

EINE THEORETISCHE KRITIK DER
SOMATISCHEN MARKER HYPOTHESE
ANTONIO DAMASIOS

Von:

Ulf Hlobil

ZUSAMMENFASSUNG

Seit einigen Jahrzehnten sind grundsätzliche Probleme des Versuchs bekannt, menschliche Handlungen kausal zu erklären. Diese Probleme wurden aber bisher innerhalb der biologischen Psychologie ignoriert. Dies lässt sich insbesondere darauf zurückführen, dass angenommen wurde, man könne diesen Problemen entgehen, indem man die alltagssprachlichen Termini (bzw. deren wissenschaftliche Präzisierungen) der kognitiven Psychologie durch *termini technici* der Neurobiologie ersetzt. In der vorliegenden Arbeit wird gezeigt, dass dies nicht der Fall ist.

Hierzu wird exemplarisch die *Somatische Marker Hypothese* (SMH) Antonio R. Damasio analysiert, welche versucht Handlungen kausal zu erklären. Zunächst werden die Voraussetzungen für eine solche Analyse geschaffen. Diese bestehen in einer Darstellung der Unterscheidung zwischen empirischen und begrifflichen Relationen sowie in der Darstellung der SMH und der entsprechenden empirischen Forschung. Daran anschließend wird gezeigt, dass die SMH eine reduktive Struktur aufweist. Dies bedeutet, dass die SMH versucht, eine angebliche psychologische Gesetzmäßigkeit auf eine neurobiologische Gesetzmäßigkeit zurückzuführen. Dabei impliziert die SMH, dass das genannte psychologische Gesetz einen kausalen Zusammenhang wiedergibt. Es wird dargelegt, dass das genannte psychologische Gesetz einen begrifflichen und damit keinen kausalen Zusammenhang angibt. Hiermit ist die Unhaltbarkeit der SMH erwiesen.

Abschließend wird darauf hingewiesen, dass die vorgelegte Argumentation solch allgemeiner Natur ist, dass sie höchstwahrscheinlich auf andere Theorien der biologischen Psychologie übertragen werden kann.

Trier, den 14. 09. 2008

SCHLÜSSELWÖRTER: begriffliche Relationen • Analytizitätsproblem • Somatische Marker Hypothese • Handlungserklärung • Emotion • Reduktionismus • Kausalismus • psychologische Prädikate

INHALT

ZUSAMMENFASSUNG	3
INHALT	4
VERZEICHNISSE	8
ABBILDUNGSVERZEICHNIS	8
ABKÜRZUNGSVERZEICHNIS	8
HINWEISE ZUR ZITATIONSWEISE	8
SIGLEN	8
1. EINLEITUNG	9
1.1. ZIEL UND MOTIVATION DER UNTERSUCHUNG	9
1.2. ÜBERBLICK ÜBER DIE ARGUMENTATION DER UNTERSUCHUNG	11
1.2.1. Das Hauptargument	12
1.2.2. Das Argument für die Prämisse (P1)	14
1.2.3. Das Argument für die These (T1)	15
1.2.4. Das Argument für die These (T2)	19
1.3. ABGRENZUNG DIESER ARBEIT VON ANDERER KRITIK DER SMH	21
2. METHODISCHE GRUNDLAGEN – BEGRIFFLICHE UND EMPIRISCHE RELATIONEN ...	24
2.1. HISTORISCHES ZU BEGRIFFLICHEN RELATIONEN	24
2.1.1. Entwicklungen vor Kant	24
2.1.2. Kant	28
2.1.3. Frege	29
2.1.4. Einige Lehren aus der historischen Reflexion	30
2.2. DIE HEUTIGE DISKUSSION UM BEGRIFFLICHE RELATIONEN	30
2.2.1. Referenzbasierte Bedeutungstheorien	31
2.2.1.1. Carnap	31
2.2.1.2. Quine	34
2.2.1.3. Entwicklungen nach Quine	36
2.2.1.3.1. Boghossian	36
2.2.1.3.2. Putnam, Kripke und der semantische Externalismus	39
2.2.1.3.3. Analytizität und die Semantik möglicher Welten	43
2.2.1.4. Zusammenfassung zum referenzbasierten Ansatz	48
2.2.2. Gebrauchsbasierte Bedeutungstheorien	49
2.2.2.1. Wittgenstein	49
2.2.2.1.1. Analytizität und Konvention	51
2.2.2.1.2. Apriorität, Notwendigkeit und Semantik	53

	Inhalt	5
2.2.2.1.3.	Der ontologische Status von Regeln	56
2.2.2.2.	Der Inferentialismus	58
2.2.2.2.1.	Grundlagen des Inferentialismus	58
2.2.2.2.2.	Der Inferentialismus und die analytisch-synthetisch Unterscheidung	60
2.3.	EINE GANGBARE UNTERSCHIEDUNG ZWISCHEN EMPIRISCHEN UND BEGRIFFLICHEN RELATIONEN	62
2.3.1.	Adäquatheitskriterien einer Unterscheidung	62
2.3.2.	Mehrere Unterscheidungen	63
2.3.3.	Eine Skizze des Sprachgebrauchs	65
2.4.	ERGEBNIS.....	68
3.	DARSTELLUNG DER THEORIE DER SOMATISCHEN MARKER.....	71
3.1.	DER ENTDECKUNGSZUSAMMENHANG DER SMH	71
3.1.1.	Historischer Rahmen der SMH – Biopsychologie der Emotionen	71
3.1.2.	Phineas P. Gage und seine Leidensgenossen.....	74
3.2.	SYSTEMATISCHE DARSTELLUNG DER SMH	76
3.2.1.	Was ist ein somatischer Marker?.....	77
3.2.2.	Wie beeinflussen somatische Marker die Entscheidungsfindung?	80
3.3.	EMPIRISCHE ÜBERPRÜFUNG DER SMH	86
3.3.1.	Läsionsstudien.....	87
3.3.1.1.	Somatische Reaktionen auf emotional besetzte Stimuli.....	87
3.3.1.2.	Befunde mittels neuropsychologischer Tests – Die Iowa Gambling Task	88
3.3.2.	Belege mittels bildgebender Verfahren	90
3.3.3.	Psychopharmakologische Daten.....	91
3.3.4.	Zwei empirische Herausforderungen an die SMH	92
3.3.4.1.	Eine erste alternative Erklärung der Befunde mittels IGT	92
3.3.4.2.	Die kognitive Durchschaubarkeit der IGT	93
3.4.	ZUSAMMENFASSUNG	93
4.	DER PSYCHOLOGISCHE GEHALT DER SMH.....	95
4.1.	PSYCHOLOGISCHE PRÄDIKATE IN DER BIOPSYCHOLOGIE.....	96
4.1.1.	Braucht die Neurowissenschaft psychologische Prädikate?	96
4.1.2.	Wo können psychologische Prädikate in biopsychologischen Theorien auftauchen?	101
4.2.	WO VERWENDET DIE SMH PSYCHOLOGISCHE PRÄDIKATE?	103
4.2.1.	Psychologische Prädikate im Consequens der Gesetze der SMH	103
4.2.2.	Psychologische Prädikate im Antecedens der Gesetze der SMH	104
4.2.3.	Psychologische Prädikate in Brückenprinzipien	105
4.3.	DIE REDUKTIONISTISCHE STRUKTUR DER SMH.....	107
4.3.1.	Die Brückenprinzipien der SMH.....	108
4.3.1.1.	„Emotion“	108
4.3.1.2.	„Gefühl“	109
4.3.1.3.	„Erwartetes Handlungsergebnis“	110

4.3.1.4.	„Etwas emotional bewerten“	111
4.3.1.5.	„Auswahl einer Handlungsoption“	112
4.3.2.	Es muss einen genuin psychologischen Gehalt der SMH geben	114
4.4.	WELCHES IST DER PSYCHOLOGISCHE GEHALT DER SMH?	117
4.4.1.	Das neurobiologische Gesetz der SMH.....	117
4.4.2.	Anwendung der Brückenprinzipien.....	119
5.	EMOTIONALE BEWERTUNG UND HANDLUNGSERKLÄRUNG	122
5.1.	LAUT SMH ERKLÄRT (PG) HANDLUNGEN KAUSAL	122
5.2.	DIE SMH MUSS HANDLUNGSERKLÄRUNGEN GEBEN KÖNNEN	124
5.2.1.	Das Verhalten, welches auf Entscheidung beruht, ist eine Handlung	124
5.2.2.	Erklärungen von Verhaltensweisen, welche auf Entscheidungen beruhen, sind Handlungserklärungen	125
5.2.2.1.	Erklärungen sind intensionale Kontexte.....	125
5.2.2.2.	Die SMH und das Wunsch-Überzeugungs-Modell	127
5.3.	EMOTIONALE BEWERTUNGEN SIND RAHMENBEDINGUNGEN VON HANDLUNGSERKLÄRUNGEN	128
5.3.1.	Handlungserklärungen	129
5.3.1.1.	Handlungserklärungen sind Erklärungen aus Gründen	129
5.3.1.2.	Normative und motivierende Gründe	130
5.3.1.2.1.	Die Erklärungsbedingung.....	131
5.3.1.2.2.	Die Normativitätsbedingung	133
5.3.1.3.	Gründe sind Sachverhalte.....	134
5.3.1.4.	Gründe und Rahmenbedingungen	135
5.3.2.	Emotionale Bewertungen	138
5.3.2.1.	Emotionen und Werturteile	139
5.3.2.2.	Emotionen und Gründe	143
5.3.2.3.	Bewertungen	146
5.4.	RAHMENBEDINGUNGEN UND KAUSALERKLÄRUNGEN	147
5.4.1.	Der Kausalismus	147
5.4.1.1.	Davidson's challenge	148
5.4.1.2.	Weitere Argumente für den Kausalismus.....	151
5.4.2.	Argumente für den Intentionalismus	154
5.4.2.1.	Das Argument aus der begrifflichen Abhängigkeit	155
5.4.2.2.	Das Argument aus dem Externalismus.....	159
5.4.2.3.	Das Argument aus der Normativität.....	165
5.4.2.4.	Das Argument aus der Ontologie	170
5.5.	ZUSAMMENFASSUNG	175
6.	ERGEBNISSE UND DISKUSSION.....	177
6.1.	ERGEBNISSE.....	177
6.2.	MÖGLICHE EINWÄNDE.....	178

	Inhalt	7
6.2.1.	Modus tollens vs. modus ponens	178
6.2.2.	Begrifflicher oder teleologischer Zusammenhang?	179
6.2.3.	Epiphänomenalismus	180
6.3.	WAS BLEIBT VON DER SMH?	181
6.4.	VERALLGEMEINERBARKEIT DER ERGEBNISSE	183
LITERATUR	186

VERZEICHNISSE

Abbildungsverzeichnis

Abbildung 1: Gültigkeitsnachweis für das Hauptargument.	13
Abbildung 2: Formale Darstellung des Arguments für die These (T1).	17
Abbildung 3: Gültigkeitsnachweis für das Argument für die These (T1).....	18
Abbildung 4: Gültigkeitsnachweis für das Argument für die These (T2).....	20
Abbildung 5: Darstellung der wichtigsten Annahmen der SMH	85
Abbildung 6: Schematische Darstellung der Struktur der SMH	121

Abkürzungsverzeichnis

ECS:	potenziell emotionsauslösender Stimulus (<i>emotionally competent stimulus</i>)
IGT:	Iowa-Glücksspielaufgabe (<i>Iowa Gambling Task</i>)
SI:	primärer somatosensorischer Kortex
SII:	sekundärer somatosensorischer Kortex
SMH:	Somatische Marker Hypothese (<i>somatic marker hypothesis</i>)
VMPFC:	ventromedialer präfrontaler Kortex

HINWEISE ZUR ZITATIONSWEISE

Die Zitationsweise basiert auf dem *Publication Manual of the American Psychological Association*, wobei die gesamte Zitationsweise an das Deutsche angepasst wurde. Die Hervorhebungen in Zitaten sind stets aus dem Original übernommen. Innerhalb von Zitaten wurden Auslassungen und Einschübe des Verfassers der vorliegenden Arbeit durch eckige Klammern „[...]“ kenntlich gemacht.

Siglen

KANT:

KrV:.....Kritik der reinen Vernunft

WITTGENSTEIN:

BB: Das Blaue Buch

BPP I: Bemerkungen über die Philosophie der Psychologie: Band 1

BPP II: Bemerkungen über die Philosophie der Psychologie: Band 2

PU:..... Philosophische Untersuchungen

ÜG: Über Gewißheit

1. EINLEITUNG

1.1. Ziel und Motivation der Untersuchung

Warum handeln Menschen so, wie sie es tun? Dies ist eine der Grundfragen der Psychologie. Doch um eine Antwort auf diese Frage geben zu können, muss zunächst die Frage geklärt werden, welche Art von Antwort wir hier akzeptieren sollten. Anders ausgedrückt: Wie können wir Handlungen erklären? Welche Art von Erklärungen sind Handlungserklärungen? Innerhalb der Psychologie, der Philosophie und verschiedener Sozialwissenschaften gibt es daher seit Langem eine intensive Diskussion darüber, ob sich die erfolgreiche Erklärung von Handlungen als ein Typ kausaler Erklärungen verstehen lässt (vgl. z. B. Anscombe, 2000, 2005; Brandtstädter, 1987, 2007; Brandtstädter & Sturm, 2004; Dancy, 2000; Davidson, 1998d; Mele, 1992; Schueler, 2003; Schon, 2005; Stoutland, 1970, 1998, 2005; von Wright, 1984). Die vorliegende Arbeit argumentiert für die These, dass Handlungserklärungen keine Kausalerklärungen sind. Hieraus ergibt sich unmittelbar der Grundgedanke der vorliegenden Untersuchung: *Wenn Handlungserklärungen keine Kausalerklärungen sind, ist nicht zu verstehen, wie eine Wissenschaft wie die biologische Psychologie Handlungen adäquat erklären kann.*

Um diesen Grundgedanken mit Inhalt zu füllen und ihn zu prüfen, will diese Arbeit eine begriffliche Analyse einer neurobiologischen Theorie menschlichen Handelns geben. Bei der angesprochenen Theorie handelt es sich um die *Somatische Marker Hypothese* (SMH), welche im vergangenen Jahrzehnt insbesondere durch die Forschergruppe um Antonio R. Damasio vertreten wurde (vgl. z. B. Bechara & Damasio, 2005; Bechara, Damasio & Damasio, 2000; Bechara, Damasio, Damasio & Lee, 1999; Damasio, 1996, 2006; Damasio, Tranel & Damasio, 1991). Diese Theorie will menschliches Entscheidungsverhalten erklären. Kern der Theorie ist die Annahme, dass Entscheidungsverhalten wesentlich durch die emotionale Bewertung von Handlungsoptionen und ihren Folgen bestimmt wird. Die Motivation einer solchen Analyse ergibt sich aus den vermeintlichen jüngsten Fortschritten der kognitiven Neurowissenschaften. Seit einigen Jahren beansprucht die biologische Psychologie nicht mehr, bloß relativ einfaches menschliches Erleben und Verhalten, wie beispielsweise Reflexe, Wahrnehmung, einfache Greifbewegungen oder Gedächtnisleistungen, erklären zu können. Das Verhalten, welchem zurzeit große Aufmerksamkeit geschenkt wird, ist zielgerichtetes, intentionales Verhalten, welches wir im Allgemeinen als „rational“ bezeichnen (vgl. Grafton & Ivry, 2004; Kandel, Schwartz & Jessel, 1996, S. 325f.). Solches

intentionale Verhalten nennen wir auch Handeln bzw. Handlung. Die biopsychologische Forschung bemüht sich also derzeit darum, menschliches Handeln erklären zu können. Damit stellt sich die Frage nach dem Wesen der Erklärung menschlichen Handelns im Zusammenhang mit der jüngsten neurowissenschaftlichen Forschung.

Die SMH wurde hierzu aus zwei Gründen als Beispiel einer neurowissenschaftlichen Theorie menschlichen Handelns ausgewählt: Erstens beansprucht die SMH, Verhalten zu erklären, welches auf Entscheidungen beruht. Wir sprechen aber nur dort von Entscheidungen, wo es eine Wahl zwischen Alternativen gibt, d. h. dort, wo das Verhalten sich potenziell unterschiedlich ausrichten kann. Es handelt sich hier also um Verhalten, das in einem relevanten Sinn intentional ist, und somit als Handlung betrachtet werden kann, d. h., die SMH ist tatsächlich eine Theorie menschlichen Handelns. Zweitens erfreut sich die SMH einer sehr breiten Rezeption.¹ Es finden sich weit über fünfzig empirische Untersuchungen zu dieser Theorie (für einen Überblick s. Dunn, Dalgleish & Lawrence, 2006); und es wurde aus den unterschiedlichsten Fachrichtungen auf sie Bezug genommen (vgl. Cromby, 2007; Gluck, 2007; Stamm, 2007). Die breite Rezeption hat dazu geführt, dass die SMH bereits in Einführungslehrbüchern in die Neurowissenschaft besprochen wird (z. B. Kandel & Kupfermann, 1996, S. 609; Martin, 2006, S. 203f.). Es handelt sich also um eine einschlägige, breit rezipierte, fruchtbare und aktuelle neurobiologische Theorie menschlichen Handelns.

Die vorliegende Arbeit will Erklärungen menschlichen Handelns kritisieren, welche solche Erklärungen ausschließlich als kausale verstehen. Damit reiht sich die Arbeit in einen breiteren Rahmen ein, welchen man als die „Untersuchung des Analytizitätsproblems in der psychologischen Forschung“ bezeichnen kann. Dieses Problem zeigt sich darin, dass in der Psychologie immer wieder Ergebnisse als das Resultat empirischer Forschung präsentiert werden, welche sich bei näherer Betrachtung als nicht-empirische Sätze erweisen (vgl. Brandtstädter & Sturm, 2004; Smedslund, 1988, 1997). Kann es beispielsweise tatsächlich als ein sinnvolles Ergebnis empirischer Forschung betrachtet werden, wenn bezüglich der Ursachen der Ablehnung von Schülern durch ihre Mitschüler festgestellt wird, es sei „vor allem die mangelnde Sympathie, die [...] zu ihrer Ablehnung führt“ (Hascher & Baillod, 2004, S. 150)? Sollten wir hier nicht sagen, dass Ablehnung teilweise im Fehlen von Sympathie *besteht*? Solche Forschung ähnelt dem Versuch, empirisch nachzuweisen, dass

¹ Am 12. 08. 2008 listet der *Science Citation Index* 114 Einträge unter dem Stichwort „*somatic marker hypothesis*“. Damasio wird vom *Institute for Scientific Information* unter die „*highly cited researchers*“ gezählt; es beziffert die Zahl seiner Publikationen mit 392. Man kann hier also gerechtfertigterweise von einer enormen akademischen Bedeutung sprechen.

Rüden Hunde sind. Es handelt sich hier um eine begriffliche Beziehung und nicht um eine kausale. Die Psychologie steht also vor der Aufgabe, empirische und begriffliche Beziehungen zu unterscheiden. Bezogen auf das menschliche Handeln nimmt dieses Problem eine spezifische Form an: Handlungserklärungen sind Erklärungen aus Gründen; Gründe aber stehen mit demjenigen, was sie begründen, nicht in einer kontingenten Beziehung (vgl. Holzkamp, 1986). Dieselbe Abhängigkeit lässt sich auch auf andere Weise bemerken: Wir sprechen nicht von Handlungen, sofern keine Gründe oder Absichten zu erkennen sind. Umgekehrt sprechen wir von Gründen und Absichten nur dort, wo es (zumindest vermeintlich) etwas gibt, was beabsichtigt oder begründet werden kann. Anders ausgedrückt: Wir lernen nicht durch Erfahrung, dass wir beabsichtigen müssen, zum Supermarkt zu gehen, um dort hin zu gelangen. Es scheinen hier begriffliche Abhängigkeiten vorzuliegen, welche es fraglich erscheinen lassen, ob Handlungserklärungen aus Gründen bzw. Absichten empirisch gehaltvoll sind. Zwar scheinen Handlungserklärungen, anders als viele analytischen Sätze, in gewisser Hinsicht informativ, aber ein Blick auf die Logik zeigt uns, dass analytische Zusammenhänge durchaus nicht immer transparent sein müssen (vgl. Brandtstädter & Sturm, 2004, S. 18). Die Frage nach dem empirischen Gehalt von Handlungserklärungen ist also durchaus angemessen. Unangebracht dagegen ist sicherlich die *„relaxed, if not neglectful, attitude toward these traditional controversies“*, welche Jochen Brandtstädter (2007, S. 51) seinen Fachkollegen innerhalb der Psychologie zuschreibt. Im Folgenden will ich einen Überblick über die argumentative Struktur der vorliegenden Untersuchung geben.

1.2. Überblick über die Argumentation der Untersuchung

Die argumentative Struktur der vorliegenden Arbeit besteht in einer Reihe ineinander verwobener Argumente. Dabei gibt es ein Argument höchster Stufe, welches ich als „Hauptargument“ bezeichnen will. Dieses Hauptargument stellt eine Widerlegung der SMH dar; seine Konklusion besteht also in der Behauptung der Falschheit der SMH. Die anderen Argumente dienen der Untermauerung der Prämissen des Hauptarguments. In diesem Überblick will ich ausschließlich das Hauptargument und die Argumente der nächst niedrigeren Stufe besprechen. Da das Hauptargument drei Prämissen besitzt, gibt es drei solcher Argumente der nächst niedrigeren Stufe. Formale Gültigkeitsnachweise werden für das Hauptargument sowie für die Argumente für zwei seiner Prämissen gegeben. Für die erste Prämisse des Hauptarguments wird nicht auf formaler Ebene argumentiert, auf die Gründe hierfür

werde ich zurückkommen.² Ich möchte zunächst das Hauptargument besprechen, um anschließend einen Überblick über die Argumente für dessen Prämissen zu geben.

1.2.1. Das Hauptargument

Wie bereits erwähnt, kommt das Hauptargument zu der Konklusion, dass die SMH falsch ist. Ausgangspunkt ist hierbei die Feststellung, dass jeder Satz entweder einen begrifflichen oder einen empirischen Zusammenhang wiedergibt und es sich hier um ein ausschließendes „oder“ handelt. Da eine kausale Beziehung eine empirische Beziehung ist, kann ein Satz, welcher einen begrifflichen Zusammenhang wiedergibt, keine kausale Relation behaupten. (Diese Prämisse werde ich mit (P1) bezeichnen.) Es wird sich nun herausstellen, dass sich der psychologische Kerngehalt der SMH als ein Satz formulieren lässt – nennen wir ihn „(PG)“³ (für ‚psychologischer Gehalt‘) –, von welchem die SMH annehmen muss, dass er eine kausale Relation angibt. (Die These, dass die SMH die Behauptung dieses Kausalverhältnisses impliziert, nennen ich (T1).) Des Weiteren wird sich aber zeigen, dass der Satz (PG) kein kausales, sondern ein begriffliches Verhältnis wiedergibt. (Diese These heiße (T2).) In Abbildung 1 wird ein formaler Beweis der Gültigkeit des Hauptarguments gegeben.⁴

Nun lässt sich die Struktur der vorliegenden Arbeit sehr einfach beschreiben: Die erste Prämisse des Hauptarguments, d. i. (P1), soll in Kapitel 2 verteidigt werden. In Kapitel 3 wird die SMH und der Stand der zugehörigen empirischen Forschung dargestellt. Hiermit werden die Grundlagen dafür gelegt, in Kapitel 4 für die Prämisse (T1) zu argumentieren. Letzteres wird durch das Herausarbeiten eines genuin psychologischen Gehalts der SMH geschehen. Dieser genuin psychologische Gehalt kann als der – bereits erwähnte – Satz (PG) formuliert werden. In Kapitel 5 wird gezeigt, dass die SMH annehmen muss, dass (PG) ein kausales Verhältnis wiedergibt, und dass diese Annahme falsch ist, da (PG) eine begriffliche Beziehung angibt (d. i., es gilt (T2)). Hiermit ist die SMH gemäß dem Hauptargument widerlegt. Kapitel 6 zieht diese Schlussfolgerung und untersucht mögliche Einwände sowie die Möglichkeit, das gewonnene Ergebnis zu verallgemeinern.

² s. Kapitel 1.2.2.

³ Für die spätere Ausformulierung s. S. 97 unten.

⁴ Aus Gründen der Eleganz wurde für die Beweise die Methode der logischen Tableaus verwendet. Für Hinweise bezüglich dieser Methode sei auf das *Handbook of tableau methods* (D'Agostino, Gabbay, Hähnle & Posegga, 1999) verwiesen. Ich vernachlässige in meiner Darstellung das Problem der Quantifizierung über Propositionen.

Hauptargument:

Die nicht-logischen Konstanten stehen für:

$B(p)$: p gibt eine begriffliche Beziehung wieder.

$K(p)$: p gibt eine kausale Beziehung wieder.

SMH: Somatische Marker Hypothese.

PG: psychologischer Gehalt der Somatischen Marker Hypothese.

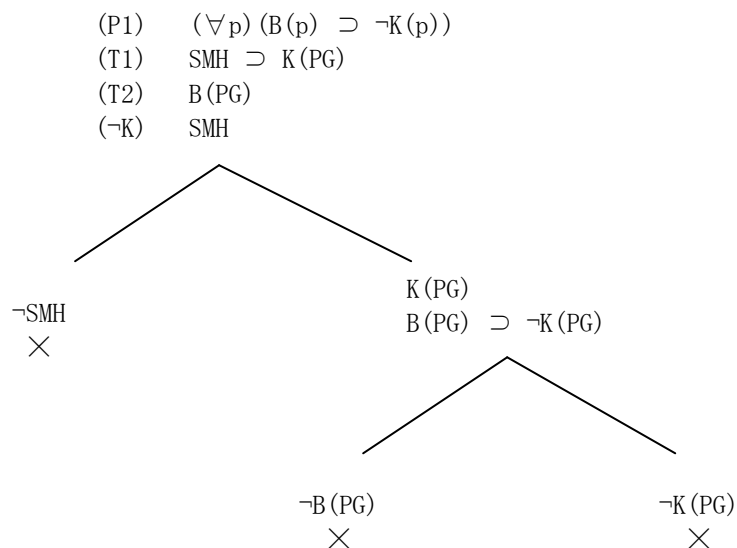
(P1) $(\forall p) (B(p) \supset \neg K(p))$

(T1) $SMH \supset K(PG)$

(T2) $B(PG)$

(K) $\neg SMH$

Gültigkeitsprüfung durch Konsistenzprüfung der Satzmenge, welche aus den Prämissen und der Negation der Konklusion besteht:



Der Baum ist geschlossen, also ist das Argument gültig.

Abbildung 1: Gültigkeitsnachweis für das Hauptargument.

1.2.2. Das Argument für die Prämisse (P1)

Es wird, wie bereits erwähnt, kein formalisiertes Argument für (P1) gegeben. Dies hat mehrere Gründe: Erstens ist die zentrale Frage, welche sich hinter (P1) verbirgt, die Frage, was begriffliche Relationen sind. Dies wiederum ist eine Kernfrage der Semantik. Oft wird diese Frage übersetzt in die Frage „Wie können wir semantische Relationen formal fassen?“. Es wurden unterschiedliche Vorschläge hierzu gemacht, doch konnte sich bis heute keiner dieser Vorschläge endgültig durchsetzen. Um die Frage, was eine begriffliche Relation sei, formal zu fassen, bedarf es jedoch eben einer solchen formalen Semantik. Um ein formalisiertes Argument für (P1) zu geben, müsste man sich also einer dieser formalen Semantiken bedienen. Dies bereitet jedoch zwei Probleme. Zum einen beginge man hierdurch, im gegenwärtigen Kontext, eine *petitio principii*, da man dabei eine Semantik voraussetzen müsste. D. h., man müsste bereits geklärt haben, was eigentlich in Frage steht, nämlich was begriffliche und empirische Relationen unterscheidet. Und zum Zweiten wäre der hierfür notwendige formale Apparat so komplex und schwerfällig, dass es den Rahmen der gegenwärtigen Arbeit sprengen würde, die Gültigkeit von (P1) mittels seiner formal zu beweisen. Ein dritter Grund liegt in der Struktur des Arguments selbst. Das Kapitel 2 bietet im Wesentlichen einen historischen Überblick über die Diskussion bezüglich begrifflicher Relationen. Ich hoffe, durch eine solche historische Reflexion auf die methodischen Grundlagen dieser Arbeit die These plausibel zu machen, dass begriffliche und empirische Relationen streng unterschieden werden müssen. Doch die Überzeugungskraft dieses Unternehmens hängt wesentlich davon ab, zu sehen, dass historische und zeitgenössische Alternativen zu einer solchen Unterscheidung sich als unhaltbar erweisen haben. Ein solches 'historisches Argument' lässt sich kaum in das Korsett der formalen Logik zwängen. Trotzdem will ich einige vorausblickende Hinweise zum Inhalt dieses Kapitels geben.

Im Wesentlichen versuche ich in Kapitel 2, eine pragmatische Bedeutungstheorie gegenüber einer wahrheitsfunktionalen/referenzbasierten Bedeutungstheorie zu verteidigen. Erstere ist mit Namen wie Wittgenstein, Austin, Hacker oder Brandom verbunden, die zweite mit Namen wie Carnap, Kripke oder Lewis. Die entscheidende Schwäche einer wahrheitsfunktionalen/referenzbasierten Bedeutungstheorie besteht darin, dass ihre grundlegenden Begriffe, wie Referenz oder Wahrheit, unanalysiert bleiben müssen. Eine normalsprachliche Gebrauchstheorie der Bedeutung hingegen geht nicht von einem Grundlegungsverhältnis aus, sondern klärt jeden Begriff 'situationsbedingt' anhand seiner Einbettung in unsere alltägliche Sprache. Darüber hinaus stellt sich für eine wahrheitsfunktionale/referenzbasierte Be-

deutungstheorie seit jeher die Frage, ob die Bedeutung sprachlicher Zeichen in abstrakten Entitäten oder in empirischen Tatsachen besteht. Bei ersterer Antwort schließt sich die weitere Frage an, wie wir diese Entitäten erfassen können, bei der zweiten ergibt sich die Frage, wie diese empirischen Tatsachen ihren Charakter von Notwendigkeit erlangen. Eine Gebrauchstheorie kann diese beiden Probleme umgehen.

Eine weitere wesentliche Aufgabe des Kapitels 2 besteht darin, die heute weitverbreitete Annahme zu entkräften, Willard V. O. Quine habe mit seinem Aufsatz *Two Dogmas of Empiricism* (1964) gezeigt, dass es keine sinnvolle Unterscheidung zwischen empirischen und begrifflichen Relationen geben könne. Tatsächlich wurden seit der Veröffentlichung des Aufsatzes überzeugende Gegenargumente vorgebracht, sodass es heute als gesichert betrachtet werden kann, dass Quines Argumente zwar vielleicht die Position Rudolf Carnaps (Quines eigentliches Ziel) trafen, dabei aber kein allgemeingültiges Ergebnis zutage förderten.

