Di Paolo 2009, 19; Thompson & Stapleton 2009, 26), sodass sich die Frage, ob kognitive Prozesse auf das Gehirn beschränkt sind, den gesamten Körper umfassen oder sich in Teile der Umwelt hinaus erstrecken, schlicht nicht stellt. Auch hier gilt: Es ist ohne Zweifel philosophisch hilfreich, derartige Unterscheidungen auf theoretischer Ebene klar auszubuchstabieren (vgl. auch Rowlands 2009) und zu verstehen, was den Enaktivismus zur Zurückweisung einer Lokalisierung des Kognitiven treibt. Die entscheidende Frage ist aber eben wiederum, wie genau wir entscheiden sollen, ob wir es in einem konkreten Fall etwa mit einem räumlich lokalisierten erweiterten kognitiven Prozess oder einem räumlich nicht lokalisierbaren enaktiven Prozess zu tun haben.

Solange solche Abgrenzungen nicht begründet vorzunehmen sind, wird nicht nur die seit einiger Zeit mit beeindruckender Vehemenz geführte innerphilosophische Debatte um die Frage, ob der Geist „im Kopf“ ist, obsolet, die philosophische Diskussion setzt sich dann auch dem Vorwurf aus, brotlose Metaphysik um ihrer selbst willen zu betreiben, und das ist insbesondere auf einem Gebiet prekär, auf dem gemeinhin wie auf kaum einem anderen der Anspruch propagiert wird, mit der Philosophie am Puls der alltäglichen Forschungspraxis einer empirischen Einzelwissenschaft zu sein und ihr neue Impulse vermitteln zu können (z.B. Clark 2008).

### **3 Perspektiven und Aufgaben einer Philosophie der Kognition**

Der in Abschnitt 2 versuchte skizzenhafte Überblick über einige offene Fragen mit Blick auf die zentrale Leitfrage der Philosophie der Kognition, welche Teile eines Systems und seiner Umwelt auf welche Weise zu seinem intelligenten Verhalten beitragen, kann noch nicht einmal andeutungsweise das Spektrum an spannenden Fragen abdecken, das diese Disziplin bereithält (vgl. Stephan und Walter 2013 für einen Überblick sowie Abs. 3.3 für weiterführende Literatur). Da es in den Beiträgen dieses Bandes primär um wissenschaftsphilosophische Fragen zu ontologischen bzw. explanatorischen Aspekten einer Einzelwissenschaft wie der Kognitionswissenschaft gehen sollte, blieben zum Beispiel Detailfragen, die im Zuge der Erforschung einzelner kognitiver Leistungen wie Wahrnehmung, Erinnerung oder Handlungssteuerung und -kontrolle die Philosophie der Kognition maßgeblich gestaltet haben und gestalten werden (z.B. Margolis et al. 2012), ebenso ausgeblendet, wie etwa auch ethische Fragen, welche die Arbeit in der Kognitionswissenschaft selbst bzw. ihre Anwendungen betreffen (z.B. Choudhury und Slaby 2012; Levy 2010; Nagel 2010). Zwei Fragen sollen hier jedoch abschließend in Form eines „subjektiven Ausblicks“ auf mögliche zukünftige Entwicklungen der Philosophie der Kognition noch kurz zur Sprache kommen. Es geht dabei zum einen um allgemeine Überlegungen zu einer Wissenschaftstheorie der Kognitionswissenschaft zweiter Generation (s. Abs. 3.1) und zum anderen um den grundsätzlichen Skopus der Kognitionswissenschaft bzw. der in Abschnitt 2 skizzierten situierten Positionen (s. Abs. 3.2).