1.2.3. Das Argument für die These (T1)

Die These (T1) besagt: Wenn die SMH wahr ist, so ist es auch wahr, dass der genuin psychologische Gehalt der SMH, d. i. (PG), eine kausale Relation wiedergibt. Zu diesem Ergebnis gelangt man mittels fünf Prämissen. Die erste Prämisse (T1.1) ist dabei die komplexeste: Gegeben sei eine Theorie t . Diese Theorie t wird dabei als eine Menge von Sätzen verstanden. Nun sei diese Theorie eine neurobiologische Theorie (dies entspricht NBG(t) im Formalismus), d. h., die Theorie beinhalte mindestens eine neurobiologische Gesetzmäßigkeit. Des Weiteren gebe es zwei Sätze p und q , welche durch t impliziert werden. Beide dieser Sätze verwenden psychologische Prädikate (dies sei durch $\phi(p)$ ausgedrückt); dabei bildet einer dieser Sätze einen Teil des Antecedens der neurobiologischen Gesetzmäßigkeit der Theorie t (im Formalismus steht hierfür $A(p,t)$), der andere bildet einen Teil des Consequens dieses Gesetzes (dies wird durch $C(q,t)$ wiedergegeben). Nun besagt (T1.1), dass es in diesem Fall einen Satz r gibt, so dass r eine rein psychologische Gesetzmäßigkeit darstellt (im Formalismus wird dies durch $\Psi(r)$ ausgedrückt), welche durch t und bestimmte Brückenprinzipien von t (diese seien mit BP_t bezeichnet) impliziert wird. In diesem Fall ist das psychologische Gesetz (also r) genauso eine kausale Gesetzmäßigkeit wie das neurobiologische Gesetz, von welchem es abgeleitet ist. Der Grundgedanke hierbei ist sehr einfach: Die biologische Psychologie bzw. die kognitive Neurowissenschaft versprechen uns Erkenntnisse, welche den Namen „psychologische Erkenntnisse“ verdienen. Um dies zu leisten, müssen solche Theorien Sätze implizieren, welche psychologische Prädikate enthalten. Wenn nun sowohl die Antecedensseite, als auch die Consequensseite der von der Theorie behaupteten Gesetz-

mäßigkeit psychologische Prädikate enthalten, so lassen sich Konditionale formulieren, welche kein neurowissenschaftliches Fachvokabular mehr enthalten, sondern sich ausschließlich auf der Ebene der Beschreibung ganzer Personen bewegen (vgl. Bennett & Hacker, 2003). Lässt sich ein solcher Satz mittels geeigneter Brückenprinzipien aus der neurobiologischen Theorie ableiten, so handelt es sich hierbei um ein psychologisches Gesetz. Dieses Gesetz muss nun als eine bloße Umformulierung des ursprünglichen neurobiologischen Gesetzes betrachtet werden; daher muss es – ebenso wie das ursprüngliche Gesetz – eine kausale Relation angeben. Die SMH ist nun genau eine solche neurobiologische Theorie; und das psychologische Gesetz, welches sich aus ihr gewinnen lässt, ist der Satz (PG). Dies besagen die Prämissen (T1.2) und (T1.3). Die Prämisse (T1.4) fügt dem die Trivialität hinzu, dass kausale psychologische Gesetze Sätze sind, welche eine kausale Relation wiedergeben. Und (T1.5) behauptet, dass sich aus der SMH tatsächlich die erforderlichen Brückenprinzipien gewinnen lassen. Eine Formalisierung des Arguments findet sich in Abbildung 2 und der zugehörige Gültigkeitsbeweis in Abbildung 3.

Argument für (T1):

Die nicht-logischen Konstanten stehen für:

- $NBG(t)$: t ist eine neurobiologische Theorie.
 $\phi(p)$: p enthält ein psychisches Prädikat.
 $A(p, t)$: p ist Teil des Antezedens eines Gesetzes von t .
 $C(q, t)$: p ist Teil des Consequens eines Gesetzes von t .
 BP_t : Brückenprinzipien von t .
 $\Psi(p)$: p ist ein Konditional auf rein psychologischer (personaler) Ebene.
 $KPG(p)$: p ist ein kausales, psychologisches Gesetz.
 $K(p)$: p gibt eine kausale Beziehung wieder.
 SMH : somatische Marker Hypothese.
 PG : psychologischer Gehalt der Somatischen Marker Hypothese.

- (T1. 1) $(\forall t) ((NBG(t) \wedge (\exists p) (\exists q) ((t \supset p \wedge q) \wedge (\phi(p) \wedge \phi(q)) \wedge A(p, t) \wedge C(q, t))) \supset (\forall r) (((t \wedge BP_t \supset r) \wedge \Psi(r)) \supset KPG(r)))$
 (T1. 2) $NBG(SMH) \wedge (\exists p) (\exists q) ((SMH \supset p \wedge q) \wedge (\phi(p) \wedge \phi(q)) \wedge A(p, SMH) \wedge C(q, SMH))$
 (T1. 3) $(SMH \wedge BP_{SMH} \supset PG) \wedge \Psi(PG)$
 (T1. 4) $(\forall p) (KPG(p) \supset K(p))$
 (T1. 5) $SMH \supset BP_{SMH}$

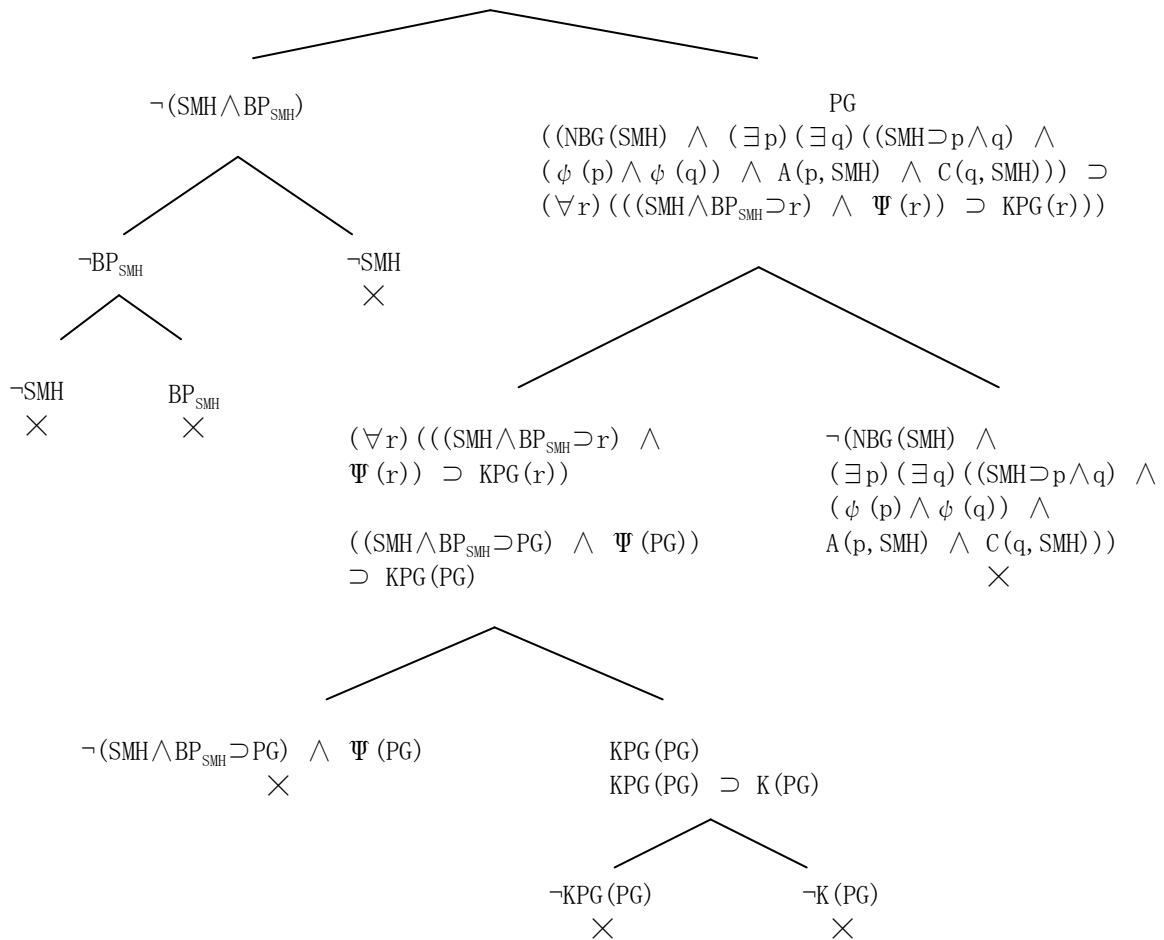
 (T1) $SMH \supset K(PG)$

Abbildung 2: Formale Darstellung des Arguments für die These (T1).

Argument für (T1):

Gültigkeitsprüfung durch Konsistenzprüfung der Satzmenge, welche aus den Prämissen und der Negation der Konklusion besteht:

- (T1. 1) $(\forall t) ((NBG(t) \wedge (\exists p)(\exists q)((t \supset p \wedge q) \wedge (\phi(p) \wedge \phi(q)) \wedge A(p, t) \wedge C(q, t))) \supset (\forall r)((t \wedge BP_t \supset r) \wedge \Psi(r)) \supset KPG(r)))$
- (T1. 2) $NBG(SMH) \wedge (\exists p)(\exists q)((SMH \supset p \wedge q) \wedge (\phi(p) \wedge \phi(q)) \wedge A(p, SMH) \wedge C(q, SMH))$
- (T1. 3) $(SMH \wedge BP_{SMH} \supset PG) \wedge \Psi(PG)$
- (T1. 4) $(\forall p)(KPG(p) \supset K(p))$
- (T1. 5) $SMH \supset BP_{SMH}$
- ($\neg T1$) $SMH \wedge \neg K(PG)$
- SMH
 $\neg K(PG)$
 $\Psi(PG)$



Der Baum ist geschlossen, also ist das Argument gültig.

Abbildung 3: Gültigkeitsnachweis für das Argument für die These (T1).

Es sei bereits an dieser Stelle darauf hingewiesen, dass ich versucht habe, die gesamte Argumentation so zu gestalten, dass sie von möglichst wenigen Annahmen bezüglich inter-theoretischer Reduktionen oder bezüglich des Begriffs der Kausalität abhängt. Dass dies nicht durchgängig möglich war, kann nicht überraschen. Es sollte aber klar sein, dass ich die Begriffe „Brückenprinzip“ und „kausales Verhältnis“ hier wesentlich weiter gebrauche, als dies vielfach üblich ist. Unter „Brückenprinzip“ verstehe ich jede Art von Verhältnis, welches es gestattet, Aussagen einer Wissenschaft aus Aussagen einer anderen zu gewinnen, egal ob dieses Verhältnis als apriorisch, empirisch oder als Abbildungsverhältnis gedacht wird (vgl. hierzu auch Schon, 2005, S. 27ff.). Was den Begriff „kausales Verhältnis“ betrifft, will ich mich ebenfalls auf keine einschlägige Theorie festlegen. Ich setze allerdings tatsächlich voraus, dass der Kausalbegriff aufs Innigste mit dem Gesetzesbegriff verbunden ist. Dies trifft meines Erachtens für alle gängigen Theorien der Kausalität⁵ zu. Allerdings will ich keineswegs behaupten, dass eine solche Analyse des Kausalbegriffs richtig ist. Tatsächlich erachte ich es für erstrebenswert, zumindest die Möglichkeit eines teleologischen Kausalbegriffs offen zu lassen. Es ist jedoch offensichtlich, dass die kognitive Neurowissenschaft diese Möglichkeit ausschließt, da die Annahme von Finalursachen die Grundlagen dieser Wissenschaft in Frage stellen würde. Daher soll mit dem Begriff „Kausalität“ im Rahmen dieser Arbeit stets Effizienzkausalität im Sinne der genannten Theorien gemeint sein.

1.2.4. Das Argument für die These (T2)

Die These (T2) besagt, dass der Satz (PG) ein begriffliches Verhältnis wiedergibt. Das Argument hierfür ist sehr einfach: Wenn (PG) kein begriffliches Verhältnis angibt, so muss der Satz eine kausale Relation behaupten. (Diese These heiße (T2.1).) (PG) gibt aber keine kausale Relation wieder. (Diese These heiße (T2.2).) Also muss (PG) eine begriffliche Relation angeben (d. i., (T2) ist wahr). Der Vollständigkeit halber wird ein formaler Beweis dieses schlichten *modus tollens* in Abbildung 4 gegeben.

⁵ Hiermit meine ich insbesondere Analysen des Kausalbegriffs, welche auf dem deduktiv-nomologischen Erklärungsschema beruhen (vgl. Popper, 1973, S. 31), sowie die kontrafaktische Analyse der Kausalität (vgl. Kim, 2007, S. 233; Lewis, 1973).

Argument für (T2):

Die nicht-logischen Konstanten stehen für:

$B(PG)$: p gibt eine begriffliche Beziehung wieder.

$K(PG)$: p gibt eine kausale Beziehung wieder.

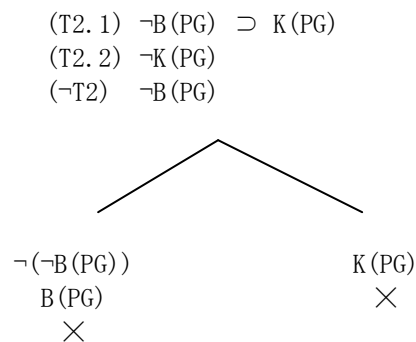
PG : psychologischer Gehalt der Somatischen Marker Hypothese

(T2. 1) $\neg B(PG) \supset K(PG)$

(T2. 2) $\neg K(PG)$

(T2) $B(PG)$

Gültigkeitsprüfung durch Konsistenzprüfung der Satzmenge, welche aus den Prämissen und der Negation der Konklusion besteht:



Der Baum ist geschlossen, also ist das Argument gültig.

Abbildung 4: Gültigkeitsnachweis für das Argument für die These (T2).

Die erste Prämisse dieses Schlusses ergibt sich aus dem soeben angesprochenen weiten Kausalitätsbegriff. Es ist der Anspruch der kognitiven Neurowissenschaft, dass ihre Erklärungen kausaler Natur sind. Es sollen eben solche Erklärungen ausgeschlossen werden, welche auf allgemeingültigen, kontrafaktische Konditionalsätze unterstützenden Konditionalen beruhen, aber nicht kausaler Natur sind, wie etwa teleologische oder vitalistische Erklärungen. Damit ist aber auch klar, dass ein allgemeingültiges, kontrafaktische Konditionalsätze unterstützendes Konditional als die Angabe eines kausalen Verhältnisses betrachtet werden muss. Natürlich ergeben sich hier die Probleme der supervenienten Kausalbeziehungen und der Scheinkorrelationen. Aber hier ist die kognitive Neurowissenschaft selbst bestrebt, diese Probleme zu lösen bzw. zu umgehen, da sie sich im Allgemeinen nicht dem Vorwurf des Epiphänomenalismus aussetzen will (vgl. Kim, 2007; Loewer, 2007).

Der Leser wird sofort bemerkt haben, dass sich hier eine wesentliche Vereinfachung des Argumentationsgangs anbietet. Setzt man Prämisse (T2.2) in das Hauptargument ein, so folgt die Falschheit der SMH unter Zuhilfenahme von (T1) unmittelbar *per modus tollens*. Die Prämissen (P1) und (T2) wären damit überflüssig. Warum habe ich die vorliegende Arbeit nicht auf der vereinfachten Argumentationsstruktur aufgebaut? Hierfür gibt es zwei Gründe: Erstens übernimmt das Kapitel 2 nicht nur die Rolle, welche die Formalisierung der Prämisse (P1) zuweist, sondern ist bestrebt, die methodischen Grundlagen der vorliegenden Arbeit zu klären. Hierauf sollte nicht verzichtet werden, selbst wenn es die Folgerichtigkeit der Argumentation erlaubt. Zweitens ist die Wahrheit der Prämisse (T2.2) nur auf der in Kapitel 2 erarbeiteten Grundlage einsichtig. Daher wäre die Argumentation ohne die redundanten Prämissen durchaus unsystematisch und weniger überzeugend. Nichtsdestoweniger lässt diese Redundanz erkennen, dass die Argumentationsstruktur nicht unmittelbar von einer bestimmten Bedeutungstheorie abhängt. Wenn sich auf einer alternativen Grundlage ebenfalls zeigen lässt, dass Handlungserklärungen keine Kausalerklärungen sind, lässt sich die Falschheit der SMH auch ohne die Thesen des Kapitels 2 belegen. Angesichts der Vielfalt der Argumente, welche gegen die These spechen, Handlungserklärungen seien kausaler Natur, erscheint die Existenz solcher Alternativen durchaus plausibel.

1.3. Abgrenzung dieser Arbeit von anderer Kritik der SMH

Da die SMH bereits aus unterschiedlichen Perspektiven kritisiert wurde, möchte ich die vorliegende Arbeit gegen diese anderen Zugänge abgrenzen. Zunächst gilt es, die empirische von der theoretischen Kritik zu unterscheiden. Es wurde verschiedentlich versucht, die SMH auf

der Grundlage empirischer Befunde zu kritisieren (z. B. Chiu & Lin, 2007; Lin, Chiu, Lee & Hsieh, 2007; Maia & McClelland, 2004; Maia & McClelland, 2005). Obwohl eine solche Kritik durchaus Zweifel an der Richtigkeit der SMH aufkommen lassen kann,⁶ ist ihre Tragweite doch sehr begrenzt. Sie führt nicht zu einem Überdenken der theoretischen Grundlagen des Vorhabens, Handlungen neurobiologisch zu erklären, sondern zu einer Adjustierung einzelner Konstrukte und Parameter der SMH. Eine solche Kritik kann kaum auf andere Theorien verallgemeinert werden. Dagegen ist die hier vorgelegte Kritik rein theoretischer Natur, sie beansprucht Probleme aufzuzeigen, welche dem gesamten Ansatz zugrunde liegen, Handlungen neurobiologisch zu erklären. Die SMH wird dabei lediglich als ein Beispiel betrachtet. Die vorliegende Arbeit formuliert also eine wesentlich breitere und allgemeinere Kritik. Durch die Orientierung am Beispiel der SMH soll dabei die nötige Nähe zur empirischen Forschung gesichert werden.

Die SMH wurde aber auch bereits aus theoretischer Perspektive kritisiert. Die vorliegende Arbeit beinhaltet sowohl Übereinstimmungen als auch Unterschiede zu diesen kritischen Arbeiten. Andrew Gluck (2007) kritisierte die SMH insbesondere aus einer historischen und subjektphilosophischen Perspektive. Die vorliegende Arbeit geht dagegen von einem sprachanalytischen und wissenschaftstheoretischen Ansatzpunkt aus. Während Gluck wichtige Probleme korrekt diagnostiziert, betrachte ich seine eigene ontologische Position als unhaltbar. Eine explizit wissenschaftstheoretische Kritik legte Giovanna Colombetti (2008) vor. Doch obwohl dabei grundlegende Unklarheiten, Ungenauigkeiten und Äquivokationen der SMH überzeugend dargestellt werden, geht Colombetti nicht auf die Frage ein, welche Form adäquate Erklärungen menschlicher Handlungen aufweisen müssen. Eine sprachanalytisch orientierte Kritik wurde von Max R. Bennett und Peter M. S. Hacker vorgelegt (s. Bennett & Hacker, 2003, S. 199ff.). An dieses Werk wird sich die vorliegende Arbeit eng anschließen. Doch auch Bennett und Hacker ziehen keine direkte Verbindung zwischen Handlungserklärungen und der SMH, sondern konzentrieren sich auf Damasio's Theorie der Emotionen. Mit dem Werk Bennetts und Hackers verbindet die vorliegende Arbeit insbesondere die Orientierung an der Spätphilosophie Wittgensteins.

Da auch nach der Publikation der genannten Kritik zahlreiche Arbeiten unter dem Einfluss der SMH veröffentlicht wurden und die Vertreter dieser Theorie sogar beanspruchen, aus der SMH valide Folgerungen bezüglich des Rechtssystems und der Ethik ableiten zu können (vgl. Burns & Bechara, 2007; Damasio, 2005), erscheint eine erneute kritische Prüfung durchaus

⁶ s. Kapitel 3.3.4.

geboten. Hierzu wird die SMH erstmals unter einer explizit handlungstheoretischen Perspektive analysiert.

2. METHODISCHE GRUNDLAGEN – BEGRIFFLICHE UND EMPIRISCHE RELATIONEN

Der zentrale Ansatzpunkt der vorliegenden Untersuchung besteht die Unterscheidung zwischen empirischen und begrifflichen Relationen. Diese Unterscheidung bietet die methodologische Grundlage der in den anschließenden Kapiteln folgende Kritik der SMH. Was hier mit „begriffliche Relationen“ angesprochen ist, wurde in der Tradition seit Immanuel Kant oft als „analytisches Urteil“ bezeichnet. Folgt man vorläufig dieser Terminologie, so setzt die vorliegende Arbeit also die Möglichkeit einer Unterscheidung zwischen synthetischen und analytischen Urteilen voraus. Nun ist aber die Möglichkeit einer solchen Unterscheidung seit der Mitte des zwanzigsten Jahrhunderts einer einflussreichen Kritik unterzogen worden. Entscheidend war hierbei die Veröffentlichung des Aufsatzes *Two Dogmas of Empiricism* (Quine, 1964). Während die Unterscheidung zwischen synthetischen und analytischen Urteilen zum klassischen Methodenkanon der sprachanalytischen Philosophie zählte, besteht heute kein Konsens bezüglich einer solchen methodologischen Festlegung (vgl. Harman, 1996, S. 396). Eine Arbeit wie die vorliegende kann also im heutigen Kontext die Möglichkeit einer solchen Unterscheidung nicht unkritisch voraussetzen. Daher besteht das Ziel dieses Kapitels darin, die Möglichkeit einer Unterscheidung zwischen empirischen und begrifflichen Zusammenhängen zu verteidigen. Ich will also versuchen zu zeigen, dass die Leugnung der Möglichkeit der angesprochenen Unterscheidung heute durchaus zweifelhaft erscheint. Die Tragweite der Leugnung einer solchen Unterscheidung wird oft unterschätzt. Auch kann die Diskussion um analytische Urteile in philosophischen Fachkreisen bis heute nicht als entschieden gelten. Ich will zunächst einige geschichtliche Hintergründe der angesprochenen Unterscheidung darlegen, um dann näher auf die heutige Debatte einzugehen. Dabei will ich die zwei Hauptströmungen der heutigen Diskussion getrennt darstellen. Das Kapitel endet mit dem Vorschlag einer Formulierung der angestrebten Unterscheidung.

2.1. Historisches zu begrifflichen Relationen

2.1.1. Entwicklungen vor Kant

Wie die Geschichte fast jedes philosophischen Themas, so kann auch die Geschichte der Unterscheidung zwischen begrifflichen und empirischen Relationen bei Sokrates und Platon

begonnen werden. Sokrates fordert von seinen Gesprächspartnern in den platonischen Dialogen stets eine strikt allgemeine Charakterisierung des Gesprächsgegenstandes, Beispiele und Geschichten genügen ihm nicht. Er fordert die Angabe begrifflicher Zusammenhänge (vgl. Euthyphron 6e, Platon, 1957, S. 184). Sein Grund dafür ist folgender: Weiß jemand z. B., dass Bello ein Rüde ist, so kann er sich auch sicher sein, dass Bello ein Hund ist. Dagegen kann er sich nicht sicher sein, dass Bello gerne Stöcke apportiert. Es scheint hier also einen Unterschied zu geben: Der Zusammenhang zwischen den Ausdrücken „ist ein Rüde“ und „ist ein Hund“ erlaubt einen logischen Schluss. Wer, wie Sokrates, an der Erkenntnis der Wahrheit durch logisches Schließen interessiert ist, hat daher ein natürliches Interesse daran, diesen Zusammenhängen auf den Grund zu gehen. Dabei ist ein naheliegender Gedanke, dass der erwähnte Zusammenhang etwas mit der Definition der Begriffe „Hund“ und „Rüde“ zu tun hat. Und so heißt es bei Aristoteles:

„Sokrates suchte aber mit gutem Grund das Was, denn er sucht Schlüsse zustande zu bringen, doch das Prinzip der Schlussfolgerung ist das Was. [...] Zweierlei nämlich ist es, was man Sokrates mit Recht zuschreiben dürfte: die Begründung durch Heranführung und das allgemeine Definieren; dies beides nämlich betrifft das Prinzip der Wissenschaft“ (Metaphysik 1078b: 23-29, Aristoteles, 2000, S. 334).

Logische Schlüsse zeichnen sich offensichtlich durch eine bestimmte Art von Notwendigkeit aus. Und diese Notwendigkeit ist offenbar in dem Zusammenhang von Ausdrücken wie „ist ein Hund“ und „ist ein Rüde“ enthalten. Eine der Hauptaufgaben von Platons Ideenlehre war es, diese Notwendigkeit zu erläutern. Man kann hier von begrifflichen Beziehungen sprechen. Wenn solche begrifflichen Beziehungen notwendig und erfahrungsunabhängig bestehen, dann muss ihnen nach Platon auch eine Realität entsprechen (vgl. Politeia V, 475e-480a, Platon, 1994, S. 380ff.). „*The eternal and intelligible Forms were to provide rational thought with objects of knowledge*“ (Cornford, 2000, S. 7). Für Platon war klar, dass diese Formen weder psychischer noch sprachlicher Natur sein können und keinesfalls der Erfahrungswelt angehören (vgl. Ros, 1989, S. 56ff.). Denn all das widerspräche der Notwendigkeit und Objektivität begrifflicher Zusammenhänge. Platon steckte den Rahmen für jede Theorie begrifflicher Relationen ab. Eine solche Theorie muss den Charakter von Notwendigkeit und Objektivität begrifflicher Relationen erklären können.

Bei Aristoteles führt dies zur Unterscheidung von Substanz (οὐσία) und Akzidens (συμβεβηκός). „Akzidens heißt das, das an etwas vorhanden ist und der Wahrheit gemäß von ihm ausgesagt werden kann, jedoch nicht mit Notwendigkeit und nicht in der Regel“ (Meta-

physik 1025a, 14-15, Aristoteles, 2000, S. 153). Hingegen wird „das Was-es-ist-dies-zu-sein, dessen Begriff Definition ist, als Wesen (οὐσία) jedes Einzelnen bezeichnet“ (Metaphysik, 1017b, 21, Aristoteles, 2000, S. 127). Wesensbestimmungen sagen etwas über eine unvergängliche Form, welche sich in einem bestimmten Stoff verwirklicht. Dabei wird ausdrücklich die abgetrennte Existenz des Allgemeinen, also der Begriffe, bestritten (vgl. Metaphysik, 1040b, 26, Aristoteles, 2000). Das Allgemeine liegt in den Dingen (*universalia in rebus*). Begriffliche Relationen bestehen also nicht unabhängig von Erfahrungsgegenständen, aber sind dennoch in unvergänglichen Formen begründet und damit notwendig.

Hiermit ist explizit die Frage nach dem ontologischen Status der Begriffe aufgeworfen. Diese bildet eines der dominierenden Themen des Mittelalters. Vorbereitet durch den Neuplatonismus und durch Augustinus, verbreitete sich im Mittelalter die Auffassung, dass platonische Formen als Ideen im Geiste Gottes zu verstehen seien. So schreibt etwa Thomas von Aquin: „*Necesse est quod in mente divina sit forma ad similitudinem cuius mundus est factus*“ (Summa theologiae, I, qu.15, art. 1., Thomas, 2008). An der Frage, wie diese Formen zu verstehen seien, entzündete sich der Universalienstreit. Die entscheidende Entwicklung des Mittelalters liegt aber in Folgendem: Hier eröffnet sich für das europäische Denken erstmals die Möglichkeit begriffliche Zusammenhänge als etwas mentales zu verstehen.

„Denn in der Entdeckung der psychischen Innenwelt (als einer Welt von Gegenständen, die von der der Gegenstände der Außenwelt durch die Art, zu ihr Zugang gewinnen zu können, scharf unterschieden ist) steckte, wie sich zeigen sollte, ein bis zu jener Zeit noch nicht genutztes Entwicklungspotential. Und es ist die Nutzung dieses Potentials, die zu Lockes ‘new way of ideas’, und ganz allgemein zur neuzeitlichen Konzeption von (allgemeinen beziehungsweise abstrakten) ‘Ideen’, von (allgemeinen) ‘Vorstellungen’, von ‘Begriffen’, usw., führen sollte“ (Ros, 1990, S. 4).

Ein entscheidender Unterschied zwischen dem mittelalterlichen und dem neuzeitlichen Verständnis begrifflicher Relationen besteht darin, dass die Neuzeit Begriffe im Geist des Menschen verortete, während für das Mittelalter ihr Ort im Geiste Gottes lag. Hiermit scheint ein entscheidendes epistemologisches Problem der Antike und des Mittelalters lösbar zu werden. So wie Platon das Problem, wie die Begriffe erkennbar sein können, im Dialog *Parmenides* explizit anspricht, aber keine klare Lösung bereitzuhalten scheint (vgl. Ros, 1989, S. 237), so halten auch mittelalterliche Autoren, wie Cusanus, die Erkennbarkeit des Wesens der Dinge für problematisch. „Als urbildhafte Darstellung des unendlichen Geistes [Gottes] bleibt das Wesen der Dinge [...] der mens des Menschen unfäßlich“ (Volkmann-Schluck,

1984, S. 121). Liegen aber die Ideen oder das Wesen der Dinge im menschlichen Geist, so scheint ein Wissen darum unproblematisch. Als etwas im eigenen Geist werden begriffliche Relationen dem Menschen unmittelbar zugänglich.

„Allgemeine Wörter bezeichnen somit eine *Art* von Dingen; und zwar geschieht das bei jedem dieser Wörter dadurch, daß es als Zeichen für eine im Geist vorhandene Idee auftritt. [...] Daraus erhellt, daß die *Wesenheiten* der Arten [...] von Dingen nichts anderes sind als diese abstrakten Ideen“ (Versuch über den menschlichen Verstand, III, iii, 12, Locke, 1988, S. 17).

Es wird hier eine Möglichkeit geschaffen, klar zwischen empirischen und begrifflichen Zusammenhängen zu unterscheiden. Während empirische Relationen Beziehungen zwischen Gegenständen der Welt sind, bestehen begriffliche Relationen in Beziehungen zwischen Ideen. Wissen über die Beziehungen von Ideen kann erfahrungsunabhängig also *a priori* erworben werden. Doch dem epistemologischen Vorteil des neuzeitlichen Verständnisses steht auch ein Nachteil gegenüber: Wie lässt sich die Notwendigkeit und Objektivität von Begründung und Argumentation verständlich machen, wenn begrifflichen Zusammenhängen keine Realität außerhalb des menschlichen Geistes entspricht? Eine typische neuzeitliche Antwort auf diese Frage besteht im Hinweis auf die Natur der menschlichen Vernunft (vgl. Ros, 1990). Ein paradigmatisches Beispiel bietet Gottfried Wilhelm Leibniz. Leibniz unterscheidet zwischen Vernunftwahrheiten und Tatsachenwahrheiten:

„Die Vernunftwahrheiten sind notwendig, die Tatsachenwahrheiten zufällig. Die ursprünglichen Vernunftwahrheiten sind diejenigen, die ich mit einem Gesamtnamen *identische* nenne, weil sie, ohne uns etwas zu lehren, nur dasselbe zu wiederholen scheinen“ (Neue Abhandlungen über den menschlichen Verstand, IV, 2, § 1, Leibniz, 1996, S. 378).

Alle Vernunftwahrheiten beruhen, nach Leibniz, auf logischen Prinzipien. Man kann keine Vernunftwahrheit bestreiten, ohne in einen Widerspruch zu geraten (vgl. Copleston, 1999, S. 273). Die Objektivität und Notwendigkeit begrifflicher Relationen soll durch deren Fundierung in der Logik gesichert werden. Die Logik aber beruht auf dem Wesen unseres Geistes. Und daher gilt: „[O]ft ist die Betrachtung des Wesens der Dinge nichts anderes, als die Betrachtung des Wesens unseres Geistes“ (Neue Abhandlungen über den menschlichen Verstand, I, 1, § 21, Leibniz, 1996, S. 48). Auf diesem Hintergrund entwickelte sich Kants Verständnis des analytischen Urteils.

2.1.2. Kant

Kants Verständnis der Begriffe „*a priori*“, „*a posteriori*“, „analytisch“ und „synthetisch“, prägt diese Terminologie bis heute. Kants Bestimmung eines analytischen Urteils lautet:

„In allen Urteilen, worin das Verhältnis eines Subjekts zum Prädikat gedacht wird [...] ist dieses Verhältnis auf zweierlei Art möglich. Entweder das Prädikat B gehört zum Subjekt A als etwas, was in diesem Begriffe A (versteckter Weise) enthalten ist; oder B liegt ganz außer dem Begriff A, ob es zwar mit demselben in Verknüpfung steht. Im ersten Fall nenne ich das Urteil *analytisch*, in dem anderen *synthetisch*“ (KrV, B10/A6, Kant, 1966, S. 58).

Wichtig ist für Kant, dass analytische Wahrheiten nicht koextensiv mit notwendigen Wahrheiten sind, wie es Leibniz Vernunftwahrheiten waren. Denn es gibt nach Kant synthetische Urteile *a priori*. Und „Notwendigkeit und strenge Allgemeinheit sind [...] sichere Kennzeichen einer Erkenntnis *a priori*“ (KrV, B4, Kant, 1966, S. 52). Damit stellt sich die Frage, wie Analytizität, Notwendigkeit und Apriorität sich zueinander verhalten. Analytische Urteile sind nach Kant nicht erkenntniserweiternd. Kernelement analytischer Urteile ist ein Verhältnis zwischen Begriffen. Ein Begriff ist aber nach Kant das Vermögen, etwas nach Regeln zu subsumieren bzw. hervorzubringen (vgl. Ros, 1990, S. 67ff.). Ein begriffliches Verhältnis ist demnach ein Verhältnis zwischen Regeln. Und ein analytisches Urteil drückt das Beinhaltetsein einer Regel in einer anderen aus; ähnlich wie die Regel, nicht das Haus zu verlassen, die Regel beinhaltet, nicht ins Kino zu gehen. Aber wie können solche Regeln, wenn sie völlig subjektive, eigenpsychische Entitäten darstellen, Anspruch auf objektive Geltung beanspruchen? Wie können sie zur Überprüfung der Geltung prädikativer Aussagen dienen? Denn dies ist der Kern dessen, was eine begriffliche Relation seit Sokrates ausmachen soll. Wie bei Leibniz, so liegt auch bei Kant die Antwort hierauf im denkenden Subjekt. Er unterscheidet ein empirisches Selbstbewusstsein, welches von der Psychologie erforscht wird, von einem reinen, welches auch die „ursprüngliche Apperzeption“ genannt wird (KrV, B132, Kant, 1966, S. 175). Dabei ist „die synthetische Einheit der Apperzeption der höchste Punkt, an dem man allen Verstandesgebrauch, selbst die ganze Logik, und, nach ihr, die Transzendental-Philosophie heften muß, ja dieses Vermögen ist der Verstand selbst“ (KrV, B133, Kant, 1966, S. 177). Aber der Verstand ist eben auch „das Vermögen der Regeln“ (KrV, B171/A132, Kant, 1966, S. 209). Und so wird die objektive Geltung des Wissens um begriffliche Relationen an die ursprünglich-synthetische Einheit der Apperzeption gebunden, welche eine Voraussetzung allen Denkens überhaupt ist. Dabei ist es ein wesentliches Ver-

dienst Kants, dass er, indem er Begriffe mit Regeln in Zusammenhang brachte, den normativen Charakter begrifflicher Relationen deutlich machte. Es ist „eine fundamentale Einsicht Kants [...], dass] alles, was begrifflich gehaltvoll ist, eine *normative* Signifikanz besitzt“ (Brandom, 2001a, S. 212). Das Operieren mit Begriffen ist stets ein durch Regeln bestimmtes Tun, also ein Tun, bei welchem es ein ‘Falsch’ und ‘Richtig’ gibt. Jede gute Theorie begrifflicher Relationen sollte hierfür eine Erklärung geben können.

2.1.3. Frege

Gottlob Frege kritisiert Kants Ansätze zur Charakterisierung analytischer Urteile als psychologisch. Demgegenüber wollte Frege die Objektivität begrifflicher Relationen stärker betonen (vgl. Bar-Elli, 1996, S. 34ff.). Grundlegend ist für Frege ein Verständnis des Verhältnisses zwischen sprachlichen Zeichen und Begriffen. Zu einem solchen Verständnis gehört die Einsicht, dass „mit einem Zeichen [...] außer dem Bezeichneten, was die Bedeutung des Zeichens heißen möge, noch das verbunden zu denken [ist], was ich den Sinn des Zeichens nennen möchte, worin die Art des Gegebenseins enthalten ist“ (Frege, 1990, S. 144). Was Frege mit „Sinn“ bezeichnete, wird seit Carnap oft „Intension“ genannt; und Freges „Bedeutung“ ist heute als „Extension“, „Bezug“ oder „Referenz“ bekannt. In Opposition zur neuzeitlichen Tradition grenzt Frege den Sinn von jeder Art von Vorstellung ab. „Die Vorstellung unterscheidet sich dadurch wesentlich von dem Sinne eines Zeichens, welcher gemeinsames Eigentum von vielen sein kann und also nicht Teil oder Modus der Einzelseele ist“ (Frege, 1990, S. 146). Frege kehrt sich also von der, seit dem Mittelalter erarbeiteten, Vorstellung von Begriffen als psychischen Entitäten ab.

Eine weitere wesentliche Veränderung gegenüber Kant besteht in Freges Bestimmung analytischer Urteile. Hierzu heißt es in den *Grundlagen der Arithmetik*:

„Es kommt [...] darauf an, den Beweis zu finden und ihn bis auf die Urwahrheiten zurückzuverfolgen. Stößt man auf diesem Wege nur auf die allgemeinen logischen Gesetze und auf Definitionen, so hat man eine analytische Wahrheit, wobei vorausgesetzt wird, dass auch die Sätze mit in Betracht gezogen werden, auf denen etwa die Zulässigkeit einer Definition beruht. Wenn es aber nicht möglich ist, den Beweis zu führen, ohne Wahrheiten zu benutzen, welche nicht allgemein logischer Natur sind, sondern sich auf ein besonderes Wissensgebiet beziehen, so ist der Satz ein synthetischer“ (§ 3, Frege, 1986, S. 15).

Dieser Begriff der Analytizität weist erhebliche Ähnlichkeiten mit Leibniz Vernunftwahrheiten auf und ist bis heute wirkmächtig. Man kann den Kerngedanken so ausdrücken: Analytische Wahrheiten lassen sich allein aus den Gesetzen der Logik und Definitionen ableiten (vgl. Russell, 2008, S. 8). Da diese Idee von Rudolf Carnap übernommen wurde und Quines Kritik sich insbesondere auf Carnap bezieht, kann man Freges Konzeption als Ausgangspunkt der heutigen Diskussion um das Verständnis begrifflicher Relationen betrachten. Doch bevor ich hierauf eingehe, will ich versuchen, einige Ergebnisse zu formulieren.

2.1.4. Einige Lehren aus der historischen Reflexion

Folgende Thesen lassen sich als Ergebnisse dieser historischen Betrachtung festhalten:

- (1) Jede Theorie begrifflicher Zusammenhänge muss deren Rolle im Argumentieren Rechnung tragen. Dies schließt eine Erläuterung dessen ein, was man Objektivität oder Notwendigkeit begrifflicher Beziehungen nennen kann.
- (2) Jede solche Theorie muss die Frage beantworten, wie begriffliche Relationen vom Menschen erkannt werden können. (Dies ist ein typisches Problem platonistischer Auffassungen.)
- (3) Jede solche Theorie sollte auf den normativen Charakter begrifflicher Zusammenhänge eingehen. Es ist zu klären, in welcher Beziehung begriffliche Relationen zu Regelsystemen stehen.
- (4) Von besonderem Interesse dabei ist eine Abgrenzung begrifflicher und empirischer Zusammenhänge sowie gegebenenfalls eine Binnendifferenzierung innerhalb dieser beiden Klassen.

2.2. Die heutige Diskussion um begriffliche Relationen

Ich möchte mich bei der Darstellung der heutigen Diskussion um begriffliche Relationen auf die sogenannte sprachanalytische Philosophie beschränken, da die vorliegende Arbeit dieser Tradition verpflichtet ist und eine Auseinandersetzung mit kontinental-europäischen Ansätzen den Rahmen dieser Arbeit sprengen würde. Innerhalb der aktuellen sprachanalytischen Diskussion lassen sich zwei breite Strömungen feststellen:

„Contemporary theoreticians of meaning can be divided, with a degree of oversimplification, into those seeing the meaning of an expression as principally a matter of what the expression denotes or stands for, and those seeing it as a matter of how the expression is used“ (Peregrin, 2006, S. 1).

Die Entwicklung des ersten Ansatzes geht vom logischen Positivismus aus und führt über die `Methode der Intension und Extension` zur Ausarbeitung der Modallogik und zu aktuellen mehrdimensionalen Ansätzen innerhalb der Semantik. Dieser Ansatz geht von Begriffen wie Referenz oder Wahrheit aus und versucht auf dieser Grundlage, eine Bedeutungstheorie zu entwickeln. Daher will ich hier von referenzbasierten Bedeutungstheorien sprechen. Die zweite Strömung geht auf die Spätphilosophie Wittgensteins zurück. Hierbei wird vom Gebrauch der Worte durch kompetente Sprecher ausgegangen und versucht, den Bedeutungsbegriff rekursiv anhand einer solchen Analyse seines Gebrauchs zu klären. Die jüngste Ausarbeitung dieses Ansatzes kann im sogenannten Inferentialismus erblickt werden. Ich will in diesem Fall von gebrauchsbasierten Bedeutungstheorien reden.

2.2.1. Referenzbasierte Bedeutungstheorien

Als gemeinsame Hintergrundannahmen der Autoren unterschiedlicher referenzbasierter Bedeutungstheorien können folgende Thesen gelten:⁷ (1) Das Wesen der Begriffe besteht in der Repräsentation, sprich: darin, dass sich durch Begriffe Sachverhalte wiedergeben lassen. (2) Dies führt dazu, dass die Grundlage jedes Verständnisses begrifflicher Zusammenhänge die Extension (Referenz) ist. Die gesamte semantische Theorie wird als Funktion auf Extensionen aufgebaut, wobei Wahrheitswerte als Extensionen von Aussagen/Propositionen verstanden werden. Anschließend an eine solche semantische Theorie wird dann versucht, die Pragmatik auf diese Semantik zu gründen. (3) Die Erklärungsrichtung in den hier zusammengefassten Ansätzen verläuft stets vom Teil zum Ganzen. Propositionaler Gehalt wird durch den Gehalt der Bestandteile der Proposition erklärt. Die Güte einer Inferenz wird durch die Struktur der Sätze bestimmt, welche sie bilden. Diese Merkmale unterscheiden diesen Ansatz grundlegend von einer gebrauchsbasierten Bedeutungstheorie.

2.2.1.1. Carnap

Rudolf Carnap war einer der prominentesten Vertreter des logischen Positivismus, also einer Bewegung, welche durch eine radikale Kritik aller überkommenen Metaphysik gekennzeichnet war. Diese Kritik wurde wesentlich durch den Verifikationsgrundsatz getragen (vgl. Pears & Kenny, 1995, S. 278). Nach diesem ist der Sinn eines Satzes die Methode seiner

⁷ Für einige der besprochenen Autoren bedürften diese Punkte durchaus einiger Qualifikationen (insbesondere für Chalmers und Boghossian). Doch zum Zweck einer prägnanten Darstellung versuche ich, die Unterscheidung dessen, was ich referenzbasierte Bedeutungstheorien nenne, gegenüber den gebrauchsbasierten Bedeutungstheorien schärfer zu ziehen, als sie ist. Tatsächlich sind die Beziehungen zwischen diesen Ansätzen zahlreich und die Unterschiede nicht immer eindeutig.

Verifikation. Alle Erkenntnisansprüche, welche sich nicht durch die Methoden der empirischen Wissenschaft überprüfen lassen, stehen somit unter dem Verdacht der Sinnlosigkeit. Damit stellt sich aber die Frage, wie die Wahrheiten der Mathematik und Logik als sinnvoll betrachtet werden können (vgl. Russell, 2008, S. 8). Die Lösung dieses Problems sah Carnap darin, diese Sätze als analytisch zu betrachten. Als analytische Sätze gehören sie weder zur Metaphysik noch zur empirischen Wissenschaft, da sie keine Behauptungen über die Welt darstellen; können aber dennoch als wahr gelten, da sie Tautologien sind. Gleichzeitig soll so ihre Apriorität und Notwendigkeit erklärt werden. Sofern sich alle Wahrheiten bezüglich begrifflicher Relationen als inhaltsleere Tautologien verstehen lassen, scheinen sie dem logischen Positivismus ontologisch unproblematisch. Damit ist uns aber der logische Positivismus eine Erklärung dafür schuldig, dass und wie all diese Wahrheiten so verstanden werden können. Eine solche Erklärung versucht Carnap in *Meaning and Necessity*. Wie Frege, so war auch Carnap von der „Unzulänglichkeit der Sprache“ (Frege, 1998, S. X) überzeugt. Um wissenschaftliche Exaktheit zu erreichen, ist die Konstruktion logisch einwandfreier Kunstsprachen erforderlich. Ein umgangssprachlicher Terminus ist vage und kann nicht definiert werden, er kann nur präzisiert werden, indem ihm ein logisch exaktes Gegenstück in einer Kunstsprache gegeben wird. Diese Schärfung und „Reinigung“ eines Terminus nennt Carnap „Explikation“ (vgl. Carnap, 1972, S. 10). Auch die Frage nach dem Wesen begrifflicher Relationen soll durch Explikation geklärt werden. Termini, welche typischerweise zur Charakterisierung begrifflicher Zusammenhänge benutzt werden, sind beispielsweise: „Bedeutung“, „notwendig“, „analytisch“ und „logisch wahr“. Carnap will nun diese Termini so explizieren, dass Mathematik und Logik für ontologisch unproblematisch erklärt werden können, während die Metaphysikkritik ihre Durchschlagskraft behalten soll. Hierzu muss Carnap sie in einer idealen Kunstsprache definieren:

„Der Begriff der *L-Wahrheit* wird hier als Explikation für das definiert, was die Philosophen logische oder notwendige oder analytische Wahrheit nennen. Die Definition führt zu dem Ergebnis, dass ein Satz in einem semantischen System *L-wahr* ist, wenn und nur wenn die semantischen Regeln des Systems für die Begründung seiner Wahrheit genügen. Die Begriffe der *L-Falschheit*, *L-Implikation* und *L-Äquivalenz* werden als Explikata für logische Falschheit, logische Implikation oder Notwendigkeit beziehungsweise wechselseitige logische Implikation definiert. Ein Satz wird *L-determiniert* genannt, wenn er entweder *L-wahr* oder *L-falsch* ist, sonst wird er *L-indeterminiert* oder *tatsachenabhängig*

genannt. Der letztere Begriff ist ein Explikatum dessen, was Kant synthetische Urteile genannt hat“ (Carnap, 1972, S. 9f.).

Um diese Explikata zu bilden, geht Carnap – auf Frege zurückgreifend – von der Unterscheidung zwischen Intension und Extension eines sprachlichen Ausdrucks aus. Während die Extension eines Prädikats die Klasse der Objekte darstellt, auf welche das Prädikat korrekt angewendet werden kann, ist die Intension die Eigenschaft, für welche das Prädikat steht (vgl. Carnap, 1972, S. 21). Außerdem führt Carnap den Terminus „Zustandsbeschreibung“ ein. Eine Zustandsbeschreibung ist eine Klasse von Sätzen, „die für jeden Atomsatz⁸ entweder diesen Satz oder seine Negation enthält, aber nicht beide und keine anderen Sätze“ (Carnap, 1972, S. 11). Dann lässt sich *L*-wahr wie folgt definieren: „Ein Satz s_i ist **L-wahr** (in [dem semantischen System] S_1) =_{Def} s_i gilt in jeder Zustandsbeschreibung (in S_1)“ (Carnap, 1972, S. 13). Wendet man diese Idee auf die Intension von Prädikaten an, so ergibt sich, dass zwei Prädikate genau dann intensionsgleich sind, wenn sie in allen Zustandsbeschreibungen die gleiche Extension besitzen. Damit ist man der Explikation von „Bedeutung“, „notwendig“, „analytisch“ und „logisch wahr“ einen wesentlichen Schritt näher gekommen. Man kann jetzt sagen, dass ein Satz dann und nur dann notwendig, analytisch und logisch wahr ist, wenn er aus den semantischen Regeln der Sprache folgt.⁹ Die Unterscheidung zwischen begrifflichen und empirischen Zusammenhängen ist damit natürlich immer nur relativ zu einer Sprache gegeben; Carnap spricht hier auch von *linguistic framework*. Begriffliche Zusammenhänge zeichnen sich für den logischen Positivismus genau durch ihre Inhaltsleere aus. Es ist eine direkte Konsequenz dieser Idee, dass die Frage, welches zweier *linguistic frameworks* das korrekte ist, keine gehaltvolle Antwort haben kann. Deshalb vertrat Carnap bezüglich der Wahl eines solchen Rahmensystems ein Toleranzprinzip: Jedem Wissenschaftler sollte es erlaubt sein, das ihm am geeignetsten erscheinende Rahmensystem frei zu wählen. Fragen der Wahl eines solchen Systems seien Rahmensystem-extern.

„Framework-external deliberations are not guided by any rules of justification [...]. Carnap calls the kinds of considerations that play a role in a choice of framework ‘pragmatic’, and he insists on distinguishing them from the ‘theoretical’ reasons one might have for choosing a hypothesis within a framework“ (George, 2000, S. 2).

⁸ Ein Atomsatz ist ein „Prädikat vom Grade n , gefolgt von n Individuenkonstanten“ (Carnap, 1972, S. 6).

⁹ Zu den Regeln der Sprache gehört dabei nicht nur das Grundgerüst der Syntax und Semantik, sondern auch spezielle Bedeutungspostulate für einzelne Prädikate (also deren Definitionen).

Damit ist klar, dass in Carnaps System die Wahrheit derjenigen Sätze, welche begriffliche Zusammenhänge wiedergeben, eine Sache der Wahl ist. Anders ausgedrückt: Begriffliche Wahrheiten sind Wahrheiten durch Konvention. Sind die Konventionen in einer formal korrekten Weise gegeben, so ergibt sich dadurch die Unterscheidung zwischen begrifflichen und empirischen Relationen. Auf diese Position zielte Quines Kritik.

2.2.1.2. Quine

Das Anliegen Quines kann in einem radikalen Zu-Ende-Denken des Empirismus gesehen werden. Die einzige legitime Rechtsquelle einer Erkenntnis ist für Quine die Erfahrung. Er geht darin aber so weit, dass ihm Carnaps Erklärung analytischer Wahrheiten als metaphysisch, und damit sinnlos erscheint. *„Quine is pushing against Carnap the very demands that Carnap had pushed against the metaphysicians“* (Creath, 2004, S. 49). Der zentrale Punkt der gesamten Quine'schen Kritik besteht dabei darin, dass auch die Wissenschaft von der Sprache dem Verifikationsgrundsatz entsprechen müsse. Damit Carnaps Rede sinnvoll sei, müsse sie empirische Signifikanz besitzen. Wichtig ist hierbei insbesondere, dass Quines Kritik explizit auf die Position Carnaps zielt und fraglich ist, inwieweit sie sich verallgemeinern lässt.

Quines kritische Argumente zerfallen in zwei Typen. Die erste Kritik verlangt nach einer zirkelfreien Definition des Wortes „analytisch“, während das zweite Argument die Carnap'sche Position auf der Grundlage eines epistemologischen Holismus kritisiert. Zunächst unterscheidet Quine zwischen zwei Arten analytischer Sätze: Analytische Sätze, welche logisch wahr sind (wie „Alle unverheirateten Männer sind unverheiratet“), und solche, die nicht logisch wahr sind (wie „Alle Junggesellen sind unverheiratet“).¹⁰ Quine betrachtet die ersten, im Gegensatz zu den zweiten, als unproblematisch. Nach der klassischen Betrachtungsweise kann der zweite Typ von Sätzen in den ersten überführt werden, wenn man für bestimmte Termini Definitionen, also synonyme Ausdrücke, einsetzt. *„Just what it means to affirm synonymy [...] is far from clear“* (Quine, 1964, S. 24f.). Und hier fängt eine Kette der Abhängigkeiten an: Analytizität kann durch Synonymie erklärt werden, Synonymie durch Notwendigkeit, all diese Termini können durch semantische Regeln erklärt werden, doch dieser Rückgriff auf semantische Regeln muss selbst wiederum den Terminus „analytisch“ bemühen. Es ergibt sich also keine Möglichkeit, aus dieser Gruppe unklarer Termini ausubrechen und festen, empirischen Grund zu finden. Quine kommt zu dem Ergebnis:

¹⁰ „Logisch wahr“ bedeutet für Quine: wahr unter allen Interpretationen der nicht-logischen Satzelemente (vgl. Quine, 1964, S. 22).

„But, for all its a priori reasonableness, a boundary between analytic and synthetic statements simply has not been drawn. That there is such a distinction to be drawn at all is an unempirical dogma of empiricists, a metaphysical article of faith“ (Quine, 1964, S. 37).

Das zweite Argument Quines setzt an einem möglichen Lösungsvorschlag an. Nach dem Verifikationsgrundsatz ist die Bedeutung eines Satzes die Methode seiner empirischen Überprüfung. Daher kann ein analytischer Satz als Grenzfall verstanden werden, der immer und durch alle Erfahrungen bestätigt wird. Hiergegen argumentiert Quine, dass die Erfahrung einzelne Sätze überhaupt nicht bestätigen oder widerlegen kann. „*[O]ur statements about the external world face the tribunal of sense experience not individually but only as a corporate body*“ (Quine, 1964, S. 41). Demnach gibt es auch keine Sätze, welche immun gegen eine Revision durch Erfahrung sind. Quines epistemologischer Holismus scheint also die letzte Möglichkeit zu blockieren, innerhalb der Carnap'schen Position zu einer klaren Definition von „Analytizität“ oder verwandten Ausdrücken zu kommen.

Tatsächlich ist die Sachlage allerdings, aus folgenden Gründen, komplizierter (vgl. George, 2000, S. 4): (1) Carnap gesteht durchaus zu, dass Termini wie „analytisch“ oder „logisch wahr“ sich nur wechselseitig durch einander definieren lassen, aber er bestreitet, dass dies irgendwelche Probleme aufwirft. Denn er betrachtet „Wahrheit“ ebenfalls als einen Terminus aus eben dieser Gruppe und ist nicht bereit, „Wahrheit“ als sinnlos zu betrachten (vgl. George, 2000, S. 4f.), da ohne den Begriff der Wahrheit keine Wissenschaft möglich ist. (2) Carnap lehnt einen epistemologischen Holismus nicht ab. „*I agree that 'any statement can be held true come what may'. But the concept of an analytic statement which I take as an explicandum is not adequately characterized as 'held true come what may'*“ (Carnap, 1963, S. 921). (3) Dies führt dazu, dass Carnap zugesteht, dass die analytisch-synthetische Unterscheidung keine empirische Signifikanz besitzt. Er sagt hierzu lakonisch: „*I am surprised that Quine deems it necessary to support this view by detailed arguments*“ (Carnap, 1963, S. 917). Die Philosophie und ihre Unterscheidungen gehören eben für Carnap selbst zu den inhaltsleeren *linguistic frameworks*, welche die empirische Forschung aus pragmatischen Gründen einsetzt. Inwieweit Quine die Carnap'sche Position getroffen hat, ist also unklar (vgl. George, 2000; Gregory, 2003; O'Grady, 1999). Glücklicherweise ist für die vorliegende Arbeit keine eingehende Betrachtung dieser Schwierigkeiten notwendig. Wichtig ist im gegenwärtigen Kontext hingegen, dass Quine für eine ‚Naturalisierung‘ der Semantik und der Erkenntnistheorie plädierte. Er betrachtete diese Wissenschaften als nicht wesentlich verschieden von

empirischen Wissenschaften wie der Physik oder Chemie. Deshalb gilt auch für sie der Verifikationsgrundsatz.

2.2.1.3. Entwicklungen nach Quine

Die wesentlichen Entwicklungen nach Quine können in drei Gruppen geordnet werden: die Kritik Boghossians an Quine, die Entwicklung des semantischen Externalismus sowie die Entwicklung von mehrdimensionalen Ansätzen innerhalb der Semantik möglicher Welten.

2.2.1.3.1. Boghossian

Paul A. Boghossian versuchte in verschiedenen Aufsätzen den Begriff der Analytizität zu rehabilitieren (Boghossian, 1996, 1997, 2003a). Er unterscheidet dabei metaphysische Analytizität von epistemologischer Analytizität. Während er wie Quine erstere verwirft, hofft er, den Begriff einer epistemologischen Analytizität retten zu können. Boghossians Argumente gegen metaphysische Analytizität können folgendermaßen zusammengefasst werden: Analytische Urteile sollen, unabhängig von außersprachlichen Fakten, wahr sein – „wahr allein aufgrund ihrer Bedeutung“, wie oft gesagt wird. Wenn Wahrheit aber in der „Übereinstimmung“ mit Wahrheitsbedingungen besteht, dann müssen für die Wahrheit jedes Satzes diese Bedingungen erfüllt sein. Worin sollen aber die Wahrheitsbedingungen eines analytischen Satzes bestehen? Es bleiben nur die sprachlichen – linguistischen – Tatsachen übrig, um diese Rolle zu spielen. Linguistische Tatsachen sind aber kontingent. Analytische Sätze sollen hingegen notwendig sein. Notwendige Sätze lassen sich jedoch nicht durch kontingente begründen. Zudem sind linguistische Fakten für die meisten analytischen Sätze noch nicht einmal einschlägig. Dass ein Satz eine bestimmte Bedeutung hat, scheint nicht auszureichen, um ihn wahr zu machen.

“Are we really to suppose that, prior to our stipulating a meaning for the sentence

Either snow is white or it isn't

it wasn't the case that either snow was white or it wasn't? Isn't it overwhelmingly obvious that this claim was true *before* such an act of meaning, and that it would have been true even if no one had thought about it, or chosen it to be expressed by one of our sentences?”

(Boghossian, 1997, S. 336).

Boghossians Hauptanliegen ist es, trotz der Akzeptanz dieser Kritik, die Möglichkeit apriorischen Wissens durch Rückgriff auf Analytizität zu rechtfertigen. Hier spricht er von epistemologischer Analytizität und entwickelt diese von einem Frege'schen Ausgangspunkt.

Nach Frege sind analytische Sätze solche, die allein aus den Regeln der Logik und den Definitionen der Termini abgeleitet werden können.¹¹ Diese Konzeption will Boghossian verteidigen. Seine Strategie dabei ist es, eine *reductio ad absurdum* gegen Quine vorzubringen. Er geht davon aus, dass wenige der heutigen Philosophen Quines radikalen Thesen der 'Unbestimmtheit der Übersetzung' oder der 'ontologischen Relativität' zustimmen wollen. Diese Thesen gelten als extrem relativistisch und unplausibel. Es soll dann gezeigt werden, dass Quines Kritik an der analytisch-synthetisch Unterscheidung diese Thesen impliziert. Boghossian gibt zwei alternative Deutungen der Quine'schen Kritik bezüglich epistemologischer Analytizität. Nach einer Lesart kritisiert Quine, dass sich das Prädikat „analytisch“ nicht einwandfrei definieren lasse, und deshalb nicht gehaltvoll sei. Der zweiten Lesart zufolge behauptet Quine, dass notwendigerweise kein Satz einer normalen Sprache¹² ein analytischer ist. Gegen die erste Lesart führt Boghossian ins Feld, dass die Behauptung, dass sich „bedeutungsgleich“ nicht einwandfrei definieren lasse, die Behauptung impliziere, dass „hat die Bedeutung X“ sich ebensowenig definieren lasse, da sich mithilfe der zweiten Definition die erste erstellen ließe (vgl. Boghossian, 1997, S. 342f.). Wenn sich aber „hat die Bedeutung X“ nicht einwandfrei definieren lässt, stellt sich die Frage, wie Sprache überhaupt noch als bedeutungsvoll betrachtet werden kann. Hiermit ist man bei den radikalen Thesen der 'Unbestimmtheit der Übersetzung' und der 'ontologischen Relativität' angelangt.

Die zweite Lesart ist dadurch motiviert, dass Quine die Sinnhaftigkeit von „Analytizität“ für Kunstsprachen und logische Wahrheiten zugesteht. Quine muss daher die Ablehnung der analytisch-synthetisch Unterscheidung durch die notwendige Vagheit der Umgangssprache rechtfertigen. Es ist aber nicht ersichtlich, wie dies gelingen soll, ohne auf einen Bedeutungsholismus zurückzugreifen, welcher besagt, dass die Bedeutung aller Worte von allen unseren Überzeugungen abhängt. Ein solcher Holismus soll aber bei Quine ein Ergebnis seiner Kritik der analytisch-synthetisch Unterscheidung sein und nicht ihre Voraussetzung. Darüber hinaus beruht der Bedeutungsholismus auf dem empiristischen Verifikationsgrundsatz, dieser wird aber heute allgemein abgelehnt. Boghossian kommt zu folgendem Ergebnis:

„I'm inclined to think, therefore, that anyone who rejects radical indeterminacy of meaning must believe that a distinction between the meaning-constitutive and the non-meaning-constitutive can be drawn. The only question is how“ (Boghossian, 1997, S. 356).

¹¹ s. Abschnitt 2.1.3.

¹² Im Unterschied zu einer Kunstsprache.