#### **3.1 Eine Wissenschaftstheorie der Kognitionswissenschaft zweiter Generation**

Anders als die erste Generation ist die Kognitionswissenschaft zweiter Generation bislang keine einheitliche Disziplin mit paradigmatischen Untersuchungsgegenständen und Methoden sowie einem zumindest annähernd allgemein akzeptierten Kanon an fundamentalen Begriffen und Erklärungsmodellen (Shapiro 2013). Während die erste Generation von einer klaren und einheitlichen Vorstellung der typischen Untersuchungsgegenstände (Gedächtnis, Sprachverständnis, Planen usw.), fundamentalen Begriffen (Repräsentation, Berechnung, Syntax und

Semantik, Algorithmus usw.) und Methoden (computationale Modelle, algorithmische Analysen usw.) und Vorzügen ihres Ansatzes (im Vergleich etwa zum Behaviorismus) geprägt war und sich unter anderem überhaupt nur deshalb als eigenständige Disziplin etablieren konnte, die nicht nur die Bezeichnung „Kognitionswissenschaften“, sondern die Auszeichnung „Kognitionswissenschaft“ im Singular verdient, fallen die Antworten der zweiten Generation sehr viel heterogener aus. Die Palette vermeintlich paradigmatischer Untersuchungsgegenstände reicht von Phänomenen wie dem auditiven System von Insekten über rhythmisches Fingerwackeln, visuelle Wahrnehmung, das Laufverhalten von Kindern und ihre Leistungen beim sogenannten A-nicht-B-Fehler bis hin zur Navigation von Schiffen, Getränkedosen einsammelnden Robotern, Tetris spielenden Erwachsenen und sozialen Interaktionen, und der dabei bemühte Begriffsapparat (Affordanzen, handlungsspezifische Repräsentationen, sensomotorische Abhängigkeiten, Attraktoren in Zustandsräumen, neuronale Simulationen, *sense-making* usw.) ist ebenso divers wie die verwendeten Methoden (dynamische Modellierung, computationale Modelle, implizites Design durch Evolution, *organic computing* usw.) und die Meinungen darüber, worin genau die Vorzüge der jeweiligen Position gegenüber anderen Ansätzen bestehen (s. Abs. 2.3 und 2.4). Eine einheitliches Forschungsprogramm einer Kognitionswissenschaft zweiter Generation *als Einzelwissenschaft* ist darin bislang kaum erkennbar. Eines der zentralen Desiderata der Philosophie der Kognition, das ihre Aufnahme in den vorliegenden Band rechtfertigen würde, ist daher eine umfassende, begrifflich klare und zugleich empirisch informierte Wissenschaftstheorie der Kognitionswissenschaft zweiter Generation, die deutlich werden lässt, was sie eigentlich zu *einer Wissenschaft* von Kognition macht.

### **3.2 Von situierter Kognition zu situierter Affektivität**

Wie in Abschnitt 2 gesehen, hat sich unsere Auffassung von Kognition in den vergangenen Jahrzehnten insbesondere mit Blick auf die Rolle des Körpers, der Umwelt sowie unserer Interaktion mit ihr zum Teil erheblich gewandelt. Dabei wurde bislang nur selten der Tatsache Beachtung geschenkt, dass sich Kognition nicht losgelöst von Emotion und Motivation verstehen lässt. Die einst strikte Gegenüberstellung von Kognition einerseits und Emotion sowie Motivation andererseits lässt sich allerdings nicht länger aufrechterhalten und ist ebenso überholt (z.B. Dai und Sternberg 2004), wie die sich daraus ergebende Tendenz der frühen Kognitionswissenschaft, sich vornehmlich mit den subpersonalen Mechanismen von kognitiven Leistungen wie Planen, Problemlösen, Erinnern oder Lernen zu beschäftigen, affektive und konative Phänomene hingegen nahezu vollständig zu ignorieren: Als immer schon nicht nur denkende, sondern stets auch empfindende, wollende und wertende Wesen bewegen wir uns nämlich zu keiner Zeit in einem emotions- und motivationsfreien Raum reinen Denkens. Eine ausschließliche Beschäftigung mit kognitiven Prozessen kann aus diesem Grund niemals ein vollständiges und angemessenes Bild des Menschen zeichnen, sodass es der Kognitionswissenschaft nicht

länger nur um ein Verständnis von Kognition gehen darf. Anzustreben ist vielmehr ein umfassendes interdisziplinäres Verständnis des Verhältnisses von Kognition zu Emotion und Motivation und Phänomenen wie Bewusstsein oder (Inter-)subjektivität.