Bezeichnenderweise kreiste die sich an Boghossians Veröffentlichungen anschließende Debatte um Themen wie die implizite Definition logischer Konstanten, die Apriorität linguistischen Wissens und die Natur von Konventionen (vgl. Boghossian, 2003a; Harman, 1996; Margolis & Laurence, 2001). Diese Diskussionen treffen zwar manche Voraussetzungen, welche Boghossian – meist explizit – annimmt, aber nicht den Kern seiner Position. Zudem stellen viele der Thesen, welche gegen Boghossian starkgemacht wurden, ursprünglich Ergebnisse der Quine'schen Philosophie dar, welche selbst erst nach der Zurückweisung der analytisch-synthetisch Unterscheidung möglich wurde. Hierzu zählt die Position, dass alles Wissen über Bedeutungen, und damit über begriffliche Zusammenhänge, empirischer Natur sei (vgl. Margolis & Laurence, 2001). Es zeigt sich hier, dass die Schwierigkeit bei einem solch fundamentalen Problem großteils darin besteht, einen Ausgangspunkt zu finden, der nicht schon das Ergebnis voraussetzt. Die Diskussion um Boghossians Konzeption einer epistemologischen Analytizität kann heute noch nicht als abgeschlossen gelten (vgl. Ebert, 2005; Heitner, 2006; Philie, 2005). Doch scheint klar, dass die Herrschaft des 'Quine'schen Dogmas' gebrochen ist.

Die vielleicht wichtigsten Entwicklungen seit Quines *Two Dogmas of Empiricism* sind auf dem Gebiet der philosophischen Semantik zu finden. Mit den heutigen Bedeutungstheorien steht die Diskussion um Analytizität auf einer völlig anderen Grundlage als zu Zeiten der Quine'schen Kritik. Gillian Russell beschreibt die heutige Situation folgendermaßen:

„The Quinean camp raised a lot of problems for this picture of analyticity but in the meantime, the 'obvious' picture of meaning that supported it started to slip for relatively independent reasons. In three astonishingly influential pieces of philosophical writing, Putnam (1985) argued that meaning couldn't be both what a speaker grasped and what determined extension, Kaplan (1989) argued that what determines extension (character) and what got contributed to what a sentence said (content) came apart in the cases of indexicals and demonstratives, and Kripke (1981) argued that what determined the extension of a name or natural kind term need not be known in order for a speaker to understand the expression, nor was it what was contributed to the proposition expressed by a sentence containing one“ (Russell, 2008, S. X).

Wie sind begriffliche Relationen unter dieser neuen Perspektive zu beurteilen? Um dies zu beantworten, möchte ich zuerst, am Beispiel von Hilary Putnam und Saul A. Kripke, auf die Entwicklung des semantischen Externalismus eingehen, um danach näher auf die

Implikationen der Semantik der möglichen Welten für das Verständnis begrifflicher Relationen einzugehen.

2.2.1.3.2. Putnam, Kripke und der semantische Externalismus

Putnam reagierte schon in seiner ersten Schaffensphase auf Quines Angriff auf die analytisch-synthetisch Unterscheidung (vgl. Putnam, 1962). Während er Quines Kritik für die meisten Sätze zustimmte, versuchte er, eine Klasse von Sätzen zu isolieren, welche nicht der Quine'schen Kritik unterliegen. Sätze dieser Klassen enthalten, nach Putnam, sogenannte Ein-Kriterium-Worte. Dies sind Worte, für deren Anwendung nur ein einziges Kriterium existiert; sie werden von *Clusterbegriffen* unterschieden, für deren Anwendung es mehrere Kriterien gibt. Sätze wie „Alle Junggesellen sind unverheiratet“ sind nach Putnam deshalb analytisch, weil es nur ein Kriterium für die Anwendung des Ausdrucks „Junggeselle“ gibt, nämlich ein unverheirateter Mann zu sein. Putnams Argument ist dabei, dass sich die Bedeutung der Worte in solchen Sätzen auflöst, wenn sie negiert werden. Bei *Clusterbegriffen* ist dies nicht der Fall, weil für deren Bedeutung nicht ein einziges Kriterium konstitutiv ist (vgl. Putnam, 1962). Es ist nicht klar ob dieser Rettungsversuch analytischer Sätze tatsächlich erfolgreich war (vgl. Shirley, 1973). Doch dies ist letztlich auch nicht entscheidend, da Putnam mit der Begründung des semantischen Externalismus die Diskussion ohnehin derart in Bewegung brachte, dass seitdem neue Anforderungen an eine Erklärung begrifflicher Zusammenhänge gestellt werden müssen.

Akzeptiert man mit Putnam die Quine'sche Kritik zumindest soweit, dass in den meisten interessanten Fällen eine Unterscheidung zwischen Überzeugungen, welche für die Bedeutung unserer Ausdrücke konstitutiv sind, und solchen, welche dies nicht sind, unmöglich ist, so ergeben sich bedenkliche wissenschaftstheoretische Konsequenzen. Denn wenn wir diese Unterscheidung nicht treffen können, so scheinen alle unsere Überzeugungen konstitutiv für die Bedeutung unserer Ausdrücke zu sein.¹³ Wie können wir dann aber unsere Überzeugungen über einen Gegenstand ändern und trotzdem noch über den gleichen Gegenstand reden? Wie können wir z. B. sagen, dass die relativistische Physik der Newton'schen widerspricht, wenn die Ausdrücke „Masse“, „Beschleunigung“ etc. in beiden Systemen unterschiedliche Bedeutungen besitzen? Und wenn wir das nicht können, wie können wir dann die eine Theorie vorziehen? Diese Argumente wurden, in den 60er und 70er Jahren, auf eine

¹³ Für eine Kritik dieses Schlusses siehe Boghossians (1997, S. 355).

schlagkräftige Weise von Thomas S. Kuhn und Paul Feyerabend vorgebracht (vgl. Feyerabend, 1983; Kuhn, 1976).¹⁴ Durch diese Konsequenzen sah sich Putnam veranlasst, die Grundlagen der Semantik, so wie sie sich seit Carnap entwickelt hatten, neu zu überdenken.¹⁵ Putnam betrachtet zwei Thesen als grundlegend für diese Semantik:

„(I) Um einen Ausdruck zu verstehen, muß man sich einfach in einem bestimmten psychischen Zustand befinden [...].

(II) Die Bedeutung eines Ausdrucks (im Sinne von ‚Intension‘) bestimmt seine Extension (d. h. aus Intensionsgleichheit folgt Extensionsgleichheit)“ (Putnam, 1979, S. 27).

Und er erläutert dann sein Vorhaben wie folgt:

„Ich werde darlegen, daß diese beiden Annahmen von *keinem* Begriff, geschweige denn von irgendeinem Bedeutungsbegriff, gleichzeitig erfüllt werden. Der traditionelle Bedeutungsbegriff gründet auf einer falschen Theorie“ (Putnam, 1979, S. 27).

Um dies zu zeigen, führt Putnam ein Gedankenexperiment durch, welches zeigen soll, dass sich die Extension eines Ausdrucks bei zwei Sprechern unterscheiden kann, obwohl beide in exakt demselben psychischen Zustand sind. Die Geschichte des Gedankenexperimentes ist folgende: Es existiere ein Planet, welcher sich makroskopisch nicht von der Erde unterscheide – die Zwerde. Jeder Erdenmensch habe einen Zwerden-Zwilling, welcher dieselbe Sprache spreche. Der einzige Unterschied zwischen Erde und Zwerde bestehe darin, dass es auf der Zwerde kein H₂O gebe; alles, was auf der Erde aus H₂O-Molekülen bestehe, sei auf der Zwerde aus anderen Molekülen aufgebaut. Als Abkürzung für die komplizierte chemische Formel dieser Moleküle gelte XYZ. (Putnam wählt Wasser als Beispiel, da es sich dabei um eine ‚natürliche Art‘ handelt. Später dehnt er aber seine Thesen auf andere Prädikate aus.)¹⁶ Nun seien wir zusätzlich in das Jahr 1750 versetzt, in dem weder auf der Erde noch auf der Zwerde die chemische Formel für Wasser bekannt war. Oskar₁ sei ein Erdling, Oskar₂ ein Bewohner der Zwerde. Nun gebe es eine Situation, in welcher Oskar₁ und Oskar₂ das Wort

¹⁴ Für einen Überblick über die sich anschließende Diskussion siehe Howard Sankeys *Incommensurability – An Overview* (1999).

¹⁵ Dass diese Konsequenzen tatsächlich eine Folge des Quine’schen Bedeutungsskeptizismus sind, ist in sofern nicht verwunderlich, als es eine Anforderung an jede Theorie begrifflicher Zusammenhänge ist, deren besondere Rolle im menschlichen Begründen und Argumentieren zu erklären (s. Kapitel 2.1.4). Wer begriffliche Relationen vollständig durch empirische „wegerklären“ will, der erklärt damit auch diese besondere Rolle „weg“. „Notwendigkeit“ und „Objektivität“ haben in einer solchen Konzeption keinen Platz. Ironischerweise endet der Kampf für die „harte Wissenschaft“ und gegen die Metaphysik im Relativismus.

¹⁶ Der Ausdruck „natürliche Art“ soll Klassen von Dingen bezeichnen, deren Identität durch die Natur und nicht durch den Menschen bestimmt wird (Beispiele wären Katzen, Wasser oder Gold) (vgl. Morris, 2007, S. 94ff.). Oft werden natürliche Arten von „künstlichen Arten“ abgegrenzt, welche ihre Identität dem Menschen verdanken (wie Bremsen, Taschenlampen oder Wecker).

Wasser verwenden und dabei im gleichen psychischen Zustand sind. Dies ist durchaus plausibel, da es ja zu jener Zeit kein Wissen auf den jeweiligen Planeten gibt, welches bei den beiden Oskars zu unterschiedlichen Überzeugungen führen könnte.¹⁷ Nun würden wir aber wohl sagen, dass Oskar₁ sich mit „Wasser“ auf H₂O bezieht, Oskar₂ hingegen auf XYZ.

„Oskar₁ und Oskar₂ faßten also im Jahre 1750 den Ausdruck ‘Wasser’ verschieden auf, *obwohl sie sich im selben psychischen Zustand befanden* und obwohl ihre Wissenschaften bei ihrem damaligen Stand noch etwa fünfzig Jahre brauchten, um zu entdecken, daß sie den Ausdruck ‘Wasser’ unterschiedlich verstanden haben. Das heißt, daß die Extension des Ausdrucks ‘Wasser’ (und auch seine Bedeutung im intuitiven, vortheoretischen Sinne) *keine* Funktion allein des psychischen Zustandes des Sprechers ist“ (Putnam, 1979, S. 34).

Und Putnam kommt damit zu dem Ergebnis: „Man kann’s drehen und wenden, wie man will, Bedeutungen sind einfach nicht im *Kopf*“ (Putnam, 1979).¹⁸ Dies ist die Kernthese des semantischen Externalismus. Und hiermit ergibt sich auch eine Antwort auf Quine. Wenn die Bedeutung unserer Worte durch keine unserer Überzeugungen bestimmt wird, so wird sie *a fortiori* nicht durch alle unsere Überzeugungen bestimmt. Hiermit sind auch die Konsequenzen, welche Kuhn und Feyerabend aus der traditionellen Semantik zogen, hinfällig.

An diese Kritik schließt Putnam seinen positiven Vorschlag zur Analyse des Bedeutungsbegriffs an. Statt aber zu sagen, was die Bedeutung eines Wortes sei, will Putnam eine Normalform für die Erläuterung der Bedeutung eines Wortes angeben. Es soll also eine Standardbeschreibung für Bedeutungen gefunden werden, welche nicht voraussetzt, dass allen zu nennenden Merkmalen einer Bedeutung nur ein Objekt entspricht. Eine solche Normalform-Beschreibung soll, nach Putnam, vier Komponenten enthalten: (1) einen syntaktischen Marker, welcher die grammatischen Verbindungsmöglichkeiten des Wortes angibt; (2) einen semantischen Marker, welcher die Art des bezeichneten Gegenstandes angibt; (3) eine Beschreibung des „Stereotyps“, wobei ein Stereotyp eines Ausdrucks die Information ist, die ein Sprecher besitzen muss, um als kompetenter Benutzer dieses Ausdrucks zu gelten (vgl. Putnam, 1979, S. 68);¹⁹ und (4) die Extension des Wortes (vgl. Putnam, 1979, S. 94).

¹⁷ Wer hier prinzipiell an der Möglichkeit zweier qualitativ identischer psychischer Zustände zweifelt, der ist für Putnams Kritik ohnehin kein Gegner. Denn dieser Zweifel führt, verbunden mit den Thesen (I) und (II), zu einem radikalen Bedeutungsskeptizismus. Es ist dann ohnehin unklar, wie Kommunikation stattfinden kann, wenn doch die Worte stets etwas anderes bedeuten. Und wie will uns dann dieser Skeptiker seinen Zweifel mitteilen?

¹⁸ Für eine Kritik an Putnams Position siehe Zermach (1976).

¹⁹ Diese Information kann falsch sein, sie bietet also keine Grundlage für analytische Sätze.

Entscheidend dabei ist, dass die ersten drei Komponenten nicht die vierte festlegen. Die Fixierung der Referenz eines Ausdrucks funktioniert auf andere Weise, und diese muss dem Sprecher nicht einmal bekannt sein. Bezüglich des Problems der Bestimmung des Bezugs referiert Putnam auf die von ihm und Kripke unabhängig voneinander entwickelte 'kausale Theorie der Referenz' (vgl. Putnam, 1979, S. 63).

Die 'kausale Theorie der Referenz' will ich am Beispiel Kripkes erläutern. Kripke entwickelt, anders als Putnam, seine Theorie ausgehend von Eigennamen. Er kritisiert dabei die Theorie der Eigennamen, welche diese mit Beschreibungen ihrer Träger identifiziert. Eine solche Theorie wurde insbesondere von Frege und Russell vorgeschlagen. Namen sind nach Kripke „starre Designatoren“, welche in allen möglichen Welten den selben Gegenstand bezeichnen, unabhängig davon, welche wahren Aussagen über diesen Gegenstand getroffen werden können. „Es scheint falsch zu sein, daß wir uns einige Eigenschaften geben, die einen Gegenstand irgendwie qualitativ als einzigen herausgreifen, und daß wir auf diese Weise unsere Referenz bestimmen“ (Kripke, 1981, S. 110). Damit muss Kripke aber eine Alternative anbieten, wie die Referenz fixiert wird. Denn es „ist die Extension unauflöslich mit Wahrheit verknüpft“ (Putnam, 1979, S. 49). Und damit ist es die Referenzfixierung, die interessant ist im Zusammenhang mit begrifflichen Relationen und Begründungen. Wie sieht diese Alternative aus?

„Im allgemeinen hängt unsere Referenz nicht einfach davon ab, was wir selbst denken, sondern auch von anderen Leuten in der Gemeinschaft, der Geschichte, wie der Namen einen erreicht hat, und Dingen dieser Art. Und die Referenz erhält man, wenn man eine solche Geschichte verfolgt“ (Kripke, 1981, S. 111).

Kripke nimmt an, dass es einen ursprünglichen Taufakt gegeben hat, durch den ein Wort seine Referenz erhält. Bei späteren Verwendungen des Wortes wird die Referenz über einen kausalen Konnex des Sprechers zum ursprünglichen Taufakt hergestellt. Der Name wird dabei von einem Mitglied der Sprachgemeinschaft zum anderen weitergegeben. Was aber notwendige und hinreichende Bedingungen einer entsprechenden Kausalkette sind, lässt Kripke im Dunkeln (vgl. Kripke, 1981, S. 112ff.). Doch wie dem auch sei, entscheidend ist, dass mit dieser Entwicklung sich eine Unterscheidung zwischen notwendigen und apriorischen Wahrheiten etabliert hat. Notwendige Sätze sind, laut Kripke, solche Sätze, welche in allen möglichen Welten²⁰ wahr sind. Apriorische Sätze sind Sätze, welche ohne

²⁰ Je nach Terminologie des Autors werden auch die Ausdrücke „Zustandsbeschreibungen“ oder „mögliche Weltzustände“ verwendet.

Berufung auf Erfahrungstatsachen gerechtfertigt werden können. Doch es ist nicht sosehr diese Unterscheidung, die neu ist, sondern die Tatsache, dass Kripke und Putnam starke Argumente vorgebracht haben für die Existenz sowohl notwendiger Wahrheiten, welche nur *a posteriori* gewusst werden können, als auch kontingenter Tatsachen, die *a priori* gewusst werden können. Diese Entdeckung der Unabhängigkeit von Apriorität und Notwendigkeit wurde sehr einflussreich, denn hierhin liegt eine grundlegende Abkehr von maßgeblichen historischen Wurzeln. Carnap wollte sowohl Apriorität als auch Notwendigkeit durch Analytizität erklären. Dies wird durch deren Unabhängigkeit unmöglich. Wir stehen also jetzt vor drei Begriffen und müssen deren Verbindungen klären: (1) Analytizität, was ein linguistischer Begriff ist, der sich möglicherweise auf eine Klasse von Sätzen und ihre innere Struktur bezieht; (2) Apriorität, was ein epistemologischer Begriff ist und sich auf die Art der Begründung einer Erkenntnis bezieht; und (3) Notwendigkeit, was ein metaphysischer Begriff ist und sich auf die Wahrheit von Aussagen in allen möglichen Welten bezieht. Um diese Verhältnisse zu klären, wurde in jüngster Zeit die Semantik möglicher Welten zu mehrdimensionalen Ansätzen weiterentwickelt.

2.2.1.3.3. Analytizität und die Semantik möglicher Welten

Wie hat man über begriffliche Relationen zu denken, wenn man den semantischen Externalismus akzeptiert? „[S]ome of the leading lights of the externalist movement have applied their insights to argue that some sentences which were thought to be analytic cannot really be“ (Russell, 2008, S. 14). Nach Putnam kann „Alle Katzen sind Tiere“ nicht analytisch sein, weil wir entdecken könnten, dass Katzen Spion-Roboter vom Mars sind (vgl. Putnam, 1979). Und laut Kripke ist „Gold ist gelb“ nicht analytisch, da die gelbe Farbe sich als optische Täuschung entpuppen könnte, welche durch die Erdatmosphäre hervorgerufen wird (vgl. Kripke, 1981). David Kaplan argumentierte dafür, dass Sätze wie „Ich bin jetzt hier“ analytisch seien, da ihre Bedeutung ihre Wahrheit garantiere, aber einen kontingenten Sachverhalt wiedergebe, da es natürlich nicht notwendig sei, dass ich mich jetzt hier aufhalte (vgl. Kaplan, 1989). Und tatsächlich gibt es Philosophen, die in diesen Ergebnissen das endgültige Ende der ‘analytischen Philosophie’ – verstanden als begriffliche Analyse – sehen (vgl. Campos, 2003; Soames, 2006). Das Problem besteht hauptsächlich darin, dass der Begriff der Bedeutung, welcher die Relata in begrifflichen Relationen bezeichnen soll, immer weiter aus-

differenziert wurde. Wir können demnach mindestens vier Arten von Bedeutungen unterscheiden:²¹

- (1) *Character*: dasjenige, was ein Sprecher wissen muss, um als kompetenter Benutzer des Wortes zu gelten.
- (2) *Content*: das, was das Wort zur Aussage eines Satzes beiträgt.
- (3) *Reference Determiner*: die Bedingung, welche ein Objekt erfüllen muss, um zur Referenz des Wortes zu gehören.
- (4) *Referent / Extension*: die Klasse der Objekte, auf welche das Wort zutrifft.

In dieser Terminologie kann man die besprochenen Ergebnisse folgendermaßen darstellen: Putnams Argumente zeigten, dass *Charakter* und *Reference Determiner* nicht zusammenfallen. Kripke argumentierte, dass *Reference Determiner* und *Content* nicht dasselbe sind. Wichtig dabei ist, dass die Rolle, welche begriffliche Relationen für das Erkennen und Argumentieren spielen, nicht durch die Komponente des *Characters* gespielt werden kann. Dies ist durch die indirekte Beziehung des *Characters* zur Wahrheit bedingt. Es ist klar, dass sich hieraus ein erkenntnistheoretisches Problem bezüglich begrifflicher Relationen ergibt, denn der *Character* ist die einzige Bedeutungskomponente, bei welcher von menschlichem Wissen die Rede ist. Dieses Ergebnis kann nicht überraschen, da es genau das Ziel des neuzeitlichen Verständnisses begrifflicher Relationen war, die epistemologischen Schwierigkeiten der antiken und mittelalterlichen Konzeption zu überwinden. Eine breite Strömung der modernen Forschung hat sich aber, im Ausgang von Frege, immer weiter von einem Verständnis entfernt, in welchem begriffliche Relationen entweder mentale Entitäten sind, oder zumindest Entitäten, deren Erfassen ein mentaler Akt ist. Wir befinden uns also zurzeit in einer Situation, die derjenigen, in welcher sich die Autoren der frühen Neuzeit sahen, nicht unähnlich ist. Auch wir müssen einen neuen Weg finden, logisch-semantische, metaphysische und epistemologische Sachverhalte zueinander ins Verhältnis zu setzen, wenn wir unsere Praxis des Argumentierens und Begründens explizieren wollen. Die Option, diese Praxis nicht explizieren zu wollen, steht nicht offen, zumindest solange man an vernünftigem Meinungsaustausch interessiert ist. Ich möchte zwei Versuche einer solchen Neuintegration der jüngsten Zeit darstellen. Während David J. Chalmers sich seit einigen Jahren um eine neue Verbindung epistemologischer und metaphysischer Sachverhalte bemüht, versucht Gillian Russell eine Brücke zwischen logisch-semantischen und metaphysischen Zusammenhängen zu schlagen.

²¹ Ich schließe mich hier der Terminologie Gillian Russells an (vgl. Russell, 2008, S. 45f.).

Chalmers entwickelt eine zwei-dimensionale Semantik, wie sie in jüngerer Zeit zu anderen Zwecken auch von anderen Autoren vorgeschlagen wurde. Dabei werden jedem Ausdruck, welcher eine Extension haben kann, drei unterschiedliche Intensionen zugeordnet, eine primäre, eine sekundäre und eine zwei-dimensionale Intension. Eine primäre Intension ist eine Funktion von Szenarien auf Extensionen. Szenarien sind in Chalmers Ansatz epistemische Möglichkeiten, d. h. vollständige Beschreibungen von Weltzuständen, welche nicht *a priori* ausgeschlossen werden können. Eine sekundäre Intension ist eine Funktion von möglichen Welten auf Extensionen. Und eine zwei-dimensionale Intension ist eine Funktion von geordneten Paaren von Szenarien und möglichen Welten auf Extensionen. Die Intension komplexer Ausdrücke wird durch die Intension ihrer Bestandteile bestimmt. Bei der Äußerung eines Ausdrucks ist dessen Extension der Wert der Funktion der primären Intension im tatsächlichen Szenario der Äußerung und gleichzeitig der Wert der sekundären Intension in der tatsächlichen Welt der Äußerung. Ein Satz drückt dann eine metaphysische Wahrheit aus, wenn er in allen möglichen Welten gilt (dies ist die normale modallogische Bestimmung von Notwendigkeit). Chalmers führt nun aber über die variablen Szenarien zusätzlich eine epistemologische Notwendigkeit ein, welche mit Apriorität identisch ist. Ein Satz ist demnach *a priori*, wenn er in allen Szenarien wahr ist (vgl. Chalmers, 2006). In diesen Rahmen lassen sich Putnams und Kripkes Beispiele für notwendige Wahrheiten *a posteriori* so integrieren, dass zwar jeder dieser Sätze eine notwendige sekundäre, aber eine nicht-notwendige primäre Intension hat. Ähnlich können Kripkes kontingente Wahrheiten *a priori* als Sätze betrachtet werden, welche eine in allen Szenarien wahre primäre Intension haben, aber eine sekundäre Intension, deren Wert für manche mögliche Welten 'das Falsche' ist. Gleichzeitig lässt sich aber in diesem System die Behauptung aufstellen, dass es epistemische Möglichkeiten gibt, welche metaphysisch unmöglich sind. Die primäre Intension verhält sich also in zwei entscheidenden Punkten wie der Frege'sche 'Sinn'. Erstens bestimmt sie die Extension, und zweitens stellt sie eine Verbindung zum Wissen des Sprechers her. In Chalmers Semantik lassen sich begriffliche Zusammenhänge als Relationen zwischen primären, sekundären und zwei-dimensionalen Intensionen darstellen. Empirische Relationen hingegen sind Beziehungen zwischen den Referenten der sprachlichen Ausdrücke. Im Hintergrund steht hier ein ausgefeilter formaler Apparat, dessen Darstellung den Rahmen dieser Arbeit sprengen würde. Deshalb will ich hier nur Christian Nimtz bezüglich Chalmers Ergebnissen zu Wort kommen lassen:

„Chalmers maintains that Kant is right. There is a deep link between necessity and a priori, for a sentence is epistemically necessary if and only if it is a priori. Chalmers also holds that Frege is right. Semantics is indeed rooted in epistemology. For the identifying knowledge a competent speaker associates with her terms, as is revealed by the epistemic intension he associates with it, precisely is what determines the reference and truth-conditions of her expressions. Hence, metaphysics is not autonomous from epistemology. And neither is semantics“ (Nimtz, 2005, S. 10).

Auch Russell versucht, die unterschiedlichen Aspekte begrifflicher Zusammenhänge neu zu integrieren. Das zentrale Anliegen Russells ist dabei, eine konsistente Definition eines analytischen Satzes zu finden. Unter einem analytischen Satz versteht Russell, wie in der Tradition üblich, einen Satz, der allein durch seine Bedeutung wahr ist. Dabei geht sie von den vier oben unterschiedenen Aspekten von „Bedeutung“ aus. Als formales System bringt sie eine drei-dimensionale Semantik in Anschlag. Sie unterscheidet dabei zwischen drei unterschiedlichen Kontexten: (1) einem Einführungskontext, von welchem die Extensionen von Namen und Bezeichnungen natürlicher Arten abhängen;²² (2) einem Äußerungskontext, von welchem die Extensionen von indexikalischen Ausdrücken abhängen;²³ und (3) einem Bewertungskontext, von welchem der Wahrheitswert der resultierenden Proposition abhängt.²⁴ Der Äußerungskontext besteht dabei aus einem Quadrupel, welches sich aus Agent, Ort, Zeit und Evaluationskontext zusammensetzt. Ein solches Quadrupel lässt sich auch als „zentrierte Welt“²⁵ bezeichnen. Da in dieser zentrierten Welt bereits der Evaluationskontext enthalten ist, reduzieren sich die Kontexte wieder auf zwei. Auf diesem Hintergrund kann Russell eine erste, vorläufige Definition von einem Satz geben, der durch seine Bedeutung wahr ist.

“A sentence S is true in virtue of meaning just in case for all pairs of context of introduction and context of utterance, the proposition expressed by S with respect to those contexts is true in the context of evaluation“ (Russell, 2008, S. 71).

Diese modale Definition von Analytizität wirft einige Probleme auf, welche Russell durch Rekurs auf *Reference Determiner* – also den Fixierungsmechanismus der Extension – lösen möchte. Sie definiert dabei *Reference Determiner* folgendermaßen:

²² Hiermit werden Putnams und Kripkes Probleme bearbeitet.

²³ Hiermit werden Kaplans Probleme bearbeitet.

²⁴ Dies ist der klassische Kontext der Semantik, wie Carnap sie vorgeschlagen hatte.

²⁵ Im Englischen wird oft von *centered world* gesprochen.

“A reference determiner for an expression is a non-circular condition, such that any object which meets it satisfies the expression and does so in virtue of meeting that condition. If there are other such non-circular conditions, then the reference determiner is the one in virtue of which the others all hold“ (Russell, 2008, S. 98)

Analytische Sätze zeichnen sich, nach Russell, durch ein Beinhaltungs- oder Ausschluss-Verhältnis ihrer logischen Subjekte und logischen Prädikate aus. In bejahenden Sätzen müssen sich die jeweiligen *Reference Determiners* beinhalten, um einen analytischen Satz zu bilden, in verneinenden müssen sie sich ausschließen. „Alle Junggesellen sind unverheiratet“ ist nach Russell deshalb analytisch, weil der Referenzfixierungsmechanismus von „Junggeselle“ denjenigen von „unverheiratet“ beinhaltet. Natürlich übertragen sich bei diesem Ansatz viele Probleme der Analytizität auf den Referenzfixierungsmechanismus. Denn jetzt stellt sich natürlich die Frage, welche *Reference Determiners* sich denn tatsächlich enthalten. Russell beantwortet diese Frage nicht, sondern verweist statt dessen darauf, dass sich bisher herausgestellt habe, dass unterschiedliche Ausdrücke ihre Referenz unterschiedlich fixierten. Und die Definition analytischer Urteile könne natürlich nicht die Erforschung der unterschiedlichen Fixierungsmechanismen von Extensionen ersetzen. Russell gibt eine ausführliche Verteidigung ihres Ansatzes gegenüber der Quine’schen Kritik und untersucht die Verbindungen ihrer Definition von „Analytizität“ zu epistemologischen Fragestellungen.

“[1] If a sentence is analytic, then it is entailed by sentences which attribute reference determiners to expressions and disquotation principles [...]. Anyone who knows such propositions can reason from them, [...] to the proposition expressed by the sentence, thus obtaining a justification for belief in that proposition.

[2] What is epistemically interesting about analytic sentences is that anyone who knows the reference determiners for the expressions they contain is in a position to work out that the proposition expressed by the sentence is true, without undertaking further empirical investigation.

[3] The reference determiner attribution basis of analytic justification is distinctive in two ways. The truth of such attributions is, to a large extent, up to the linguistic community to decide. They also allow the derivation of claims about the world, from claims about language“ (Russell, 2008, S. 211).

Auch Russell verbindet also wieder, ähnlich wie Chalmers, logisch-linguistische, metaphysische und epistemologische Gesichtspunkte. Begriffliche Relationen sind auch bei

Russell Beziehungen zwischen abstrakten Entitäten, welche Funktionen von Kontexten auf Extensionen und Bedingungen für Klassenzugehörigkeiten darstellen. Empirische Relationen sind Beziehungen zwischen Referenten von Ausdrücken. Es ergibt sich also auch in Russells Konzeption ein klarer Unterschied zwischen empirischen und begrifflichen Zusammenhängen.

2.2.1.4. Zusammenfassung zum referenzbasierten Ansatz

Wir haben gesehen, dass referenzbasierte Bedeutungstheorien von einer starken Trennung begrifflicher und empirischer Zusammenhänge ausgingen, dabei aber Analytizität, Apriorität und Notwendigkeit als koextensive Termini verstanden. Begriffliche Relationen wurden als konventionelle Regelungen für die Darstellung empirischer Zusammenhänge betrachtet. In einer zweiten Phase wurde vehemente Kritik an der analytisch-synthetisch Unterscheidung geübt. Diese „naturalistische“ Bedeutungstheorie betrachtete begriffliche Zusammenhänge als eine besondere Art empirischer Sachverhalte; weshalb diese in der Terminologie der empirischen Wissenschaft zu beschreiben seien. Dabei wurde der Gebrauch der Sprache als wesentlich vom individuellen Sprecher bestimmt betrachtet. In einer dritten Phase wurde Kritik an der Vorstellung geübt, dass das Erfassen und Bestimmen begrifflicher Relationen eine rein interne Angelegenheit des erfassenden Subjekts ist. Begriffliche Relationen wurden als wesentlich abhängig von unterschiedlichen Umweltkontexten verstanden. Analytizität, Apriorität und Notwendigkeit fielen nun auseinander. Damit wurde der wissenschaftliche Nutzen der Analyse begrifflicher Relationen radikal in Frage gestellt. Denn demnach sind die wichtigen Aspekte begrifflicher Relationen für uns nur *a posteriori* erkennbar. Das Wesen der Dinge liegt demzufolge in den Dingen selbst, und worauf unsere Worte sich beziehen, können wir nicht *a priori* wissen. Als vierte Phase schließlich lässt sich eine Entwicklung ausmachen, welche wieder zu einer stärkeren Integration metaphysischer, logisch-semantischer und epistemologischer Gesichtspunkte führt. Ein fruchtbarer Weg hierzu scheint insbesondere in der Integration epistemologischer Gesichtspunkte in eine (mehrdimensionale) modallogisch aufgebaute Semantik zu liegen. Begriffliche Relationen werden in diesen Ansätzen als Funktionen auf Extensionen oder als Argumente solcher Funktionen verstanden. Empirische Relationen hingegen versteht man als Beziehungen zwischen verschiedenen Referenten sprachlicher Ausdrücke. Es sind aussichtsreiche Ansätze in dieser Richtung vorgeschlagen worden, doch wieweit diese tragen, lässt sich zur Zeit noch nicht abschätzen.

2.2.2. Gebrauchsbasierte Bedeutungstheorien

Den referenzbasierten Bedeutungstheorien stehen die gebrauchsbasierten gegenüber. Wittgenstein kann als Begründer der Gebrauchstheorien der Bedeutung betrachtet werden. Daher will ich hier zunächst Wittgensteins Position darstellen, um anschließend auf den sogenannten Inferentialismus einzugehen. Die Grundannahme, welche diese Ansätze teilen, ist, dass sich die Bedeutung sprachlicher Zeichen in deren Gebrauch konstituiert. Die primäre Rolle der Sprache wird nicht in der Repräsentation gesehen, sondern in der Kommunikation. Die Pragmatik wird als Grundlage der Semantik verstanden.

2.2.2.1. Wittgenstein

Am Anfang steht für Wittgenstein die Verwendung eines Wortes durch eine Sprachgemeinschaft. Will man die Bedeutung eines Wortes kennenlernen, so ist es hilfreich, nach seiner Verwendung zu fragen. Das Wort „Bedeutung“ bildet hier keine Ausnahme:

„Man kann für eine *große* Klasse von Fällen der Benützung des Wortes ‘Bedeutung’ – wenn auch nicht für *alle* Fälle seiner Benützung – dieses Wort so erklären: Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache“ (PU § 43, Wittgenstein, 1997b, S. 262).

Es ist Wittgenstein zufolge daher ein Fehler aller referenzbasierten Bedeutungstheorien nach einem abstrakten oder empirischen Gegenstand zu suchen, der mit Recht als Bedeutung bezeichnet werden kann – etwa einer Funktion von möglichen Welten auf Wahrheitswerte.

„Der Fehler, zu dem wir neigen, könnte folgendermaßen ausgedrückt werden: Wir suchen nach dem Gebrauch eines Zeichens, aber wir suchen nach ihm, als ob er ein Gegenstand wäre, der mit dem Zeichen in Koexistenz ist“ (BB, Wittgenstein, 1997a, S. 20f.).

„Gebrauch“ versteht Wittgenstein hier aber nicht als bloßes Auftauchen eines Wortes zu unterschiedlichen Zeiten in unterschiedlichen Kontexten. Mit Gebrauch ist die von einer Regel bestimmte Verwendung gemeint. Es sind also die Regeln der Verwendung von Worten, welche ihre Bedeutung bestimmen. Der richtige Gebrauch *folgt* daher nicht aus der Bedeutung (vgl. Glock, 2000, S. 123). Regeln stellen Maßstäbe dar, anhand derer die Verwendung eines Ausdrucks als richtig oder falsch bewertet werden kann. Die Gesamtheit solcher Regeln fasst Wittgenstein unter dem Begriff „Grammatik“ zusammen. Daher nennt Wittgenstein auch Sätze, die Regeln des Gebrauchs bestimmter Ausdrücke wiedergeben, „grammatische Sätze“.

„Wittgenstein unterscheidet zwischen ‚empirischen‘ und ‚grammatischen Sätzen‘, Sätzen, die typischerweise dazu verwendet werden, eine Regel auszudrücken [...]. Diese Unter-

scheidung stützt sich nicht auf die sprachliche Form – ein grammatischer Satz braucht nicht ein metasprachlicher Satz darüber zu sein, wie ein Ausdruck verwendet wird. Was zählt, ist ob wir den Satz als einen Maßstab der Richtigkeit *gebrauchen*. Der Kontrast zwischen grammatischen und empirischen Sätzen ist einer zwischen den *Regeln für* unsere Sprachspiele und *Züge in* unseren Sprachspielen, die in Übereinstimmung mit den Regeln ausgeführt werden“ (Glock, 2000, S. 155).

Grammatische Sätze sind demnach notwendig, weil ihre Negation ein Verstoß gegen die Regel der Sprache wäre. Ich kann mit *Notwendigkeit* von „Bello ist ein Rüde“ zu „Bello ist männlich“ übergehen, weil der Satz „Rüden sind männlich“ eine Regel für die Verwendung der Worte „Rüde“ und „männlich“ ist (vgl. Hacker, 1996, S. 215). Die Regeln der Sprache bestimmen, was eine sinnvolle Aussage ist, und also auch, was als die Wiedergabe eines empirischen Sachverhalts gelten kann. Ein Satz wie „Bello ist ein weiblicher Rüde“ verletzt die Regeln unserer Sprache und ist damit Unsinn. Begriffliche Relationen stellen also die Standards dar, auf welche wir rekurren, wenn wir beurteilen, ob Ausdrücke sinnvoll sind. „[L]ogical necessity is explained by reference to the distinction between sense and nonsense which we draw by means of our norms of representation“ (Glock, 1992, S. 156). Wer diese Standards systematisch missachtet, kann nicht als kompetenter Sprecher der Sprache gelten, welche durch diese Standards konstituiert wird. Für Wittgenstein gibt es also eine enge Verbindung zwischen epistemologischen, metaphysischen und logisch-semantischen Gesichtspunkten. Der metaphysische Begriff der Notwendigkeit darf nicht platonisch verstanden werden, sondern als ein Zug unserer Sprache. Und wir wissen um das, was notwendig ist, indem wir eine Sprache sprechen können. Mit dem Erlernen einer Sprache erlernen wir die Regeln, welche in grammatischen Sätzen ausgedrückt werden. Und diese sind notwendig, genau, weil sich in ihnen die Regeln des Sprachgebrauchs ausdrücken.

Nun scheint dieser Gedankengang eine ganze Reihe von Problemen aufzuwerfen: (1) Heißt das eben Gesagte nicht einfach, dass notwendige Sätze allein aufgrund ihrer Bedeutung wahr sind, wobei diese Bedeutungen durch Konvention festgelegt werden? Dies wäre die traditionelle, Carnap'sche Definition von Analytizität. Und hat nicht Quine gezeigt, dass es keine analytischen Sätze geben kann und sich keine Wahrheiten auf Konventionen gründen lassen? (2) Hat nicht Kripke gezeigt, dass die Ausdrücke „Apriorität“, „Notwendigkeit“ und „Analytizität“ nicht nur unterschiedliche Bedeutung haben, sondern auch, dass sie nicht einmal koextensiv sind? Wie lassen sich Kripkes und Putnams Beispiele für notwendige Wahrheiten *a posteriori* und kontingente Wahrheiten *a priori* nach Wittgenstein verstehen? (3)

Wenn sich die Regeln, welche sich in grammatischen Sätzen ausdrücken, in empirischen Tatsachen erschöpfen, wie können sie dann notwendig sein und sich von empirischen Relationen unterscheiden? Wenn sie sich aber nicht in empirischen Tatsachen erschöpfen, wie kann Wittgenstein dann dem Platonismus entgehen? Was sind diese Regeln dann?

Ich will im Folgenden einige Gründe anführen, warum die genannten Probleme nicht so schwerwiegend sind, wie sie vielleicht scheinen.

2.2.2.1.1. Analytizität und Konvention

Wittgenstein verwendet in seiner späten Philosophie den Terminus „analytisch“ nicht. Seine Gründe dafür scheinen insbesondere darin zu liegen, dass nach seiner Auffassung der logische Positivismus ganz verschiedene Sätze als „analytisch“ bezeichnete, welche unterschieden werden sollten. Wittgenstein sieht wesentliche Unterschiede zwischen logischen Wahrheiten, mathematischen Wahrheiten und klassischen analytischen Sätzen. Logische Wahrheiten sind Grenzfälle sinnvoller Sätze. Sie sagen nichts, da sie ihrer Struktur nach keine Möglichkeiten ausschließen; sie sind leere Tautologien. Mathematische Sätze dagegen sind Regeln der möglichen Umformungen empirischer Sätze in andere empirische Sätze. Sie sind aber gleichzeitig in ein System eingebettet, welches die Generierung neuer Regeln erlaubt. Das, was klassischerweise unter „analytischen Sätzen“ verstanden wird, sind hingegen Normen für die Bestimmung dessen, was ein sinnvoller Satz ist (vgl. Hacker, 1996, S. 214ff.). Was als eine solche Norm zu betrachten ist, hängt nicht von der Form des Satzes (dem *typ*), sondern von seiner Verwendung in einer Situation (dem *token*) ab. Wichtig ist, dass Wittgenstein nicht denkt, dass es die Bedeutungen der Worte in analytischen Sätzen sind, welche diese wahr machen.

„Wittgenstein, unlike the members of the Vienna Circle, did not explain so-called analytic truth by reference to type-sentences which are either (instances of) laws of logic or reducible to a law of logic by the substitution of synonyms for constituent expressions in accordance with definitions. Nor did he clarify the nature of so-called necessary truth by arguing that they are *consequences* of the meaning (definition) of their constituent expressions“ (Hacker, 1996, S. 213).

Notwendige Wahrheiten folgen, nach Wittgenstein, nicht aus der Bedeutung von Ausdrücken, sondern konstituieren diese (teilweise). Das Aufgeben dieser Sätze ist eine Veränderung der Bedeutung von mindestens *einem* der darin auftauchenden Ausdrücke (vgl. Glock, 1992, S. 163). Und da Wittgenstein nicht behauptet, dass grammatische Sätze *aufgrund* der Bedeutung

ihrer Bestandteile wahr sind, muss er auch nicht behaupten, dass sie aufgrund von Konventionen wahr sind, welche den Bestandteilen ihre Bedeutung geben. Quines gegen Carnap gerichtete Kritik an Wahrheiten durch Konvention (vgl. Quine, 1976) kann also hier nicht treffen. Zwar beruht die Sprache auf Konventionen,²⁶ doch folgt daraus nicht, dass grammatische Sätze aufgrund von Konventionen wahr sind (vgl. Glock, 2008).

Ein weiterer Punkt, welcher beachtet werden muss, ist, dass Wittgenstein wohl nicht von einer in *allen* Fällen *eindeutigen* Unterscheidung zwischen empirischen und grammatischen Sätzen ausgeht:

„Es ist aber keine scharfe Grenze zwischen methodischen Sätzen und Sätzen innerhalb einer Methode.

Aber müßte man dann nicht sagen, daß es keine scharfe Grenze gibt zwischen Sätzen der Logik und Erfahrungssätzen? Die Unschärfe ist eben die der Grenze zwischen *Regel* und Erfahrungssatz.

Hier muß man, glaube ich, daran denken, daß der Begriff ‘Satz’ selbst nicht scharf ist.

Ich sage doch: Jeder Erfahrungssatz kann umgewandelt werden in ein Postulat – und wird dann eine Norm der Darstellung. Aber auch dagegen habe ich ein Mißtrauen. Der Satz ist zu allgemein. Man möchte sagen ‘Jeder Erfahrungssatz kann, theoretisch, umgewandelt werden ...’, aber was heißt hier ‘theoretisch’?“ (ÜG §118ff., Wittgenstein, 1970, S. 83f.).

Hier sind zwei wichtige Punkte angesprochen: (1) Die Unterscheidung zwischen empirischen und grammatischen Sätzen ist *diachron* unscharf insofern, als ein empirischer Satz zu einem grammatischen werden kann (und umgekehrt). Wir können als Reaktion auf neue Erfahrungen die Regeln unserer Sprache ändern; doch ändern wir damit auch die Bedeutung unserer Ausdrücke. (2) Die Unterscheidung ist *synchron* unscharf insofern, als nicht immer feststeht, was als Regel unserer Sprache zu verstehen ist. Aber hier macht Wittgenstein den Hinweis, dass genau insoweit wie dies unklar ist, auch unklar ist, was als sinnvolle Äußerung und als Mitteilung zu verstehen ist – was als „Satz“ gilt.²⁷ Aus der Feststellung, dass die

²⁶ Dies drückt Wittgenstein dadurch aus, dass er die Regeln der Grammatik „autonom“ oder „willkürlich“ nennt (vgl. PU § 372, Wittgenstein, 1997b, S. 398).

²⁷ Dies trifft sich mit einem ähnlichen Punkt der berühmten Kritik, welche Paul H. Grice und Peter F. Strawson an Quines „Two Dogmas“ üben: „If talk of sentence-synonymy is meaningless, then it seems that talk of sentences having meaning at all must be meaningless too“ (Grice & Strawson, 1956, S. 146). Ein ähnliches Argument wurde auch von Paul Boghossian vorgebracht (vgl. Boghossian, 1994, S. 115). Der Kern dieser Argumente scheint mir zu sein, dass ein konsequenter Bedeutungs-Skeptizismus genau so selbstwidersprüchlich ist wie jeder konsequente Skeptizismus. Und sofern eine Theorie überhaupt etwas Interessantes zu Bedeutungen zu sagen hat, hat sie auch etwas zu Verhältnissen zwischen Bedeutungen zu sagen. Die Frage ist eben nicht, ob es analytische

Unterscheidung zwischen empirischen und grammatischen Sätzen nicht immer klar ist, folgt nicht, dass diese Unterscheidung sinnlos ist (vgl. Glock, 1992, S. 171). Hierin zeigt sich Wittgensteins Kritik an Freges Diktum der Bestimmtheit des Sinns. „Eine Umgrenzung, die ein Loch hat, ist so gut, wie *gar keine*. – Aber ist das denn wahr?“ (PU § 99, Wittgenstein, 1997b, S. 295). Der Hinweis auf Grenzfälle ist keine Widerlegung der Existenz einer Grenze. Es scheint sich also weder daraus, dass die Sprache auf Konventionen beruht, noch aus einer Kritik der analytisch-synthetisch Unterscheidung ein Problem für Wittgensteins Position zu ergeben.

2.2.2.1.2. Apriorität, Notwendigkeit und Semantik

Haben Kripke und Putnam nicht gezeigt, dass Epistemologie, Metaphysik und Semantik auseinandergehalten werden müssen?²⁸ Wie wir bereits gesehen haben, gibt es auch innerhalb der referenzbasierten Bedeutungstheorien Bedenken an einer zu scharfen Trennung dieser Bereiche,²⁹ doch wie stellen sich diese Unterscheidungen im Wittgenstein'schen Rahmen dar? Es ist klar, dass Wittgenstein versucht, metaphysische Begriffe durch logisch-semantische zu klären. Denn „[d]as Wesen ist in der Grammatik ausgesprochen“ (PU § 371, Wittgenstein, 1997b, S. 398). Und „Notwendigkeit“ soll durch die Regel der Sprache geklärt werden. Es wäre aber ein Missverständnis, Wittgenstein den Versuch zu unterstellen, modale Begriffe auf normative zu *reduzieren*, wie jüngst Javier Kalhat (2008). Wittgenstein will modale Begriffe nicht auf andere reduzieren, sondern eine Alternative zu *de re*-Notwendigkeiten anbieten; also zu Notwendigkeiten, welchen eine 'übergeordnete' Realität entsprechen soll. Hier weist Wittgensteins Position erhebliche Ähnlichkeiten zu derjenigen der logischen Positivisten auf (vgl. Glock, 2008, S. 44ff.). Der Satz „Rüden sind männlich“ ist, nach Wittgenstein, nicht deshalb wahr, weil ihm eine Wirklichkeit entspricht; egal ob mit Wirklichkeit platonische Ideen, die allgemeinsten Züge der empirischen Realität oder sonst etwas gemeint ist. Dem widerspricht Kripkes und Putnams Behauptung, dass die Bezeichnungen natürlicher Arten starre Designatoren seien. Denn starre Designatoren bezeichnen in allen Welten dieselbe natürliche Art, unabhängig davon, welche Eigenschaften die Art in unterschiedlichen Welten besitzt. Demnach kann es eine Welt geben, in der Rüden weiblich sind. Und dies soll mehr

Urteile gibt, sondern wie wir am besten über die Bedeutung unserer sprachlichen Ausdrücke sprechen. Wittgenstein sieht klar, dass es, solange es überhaupt eine Grenze zwischen Sätzen und Nicht-Sätzen, also zwischen Sinnhaftem und Sinnlosem (in einem semantischen Sinn) gibt, jene Sätze, welche diese Grenze markieren, als Regeln unserer Sprache verstanden werden können.

²⁸ s. Kapitel 2.2.1.3.2.

²⁹ s. Kapitel 2.2.1.3.3.

besagen, als dass wir das Wort „Rüde“ anders verwenden könnten. Die wahre Essenz des Rüden könnten wir, nach Kripke und Putnam, nur durch empirische Forschung und nicht *a priori* erkennen. Putnam versucht dieser *prima facie* unwahrscheinlichen Überlegung dadurch Plausibilität zu verleihen, dass er erstaunliche aber dennoch mögliche wissenschaftliche Entdeckungen konstruiert. Wir könnten entdecken, dass Katzen Spion-Roboter vom Mars und Bleistifte Lebewesen seien (vgl. Putnam, 1979, S. 57).³⁰ Wenn dies zugegeben wird, dann könnten wir natürlich auch herausfinden, dass Rüden weibliche Tauben sind, die von Außerirdischen in die Kostüme männlicher Hunde gesteckt wurden, um uns Freundschaft zu heucheln.³¹ Dies scheint absurd und es gibt starke Gegenargumente:

(1) Es ist unklar, was wir in solchen Situationen sagen würden. Sollten wir bei den beschriebenen Tauben von Rüden reden? Sollten wir daraus folgern, dass manche Hunde Tauben sind? Für diese Fälle scheint unsere Sprache keine Regeln bereitzuhalten. Das ist aber nicht verwunderlich, da solche Regeln wohl höchst überflüssig wären. Die Regeln unserer Sprache können sich als Reaktion auf Erfahrungen verändern, und bei extremen Erfahrungen verlieren gewisse Regeln ihren Sinn,³² aber dies zeigt nicht, dass es keine Regeln gibt.

(2) Aus der beschriebenen Position ergäbe sich, dass wir kaum wissen können, was wir mit einem Ausdruck unserer Muttersprache meinen. „*Indeed, a linguistic community may use an expression for millennia without finding out what it really means, and without realizing that they do not know what it really means*“ (Hacker, 1996, S. 251). Dies scheint einer *reductio ad absurdum* gleichzukommen. Denn woher wissen Kripke und Putnam, was sie sagen?

(3) Es ist zweifelhaft, ob das wissenschaftliche Vokabular, in welchem die Essenz der natürlichen Arten letztendlich zu beschreiben wäre, so eindeutig bestimmbar ist, dass nur ein einziges keine Alternativen zulassendes Vokabular in Frage kommt. Dies scheint insbesondere deshalb fragwürdig, weil unsere Wissenschaft auch durch unsere Interessen geprägt wird. Doch wenn hier Alternativen möglich sind, was bezeichnet dann der Ausdruck „natürliche Art“?

(4) Es können durchaus Zweifel an der Konsistenz der kausalen Theorie der Referenz vorgebracht werden, welche von Kripke und Putnam als Ersatz für eine Referenzbestimmung mittels wesentlicher Merkmale vorgebracht wurde. Insbesondere ist die Frage zu klären, wie

³⁰ Putnam dehnt die Betrachtung von natürlichen Arten auf Artefakte aus, da ja die Erkenntnis darum, welche Dinge einer natürlichen Art angehören, ebenso wenig *a priori* möglich ist, wie jene von den notwendigen Eigenschaften bestimmter natürlicher Arten.

³¹ Eine solche Position vertritt z. B. Manuel Campos (2003).

³² Vielleicht würden auch Regeln des Schachs ihren Sinn verlieren, wenn jemand eine so geniale Zugkombination fände, dass man mittels dieser immer gewönne. Hier spielt die teleologische Einbettung von Regeln eine wichtige Rolle (vgl. hierzu Ertz, in Vorbereitung).

diejenigen kausalen Ketten bestimmt werden, welche die Referenz eines Ausdrucks festlegen. Außerdem ist zu fragen, wie die Identitätsrelation mit dem getauften Beispielobjekt bestimmt wird, welche Objekte erfüllen müssen, um zur Referenz eines Terminus zu zählen. Dieses Beispielobjekt, über welches die Referenz fixiert werden soll, spielt solange keine Rolle für die Sprache, wie diese Identitätsrelation ungeklärt ist. Denn was sollte dieses Objekt erklären (selbst wenn es historisch identifiziert werden könnte), solange unklar ist, in welcher Hinsicht es als Vergleichsmaßstab dienen kann? Ist aber diese Identitätsrelation geklärt, dann sind die wesentlichen Eigenschaften der natürlichen Art selbst geklärt und das Objekt damit überflüssig (vgl. Hacker, 1996, S. 252).

Wenn all diese Kritik zutrifft, scheint es um die Verbindung zwischen metaphysischen und grammatischen Bestimmungen nicht so schlecht bestellt zu sein, wie es zunächst aussah. Es kann durchaus sein, dass sich gewisse systematische Variationen der Bedeutung eines Ausdrucks mit seinem Einführungskontext finden lassen, es kann sogar sein, dass nicht alle relevanten Fakten dieses Kontextes einem Sprecher bekannt sein müssen, um als kompetent zu gelten, aber dies allein widerlegt Wittgensteins Position nicht. Denn es wird durch diese Position nicht impliziert, dass die Regeln der Verwendung eines Wortes nicht den Rückgriff auf extralinguistische Fakten einschließen können (z. B. bei Personalpronomina). Eine linguistische Erklärung von Notwendigkeit scheint also immer noch gangbar. *„By contrast, I remain puzzled by the idea of de re metaphysical necessities independent of our conceptual apparatus, which nonetheless we somehow manage to get into the haircrosses of our intellectual periscopes“* (Glock, 2008, S. 46).

Wie steht es aber um die Unterscheidung zwischen epistemologischen und logisch-semanticen Begriffen? Gibt es nicht vielleicht doch notwenige Wahrheiten, von welchen wir nichts wissen? Tatsächlich macht Wittgenstein einen Unterschied zwischen epistemologischen und logisch-semanticen Zusammenhängen. Während der Begriff der „Regel“ in den logisch-semanticen Kontext gehört, stehen die Begriffe „Gewissheit“ und „Zweifel“ in einem erkenntnistheoretischen Zusammenhang. Wenn es aber diesen Unterschied zwischen epistemologischen und logisch-semanticen Zusammenhängen gibt, dann kann sich hieraus eine Kritik an Wittgenstein ergeben. Denn vielleicht können wir die Regeln der Sprache gar nicht *a priori* erkennen, was sowohl Naturalisten wie auch essenzialistische semantiche Externalisten behaupten. In diesem Fall müssten wir an der Bedeutung der Worte unserer Sprache zweifeln und versuchen empirisch zu klären, welche Sätze als grammatische zu gelten hätten. Wittgenstein zieht diese Möglichkeit in Betracht, verwirft sie aber.

„Ich weiß nicht, ob das eine Hand ist.‘ Weißt du aber, was das Wort ‘Hand’ bedeutet? Und sag nicht ‘Ich weiß, was es jetzt für mich bedeutet’. Und ist das nicht eine Erfahrungstatsache, daß *dies* Wort *so* gebraucht wird?

Und hier ist es nun sonderbar, daß, wenn ich auch des Gebrauchs der Worte ganz sicher bin, keinen Zweifel darüber habe, ich doch keine *Gründe* für meine Handlungsweise angeben kann. Versuchte ich’s, so könnte ich 1000 geben, aber keinen, der so sicher wäre, wie eben das, was sie begründen sollen“ (ÜG § 306f., Wittgenstein, 1970, S. 80f.).

Der Zweifel hat nur dort einen Platz, wo es gewissere Gründe gibt als das Bezweifelte. Und die Frage muss sein, ob uns die Naturalisten und Externalisten solche Gründe für einen Zweifel an der Bedeutung unserer Worte gegeben haben. Ist eine Geschichte darüber, was wir sagen würden, wenn auf einem entfernten Planeten die Substanz XYZ gefunden würde, tatsächlich ein Grund, an der Bedeutung des Wortes „Wasser“ zu zweifeln? Ich denke nicht. Und verlässt sich nicht der Naturalist und Externalist selbst auf die Bedeutung seiner Worte? Wir wissen sehr gut, was wir „die Bedeutung eines Wortes wissen“ nennen. Und dabei beziehen wir uns vornehmlich auf sprachliche Kompetenz, nicht auf das Wissen um kausale Ketten der Wort-Weitergabe oder linguistische Feldforschung. „[W]enn jemand mich fragte, wie die Farbe auf deutsch heiße, und ich sag es ihm, und er fragt mich ‘Bist du sicher?’, so werde ich antworten: ‘Ich *weiß* es; Deutsch ist meine Muttersprache’“ (ÜG § 528, Wittgenstein, 1970, S. 137). Ein Gegner könnte einwenden, es sei eine *petitio principii* hier auf die Regel der Sprache zu rekurrieren, welche sagt, was wir „die Bedeutung eines Wortes wissen“ nennen. Hierbei handelt es sich ja gerade um einen grammatischen Satz. Aber hierauf lässt sich fragen, wie mich der Opponent überzeugen will, wenn wir keine gemeinsame Sprachregeln akzeptieren. Und wenn der Opponent nun auf die apriorische Geltung der Logik rekurriert, dann stellt sich die Frage, wie er diese begründet und gleichzeitig andere grammatische Sätze von einer solchen Begründung ausschließen kann (vgl. Field, 2006). Die Differenz zwischen erkenntnistheoretischen und logisch-semantischen Zusammenhängen scheint also zu keinen entscheidenden Problemen für die Position Wittgensteins zu führen.

2.2.2.1.3. Der ontologische Status von Regeln

Es scheint sich für Wittgenstein ein Problem aus der Frage zu ergeben, was man unter den Regeln, welche sich in grammatischen Sätzen ausdrücken lassen, verstehen soll. Woher nehmen die Regeln ihre Autorität?

„‘Wie kann ich einer Regel folgen?’ – wenn das nicht eine Frage nach den Ursachen ist, so ist es eine nach der Rechtfertigung dafür, daß ich *so* nach ihr handle. Habe ich die Begründungen erschöpft, so bin ich nun auf dem harten Felsen angelangt, und mein Spaten biegt sich zurück. Ich bin dann geneigt zu sagen: ‘So handle ich eben’“ (PU § 217, Wittgenstein, 1997b, S. 350).

Aber dass ich *so* handle, scheint doch ein empirisches Faktum zu sein, ebenso wie es ein empirisches Faktum ist, dass unsere Sprachgemeinschaft *so* handelt, oder dass wir eine gewisse Lebensform teilen. Und es stellt sich die Frage, wie ein empirisches Faktum eine logische Notwendigkeit begründen können soll. „*Suppose we explain $\Box p$ by appealing to q . The facts cited in the explanation are themselves either contingent or necessary. But they cannot be contingent, since then the necessary truth to be explained would not be necessary*“ (Hanks, 2008, S. 130).³³ Die Regeln scheinen also in mehr bestehen zu müssen als in kontingenten Tatsachen bezüglich menschlichen Verhaltens.³⁴ Der einzige Ausweg hieraus scheint eine Form des Platonismus zu sein (vgl. Wright, 2007, S. 485); diesen lehnt Wittgenstein aber offensichtlich ab.³⁵

Für Wittgenstein ist der zentrale Zug notwendiger Sätze ihre normative Funktion. Und wenn dies erkannt ist, stellt sich die Frage, warum es hilfreicher sein sollte, platonischen Ideen Normativität zuzugestehen, als einer menschlichen Praxis diese zuzugestehen. „Einer Regel folgen, eine Mitteilung machen, einen Befehl geben, eine Schachpartie spielen sind *Ge- pflogenheiten* (Gebräuche, Institutionen)“ (PU § 199, Wittgenstein, 1997b, S. 344). Wir brauchen aber schon zur Beschreibung solcher Gebräuche normatives Vokabular und können sie nicht als „rohe empirische Fakten“ auffassen. Und dass eine bestimmte Institution eine kontingente empirische Tatsache darstellt, heißt nicht, dass deren Regeln in jeder Hinsicht kontingent sind. Dass die Regeln des Schachs zufällig sind, heißt nicht, dass ich beim Schachspiel ziehen kann, wie ich will. Die Regeln bestimmen, was innerhalb einer Praxis „richtig“ heißt, und dieses ‘richtig’ lässt nichts an ‘Härte’ zu wünschen übrig. „Wir müssen zwischen der Notwenigkeit innerhalb des Systems und der Notwendigkeit des gesamten Systems unterscheiden“ (Wittgenstein, 1978, S. 294).

³³ Diese Kritik bringen auch Putnam (1987) und Kalhat (2008) vor. (Das Kästchen „ \Box “ steht für den modal-logischen Operator der Notwendigkeit.)

³⁴ Dieses Problem ist nicht spezifisch für Wittgenstein (vgl. Hanks, 2008).

³⁵ In den *Vorlesungen über die Grundlagen der Mathematik* sagt Wittgenstein z. B. „Ich werde immer wieder zu zeigen versuchen, daß man viel besser daran täte, das, was man als mathematische Entdeckung bezeichnet, eine mathematische Erfindung zu nennen“ (Wittgenstein, 1978, S. 23).

„[F]ar from the ‘the hardness of the logical ‘must’` being especially problematic for Wittgenstein, it is actually *much* simpler to present a successful account of logical necessity on Wittgenstein’s ‘grammatical conception’ than it could *possibly* be through the invocation of metaphysics“ (O’Neill, 2001, S. 27).

Die Wittgenstein’sche Bedeutungstheorie scheint also auf die vorgebrachte Kritik gute Antworten bereit zu halten. Doch hat es auch auf dem Gebiet der gebrauchsbasierten Bedeutungstheorien neue Entwicklungen gegeben.