Angesichts der in Abschnitt 2.2 skizzierten Überlegungen zur Situiertheit von Kognition muss man in diesem Zusammenhang insbesondere fragen, ob auch im Hinblick auf emotionale bzw. affektive Phänomene im Allgemeinen ein ähnlicher Trend weg von gehirnzentrierten Ansätzen und hin zu einer Einbeziehung von Körper, Umwelt und Interaktion ansteht (Stephan et al. 2014; Wilutzky et al. 2011, 2013). Was für das Kognitive gilt, gilt *mutatis mutandis* nämlich auch für das Emotionale: Ebenso wenig, wie wir isolierte Denker sind, sind wir quasi-cartesianische „Bewerter“, deren Emotionalität im Zuge wiederholter Wahrnehmen-Bewerten-Fühlen-Handeln-Zyklen vollkommen „fleischlos“ zwischen ihrem Wahrnehmen und ihrem Handeln in der Welt vermittelt. Die These einer engen Koppelung von Gehirn, Körper und Umwelt ist für Emotionen in gewisser Weise sogar noch plausibler als für Kognition: Während eine Einbeziehung des Körpers und der Umwelt in Theorien des Kognitiven von wenigen historischen Vorläufern abgesehen (z.B. Dewey, Heidegger, Vygotskij) eine noch junge Entwicklung ist, wurde in der Emotionsforschung nur während der kurzen Blütezeit radikal kognitivistischer Emotionstheorien (z.B. Nussbaum 2001; Solomon 1976) vorübergehend vergessen, dass affektive Phänomene nahezu ausnahmslos immer schon auch als situiert begriffen wurden.

Insbesondere gilt dies für die Beteiligung des Körpers. Als sich Ende der 1980er Jahre in der Kognitionswissenschaft die Vorstellung durchzusetzen begann, dass die Morphologie, Biologie und Physiologie eines Akteurs eine wesentliche Grundlage seiner kognitiven Fähigkeiten bilden, widersprach dies in der Tat der herrschenden Vorstellung von Kognition als substratneutral spezifizierbarer Informationsverarbeitung. Dagegen ist die Behauptung, dass Emotionen maßgeblich durch die Details unserer körperlichen Verfasstheit mitbestimmt werden, ideengeschichtlich ein alter Hut (man denke etwa nur an Aristoteles' Bemerkung in *de anima* (403a/b), dass sich Zorn für den Naturforscher als Sieden heißen Blutes in der Herzgegend darstellt; Wilutzky et al. 2011). Die Vorstellung, dass der Körper kein bloßes Medium ist, mit dessen Hilfe wir die Welt wahrnehmen und internen Emotionen äußerlichen Ausdruck verleihen, sondern selbst in einem nichttrivialen Sinne zu diesen Emotionen beiträgt, ist mithin weniger kontrovers als die entsprechende Verkörperlichungsthese im Bereich des Kognitiven: Von wenigen kognitivistischen Emotionstheorien abgesehen war immer klar, dass Emotionen von der spezifischen körperlichen Verfasstheit eines Subjekts auf eine nichttriviale Weise mitbestimmt werden, die durchaus in dieselbe Richtung geht wie das, was in der Philosophie der Kognition gegenwärtig als „*embodiment*“ diskutiert wird. Schon allein aus diesem Grund ist es angezeigt, zukünftig intensiver der Frage nachzugehen, ob, und wenn ja wie, sich diese beiden Debatten gegenseitig befruchten können.