2.2.2.2. Der Inferentialismus

Der Inferentialismus kann als jüngste Erneuerung der Gebrauchstheorie der Bedeutung verstanden werden. Er wurde insbesondere von Robert B. Brandom vertreten. Dieser folgt hierbei dem Grundsatz Wilfrid Sellars’, dass „das Begreifen eines Begriffs im Beherrschen des Gebrauchs eines Wortes besteht“ (Brandom, 2001a, S. 16).

2.2.2.2.1. Grundlagen des Inferentialismus

Der Inferentialismus spricht dem Gebrauch eines Wortes in Aussagesätzen eine herausragende Rolle zu. Dies ist darin begründet, dass unter dem Gebrauch eines Wortes in erster Linie sein Gebrauch in Schlussfolgerungen gemeint ist. „Über Begriffe zu reden heißt, über Rollen in Begründungszusammenhängen zu reden“ (Brandom, 2001a, S. 22). Daraus ergibt sich auch, dass die Bedeutung von Sätzen der Bedeutung von Worten im Erklärungszusammenhang vorgeordnet ist. Der inferentialistische Grundbegriff ist der der „Schlussfolgerung“.³⁶ Durch diesen soll erklärt werden, was propositionalen Gehalt ausmacht. „*The idea is to understand propositional content as what can both serve as and stand in need of reasons, where the notion of a reason is understood in terms of inference*“ (Brandom, 2007, S. 654). Aus der Erklärung der Bedeutung von Sätzen soll sich dann schließlich eine Erklärung der Bedeutung von Worten ergeben. Brandom nennt diese Erklärungsrichtung *top-down* und kontrastiert sie mit anderen Ansätzen, welche er als *bottom-up* charakterisiert; z. B. der Position Jerry Fodors. Es ergibt sich eine Position, welche besagt, dass die Bedeutung von Worten primär als die Rolle verstanden werden kann, welche diese Worte in Begründungen

³⁶ Brandom versteht allerdings „Schlussfolgerung“ in einem weiten Sinne. Dieser Begriff beinhaltet sowohl das, was er „material gute Inferenzen“ nennt, also Schlussfolgerungen welche nicht ohne Weiteres logisch beweisbar sind (z. B. „Wenn A östlich von B liegt und B östlich von C, dann liegt A östlich von C“) (vgl. Brandom, 2007, S. 657), als auch Schlussfolgerungen, welche empirische Sachverhalte oder intentionale Handlungen als eines ihrer Glieder enthalten (z. B. das Berichten von Beobachtungen, oder den sogenannten ‘praktischen Schluss’) (vgl. Brandom, 1994, S. 222).

spielen können. Begriffliche Relationen wären demnach Relationen zwischen Rollen in Begründungszusammenhängen. Die Notwendigkeit begrifflicher Relationen erklärt Brandom, wie Wittgenstein, über deren normativen Status. Die Normativität menschlicher Praxis wird dabei als grundlegend betrachtet. „*Social practices are implicitly normative in a way that mere behavioral regularities are not*“ (Brandom, 1994, S. 625). Das soziale Spiel des Gebens und Verlangens von Gründen, in dem sich begriffliche Relationen konstituieren, ist wesentlich durch Berechtigungen und Festlegungen geprägt, welchen die Teilnehmer dieser Praxis verpflichtet sind. Mit einer Behauptung legt sich ein Sprecher explizit auf einen bestimmten Inhalt fest und darauf Gründe für seine Behauptung zu geben, wenn sie ernsthaft in Frage gestellt wird. Implizit legt sich der Sprecher aber auch auf die Folgen fest, die aus seiner Behauptung gezogen werden können (vgl. Brandom, 2001a, S. 64). Sprecher sind zu gewissen Behauptungen berechtigt, z. B. aufgrund von empirischer Evidenz. Diese Berechtigung kann sich auf andere Behauptungen übertragen. Sind zwei Behauptungen inkompatibel, so schließt die Festlegung auf eine die Berechtigung zur anderen aus. „*Competent linguistic practitioners keep track of their own and each other's commitments and entitlements*“ (Brandom, 1994, S. 142). Brandom will nun auf der Grundlage dieser normativ beschriebenen Praxis eine Semantik aufbauen, welche in der Lage ist, die Objektivität begrifflicher Relationen zu erklären, und deutlich machen kann, wie Begriffe objektive Tatsachen wiedergeben können.

„Der springende Punkt [...] besteht darin, daß die *Objektivität* des propositionalen Gehalts – die Tatsache, daß wir, wenn wir behaupten, daß das Stoffmuster rot ist, nichts darüber sagen, wer irgend etwas angemessen behaupten könnte, oder darüber, wer auf was festgelegt oder wozu berechtigt ist, sondern statt dessen etwas sagen, was wahr sein könnte, selbst wenn es niemals rationale Wesen gegeben hätte – daß diese Objektivität ein Merkmal ist, das wir als eine Struktur der Festlegungen und Berechtigungen verständlich machen können, die den Gebrauch von Sätzen gliedern: als eine Struktur von weitgefaßten Normen, denen die Behauptungspraxis, das Spiel des Gebens und Nehmens von Gründen, unterworfen ist“ (Brandom, 2001a, S. 263).

Die Strategie besteht dabei im Wesentlichen darin, aus pragmatischen Grundbegriffen klassische semantische Begriffe, wie „Referenz“, zu gewinnen. Danach wird eine Analyse der pragmatischen Rolle von *de re*-Zuschreibungen propositionaler Inhalte gegeben. Auf diesem Hintergrund kann erklärt werden, wie propositionaler Inhalt nicht an die Perspektive einzelner Sprecher gebunden ist. Der Unterschied zwischen *de re*- und *de dicto*-Zuschreibungen wird zur Grundlage der Unterscheidung zwischen Objektivität und Subjektivität gemacht.

Angenommen, der *A* schreibt dem *B* die Behauptung *p* zu. Bei einer *de dicto*-Zuschreibung drückt *A* die Festlegungen aus, welche *B* laut *B* mit *p* eingegangen ist. Bei einer *de re*-Zuschreibung hingegen drückt *A* damit die Festlegungen aus, welche *B* durch *p* laut *A* eingegangen ist. Für *A* ergibt sich ein Unterschied zwischen dem, auf was *B* meint festgelegt zu sein, und dem, worauf *B* (laut *A*) objektiv festgelegt ist. Der Unterschied zwischen verschiedenen Perspektiven erlaubt also die Einführung des Begriffs „Objektivität“ in pragmatischem Vokabular (vgl. Brandom, 1994, S. 592ff.).

Brandom beansprucht, sowohl die Notwendigkeit wie auch die Objektivität begrifflicher Relationen aus dem Gebrauch von Worten in Begründungszusammenhängen erklären zu können. Es lässt sich natürlich einwenden, dass Brandom nur zu diesem Ergebnis kommen kann, da er seiner Erklärungsbasis, der menschlichen Praxis, entsprechend viel Normativität ‚zugesteht‘ und seinen Begriff der Schlussfolgerung entsprechend aufweicht. Aber dies ist solange kein Gegenargument, wie nicht gezeigt ist, dass es entweder eine ‚sparsamere‘ Lösung gibt, oder dass diese Voraussetzungen inkonsistent sind.

Welche Perspektive auf die analytisch-synthetisch Unterscheidung ergibt sich nun aus dem Inferentialismus?

2.2.2.2.2. Der Inferentialismus und die analytisch-synthetisch Unterscheidung

Fodor und Lepore kritisieren den Inferentialismus explizit dafür, eine Unterscheidung zwischen analytischen und synthetischen Sätzen vorauszusetzen (vgl. Fodor & Lepore, 1991). Die Frage nach analytischen Urteilen stellt sich für den Inferentialismus dabei in der Form der Frage, welche Schlussfolgerungen für die Bedeutung eines Ausdruckes konstitutiv sind. Wenn alle Schlussfolgerungen, in welchen ein Ausdruck auftaucht, für dessen Bedeutung konstitutiv sind, auch solche bloß empirischen Inhalts, dann scheint es keinen Unterschied zwischen empirischen und begrifflichen Zusammenhängen zu geben. Die Frage ist also: Besitzt der Schluss von „In gleichen Zeiten überstreicht der Fahrstrahl zwischen Planet und Sonne gleiche Flächen (2. Keplersches Gesetz)“ und „Der Mars war gestern an der Stelle *x* und heute an der Stelle *y*“ auf „Der Mars wird sich morgen an der Stelle *z* befinden“ laut Inferentialismus einen anderen Status als der Schluss von „Bello ist ein Rüde“ auf „Bello ist männlich“?

Eine fruchtbare Antwort auf diese Frage gibt Jaroslav Peregrin. Er greift dabei auf den Begriff der Regel zurück:

„The point is that the idea behind inferentialism is that it is *us*, the speakers, who furnish expressions, and consequently languages, with their inferential power – we treat the

statements as inferable one from another (perhaps by taking one to be committed to the former whenever she is committed to the later) and as incompatible with each other. The idea is that we have a finite number of rules and that a statement is inferable from a set of other statements if it can be derived from them with help of the rules“ (Peregrin, 2006, S. 19).

Es sollten demnach diese Regeln des Schlussfolgerns sein, welche bedeutungskonstitutiv sind, und nicht jede einzelne Schlussfolgerung. Und Peregrin sieht dabei deutlich, dass Quines Lehre von der Unhaltbarkeit der analytisch-synthetisch Unterscheidung nicht unkritisch übernommen werden darf. Diese Kritik beruht ja zu wesentlichen Teilen auf Quines Verständnis eines Empirismus,³⁷ welches der Inferentialismus nicht teilen muss.

„However, the fact that there is no sharp dividing line between the ‘meaning-constitutive’ and the other rules related to a word [...] should not lead us to conclude that this distinction makes no sense whatsoever. We do, as a matter of fact, tend to see words as having something as meaning, and we do tend to divide rules we learn regarding the employment of a word into general, *sine-qua-non* rules and casual ones which we may fail to know without being thereby taken as ignorant of its meaning“ (Peregrin, 2007, S. 13).

Dass der Inferentialismus auf eine Unterscheidung zwischen bedeutungskonstitutiven und nicht-bedeutungskonstitutiven Schlussfolgerungen verpflichtet ist, scheint sich schon aus seiner Betonung der Normativität von Bedeutungen zu ergeben. Denn eine Norm hat wohl nur Sinn, sofern etwas durch sie normiert wird. Es muss eine Unterscheidung zwischen einer Norm und dem, worauf sie angewendet wird, geben. Und diese Unterscheidung ist natürlich dieselbe, wie jene zwischen Regel und Spielzug, welche für Wittgenstein eine so große Rolle spielt. Anders als Fodor und Lepore (1991) denke ich, dass die Verpflichtung des Inferentialismus auf eine solche Unterscheidung kein Argument gegen ihn darstellt. Es ist meines Erachtens genau ein Adäquatheitskriterium für jede Bedeutungstheorie, den Unterschied zwischen der Bedeutung von Worten und ihrer Anwendung in beliebige neue Situationen verständlich zu machen. Denn wer diese Unterscheidung nicht erklären kann, der kann auch nicht verständlich machen, was die Bedeutung unserer Ausdrücke ausmacht.³⁸ Die Betonung von Regeln statt von Konventionen ermöglicht es dem Inferentialismus, dabei nicht der Quine’schen Kritik an dem Begriff analytischer Sätze des logischen Positivismus zu ver-

³⁷ s. Kapitel 2.2.1.2.

³⁸ Ähnliche Positionen wurden sowohl von Grice & Strawson (1956), als auch von Boghossian (1997) vertreten (s. Fußnote 27).

fallen. Inwieweit der Inferentialismus aber ein fruchtbares Forschungsprogramm darstellt und den Herausforderungen, welche an eine zeitgenössische Bedeutungstheorie gestellt werden, gerecht werden kann, muss die zukünftige Forschung auf diesem Gebiet zeigen. Entscheidend wird dabei unter anderem der Aufbau eines potenten formalen Apparats sein, um eine Alternative zur Semantik möglicher Welten bieten zu können. Peregrin hat hier einige Ansatzpunkte aufgezeigt (vgl. Peregrin, 2006). Hieraus sollten sich auch entsprechende Lösungen für die Kontextabhängigkeit von Bedeutungen, wie sie von Kaplan und Kripke aufgezeigt wurde, ergeben.

2.3. Eine gangbare Unterscheidung zwischen empirischen und begrifflichen Relationen

Die vorliegende Arbeit setzt voraus, dass die Unterscheidung zwischen empirischen und begrifflichen Relationen nicht nur konsistent gezogen werden kann, sondern auch von grundlegender Bedeutung für die empirischen Wissenschaften, insbesondere die Psychologie, ist. Denn in der Psychologie ist das Erkennen begrifflicher Abhängigkeiten besonders schwierig, wie die folgenden Kapitel zeigen werden. Wie kann diese Voraussetzung auf dem Hintergrund der bisherigen Ausführungen gerechtfertigt werden? Zur Beantwortung dieser Frage möchte ich im Folgenden eine Unterscheidung zwischen empirischen und begrifflichen Relationen darstellen, welche ich zum derzeitigen Stand der Forschung für gangbar erachte und als methodische Grundlage der vorliegenden Arbeit nutzen will.

2.3.1. Adäquatheitskriterien einer Unterscheidung

Was sind die entscheidenden Gütekriterien einer solchen Unterscheidung? Die Aufgabe von Aussagen über begriffliche Relationen besteht darin, verschiedene Aspekte unserer Praxis des Sprechens, Begründens und Argumentierens explizit zu machen. Folgende Aspekte sind meines Erachtens besonders zentral:

- (1) Die Ausrichtung der Argumentation auf Wahrheit, also auf das Ziel der objektiven Gültigkeit. Das Bezweifeln der Möglichkeit einer solchen objektiven Gültigkeit unserer Erkenntnis führt in einen widersprüchlichen Skeptizismus. Also sollte diese Möglichkeit objektiver Gültigkeit einer Erklärung zugänglich sein; wobei sie aufs Engste mit der Fähigkeit verbunden ist, mit begrifflichen Zusammenhängen umzugehen.
- (2) Der Aspekt von Notwendigkeit, welcher in unseren Argumentationen eine zentrale Rolle spielt. Ziel einer Argumentation ist es, dem Gesprächspartner zu vermitteln, dass gewisse Be-

hauptungen *notwendigerweise* wahr sein *müssen*, wenn andere wahr sind. Und die Möglichkeit solcher Übergänge, von einer Behauptung zur anderen, scheint nicht so sehr eine Folge empirischer Tatsachen zu sein, sondern unseres begrifflichen Erfassens dieser Tatsachen.

Über diese grundlegenden Aufgaben des Begriffs begrifflicher Relationen hinaus ergeben sich aus der oben dargestellten Diskussion einige weitere Kriterien, welchen eine Theorie begrifflicher Relationen genügen muss.

(3) Eine Erklärung begrifflicher Relationen unter Rückgriff auf Konventionen, wie sie der logische Positivismus gab, ist seit Quines Kritik nicht mehr möglich. Seit dieser Kritik steht fest, dass der Begriff der Analytizität einer besonderen Klärung bedarf.

(4) Es ist seit Kripke, Putnam und Kaplan klar, dass die Bedeutungen sprachlicher Ausdrücke nicht ausschließlich von Fakten bezüglich des jeweiligen Sprechers abhängen. Sofern sich begriffliche Relationen in den Bedeutungen sprachlicher Zeichen ausdrücken, scheinen also auch diese nicht allein mit Aussagen über Fakten bezüglich des jeweiligen Sprechers erklärt werden zu können. Es stellt sich erstens die Frage, inwieweit dabei auf soziale und inwiefern auf natürliche Fakten rekuriert werden muss; und zweitens ist unklar, ob ein Externalismus ohne den historisch damit verbundenen Essenzialismus möglich ist.

(5) Eine Klärung dessen, was begriffliche Relationen sind, muss verständlich machen können, inwiefern diese epistemisch zugänglich sind. Dies hängt besonders mit der in Punkt (4) angesprochenen, Bedeutung sozialer und natürlicher Fakten für begriffliche Relationen zusammen.

(6) Es ist zu klären, in welcher Beziehung begriffliche Relationen zu metaphysischen (modalen) Begriffen stehen. Sind hier mehrere Formen von Notwendigkeit in Anschlag zu bringen? Und inwiefern entsprechen begriffliche Relationen irgendwelchen Tatsachen?

2.3.2. Mehrere Unterscheidungen

Grundlegend für ein Verständnis begrifflicher Relationen ist meines Erachtens die Überlegung, dass nicht alle dieser sechs Kriterien von *einer* Unterscheidung erfüllt werden müssen. Tatsächlich denke ich, dass Aspekte der Regelgeleitetheit unseres Sprachgebrauchs von solchen des epistemischen Zugangs und des Bezugs sprachlicher Ausdrücke auf Außer-sprachliches getrennt betrachtet werden können. Die generelle Strategie, welche ich hier verfolgen will, ist es, den Wittgenstein'schen Begriff des grammatischen Satzes von dem eines analytischen Satzes zu unterscheiden, welcher eher nach dem Muster Russells verstanden werden sollte. Eine Brücke zwischen diesen Begriffen kann durch Brandoms Überlegungen geschlagen werden. Sollte dieser Weg gangbar sein, so könnte eine neue Verbindung

zwischen epistemischen, metaphysischen und logisch-semantischen Begriffen hergestellt werden. Die Grundüberlegung ist folgende: Wenn wir, wie Wittgenstein und Brandom, bei unseren Überlegungen von der normativ strukturierten Praxis des Sprachgebrauchs ausgehen, so scheinen sich Lösungen für die Punkte (1)-(3) zu ergeben. Grammatische Sätze sind notwendig, weil sie Regeln ausdrücken, und objektiv insoweit, als diese Regeln für alle Sprecher der entsprechenden Sprache gelten. Grammatische Sätze sind nicht wahr aufgrund von Konventionen, sondern aufgrund ihres normativen Status. Damit entstehen keine Probleme aus Punkt (3) (vgl. Glock, 2008). Die Punkte (4)-(6) sollen durch den Begriff des analytischen Satzes gelöst werden. Brandom hat gezeigt, wie auf der Basis einer Beschreibung normativ geordneter sprachlicher Praxis die Begriffe „Referenz“ und „Wahrheit“ gewonnen werden können (vgl. Brandom, 1994, S. 275ff.). Stehen uns diese Begriffe aber zur Verfügung, so können wir den formalen Apparat der Semantik möglicher Welten nutzen. Wie in einer dreidimensionalen Semantik möglicher Welten der Begriff des analytischen Satzes bestimmt werden kann, wurde von Russell ausgeführt (vgl. Russell, 2008). Da dies über den Mechanismus der Referenzfixierung geschieht, können dabei die Probleme des Externalismus berücksichtigt werden. Damit wäre Punkt (4) gelöst, ohne dabei eine essenzialistische Position vertreten zu müssen. Punkt (5) und (6) beziehen sich auf das Verhältnis von grammatischen Sätzen, analytischen Sätzen und dem Wissen um diese. Dabei ist aber der epistemische Zugang zu grammatischen Sätzen insofern unproblematisch, als in solchen Sätzen schlicht die Bedingungen sprachlicher Kompetenz ausgedrückt werden. Eine Sprecherin besitzt (zumindest implizit) genau *soviel* Wissen um die Zusammenhänge, welche sich in grammatischen Sätzen ausdrücken, als sie im Gebrauch der darin vorkommenden Ausdrücke kompetent ist. Andererseits müssen die relevanten Fakten bezüglich der Fixierung der Referenz keineswegs immer durchsichtig sein. Hierbei können soziale und natürliche Fakten eine Rolle spielen, wie Kripke, Putnam und Kaplan es verlangen. Die Verbindung zwischen Sprachgebrauch und Referenzfixierung sollte allerdings ausschließen, dass sich grammatische Sätze als falsch erweisen. Die Regeln des Sprachgebrauchs könnten somit als Bedingungen betrachtet werden, welchen die Referenzfixierung genügen muss. Damit würde sich auch die absurde Konsequenz des Externalismus vermeiden lassen, dass die Bedeutung unserer Worte uns auf immer unbekannt bleiben könnte. Die Ausarbeitung eines solchen Programms würde den Rahmen dieser Arbeit sprengen, deshalb soll nur kurz auf einige Aspekte dieser Möglichkeit eingegangen werden.

2.3.3. Eine Skizze des Sprachgebrauchs

Dem hier unterbreiteten Vorschlag entspricht ungefähr folgendes Bild des Sprachgebrauchs: Wenn wir Sprache gebrauchen, so nehmen wir damit an einer menschlichen Praxis teil, welche durch Regeln bestimmt ist. Dabei müssen die Regeln selbst von dem Sprachgebrauch, welchen sie regeln, unterschieden werden. Wir besitzen hierbei die Möglichkeit, diese Regeln in unserer Sprache auszudrücken, nämlich durch grammatische Sätze. Da wir beim Ausdrücken dieser Regeln genau diejenige Sprache sprechen, in welcher diese Regeln gelten, verletzen wir durch die Verneinung dieser Ausdrücke die Regeln des Sprachgebrauchs. Grammatische Sätze bestimmen, was sinnvolle Sätze sind und was als Darstellung eines Sachverhalts gelten kann. Damit sind grammatische Sätze notwendig (vgl. Glock, 2008). Die Regeln des Sprachgebrauchs können sich aber ändern. Damit verändern sich auch die Bedeutungen sprachlicher Ausdrücke. Diese Änderungen können zwar durch neue Erfahrungen bedingt sein, doch dies ist keine Falsifikation grammatischer Sätze (vgl. Hacker, 1996). Die Veränderungen sind vielmehr meist durch die teleologische Einbettung der Sprache in unsere Lebensvollzüge bestimmt. Wir werden nicht gezwungen, grammatische Sätze aufzugeben, diese verlieren vielmehr ihren Sinn und Zweck. Ein Sprecher ist genau in soweit kompetent, als er die Regeln des Sprachgebrauchs (implizit) kennt. Diese Sprachkompetenzen verteilen sich ungleichmäßig über die Gesellschaft.³⁹ Auch wenn ein Sprecher gewisse Regeln nicht kennt, werden seine Äußerungen im Allgemeinen an diesen Normen gemessen. Die Regeln des Sprachgebrauchs sind also wesentlich ein soziales Phänomen und somit auch für alle Sprecher verbindlich und also objektiv.

Innerhalb unserer Sprache haben wir die Möglichkeit, die Unterschiede zwischen unseren verschiedenen Perspektiven deutlich zu machen. Und damit ergibt sich die Möglichkeit, deutlich zu machen, dass verschiedene sprachliche Ausdrücke sich auf das Gleiche beziehen. Wir können explizieren, wie wir Übergänge zwischen unterschiedlichen Perspektiven wiedergeben können. Dies geschieht durch Regeln für die Ersetzbarkeit von Ausdrücken, für das Setzen und Eliminieren von Anführungszeichen und für *de re*-Zuschreibungen propositionaler Gehalte (vgl. Brandom, 1994). Wir reden dann davon, dass gewisse sprachliche Ausdrücke sich auf etwas Außersprachliches beziehen, wenn sich bestimmte Möglichkeiten der Umformung von Sätzen ergeben. Die Referenz eines Ausdrucks besteht in dem, was wir konstant halten wollen, wenn wir von der Beschreibung aus einer Perspektive zu der aus einer anderen

³⁹ Somit kann Putnams sprachliche Arbeitsteilung (vgl. Putnam, 1979) und der sozialen Externalismus (vgl. Burge, 1986) in die hier vorgelegte Konzeption integriert werden.

Perspektive übergehen. Ausdrücke wie „Dieses Wort bezeichnet ...“ oder „Diese Aussage bezieht sich auf ...“ haben also eine sehr komplizierte Grammatik. Sie dienen einerseits dazu, mögliche Konsequenzen von Perspektivenwechsel explizit zu machen, zum anderen können sie als eine Form der Erklärung von Bedeutungen betrachtet werden. Mithilfe solcher Ausdrücke lassen sich ostensive Definitionen geben, und die Erklärung vieler, aber bei Weitem nicht aller Worte kann in die Form „Das Wort bezeichnet ...“ gebracht werden (vgl. PU § 10, Wittgenstein, 1997b, S. 242). Obwohl eine solche Vereinheitlichung der Bedeutungserklärung zu Verwirrungen führen kann, lässt sie sich in manchen Kontexten fruchtbar anwenden.

Eine solche fruchtbare Anwendung besteht in der Möglichkeit, durch diese Ausdrücke den formalen Apparat der modernen Semantik verwenden zu können. Aus den Regeln für die Verwendung bestimmter Worte und für die Verwendung von Ausdrücken wie „... bezieht sich auf ...“ ergeben sich die Regeln dafür, wann wir sinnvoll davon sprechen können, dass sich ein sprachlicher Ausdruck auf etwas bezieht. Diese Regeln können als Regeln der Referenzfixierung bezeichnet werden. Die Regeln der Referenzfixierung ergeben sich also aus den grammatischen Sätzen einer Sprache und sind damit selbst grammatische Sätze. Dies heißt aber nicht, dass die Referenz jedes Wortes allein durch die Regeln des Sprachgebrauchs festgelegt wird. Solche Regeln können ja selbst auf bestimmte Kontexte Bezug nehmen. So können die Regeln zur Fixierung der Referenz des Ausdrucks „du“ besagen, dass sich dieses Wort auf den Adressaten des ‚Äußerungs-Tokens‘, in dem es vorkommt, bezieht. Die unterschiedlichsten natürlichen und sozialen Fakten können also bei der Referenzfixierung eine Rolle spielen. Wie die Regeln der Referenzfixierung für einzelne Ausdrücke aussehen, ist hiermit natürlich noch nicht gesagt. Auf dieser Grundlage lässt sich ein Begriff der Analytizität formulieren, wie Russell (2008) gezeigt hat. Analytische Sätze können als diejenigen verstanden werden, deren Wahrheitswert allein durch die Regeln der Referenzfixierung als „wahr“ festgelegt ist. Und da diese Regeln grammatische Sätze darstellen, grammatische Sätze aber notwendig sind, geben analytische Sätze notwendige Verhältnisse zwischen Extensionen wieder.

Dieses Vorgehen hat vier unmittelbare Folgen: (1) Nicht alle grammatischen Sätze sind analytische Sätze. Dies ist der Fall, weil erstens nicht alle grammatischen Sätze der Referenzfixierung dienen, und zweitens manche grammatischen Sätze Regeln für die Verwendung von Worten darstellen, welchen nicht ohne Weiteres eine Referenz zugeordnet werden kann. (2) Nicht alle analytischen Sätze sind grammatische Sätze. Ein Satz wie „Ich bin hier“⁴⁰ ist allein

⁴⁰ Hier will ich von Ansagen auf Anrufbeantwortern und Ähnlichem absehen.

aufgrund der Regeln der Referenzfixierung wahr (vgl. Russell, 2008, S. 106). Und diese Regeln werden in grammatischen Sätzen ausgedrückt. Aber deshalb ist die Aussage, dass ich hier bin, noch kein grammatischer Satz. Natürlich kann sie als Ausdruck einer Regel gebraucht werden, wie fast alle Sätze, aber sie hat auch eine Verwendungsweise als normale Mitteilung. Der Satz „Ich bin hier“ drückt eine kontingente, empirische Tatsache und keine Regel des Sprachgebrauchs aus und ist dennoch analytisch. (3) Es kann keine analytischen Sätze geben, welche den Regeln des Sprachgebrauchs widersprechen. Dies folgt schon daraus, dass eine Lautfolge, welche den Regeln des Sprachgebrauchs widerspricht, kein sinnvoller Satz ist. Sofern wir „Katzen sind Tiere“ als einen grammatischen Satz betrachten, widerspricht dies dem Externalismus Putnams, nach welchem wir entdecken könnten, dass Katzen Roboter-Spione sind. Doch dieser Widerspruch ist erwünscht. Wenn wir die Referenzfixierung in solchem Maß von Kontextfaktoren, welche wir prinzipiell nicht endgültig bestimmen können, abhängig machen, dass die Regeln der Referenzfixierung nichts als Referenten eines Terminus ausschließen, so kommt das einer Abschaffung dieser Regeln gleich. Und wenn alles unter einen Begriff fallen kann, so wird der Begriff der Referenz sinnlos. Der semantische Externalismus ist nur solange haltbar, als er die Kontexte, welche Referenzfixierungen mitbestimmen, nicht außerhalb der Grenzen unseres epistemischen Zugangs sucht. Und dies wird innerhalb meines Vorschlags gewährleistet, indem grammatische Sätze als die Grundlage analytischer Sätze betrachtet werden. Sprich: Externalismus ohne Essenzialismus ist möglich. (4) Alle grammatischen Sätze können *a priori* erkannt werden, manche analytischen hingegen können dies nicht. Grammatische Sätze können allein aufgrund der Kenntnis einer Sprache gewusst werden, da der korrekte Sprachgebrauch im Befolgen der Regeln des Sprachgebrauchs besteht. Um Regeln zu befolgen, muss man aber zumindest implizit um sie wissen. Dieses Wissen wird in grammatischen Sätzen expliziert. Bei analytischen Sätzen hingegen kann der Fall eintreten, dass zwei Namen *A* und *B* unter solchen Umständen eingeführt wurden, dass es nach den Regeln der Referenzfixierung unmöglich ist, dass die Namen sich auf Unterschiedliches beziehen. Dann ist der Satz „Der *A* ist der *B*“ analytisch.⁴¹ Ein Sprecher, der nichts von den Einführungskontexten der Namen weiß, kann dies aber nicht wissen. Dieses Ergebnis kollidiert nicht mit Punkt (3), da es bei einem Namen prinzipiell möglich sein muss, seinen Träger zu identifizieren. Das Wissen um die Referenzfixierung von Namen ist prinzipiell zugänglich, doch muss nicht jeder Sprecher dieses Wissen auch tatsächlich besitzen.

⁴¹ Ich will hier von Fällen absehen, in welchen Namen nicht referieren.

2.4. Ergebnis

Die Aufgabe dieses Kapitels war es, die Unterscheidung zwischen begrifflichen und empirischen Relationen zu rechtfertigen und zu klären. Was lässt sich als Ergebnis festhalten? Begriffliche Relationen spielen in unserem Begründen und Argumentieren eine besondere Rolle. Beim Sprachgebrauch können wir zwei unterschiedliche Fehler machen. Zum einen können wir etwas Falsches sagen, dann behaupten wir von empirischen Relationen, welche nicht bestehen, dass sie bestehen. Andererseits können wir Dinge sagen, die keinen Sinn machen, dann haben wir *überhaupt nichts* gesagt. Hätte Quine recht, dass es nur eine Art von Richtigkeit und Falschheit beim Sprachgebrauch gibt, so könnten wir diese Unterscheidung nicht machen. Es gäbe dann keinen Unterschied dazwischen etwas zu sagen, was nicht der Fall ist, und Unsinn zu reden. Wenn jeder falsche Satz aber Unsinn ist, der nichts sagt, wie kann er dann einem anderen Satz widersprechen? Wie kann er dann falsch sein? Und wenn es überhaupt keine sinnlosen Sätze gibt, worin besteht dann Quines Kritik an dem Satz „Analytische Sätze unterscheiden sich von synthetischen“? Es ist klar, dass Quine diesen Satz nicht einfach für falsch hält. Eine schlichte Ablehnung der Unterscheidung zwischen begrifflichen und empirischen Relationen kommt also nicht in Frage.