Etwas anders sieht es im Hinblick auf die Rolle der Umwelt aus. Zwar wurde auch hier stets gesehen, dass es sich bei Emotionen im Normalfall um Reaktionen auf Veränderungen in der Umwelt handelt. Nach traditioneller Ansicht war die Umwelt aber nicht an den Emotionen selbst beteiligt oder gar ein Teil davon, sondern diente lediglich als Inputgeber und Outputempfänger, sodass jene Phänomene, die in der Philosophie der Kognition gegenwärtig unter Stichworten wie situative Einbettung, Erweiterung, Verteilung oder Enaktivismus diskutiert werden, dabei bislang nicht gezielt in den Blick gerieten. Eine Übertragung der entsprechenden Überlegungen auf den Bereich des Emotionalen könnte dort jedoch ebenfalls einen fruchtbaren Perspektivwechsel bewirken (Wilutzky et al. 2011): Ebenso wie im Zuge situierter Ansätze in der Kognitionswissenschaft zum Beispiel die Echtzeitnavigation in belebten Umgebungen das Schachspiel als typischen Untersuchungsgegenstand ergänzt hat, sollte in der Emotionsforschung der bislang dominierende Individualismus zugunsten eines facettenreicheren pluralistischen Ansatzes aufgegeben werden. Neben klassisch individualistischen (Basis-)Emotionen wie Ekel, Ärger, Furcht oder Freude sollte dabei insbesondere auch eine viel breiter angelegte Palette affektiver Phänomene in den Blick geraten, die zum Teil deutlich stärker von einer Interaktion eines Individuums mit seiner natürlichen und sozialen Umwelt abhängen (Colombetti und Stephan 2013). Wenn wir nicht länger ausschließlich die Emotionen eines einsamen Savannenläufers zum Paradigma machen, der auf eine Schlange trifft und nach einer internen Bewertung der Situation eine Angstreaktion zeigt, sondern auch die dynamische und reziproke Interaktion von Musikern mit ihrem Publikum, Neugeborenen mit ihren Bezugspersonen oder Pubertierenden im Konflikt mit ihren Eltern sowie affektive Phänomene betrachten, die interagierende soziale Akteure als Träger haben, die zum Beispiel im Rahmen eines Vorstellungsgesprächs als Kollektiv eine eisige Atmosphäre erzeugen, wird das Potenzial inhaltlicher Anknüpfungspunkte zur Debatte um situierte Kognition viel augenfälliger (Griffiths und Scarantino 2009). Am Ende könnte die Emotionsforschung damit in eine allgemeinere Affektwissenschaft eingebunden werden, die sich in ähnlich interdisziplinärer Weise mit affektiven Phänomenen und deren Situietheit befasst, wie die Kognitionswissenschaft zweiter Generation mit kognitiven Phänomenen und ihrer Situietheit (Wilutzky et al. 2013).

### **3.3 Literaturempfehlungen**

Den umfassendsten deutschsprachigen Überblick über das gesamte Feld der Kognitionswissenschaft verschafft derzeit das Handbuch von Stephan und Walter (2013); Margolis et al. (2012) ist ein aktuelles englischsprachiges Handbuch mit stärker philosophischem Einschlag, das dafür deutlich weniger Breite in der Kognitionswissenschaft allgemein bietet. Als Wörterbücher mit kürzeren Einträgen zu sehr vielen Stichworten aus allen Bereichen der Kognitionswissenschaft bieten sich in deutscher Sprache Strube (1996) sowie in englischer Sprache Nadel (2005) an, auch wenn Ersteres inzwischen naturgemäß nicht mehr ganz auf der Höhe der Zeit

ist. Unter den eher historisch orientierten Einführungen in die Philosophie der Kognition ist Boden (2006) das deutlich umfangreichste, aber auch das langatmigste Werk. Gardner (1985) und Brook (2007) zeichnen beide die Ursprünge der Kognitionswissenschaft nach; Ersterer konzentriert sich dabei stärker auf die Entwicklungen, die zur sogenannten „kognitiven Revolution“ gegen Mitte des 20. Jahrhunderts führten, Letzterer deckt die historischen Vorläufer in der Philosophie und der Psychologie zwischen Descartes und Sigmund Freud ab. Die aktuellste deutschsprachige Einführung in die Philosophie der Kognition, insbesondere mit Blick auf situierte Ansätze, ist Walter (2014); dort findet sich auch eine kommentierte Bibliografie zur Philosophie der Kognition. Lenzen (2002) bietet eine etwas ältere und allgemeiner gehaltene, aber immer noch lesenswerte Einführung in die Kognitionswissenschaft, allerdings mit deutlichem Schwerpunkt auf der Künstliche-Intelligenz-Forschung, während die Neurowissenschaften und die Psychologie im Wesentlichen unberücksichtigt bleiben. Fingerhut et al. (2013) ist eine sehr nützliche Sammlung von deutschen Übersetzungen einiger zentraler Beiträge zur Debatte um situierte Kognition. Im englischsprachigen Bereich bieten Sobel und Li (2013) einen guten Überblick über die Kognitionswissenschaft im Allgemeinen; Bermúdez (2010) hat einen stärker philosophischen Einschlag, deckt aber situierte Ansätze nur unzureichend ab; Shapiro (2011) legt den Schwerpunkt auf situierte Ansätze, hat dafür aber weniger zum Computermodell und zum Konnektionismus zu sagen.

## Literatur

Adams, Fred und Kenneth Aizawa (2008). *The Bounds of Cognition*. Malden: Blackwell.

Beach, King (1993). Becoming a bartender. *Applied Cognitive Psychology* 7, 191–204.

Bechtel, William (1998). Representations and cognitive explanations. *Cognitive Science* 22, 295–318.

Bermúdez, José (2010). *Cognitive Science: An Introduction to the Science of the Mind*. Cambridge: Cambridge University Press.

Boden, Magaret (2006). *Mind as Machine: A History of Cognitive Science*, 2 Bde. Oxford: Oxford University Press.

Brook, Andrew (Hg.) (2007). *The Prehistory of Cognitive Science*. Basingstoke: Palgrave Macmillan.

Brooks, Rodney (1991). Intelligence without representation. *Artificial Intelligence* 47, 139–159.

Chemero, Anthony (2009). *Radical Embodied Cognitive Science*. Cambridge (Mass.): MIT Press.

Choudhury, Suparna und Jan Slaby (Hg.) (2012). *Critical Neuroscience*. London: Routledge.

Clark, Andy (2001). *Mindware*. Oxford.

Clark, Andy (2008). *Supersizing the Mind*. Oxford: Oxford University Press.

- Clark, Andy und David Chalmers (1998). The extended mind. *Analysis* 58, 7–19.
- Clark, Andy und Josefa Toribio (1994). Doing without representing. *Synthese* 101, 401–431.
- Colombetti, Giovanna und Achim Stephan (2013). Affektwissenschaft (*affective science*). In: Achim Stephan und Sven Walter (Hg.). *Handbuch Kognitionswissenschaft*. Stuttgart: Metzler, 501–510.
- Dai, David und Robert Sternberg (Hg.) (2004). *Motivation, Emotion, and Cognition*. Mahwah: Lawrence Erlbaum.
- Di Paolo, Ezequiel (2009). Extended life. *Topoi* 28, 9–21.
- Edelman, Shimon (2003). But will it scale up? *Adaptive Behavior* 11, 273–275.
- Fingerhut, Jörg, Rebecca Hufendiek und Markus Wild (Hg.) (2013). *Philosophie der Verkörperung*. Frankfurt a.M.: Suhrkamp.
- Fodor, Jerry (1975). *The Language of Thought*. Cambridge (Mass.): MIT Press.
- Fodor, Jerry (1983). *The Modularity of Mind*. Cambridge (Mass.): MIT Press.
- Gallagher, Shaun (2008). Are minimal representations still representations? *International Journal of Philosophical Studies* 16, 351–369.
- Gardner, Howard (1985). *The Mind's New Science: A History of the Cognitive Revolution*. New York: Basic Books.
- Goldmann, Alvin (2012). A moderate approach to embodied cognitive science. *Review of Philosophy and Psychology* 3, 71–88.
- Griffiths, Paul und Andrea Scarantino (2009). Emotions in the wild. In: Philipp Robbins und Murat Aydede (Hg.). *The Cambridge Handbook of Situated Cognition*. Cambridge: Cambridge University Press, 437–453.
- Haken, Hermann, Scott Kelso und Heinz Bunz (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics* 51, 347–356.
- Hinton, Geoffrey, James McClelland und David Rumelhart (1986). Distributed representations. In: David Rumelhart, James McClelland und PDP Research Group (Hg.). *Parallel Distributed Processing*, Bd. 1. Cambridge (Mass.): MIT Press, 77–109.
- Hurley, Susan (1998). *Consciousness in Action*. Cambridge (Mass.): Harvard University Press.
- Hurley, Susan (2008). The shared circuits model. *Behavioral and Brain Sciences* 31, 1–58.
- Hutchins, Edwin (1995). *Cognition in the Wild*. Cambridge (Mass.): MIT Press.
- Hutchins, Edwin (2014). The cultural ecosystem of human cognition. *Philosophical Psychology* 27, 34–49.
- Hutto, Dan und Eric Myin (2013). *Radicalizing Enactivism*. Cambridge (Mass.): MIT Press.
- James, William (1890). *The Principles of Psychology*. New York: Holt.
- Kelso, Scott (1995). *Dynamic Patterns*. Cambridge (Mass.): MIT Press.
- Kyselo, Miriam (2011). Locked-in syndrome and BCI. *Neuroethics* 6, 579–591.
- Lakoff, George und Mark Johnson (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.