Aber dies schließt natürlich nicht aus, dass begriffliche Relationen enger und komplizierter mit empirischen Tatsachen verflochten sind, als man in der Tradition annahm. Unterschiedliche Spielarten des Externalismus haben auf die Abhängigkeit begrifflicher Verhältnisse von empirischen Kontextfaktoren hingewiesen. Daher steht die Forschung vor der Aufgabe, ein kohärentes Bild begrifflicher Zusammenhänge zu entwerfen, welches unterschiedlichen Kontextfaktoren Rechnung trägt. Zudem stellen sich Fragen nach dem epistemischen Zugang zu diesen Zusammenhängen und ihrer Beziehung zur Metaphysik. Eine Bearbeitung dieser Fragestellungen kann in dieser Arbeit nicht geleistet werden. Dennoch wurde auf die zwei großen und erfolgsversprechenden Forschungstraditionen in diesem Feld hingewiesen. Während Forscher in der Tradition des Wiener Kreises versuchen, auf den Begriffen der Referenz und der Wahrheit eine tragfähige Semantik aufzubauen, welche die Natur begrifflicher Relationen klären kann, entwickelte sich, vornehmlich in England, eine Tradition, welche das Wesen begrifflicher Zusammenhänge durch eine Analyse des Sprachgebrauchs zu klären versucht. Zwischen diesen Traditionen bilden sich in jüngster Zeit immer mehr Verbindungen, und auch ich habe in diesem Kapitel für eine solche Synthese argumentiert.

Sätze, welche begriffliche Relationen wiedergeben, werden demnach als Ausdrücke der Regeln des Sprachgebrauchs verstanden. Im Sinne Wittgensteins kann man zwischen ver-

schiedenen Normen des Sprachgebrauchs unterscheiden. Logische Wahrheiten sagen nichts und sind Grenzfälle sinnvoller Sätze, mathematische Wahrheiten sind Regeln für die Umformung empirischer Sätze und in ein umfassendes System wechselseitiger Abhängigkeiten eingegliedert, und das, was in der Tradition als „analytische“ Sätze betrachtet wurde, sind Sätze, welche Regeln für die Verwendung von Worten ausdrücken; sie sind teil-konstitutiv für deren Bedeutung. Aus diesem Ansatz kann, wie Brandom gezeigt hat, ein präziser Begriff der Referenz und Wahrheit gewonnen werden. Dabei wird insbesondere auf die bedeutungskonstitutive Funktion von Schlussfolgerungen aufmerksam gemacht. Ist das klassische semantische Grundvokabular verfügbar, so lassen sich unterschiedliche formale Apparate nutzen. Diese können zur Definition des Begriffs der Analytizität verwendet werden. Damit wird eine Brücke geschlagen zwischen Regeln, wie sie bei Wittgenstein im Zentrum der Überlegungen stehen und der Referenz von Ausdrücken, welche für andere Bedeutungstheorien zentral ist. Zur Fixierung der Referenz werden Regeln des Sprachgebrauchs und verschiedene Informationen über Kontextfaktoren herangezogen. Damit besteht ein vermittelter Zusammenhang zwischen der Bedeutung von Ausdrücken (welche dem Frege'schen Sinn nicht ganz unähnlich ist)⁴² und der Fixierung ihrer Referenz. In der hier vorgeschlagenen Konzeption wird zudem, durch die Normativität und epistemische Zugänglichkeit der Regeln des Sprachgebrauchs, die Notwendigkeit, Objektivität und Wissbarkeit begrifflicher Relationen verständlich. Manchem mag ein solches Ergebnis zu optimistisch erscheinen, doch muss man sich meines Erachtens bei skeptischen⁴³ Gegenargumenten stets fragen, ob nicht die skeptische Konklusion solcher Argumente einen Grund darstellt, an ihren Prämissen zu zweifeln. Sofern das Ergebnis solcher Argumente die Unmöglichkeit rationaler Argumentation ist, scheint mir diese umkehrte Schlussrichtung angebrachter. Und sowohl der Quine'sche Holismus und Naturalismus als auch der Putnam'sche Essenzialismus geraten mitunter gefährlich nahe an dieses Ergebnis.

Was implizieren diese Ergebnisse für die psychologische Forschung? Wohl zuallererst, dass es ein Merkmal der Qualität dieser Forschung ist, klar zwischen begrifflichen Notwendigkeiten und empirischen Hypothesen zu unterscheiden (vgl. Brandtstädter & Sturm, 2004; Holzkamp, 1986; Smedslund, 1988, 1997). Es gibt damit prinzipiell verschiedene Möglichkeiten, eine psychologische Theorie zu kritisieren; neben einer Kritik auf der Grundlage

⁴² Natürlich unterscheiden sich Regeln des Sprachgebrauchs vom Frege'schen Sinn dadurch, dass Letzterer sich nicht in einer normativ gegliederten Praxis konstituieren muss und die Regeln des Gebrauchs eine „Bestimmtheit des Sinns“ weder bieten können noch erfordern.

⁴³ Hiermit soll die Position bezeichnet sein, welche besagt, dass es keinen Unterschied zwischen begrifflichen und empirischen Zusammenhängen gibt.

empirischer Daten oder der Darstellung einer sparsameren Erklärung muss jede psychologische Theorie klar machen können, welches ihre empirischen Hypothesen sind. Entpuppt sich der Inhalt einer Theorie entweder als begrifflich notwendig oder widersprüchlich, so ist damit die Güte der Theorie in Frage gestellt. Aussagen wie „Niemand erinnert sich an die Zukunft“ oder „Wer über kein Gedächtnis verfügt, kann keine Reue empfinden“ können, trotz aller Kritik an der analytisch-synthetisch Unterscheidung, nicht als empirische Hypothesen betrachtet werden. Ebenso wenig können Sätze wie „Manche glücklichen Menschen erleben ausschließlich und ununterbrochen negative Emotionen“ oder „Manche Menschen können die Absicht haben, in der Vergangenheit Handlungen auszuführen“, auch bei noch soviel angeblicher empirischer Bestätigung⁴⁴ als aufregende neue Forschungsergebnisse betrachtet werden. Im Folgenden will ich eine solche Kritik an der Theorie der somatischen Marker versuchen. Doch bevor ich zu dieser Kritik komme, will ich die Theorie und die sie stützenden empirischen Ergebnisse darstellen.

⁴⁴ Natürlich wissen wir noch nicht einmal, was hier unter empirischer Bestätigung zu verstehen sei.

3. DARSTELLUNG DER THEORIE DER SOMATISCHEN MARKER

In diesem Kapitel möchte ich die Theorie der somatischen Marker (SMH) darstellen und den derzeitigen Forschungsstand zu dieser Theorie überblicksartig wiedergeben. Diese Theorie wurde erstmals 1991 von Antonio R. Damasio, Daniel Tranel und Hanna Damasio in dem Aufsatz *Somatic Markers and the Guidance of Behaviour* aufgestellt und wird seitdem weiter ausgearbeitet (vgl. z. B. Gonzalez, Bechara & Martin, 2007; Naqvi, Rudrauf, Damasio & Bechara, 2007; Noel, Bechara, Dan, Hanak & Verbanck, 2007a; Weller, Levin, Shiv & Bechara, 2007). Ziel dieser Darstellung ist es, eine sichere Grundlage für die nachfolgende kritische Analyse dieser Theorie zu bieten. Hierzu muss die Darstellung folgenden Kriterien genügen: (1) Sie muss einen fruchtbaren Boden für die spätere Analyse bereitstellen, d. h., sie muss die später entscheidenden Aspekte besonders deutlich wiedergeben. (2) Die Darstellung muss exegetisch korrekt sein, d. h. die Intentionen der Autoren treffen. (3) Außerdem muss die Darstellung ein maximales Wohlwollen den Autoren gegenüber zeigen.⁴⁵ Um dem zweiten Kriterium gerecht zu werden, will ich zuerst den Entdeckungszusammenhang der Hypothese der somatischen Marker⁴⁶ darstellen, um danach den Begründungszusammenhang wiederzugeben (vgl. Popper, 1973, S. 65). Die Darstellung des Begründungszusammenhangs teilt sich in eine systematische Darstellung der SMH und eine Besprechung der empirischen Befunde.

3.1. Der Entdeckungszusammenhang der SMH

3.1.1. Historischer Rahmen der SMH – Biopsychologie der Emotionen

Um die Theorie der Forschergruppe⁴⁷ um Damasio zu erläutern, ist es hilfreich, dem Leser einige Fakten aus der Geschichte der biologischen Psychologie der Emotionen in Erinnerung zu rufen. Charles Darwin kann als Begründer einer modernen, biologischen Perspektive auf Emotionen betrachtet werden. Er nahm an, dass der Ausdruck von Emotionen hauptsächlich

⁴⁵ Die Darstellung muss also dem *Principle of Charity* genügen (vgl. hierzu Davidson, 1998b; Quine, 1980, S. 115; Wilson, 1959).

⁴⁶ Obwohl die Autoren ihren Ansatz als „Hypothese“ (vgl. Bechara, Damasio, Tranel & Damasio, 2005; Tranel, Bechara & Damasio, 1999) bezeichnen, referiere ich auf ihre Position auch als „Theorie“. Mir scheint die Unterscheidung zwischen Theorie und Hypothese in diesem Fall aus folgenden Gründen nicht sinnvoll: (1) Das intendierte Anwendungsgebiet der SMH ist sehr umfassend. (2) Es liegt ein beträchtliches Maß an gezielter empirischer Forschung vor. (3) Es wurden bei dieser Forschung regelmäßig Vorhersagen aus der SMH abgeleitet, was ihre Entfernung von der Beobachtung (Theoretizität) unterstreicht (für die Unterscheidung zwischen Hypothesen und Theorien vgl. Mouly, 1970, S. 111).

⁴⁷ Diese Gruppe wird im Folgenden auch als „Iowa-Gruppe“ bezeichnet.

der Kommunikation diene und sich auf Grund dieser, Selektionsvorteile mit sich bringenden, Funktion evolutionär entwickeln konnte (vgl. Darwin, 1965).

Unabhängig voneinander entwickelten William James und Carl G. Lange die erste biologische Theorie der Entstehung von Emotionen und damit auch der Erklärung des erlebten Gefühls (vgl. Lange, 1887; William, 1884). Diese Theorie ist als James-Lange-Theorie bekannt geworden. Sie besagt, dass Gefühle nicht, wie allgemein angenommen, einfach durch die Wahrnehmung einer emotional bedeutsamen Situation entstehen, woraufhin dann die physiologischen Reaktionen folgen; sondern, dass eben jene körperlichen Reaktionen das subjektive Gefühl erst entstehen lassen. „Der Kern der James-Lange-Theorie ist, daß die Emotion durch Wahrnehmung von Reaktionen des eigenen Körpers, insbesondere von solchen des vegetativen Nervensystems, entstehen“ (Pinel, 1997, S. 460).

Walter Cannon kritisierte diese Theorie in den zwanziger Jahren des zwanzigsten Jahrhunderts und schlug eine Alternative vor (vgl. Cannon, 1927), die später von Philip Bard (vgl. Bard, 1934) weiter ausgearbeitet wurde und unter dem Namen Cannon-Bard-Theorie bekannt wurde. Diese Theorie besagt, dass es keine kausale Verbindung zwischen dem emotionalen Ausdruck im somatischen und vegetativen Nervensystem und dem subjektiven Empfinden gibt. Es wird angenommen, dass beide Ergebnisse unabhängig voneinander vom emotionalen Stimulus hervorgerufen werden. Cannon dachte, die physiologischen Veränderungen stellten einen Alarmstatus des Organismus dar, der dazu diene, möglicherweise benötigte Energie-reserven bereitzustellen, und nicht spezifisch genug zur Erklärung des subjektiven Empfindens sei. Biologisch ausgedrückt besagt die Theorie, dass sich die Verarbeitung eines emotionalen Stimulus im Thalamus in zwei Wege trennt: einer führt zum Kortex und ist die Ursache für das subjektive emotionale Erleben, der zweite stellt eine Verbindung mit dem Hypothalamus dar, über welchen die körperlichen Reaktionen erzeugt werden (vgl. LeDoux, 1998).

In den 60er Jahren schlugen Stanley Schachter und Jerome Singer eine andere Alternative vor (vgl. Schachter & Singer, 1962). „*Riding the tide of the cognitive revolution, which had begun to penetrate deep into the heart and soul of most areas of psychology by this time, they argued that cognition (thoughts) fill the gap between the nonspecificity of feedback and the specificity of felt experiences.*“ (LeDoux, 1998, S. 47). Sie behaupteten, dass Gefühle durch die kognitive Bewertung der wahrgenommen physiologischen Erregung entstehen. Dieser kognitive Ansatz wird bis heute weiter verfolgt und hat viele Variationen erfahren. Kritisiert wurde der kognitive Ansatz in den 80er Jahren von Robert Zajonc, der zeigte, dass affektive Reaktionen ohne die bewusste Wahrnehmung des Stimulus möglich sind. Deshalb dachte er, dass un-

bewusste (nicht-kognitive) Prozesse einen wesentlichen Anteil an der Entstehung von Gefühlen haben müssten (vgl. Zajonc, 1980).

Doch außerhalb dieser genuin psychologischen Diskussionen entwickelte James Papez 1937 ein biologisches Modell der Emotionen (vgl. Papez, 1937). Papez ging davon aus, dass die Reizinformation vom Thalamus aus auf zwei getrennten Wegen verarbeitet wird, von denen einer für kognitive und der andere für emotionale Verarbeitung zuständig ist. Der für Emotion zuständige Bereich wurde als Papez-Kreis bezeichnet, danach, so Papez, fließe die Information von den sensorischen Kortexen zum *Gyrus cinguli*, der für bewusste Gefühle verantwortlich sei, weiter zum *Hippocampus*. Von dort gelange sie zum Hypothalamus, der auch direkt vom Thalamus gespeist werde und der die emotionalen Reaktionen des autonomen Nervensystems kontrolliere. Von dort wiederum bestehe eine Verbindung über den anterioren Thalamus zurück zum cingulären Kortex (vgl. Rosenzweig, Leiman & Breedlove, 1999).

Unabhängig von Papez veröffentlichten Heinrich Klüver und Paul Bucy im gleichen Jahr ihren ersten Bericht über ein Phänomen, das als Klüver-Bucy-Syndrom bekannt wurde (vgl. Klüver & Bucy, 1937). Dieses Phänomen wurde zwar schon von Brown und Schäfer 1888 entdeckt (vgl. Brown & Schäfer, 1888), fand jedoch damals keine Beachtung. Das Klüver-Bucy-Syndrom kann bei Affen durch die bilaterale Entfernung des anterioren Temporallappens hervorgerufen werden. Der Zustand dieser Affen zeichnet sich vor allem durch undifferenziertes, untypisches Fressverhalten, gesteigertes Sexualverhalten, Wahl unangemessener Sexualobjekte, verstärkte orale Exploration (auch schon bekannter Gegenstände), Zahmheit und den Verlust von Angstreaktionen aus (vgl. Kolb & Whishaw, 1996).

Paul MacLean versuchte 1949 eine umfassende Synthese der bisherigen Theorien auf Grundlage der von Papez aufgestellten Hypothesen (vgl. MacLean, 1949). Die zentrale Struktur dieser Theorie war das limbische System (*Hippocampus*, *Gyrus cinguli*, *Hypothalamus*, *anterioren Thalamus*, *Fornix*, *Gyrus parahippocampalis*, *Gyrus dentatus*), das nach MacLean den phylogenetisch ältesten Teil des Gehirns ausmacht; im Gegensatz zum Neokortex, dem jüngsten Teil der ihm zufolge mit Rationalität und Kognition verbunden ist (vgl. LeDoux, 1998).

Die SMH kann in diesem Rahmen als ein Versuch betrachtet werden, die beschriebene Forschung zu den biologischen Grundlagen der Emotionen fruchtbar zu machen für eine biopsychologische Betrachtung von Entscheidungsverhalten (vgl. Damasio, 2006, S. 129). „Damasio extended the James-Lange theory with Somatic Marker Hypothesis (SMH) to not only explain the generation of subjective feeling but also the factors guiding decision making“ (Chiu & Lin, 2007, S. 2). Darüber hinaus kann die SMH auch die Positionen von Schachter

und Singer, Zajonc und LeDoux integrieren (vgl. Bechara, Damasio & Damasio, 2000; Colombetti & Thompson, 2007). Manche Autoren sprechen deshalb von einer „James-Lange-Schachter-Damasio-Theorie“ (Kandel & Kupfermann, 1996, S. 618). Die SMH versucht jedoch auch Schwächen dieser älteren Theorien zu überwinden. Aus Sicht der SMH überbetonen Schachter und Singer die Kognition, während LeDoux die Beteiligung des Neokortex bei der Entstehung von Emotionen vernachlässigt.⁴⁸ Bevor ich genauer auf die Theorie eingehe, möchte ich noch eine weitere historische Quelle der SMH darstellen: die Forschung zu Patienten, welche Schädigungen des präfrontalen Kortex erlitten haben.

3.1.2. Phineas P. Gage und seine Leidensgenossen

„One observation was key to the development of the somatic marker hypotheses: Although patients with ventromedial prefrontal damage were capable of performing normally on nearly every neuropsychological test they were asked to take, and obviously had preserved conventional intellectual functions, their ability to express emotion and experience feeling relative to complex personal and social situations, for example, as in the expression and experience of embarrassment, was compromised“ (Tranel, Bechara & Damasio, 1999, S. 1048).

Solche Patienten mit Schädigungen des ventromedialen präfrontalen Kortex (VMPFC)⁴⁹ bilden einen Hauptgegenstand der Forschungen der Iowa-Gruppe, und bei der Mehrzahl der empirischen Arbeiten zur SMH wurden solche Patienten untersucht. Der erste Patient, von dessen Verletzung des VMPFC wir eine wissenschaftliche Dokumentation besitzen, hieß Phineas P. Gage. Der Fall wurde erstmals im Jahr 1884 von John M. Harlow beschrieben (vgl. Harlow, 1999). Gage war Vorarbeiter bei der *Rutland & Burlington Railroad* und erlitt bei Sprengarbeiten am 13. September 1848 bei *Cavendish* in *New England* eine schwere Kopfverletzung.⁵⁰ Dabei durchbohrte ein ca. 1 Meter langer Eisenstab seinen Kopf. Der Eisenstab trat unterhalb des linken Backenknochens in den Schädel ein, zerstörte das linke Auge und verließ ihn am vorderen Bereich des Schädels bei der Mittellinie, nahe der Verbindung zwischen *Sutura coronalis* und *Sutura sagitalis* (vgl. Harlow, 1999, S. 281). Gages Schädel und der

⁴⁸ Diese Schwächen entsprechen dem, was Richard J. Davidson „*sin 1*“ und „*sin 2*“ der biopsychologischen Emotionsforschung nennt (vgl. Davidson, 2003).

⁴⁹ Die Autoren verstehen hierunter die medialen und verschiedene laterale Bereiche des orbitofrontalen Kortex, welche Brodmann-Areal 25, untere Teile von 24, 32, mediale Bereiche von 11, 12 und 10, sowie die angrenzende weiße Substanz einschließen (vgl. Colombetti & Thompson, 2007, S. 51).

⁵⁰ Für eine ausführliche Darstellung des Falls und seiner Folgen für die Neurowissenschaften s. Macmillan (2000).

Eisenstab sind erhalten. Hanna Damasio nutzte diese, um durch eine Computersimulation plausibel zu machen, dass ausschließlich der präfrontale Kortex Gages durch den Unfall verletzt wurde (vgl. Damasio, 2006, S. 33; Damasio, Grabowski, Frank, Galaburda & Damasio, 1994).

Berühmt wurde der Fall zum einen dadurch, dass Gage den Unfall tatsächlich überlebte, zum anderen durch die bemerkenswerten Veränderungen in Gages Verhalten und seiner Persönlichkeit, welche nach dem Unfall eintraten. Harlow beschrieb diese Veränderung wie folgt:

“Previous to his injury, although untrained in the schools, he possessed a well-balanced mind, and was looked upon by those who knew him as a shrewd, smart businessman, very energetic and persistent in executing all his plans of operation. In this regard his mind was radically changed, so decidedly that his friends and acquaintances said he was ‘no longer Gage’” (Harlow, 1868, S. 340).

Beachtenswert an diesen Veränderungen ist besonders, dass sie nicht auf andere, grundlegendere Veränderungen zurückzuführen waren, da alle grundlegenden körperlichen und geistigen Funktionen erhalten schienen (vgl. Bechara & Damasio, 2005, S. 338; Harlow, 1999). Die Iowa-Gruppe betrachtet diese Veränderungen vornehmlich als eine Störung des Entscheidungsverhaltens. Gage wird als unfähig betrachtet, seine eigenen Interessen zu verfolgen sowie rationale Entscheidungen zu treffen, besonders im persönlichen und sozialen Bereich:

„The alterations in Gage’s personality were not subtle. He could not make good choices, and the choices he made were not simply neutral. They were not the reserved or slight decisions of someone whose mind is diminished and who is afraid to act, but were instead actively disadvantageous. One might venture that either his value system was now different, or, if it was still the same, there was no way in which the old values could influence his decisions“ (Damasio, 2006, S. 11).

Die Iowa-Gruppe folgert hieraus, dass ein Zusammenhang zwischen VMPFC und Entscheidungsverhalten bestehen muss (vgl. Bechara & Damasio, 2005, S. 337). Einen anderen Patienten, den Fall EVR⁵¹ (vgl. Eslinger & Damasio, 1985), welchen die Iowa-Gruppe als entscheidend für die Entwicklung der SMH ansieht (vgl. Damasio, Tranel & Damasio, 1991, S. 217ff.), betrachtet Damasio als direkt analog zu dem Fall Gage. Auch EVR erlitt eine

⁵¹ In Damasio (2006) wird der Deckname Elliot für diesen Patienten verwendet.

Schädigung des VMPFC und zeigte anschließend starke Persönlichkeitsveränderungen und Beeinträchtigungen des Entscheidungsverhaltens.

„I was now certain that Elliot had a lot in common with Phineas Gage. Their social behaviour and decision-making defect were compatible with a normal social-knowledge base, and with preserved higher-order neuropsychological functions such as conventional memory, language, basic attention, basic working memory and basic reasoning“ (Damasio, 2006, S. 51).

Die SMH kann in diesem Rahmen als ein Versuch betrachtet werden, folgende drei Befunde zu erklären: (1) Verletzungen des präfrontalen Kortex ziehen sowohl Beeinträchtigungen im Entscheidungsverhalten bezüglich des persönlichen und sozialen Bereichs als auch Veränderungen im Gefühlsleben nach sich. (2) Wenn Probleme im Entscheidungsverhalten und im Gefühlsleben bei anderweitig unauffälligem neuropsychologischem Profil auftauchen, so liegt in der Regel eine Schädigung des VMPFC vor. (3) Erstrecken sich die Schädigungen auch auf den dorsalen und lateralen Bereich des frontalen Kortex, so liegt in der Regel auch eine Beeinträchtigung in Funktionen wie Aufmerksamkeit und Arbeitsgedächtnis vor (vgl. Damasio, 2006, S. 61). Die SMH setzt dazu Emotionen und rationales Entscheiden, innerhalb eines neurowissenschaftlichen Rahmens, in Beziehung. Es treffen sich hier historische Entwicklungslinien, welche sich zum einen aus der Erforschung der biologischen Grundlagen der Emotionen speisen, sich zum anderen aber aus Befunden zu Verletzungen des frontalen Kortex herleiten.

3.2. Systematische Darstellung der SMH

Die SMH wurde wiederholt von Mitgliedern der Iowa-Gruppe zusammenfassend dargestellt (z. B. Bechara, 2004; Bechara & Damasio, 2005; Bechara, Damasio & Damasio, 2000; Damasio, 1996, 2006; Tranel, Bechara & Damasio, 1999). Darüber hinaus liegen verschiedene Darstellungen der SMH durch andere Autoren vor (z. B. Beer, 2007, S. 18; Dunn, Dalgleish & Lawrence, 2006; Martin, 2006, S. 203f.; Stamm, 2007, S. 101f.). Im Folgenden werden ich eine auf diesen Arbeiten aufbauende Darstellung der SMH geben. Hierzu werde ich die Darstellung der Theorie selbst von der Darstellung ihrer empirischen Überprüfung trennen.

Die SMH versucht zu erklären, wie menschliches Entscheidungsverhalten zustande kommt, welche Rolle Emotionen dabei spielen und welche biologischen Strukturen dabei welche

Funktionen übernehmen. Hierzu macht Damasio vier Voraussetzungen (vgl. Damasio, 1996, S. 1414ff.):

- (1) Menschliches Denken und Entscheiden ist von neurobiologischen Vorgängen abhängig, welche bewusst oder unbewusst ablaufen können.
- (2) Kognitionen hängen, unabhängig von ihrem Inhalt, von Prozessen wie Aufmerksamkeit, Arbeitsgedächtnis und Emotionen ab.
- (3) Denken und Entscheiden hängen von Wissen über Situationen, Akteure, Handlungsmöglichkeiten und Ergebnisse ab. Solches Wissen ist in kortikalen und subkortikalen Strukturen gespeichert. Dieses Wissen wird explizit, entweder durch unterschiedliche motorische Reaktionen oder durch Vorstellungen (*images*).
- (4) Wissen kann klassifiziert werden in vier Typen: (a) angeborenes und erworbenes Wissen bezüglich biologischer Regulationsprozesse und Körperzustände, welche sich unter anderem in Emotionen ausdrücken, (b) Faktenwissen über Dinge, Relationen, Handlungszusammenhänge, Regeln, Geschichten etc., (c) Wissen über die Verbindungen zwischen den Inhalten von (a) und (b), und (d) Wissen, welches aus der Kategorisierung von (a), (b) und (c) resultiert (vgl. Bechara, Damasio & Damasio, 2000, S. 296).

Der Kerngedanke der SMH besteht nun darin, anzunehmen, dass genau das Wissen um die Verbindung zwischen angeborenen homöostatischen Regulationsprozessen (Punkt (4a)) und Faktenwissen (Punkt (4b)) im VMPFC gespeichert ist und eine entscheidende Rolle beim Entscheidungsverhalten spielt. Das Wissen um Handlungsmöglichkeiten und Ergebnisse führt dabei zu emotionalen Prozessen, welche das Denken und Entscheiden beeinflussen. Zentral ist dabei der Begriff des somatischen Markers.

3.2.1. Was ist ein somatischer Marker?

Damasio unterscheidet zwischen Emotionen (*emotions*) und Gefühlen (*feelings*). Dabei werden Emotionen mit Veränderungen im Körper oder Gehirn identifiziert, während Gefühle Wahrnehmungen solcher Veränderungen sind (vgl. Damasio, 2006, S. 270). Innerhalb der Emotionen wiederum unterscheidet Damasio zwischen zwei Typen: primären und sekundären (vgl. Damasio, 2006, S. 131). Emotionen, welche schon kurz nach der Geburt erlebt werden können und durch feste, angeborene Stimulussituationen ausgelöst werden können, nennt Damasio „primär“. Diese Emotionen werden, nach Damasio, korrekt durch die James-Lange-Theorie beschrieben und beruhen zu größten Teil auf subkortikalen Verarbeitungsprozessen (vgl. LeDoux, 1998). Dabei löst ein Stimulus oder eine Vorstellung, vermittelt durch sub-

kortikale Prozesse (insbesondere innerhalb der Amygdala), eine körperliche Veränderung aus, welche die Emotion darstellt. Sekundäre Emotionen hingegen basieren auf Erfahrung. Stimuli oder Vorstellungen, welche wiederholt in Verbindung mit einer primären Emotion auftreten, werden selbst zu Auslösern von Emotionen, sekundären Emotionen. Es handelt sich hier also um eine klassische Konditionierung. Die Assoziationen zwischen den Stimuli und den primären Emotionen sollen im VMPFC repräsentiert sein.

„I propose that the ventromedial prefrontal cortex establishes a simple linkage, a memory in fact, between the disposition for a certain aspect of a situation [...] and the disposition for the type of emotion that in past experience has been associated with the situation“ (Damasio, 1996, S. 1415).

Doch das Auslösen dieser sekundären Emotionen bewerkstelligt der VMPFC nicht direkt, sondern indirekt über dieselben Strukturen, welche auch für primäre Emotionen verantwortlich sind, namentlich den anterioren *Gyrus cinguli*, die Amygdala und andere subkortikale Gebiete, die für Veränderungen des autonomen Nervensystems verantwortlich sind (vgl. Damasio, 2006, S. 138).

Einen Stimulus, welcher Emotionen auslösen kann, nennt Damasio einen *emotionally competent stimulus* (ECS)⁵² (Damasio, 2004, S. 50f., 2005, S. 51). Primäre und sekundäre Emotionen können durch zwei Typen von Auslösern verursacht werden. Zum einen gibt es wahrnehmungsabhängige Auslöser, welche die Iowa-Gruppe *primary inducers* nennt. Zum anderen gibt es wahrnehmungsunabhängige Auslöser (Gedanken, Vorstellungen, Erinnerungen), die *secondary inducers* heißen (vgl. Bechara & Damasio, 2005, S. 340). *Secondary inducers* sind bewusste Vorstellungen, also kognitive Inhalte. Kognitive Inhalte werden von Damasio als *images* bezeichnet und mit topografisch organisierten Repräsentationen in frühen sensorischen Kortizes⁵³ in Verbindung gebracht. Die wahrnehmungsunabhängige Produktion solcher Vorstellungsbilder⁵⁴ hängt von ihrer dispositionalen Repräsentation in höheren Assoziationskortizes ab (vgl. Damasio, 2006, S. 108).

Emotionen sind die Grundlage der Gefühle (*feelings*), welche wiederum die Grundlage somatischer Marker darstellen. Die meisten Gefühle entstehen, laut Damasio, durch

⁵² Auch Vorstellungen können als ECS wirken. Der Begriff des Stimulus ist hier also sehr weit zu verstehen.

⁵³ Hier sind insbesondere die primären sensorischen Kortizes gemeint.

⁵⁴ Der Term „Vorstellungsbilder“ soll alle Sinnesmodalitäten umfassen (s. Damasio, 2006, S. 107).