- Lakoff, George und Mark Johnson (1999). *Philosophy in the Flesh*. New York: Basic Books.
- Lenzen, Manuela (2002). *Natürliche und Künstliche Intelligenz: Einführung in die Kognitionswissenschaft*. Frankfurt a.M.: Campus.
- Levy, Neil (Hg.) (2010). *The Oxford Handbook of Neuroethics*. Oxford: Oxford University Press.
- Margolis, Eric, Richard Samuels und Stephen Stich (Hg.) (2012). *The Oxford Handbook of Philosophy of Cognitive Science*. Oxford: Oxford University Press.
- Markman, Arthur und Eric Dietrich (2000). In defense of representation. *Cognitive Psychology* 40, 138–171.
- Marr, David (1982). *Vision*. San Francisco: Freeman.
- Maturana, Humberto und Francisco Varela (1980). *Autopoiesis and Cognition*. Dordrecht: Reidel.
- Milkowski, Marcin (2013). *Explaining the Computational Mind*. Cambridge (Mass.): MIT Press.
- Miller, George (2003). The cognitive revolution. *Trends in Cognitive Sciences* 7, 141–144.
- Miller, George, Eugene Galanter und Karl Pibram (1960). *Plans and the Structure of Behavior*. New York: Holt.
- Nadel, Lynn (Hg.) (2005). *Encyclopedia of Cognitive Science*. New York: Nature Publishing Group.
- Nagel, Saskia (2010). *Ehtics and the Neurosciences*. Paderborn: mentis.
- Newell, Allen und Herbert Simon (1976). Computer science as empirical inquiry. *Communications of the ACM* 19, 113–126.
- Neisser, Ulric (1967). *Cognitive Psychology*. East Norwalk: Appleton-Century-Crofts.
- Nussbaum, Martha (2001). *Upheavals of Thought*. Cambridge: Cambridge University Press.
- Pfeifer, Rolf und Josh Bongard (2007). *How the Body Shapes the Way We Think*. Cambridge (Mass.): MIT Press.
- Piccinini, Gualtiero (2008). Computation without representation. *Philosophical Studies* 137, 205–241.
- Piccinini, Gualtiero (2012). Computationalism. In: Eric Margolis, Richard Samuels und Stephen Stich (Hg.). *The Oxford Handbook of Philosophy of Cognitive Science*. Oxford: Oxford University Press, 222–249.
- Robbins, Philipp und Murat Aydede (Hg.) (2009). *The Cambridge Handbook of Situated Cognition*. Cambridge: Cambridge University Press.
- Rowlands, Mark (2009). Enactivism and the extended mind. *Topoi* 28, 53–62.
- Rumelhart, David, James McClelland und PDP Research Group (Hg.) (1986). *Parallel Distributed Processing*, 2 Bde. Cambridge (Mass.): MIT Press.
- Rupert, Robert (2004). Challenges to the hypothesis of extended cognition. *Journal of Philosophy* 101, 389–428.
- Rupert, Robert (2009). *Cognitive Systems and the Extended Mind*. Oxford: Oxford University Press.



- Shapiro, Larry (2010). James Bond and the barking dog. *Philosophy of Science* 77, 400–418.
- Shapiro, Larry (2011). *Embodied Cognition*. London: Routledge.
- Shapiro, Larry (2013). When is cognition embodied? In: Uriah Kriegel (Hg.). *Current Controversies in Philosophy of Mind*. London: Routledge, 73–90.
- Skinner, Burrhus (1953). *Science and Human Behavior*. New York: Macmillan.
- Smolensky, Paul (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences* 11, 1–23.
- Sobel, Carolyn und Paul Li (2013). *The Cognitive Sciences: An Interdisciplinary Approach*. Los Angeles: Sage.
- Solomon, Robert (1976). *The Passions*. New York: Anchor.
- Sprevak, Mark (2010). Inference to the hypothesis of extended cognition. *Studies in History and Philosophy of Science* 41, 353–362.
- Stephan, Achim und Sven Walter (Hg.) (2013). *Handbuch Kognitionswissenschaft*. Stuttgart: Metzler.
- Stephan, Achim, Sven Walter und Wendy Wilutzky (2014). Emotions beyond brain and body. *Philosophical Psychology* 27, 65–81.
- Stewart, John, Olivier Gapenne und Ezequiel Di Paolo (2010). *Enaction: Toward a New Paradigm for Cognitive Science*. Cambridge (Mass.): MIT Press.
- Strube, Gerhard (Hg.) (1996). *Wörterbuch der Kognitionswissenschaft*. Stuttgart: Klett-Cotta.
- Sturm, Thomas und Horst Gundlach (2013). Zur Geschichte und Geschichtsschreibung der „kognitiven Revolution“ – eine Reflexion. In: Achim Stephan und Sven Walter (Hg.). *Handbuch Kognitionswissenschaft*. Stuttgart: Metzler, 7–21.
- Thelen, Esther und Linda Smith (1994). *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge (Mass.): MIT Press.
- Thompson, Evan und Mog Stapleton (2009). Making sense of sense-making. *Topoi* 28, 23–30.
- van Gelder, Tim (1995). What might cognition be, if not computation? *Journal of Philosophy* 91, 345–381.
- van Gelder, Tim (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences* 21, 615–628.
- Varela, Francisco, Evan Thompson und Eleanor Rosch (1991). *The Embodied Mind*. Cambridge (Mass.): MIT Press.
- Villalobos Marco (2013). Enactive cognitive science: revisionism or revolution? *Adaptive Behaviour* 21, 159–167.
- Walter, Sven (2010). Locked-in syndrome, BCI, and a confusion about embodied, embedded, extended, and enacted cognition. *Neuroethics* 3, 61–72.
- Walter, Sven (2014a). *Kognition*. Stuttgart: Reclam.
- Walter, Sven (2014b). Situated cognition: Some open conceptual and ontological issues. *Review of Philosophy and Psychology* ??, ??–??.

Walter, Sven (2014c). Ist der Geist im Kopf? Ist der Geist erweitert? Und vor allem: Was müssten wir wissen, um diese Fragen beantworten zu können? In: Jan Michel, Kim Boström und Michael Pohl (Hg.). *Ist der Geist im Kopf ? Die These des erweiterten Geistes in Philosophie und Wissenschaft*. Münster: mentis, ??–??.

Wilson, Rob und Andy Clark (2009). How to situate cognition. In: Philipp Robbins und Murat Aydede (Hg.). *The Cambridge Handbook of Situated Cognition*. Cambridge: Cambridge University Press, 55–77.

Wilutzky, Wendy, Sven Walter und Achim Stephan (2011). Situierete Affektivität. In: Jan Slaby, Achim Stephan, Henrik Walter und Sven Walter (Hg.). *Affektive Intentionalität*. Paderborn: mentis, 283–320.

Wilutzky, Wendy, Achim Stephan und Sven Walter (2013). Situierete Affektivität. In: Achim Stephan und Sven Walter (Hg.). *Handbuch Kognitionswissenschaft*. Stuttgart: Metzler, 552–560.