Emotionen.⁵⁵ Sie sind die Wahrnehmung der körperlichen Veränderungen, welche Damasio Emotionen nennt. „*That process of continuous monitoring, that experience of what your body is doing while thoughts about specific contents roll by, is the essence of what I call a feeling*“ (Damasio, 2006, S. 145). Wir haben, nach Damasio, zu Gefühlen einen privilegierten Zugang; sie sind private, mentale Erfahrungen (vgl. Stamm, 2007, S. 116). Emotionen dagegen, im Sinne Damasios, sind objektive Veränderungen im Soma (d. i. im muskuloskelettale System, den Viszera und Komponenten des internen Milieus (vgl. Damasio, 1996, S. 1414)). „Durch den Prozess des Fühlens werden Körperzustände für das Gehirn zugänglich“ (Stamm, 2007, S. 119). Diese Zugänglichkeit hängt von der Repräsentation der emotionalen Veränderungen des Soma in frühen somatosensorischen Rindenfeldern⁵⁶ (SI, SII), der Insula und anderen subkortikalen Strukturen ab (Damasio, 1996; vgl. Damasio, 2006, S. 65; Tranel, Bechara & Damasio, 1999; Weller et al., 2007).⁵⁷ Bis hierher lässt sich über die Entstehung von Gefühlen nach Damasio zusammenfassend Folgendes sagen: Gefühle werden durch Vorstellungsbilder verursacht, manche, aber nicht alle dieser Bilder stammen aus der Wahrnehmung. Vorstellungsbilder lösen, primär (angeboren) oder sekundär (erlernt), Reaktionen in subkortikalen (und motorischen) Bereichen aus. Diese Reaktionen verursachen somatische Veränderungen, welche die Iowa-Gruppe Emotionen nennt. Die emotionalen Körperzustände werden wiederum wahrgenommen, d. h. in frühen somatosensorischen Gebieten und der Insula repräsentiert. Diese Repräsentationen werden als Gefühle (*feelings*) bezeichnet (vgl. Damasio, 2006). Der Sachverhalt stellt sich aber etwas komplizierter dar, denn emotionale somatische Veränderungen sind keine notwendige Bedingung für Gefühle. Gefühle können zwar durch tatsächliche Veränderungen des Körpers ausgelöst werden, aber die subkortikalen Effektorstrukturen und motorischen Gebiete, welche diese Veränderungen normalerweise hervorrufen, können den Körper auch umgehen. D. h., die Strukturen, welche für Emotionen verantwortlich sind, können, ohne den Körper zu beeinflussen, Aktivitätsmuster in den frühen somatosensorischen Arealen (SI, SII) und der Insula auslösen, die denjenigen gleichen, welche durch tatsächliche Emotionen ausgelöst werden. Bezüglich dieser beiden Arten, Gefühle zu erzeugen, spricht Damasio von einem *body-loop* und einem *as-if-body-loop* (vgl. Damasio, 1996, S. 1415). Bei Aktivierung des *body-loops* tritt das somatosensorische Repräsentations-

⁵⁵ Jene Gefühle, welche aus anderen Quellen stammen, werden Hintergrundgefühle genannt. Sie bestehen im Wesentlichen in der Abwesenheit von Emotionen.

⁵⁶ Die Iowa-Gruppe verwendet den Ausdruck „somatosensorische Kortizes“ sehr eigenwillig. Es sollen hiermit insbesondere diejenigen Gebiete bezeichnet werden, in welchen sich Repräsentationen viszeraler Veränderungen finden lassen. Ich schließe mich zum Zweck meiner Kritik diesem Sprachgebrauch an.

⁵⁷ Für diese Repräsentationen nimmt Damasio eine rechtshemisphärische Dominanz an.

muster, welches ein Gefühl darstellt, mit den ihm zugehörigen somatischen Veränderungen auf, während diese somatischen Veränderungen bei Aktivierung des *as-if-body-loops* nicht auftreten (vgl. Tranel, Bechara & Damasio, 1999, S. 1048). Dabei muss ein Gefühl nicht bewusst werden, sondern kann durchaus unbewusst bleiben und wirken.

Gefühle sind, laut Iowa-Gruppe, die Grundlage der Bewertung kognitiver Inhalte. Die Bewertung eines kognitiven Inhalts bezeichnen die Autoren als „Markierung“. Ein kognitiver Inhalt wird durch eine Verbindung mit der Repräsentation eines emotionalen somatischen Zustands als „gut“ oder „schlecht“ markiert. Diese Markierung betrifft auch die Vorstellungsbilder, welche das imaginierte, zukünftige Ergebnis einer möglichen Handlung repräsentieren.

„For instance, when the somatosensory image which defines a certain emotional response is juxtaposed to the images which describe a related scenario of future outcome, and which triggered the emotional response via the ventromedial linkage, the somatosensory pattern marks the scenario as good or bad“ (Bechara, Damasio & Damasio, 2000, S. 297).

Nun lässt sich formulieren, was ein „somatischer Marker“ ist: *Die Repräsentation eines emotionalen somatischen Zustands, welche zur Bewertung eines kognitiven Inhalts führt, wird in der SMH als „somatischer Marker“ bezeichnet.* Somatische Marker sind also das neurobiologische Korrelat der emotionalen Bewertung eines Vorstellungsinhalts. Vier neurologische Strukturen sind notwendige Voraussetzungen für die Realisierung somatischer Marker: (1) der VMPFC, welcher die Assoziationen zwischen bestimmten Situationen und Repräsentationen emotionaler Körperzustände beinhaltet (via *body-loop* oder *as-if-body-loop*), (2) zentrale Effektorstrukturen des autonomen Nervensystems, die emotionale somatische Veränderungen verursachen, (3) somatosensorische Kortizes, welche Repräsentationen emotionaler somatischer Zustände wiedergeben können, und (4) die Basalganglien, der anteriore cinguläre Kortex, Kerne des Hirnstamms sowie humorale und neurale Signalwege, welche somatische Zustände dem Gehirn zugänglich machen (vgl. Bechara & Damasio, 2005, S. 344; Damasio, 1996, S. 1416).

3.2.2. Wie beeinflussen somatische Marker die Entscheidungsfindung?

Somatische Marker spielen, laut SMH, eine entscheidende Rolle bei der menschlichen Entscheidungsfindung. Die Schwierigkeiten der Patienten mit Schädigungen des VMPFC, die

oben⁵⁸ beschrieben wurden, werden auf eine gestörte Funktion der somatischen Marker zurückgeführt. Daher lautet eine Hauptthese der SMH: „*that decision making is not mediated solely by cold cognitive processes [...], and that arriving at an advantageous decision also requires affective information*“ (Weller et al., 2007, S. 962). Die SMH plädiert somit für die Einbeziehung von Emotionen in die Konzeption der praktischen Rationalität. „*Rationality is probably shaped and modulated by body signals, even as it performs the most sublime distinctions and acts accordingly*“ (Damasio, 2006, S. 200). Der SMH zufolge verfügt eine emotionslose Vernunft schlicht nicht über die Mittel, eine rationale Entscheidung in alltäglichen Lebenssituationen zu treffen (vgl. Bechara & Damasio, 2005, S. 337). „*[O]ur brains [...] must do the marvellous job with more than just pure reason*“⁵⁹ (Damasio, 2006, S. 172f.). Doch *warum* soll die ´reine Vernunft` diese Aufgabe nicht übernehmen können?

Die Unfähigkeit der menschlichen ´reinen Vernunft`, praktische Rationalität hervorzubringen, liegt, laut Iowa-Gruppe, in unseren begrenzten kognitiven Ressourcen begründet. In Entscheidungssituationen stehen wir vor unterschiedlichen Alternativen, zwischen welchen wir wählen müssen. Eine rationale Entscheidung berücksichtigt in dieser Situation die voraussichtlichen Folgen der möglichen Entscheidungen. Wären wir ausschließlich mit emotionsloser Vernunft ausgestattet, so müssten wir in solchen Fällen die voraussichtlichen Folgen jeder Alternative logisch deduzieren. Wir wären also auf „*extremely slow and laborious logical operations over many potential alternatives*“ (Tranel, Bechara & Damasio, 1999, S. 1049) angewiesen. Für ein solches logisch-berechnendes Verfahren fehlen uns aber in den allermeisten Entscheidungssituationen die zeitlichen und kognitiven Ressourcen (vgl. hierzu auch Baddeley, 2003). Emotionale Vorgänge sind genau keine logisch-berechnenden Verfahren, und hierin liegt der SMH zufolge auch ihre Stärke. Die emotionale Verarbeitung läuft ohne die begrenzenden kognitiven Ressourcen wie Aufmerksamkeit und Arbeitsgedächtnis ab. Somatische Marker stellen damit eine Möglichkeit dar, viele höchst heterogene Informationen schnell für eine Entscheidungsfindung nutzbar zu machen (vgl. Damasio, Tranel & Damasio, 1991, S. 219f.).

Zur Verdeutlichung des angenommenen Mechanismus ist es hilfreich, sich eine konkrete Situation vor Augen zu führen. Stellen wir uns also eine Person *P* vor, die mit der Möglichkeit konfrontiert ist, mit einer bekannten Person eine bestimmte Veranstaltung zu besuchen

⁵⁸ s. Kapitel 3.1.2

⁵⁹ Damasio bezieht sich an der angeführten Stelle mit dem Ausdruck *pure reason* tatsächlich auf Kant. Dies ist selbstverständlich völlig verfehlt, da nach Kant jene Vernunftleistungen, welche Damasio angeblich erforscht, eindeutig der *praktischen* Vernunft zuzurechnen sind.

(vgl. Damasio, Tranel & Damasio, 1991, S. 220). Nun läuft, der SMH zufolge, folgender Prozess ab: Wenn *P* darüber nachdenkt, was sie tun soll, so spielt sie in ihren Geist die vermuteten Folgen einer Entscheidung für oder gegen den Besuch der Veranstaltung durch. Es entstehen sehr schnell verschiedene Handlung-Ergebnis-Szenarien. „*To our consciousness, the scenarios are made of multiple imaginary scenes, not really a smooth film, but rather pictorial flashes of key images in those scenes, jump cut from one frame to another, in quick juxtapositions*“ (Damasio, 2006, S. 170). Die Vorstellungsbilder, welche diese Szenarien bilden, können nun als ECS wirken. Die Szenarien lösen also, mittels der im VMPFC gespeicherten Assoziationen, Aktivität in Strukturen wie der Amygdala und der anterioren Insula aus, welche wiederum emotionale Effektorstrukturen wie das basale Vorderhirn, den Hypothalamus, verschiedene Kerne des Hirnstamms und den anterioren cingulären Kortex aktivieren. Diese emotionalen Effektorstrukturen induzieren entweder einen entsprechenden emotionalen somatischen Zustand (*body-loop*) oder aktivieren, zusammen mit der Amygdala und dem VMPFC, direkt die frühen somatosensorischen Kortizes, in welche Gefühle (*feelings*) repräsentiert werden (*as-if-body-loop*) (vgl. Damasio, 2005, S. 51). Diese Repräsentationen in den frühen somatosensorischen Kortizes (SI, SII, Insula) können dann als somatische Marker der Handlung-Ergebnis-Szenarien fungieren. Durch diese „Markierung“ der Handlungsoptionen wird die Entscheidungsfindung beeinflusst. Je „besser“ der somatische Marker einer Handlungsalternative ist, um so höher ist die Wahrscheinlichkeit, dass diese Alternative gewählt wird. Die Wahrscheinlichkeit, dass *P* die Veranstaltung besucht, wird also umso größer, je „besser“ der somatische Marker dieses Handlung-Ergebnis-Szenarien ausfällt.

„[S]omatic markers are a special instance of feelings generated from secondary emotions. Those emotions and feelings have been connected, by learning, to predicted future outcomes of certain scenarios. When a negative somatic marker is juxtaposed to a particular future outcome the combination functions as an alarm bell. When a positive somatic marker is juxtaposed instead, it becomes a beacon of incentive“ (Damasio, 2006, S. 174).

Die SMH nimmt bezüglich der Art und Weise der Beeinflussung der Entscheidungsfindung verschiedene mögliche Wege an. Somatische Marker wirken sowohl als Kriterien der Wahl von Handlungsalternativen als auch als Mittel zur Erleichterung des logischen Denkens während einer Entscheidung. Außerdem können somatische Marker offen und bewusst, aber auch verdeckt und unbewusst wirken (vgl. Bechara & Damasio, 2005; Bechara, Damasio & Damasio, 2000; Damasio, 1996, 2006; Damasio, Tranel & Damasio, 1991). Wenn somatische Marker bewusst wirken, so wirken sie entweder als Gefühle oder durch die Veränderungen

rationaler Verarbeitungsprozesse. Treten sie als Gefühle in unser Bewusstsein ein, so werden wir uns durch sie bewusst, dass wir eine bestimmte Handlungsoption als emotional „besser“ oder „schlechter“ bewerten. Sind keine anderen Faktoren wirksam, so werden wir wahrscheinlich unser Handeln an diesen Bewertungen ausrichten. Dies trifft sich mit Damasio's Annahme, dass auch ethische und ästhetische Werte auf einer ähnlichen biologischen Grundlage beruhen (vgl. Damasio, 2005, 2007).

Die rationale Verarbeitung können somatische Marker wiederum auf zwei Arten beeinflussen: erstens durch die Allokation kognitiver Ressourcen und zweitens durch die Beschränkung des Entscheidungsraums. Die Allokation kognitiver Ressourcen soll insbesondere durch die Wirkung somatischer Marker auf die Ordnung der Verarbeitungsinhalte zustande kommen:

„Whether we conceive of reason based on automated selection, or on a logical deduction mediated by a symbolic system, or – preferably – both, we cannot ignore the problem of order. I propose the following solution: (1) If order is to be created among available possibilities, then they must be ranked. (2) If they are to be ranked, then criteria are needed (values or preferences are equivalent terms). (3) Criteria are provided by somatic markers, which express, at any given time, the cumulative preferences we have both received and acquired“ (Damasio, 2006, S. 199).

Die Einschränkung des Entscheidungsraums geschieht primär durch das schnelle Ausschließen von Handlungsoptionen, welche von vornherein ausgeschlossen und keiner Erwägung wert sind.

„Certain option-outcome pairs can be rapidly rejected and, pertinent facts can be more effectively processed. The hypothesis thus suggests that somatic markers normally help constrain the decision-making space by making that space manageable for logic-based, cost-benefit analyses“ (vgl. Damasio, 1996, S. 1415).

Dieser Mechanismus der Einschränkung des Entscheidungsraums wurde wiederholt mit dem sogenannten Rahmen-Problem⁶⁰ in Zusammenhang gebracht. Es wurde vermutet, dass die SMH eine mögliche Erklärung gibt, warum dieses Problem im menschlichen Alltag nicht auftritt⁶¹ (vgl. Gerrans, 2007; Megill & Cogburn, 2005).

⁶⁰ Als „Rahmen-Problem“ wird die Schwierigkeit bezeichnet, dass bei automatischen Deduktionen und Berechnungen das Berücksichtigen irrelevanter nicht ausgeschlossen werden kann (vgl. Lormand, 2001).

⁶¹ Dies kann nicht überraschen, da Damasio die theoretischen Grundideen der SMH vollständig von Ronald de Sousa (1987, S. 190ff.) übernommen hat. De Sousa schlägt aber seine Theorie explizit als eine Lösung des philosophischen Rahmenproblems vor. Warum Damasio, trotz der enormen Übereinstimmungen zwischen der SMH

Wenn somatische Marker hingegen unbewusst wirken, sind sie eher als Signale zu verstehen, welche eine Handlungstendenz bewirken. Sie begünstigen dann unbewusst appetitives oder aversives Verhalten. Dieses *biasing signal* (vgl. Tranel, Bechara & Damasio, 1999, S. 1048) beruht zum einen auf der unspezifischen Aktivierung der Transmittersysteme des Hirnstamms, zum anderen auf der Beeinflussung von Arealen, welche motorischen Aufgaben dienen, wie dem *Striatum* und dem anterioren supplementären/cingulären motorischen Kortex (vgl. Bechara & Damasio, 2005, S. 343). Dieser Mechanismus wird von der SMH mit dem in Verbindung gebracht, was wir im Alltag „Intuition“ nennen (vgl. Damasio, 2006, S. 187f.). Ein weiterer Effekt, welchen somatische Marker haben können, ist die Rückwirkung auf den VMPFC und die Amygdala. Hierdurch wird die Auslösung weiterer emotionaler somatischer Zustände moduliert (vgl. Bechara & Damasio, 2005, S. 343).

Zusammenfassend lässt sich sagen, dass somatische Marker die Entscheidungsfindung durch die emotionale Bewertung erwarteter Handlungs-Ergebnis-Szenarien beeinflussen. Diese Bewertung möglicher Handlungsoptionen verursacht eine Veränderung der Wahrscheinlichkeit des Ergreifens der jeweiligen Handlungsoption. Diese kausale Wirkung kann auf unterschiedliche Weise hervorgebracht werden, durch bewusste und unbewusste Prozesse, durch viszerale Veränderungen (*body-loop*) und ‘scheinbare’ viszerale Veränderungen (*as-if-body-loop*), durch spezifische und unspezifische Mechanismen. Doch wie auch immer der jeweilige Kausalnexus beschaffen ist, es ist letztlich stets ein Gefühl (*feeling*), welches die Erwartungen ‘bewertet’, die das Handeln der Person beeinflusst. Eine zusammenfassende Darstellung der, von der SMH angenommenen, Zusammenhänge findet sich in Abbildung 5.

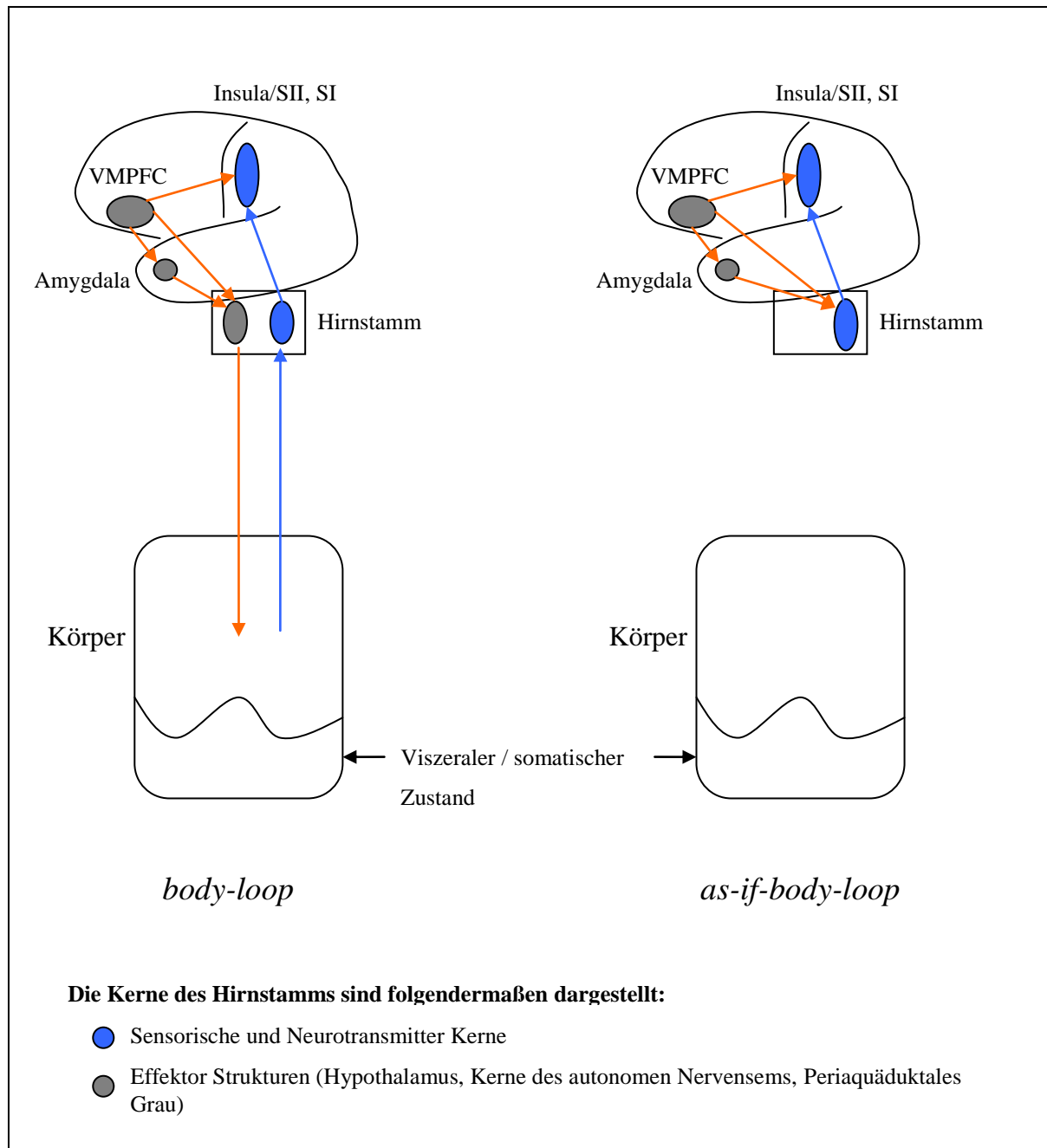


Abbildung 5: Darstellung der wichtigsten Annahmen der SMH

Die Darstellung basiert auf einer entsprechenden Abbildung in Bechara & Damasio (2005, S. 342). Die Pfeile geben die Reihenfolge der physiologischen Veränderungen an. Die oberen schwarzen Umrisse stellen das Gehirn schematisch dar, die unteren den Körper. Links ist der Mechanismus des *body-loops* dargestellt, rechts der des *as-if-body-loops*. Effektor-Strukturen sind mit roten Pfeilen wiedergegeben, sensorische Komponenten mit blauen.

3.3. Empirische Überprüfung der SMH

Die SMH hat sich als äußerst fruchtbar erwiesen. Es gibt heute eine große und ständig wachsende Anzahl empirischer Arbeiten, in welche die SMH auf unterschiedliche Weise eingeht (vgl. Dunn, Dalgleish & Lawrence, 2006). Es wurden Brücken zwischen der SMH und einer allgemeinen Theorie emotionaler und sozialer Intelligenz geschlagen (z. B. Bar-On, Tranel, Denburg & Bechara, 2003; Killgore & Yurgelun-Todd, 2007). Die SMH wurde in verschiedenen klinischen Stichproben untersucht (z. B. Blair, 2001; Haaland & Landro, 2007; Lee, Kim, Seo, Park, Jeong, Kim & Lee, 2007; Schmitt, Brinkley & Newman, 1999; Tchanturia, Liao, Uher, Lawrence, Treasure & Campbell, 2007; van Honk, Hermans, Putman, Montagne & Schutter, 2002), unter entwicklungspsychologischer Perspektive betrachtet (z. B. Denburg, Cole, Hernandez, Yamada, Tranel, Bechara & Wallace, 2007; Denburg, Tranel & Bechara, 2005; Fein, McGillivray & Finn, 2007) und auf dem Hintergrund der differenziellen Psychologie bearbeitet (z. B. Craig, 2004; Suzuki, Hirota, Takasawa & Shigemasa, 2003). Darüber hinaus wurde die SMH bei der Erforschung von Substanzabhängigkeiten eingesetzt (z. B. Gonzalez, Bechara & Martin, 2007; Johnson, Xiao, Palmer, Sun, Wang, Wei, Jia, Grenard, Stacy & Bechara, 2008; Naqvi et al., 2007; Noel et al., 2007a; Noel, Van der Linden, d'Acremont, Bechara, Dan, Hanak & Verbanck, 2007b; Verdejo-Garcia, Bechara, Recknor & Perez-Garcia, 2007). In all diesen Gebieten wurden vielfältige Ergebnisse hervorgebracht, welche potenziell Implikationen für die Weiterentwicklung der SMH mit sich bringen. Da es aber im Rahmen dieser Arbeit weder möglich noch notwendig ist, all diese Forschungsfelder zu besprechen, will ich zuerst einen Überblick über die empirischen Befunde geben, welche die Iowa-Gruppe meist zur Untermauerung der SMH anführt, um dann einige Herausforderungen der jüngsten Zeit zu besprechen. Das 'klassische' empirische Fundament der SMH bilden Läsionsstudien. Dabei wurden hauptsächlich drei unterschiedliche Variablen untersucht: (1) somatische Reaktionen auf emotional besetzte Stimuli, (2) die Testergebnisse verschiedener neuropsychologischer Tests, insbesondere eines durch die Iowa-Gruppe entwickelten Tests namens *Iowa Gambling Task* (IGT), und (3) somatische Reaktionen während der Bearbeitung der IGT. Außerdem liegen Studien mittels bildgebender Verfahren und einige neuropharmakologische Befunde vor.

3.3.1. Läsionsstudien

3.3.1.1. Somatische Reaktionen auf emotional besetzte Stimuli

Die SMH sagt voraus, dass Patienten mit Schädigungen des VMPFC veränderte somatische Reaktionen auf emotionsauslösende Stimuli zeigen. Die Kernidee bei dieser Art von Experimenten bestand darin, dass der VMPFC zwar nicht für die Auslösung primärer Emotionen notwendig ist, jedoch durchaus für das Entstehen sekundärer Emotionen benötigt wird. Da im VMPFC die assoziativen Verknüpfungen zwischen Situationen und häufig mit ihnen auftretenden Emotionen gespeichert sind, können bei Ausfall des VMPFC die Situationen keine entsprechenden emotionalen Reaktionen mehr auslösen (vgl. Damasio, 1996, S. 1417). Da bei der Entstehung der emotionalen somatischen Zustände insbesondere das autonome Nervensystem eine große Rolle spielen soll, wurde als Index der Aktivierung des autonomen Nervensystems in den meisten Experimenten der Hautleitwiderstand herangezogen. Die Hypothesen der Experimente besagten, dass Patienten mit geschädigtem VMPFC keine elektrodermale Reaktion auf emotional besetzte Stimuli (sekundäre Emotionen) zeigen, während sich bei gesunden Probanden oder anderweitig neurologisch geschädigten Patienten eine solche Reaktion finden lässt. Die Fähigkeit, elektrodermale Reaktionen als Teil einer primären Emotion zu zeigen, wurde durch die Messung des Hautleitwiderstandes während einer Schreckreaktion kontrolliert. Diese Fähigkeit sollte der SMH zufolge bei einer Schädigung des VMPFC erhalten bleiben. Diese Vorhersagen können beim derzeitigen Forschungsstand als bestätigt betrachtet werden (Damasio, Tranel & Damasio, 1990; Damasio, Tranel & Damasio, 1991; Tranel, 1994; Tranel, Damasio & Damasio, 1995). Damasio kommentiert diese Befunde folgendermaßen:

„The findings suggest that patients with bilateral ventromedial frontal damage and decision-making defects in personal and social domain, no longer have a normal ability to generate somatic responses to stimuli with an emotional component“ (Damasio, 1996, S. 1417).

Diese Befunde dienen vor allem der Stützung der Auffassung von Emotionen, welche der SMH zugrunde liegt.⁶² Zusätzlich zeigen diese Befunde, dass bei Patienten mit bilateraler Schädigung des VMPFC gleichzeitig mit einer Beeinträchtigung der Entscheidungsfindung in persönlichen und sozialen Bereichen eine veränderte emotionale Reaktion auftritt. Dies lässt

⁶² s. Kapitel 3.2.

eine Verbindung zwischen Emotion und Entscheidungsfindung zumindest plausibel erscheinen.

3.3.1.2. Befunde mittels neuropsychologischer Tests – Die Iowa Gambling Task

Die SMH kann anscheinend erklären, warum Patienten mit Verletzungen des VMPFC in einer großen Anzahl von neuropsychologischen Standardtestverfahren keine auffälligen Ergebnisse erzielen. Diese Patienten schneiden in der Regel sowohl in Intelligenztests (z. B. *Wechsler Adult Intelligence Scale*), bei Gedächtnisaufgaben (z. B. *Wechsler Memory Scale* oder *Rey-Osterrieth Complex Figure Test*), in Sprachtests (z. B. *Shipley Hartford Vocabulary Test*) als auch bei Wahrnehmungstests nicht schlechter ab als Kontrollpersonen (vgl. Damasio, Tranel & Damasio, 1991, S. 219).

„[M]any patients with prefrontal damage do not demonstrate impairment on conventional neuropsychological procedures, even on quintessential ‘executive function’ or ‘frontal lobe’ tasks such as the Wisconsin Card Sorting Test, the Trail-Making Test, and the Tower of Hanoi“ (Tranel, Bechara & Damasio, 1999, S. 1051).

Trotz dieser unauffälligen Ergebnisse sind diese Patienten in ihrem alltäglichem Leben beeinträchtigt. Die Patienten zeichnen sich durch einen flachen Affekt und durch Entscheidungen aus, welche ihren eigenen Interessen widersprechen. Daher entwickelte Antoine Bechara einen Test, welcher angeblich genau für die Defizite dieser Patienten sensitiv ist: die IGT (vgl. Bechara, Damasio, Damasio & Anderson, 1994).

Die IGT besteht in einem Kartenspiel. Vor dem Probanden werden vier Stapel mit Spielkarten auf den Tisch gelegt, sie werden mit A, B, C und D bezeichnet. Aufgabe des Probanden ist es, von einem Stapel seiner Wahl eine Karte zu ziehen. Bei jeder Karte gewinnt er Spielgeld; bei manchen Karten verliert er aber auch Spielgeld. Ziel ist es, möglichst viel Spielgeld zu gewinnen. Es ist eine entscheidende Eigenschaft des Spiels, dass der Proband der Möglichkeit kurzfristiger Gewinne widerstehen muss, um langfristig seinen Gewinn zu maximieren. Dies wird durch die unterschiedlichen Gewinn- und Verlust-Festlegungen für jeden Stapel erreicht. Karten von Stapel A und B bringen stets \$100 pro Karte, während Stapel C und D nur \$50 pro Karte einbringen. Fünf von zehn Karten aus Stapel A bringen einen Verlust zwischen \$35 und \$150 mit sich. In Stapel B ist unter zehn Karten nur eine Verlustkarte, der Verlust beträgt aber \$1250. Zieht man von Stapel C, so trifft man in fünf von zehn Fällen auf Verluste zwischen \$25 und \$75. Stapel D enthält unter zehn Karten eine mit einem Verlust von \$250.

Demzufolge sind die Stapel mit niedrigen Gewinnen (C und D) langfristig vorteilhaft, da sie geringere Verluste mit sich bringen. Der Proband zieht 100 Karten. Dabei weiß er nicht, wie viele Karten er zu ziehen hat und erhält zu Beginn ein Darlehen von \$2000 (vgl. Bechara et al., 1994; Bechara, Damasio & Damasio, 2000). Es wird erfasst, welcher Stapel bei welchem Zug gewählt wird. Die IGT wurde inzwischen in über 100 neurologischen und psychiatrischen Untersuchungen eingesetzt (vgl. Bowman, Evans & Turnbull, 2005).

Da die IGT die Entscheidungsfindung im Alltag simulieren soll, ist die Vorhersage der SMH, dass Patienten mit geschädigtem VMPFC schlechter bei dieser Aufgabe abschneiden werden als Kontrollpersonen oder anderweitig neurologische beeinträchtigte Patienten. Diese Vorhersage wurde wiederholt bestätigt (vgl. Bechara et al., 1994; Bechara, Damasio & Damasio, 2003; Bechara et al., 1999; Bechara, Damasio, Tranel & Damasio, 1997; Bechara, Tranel, Damasio & Damasio, 1996). Da die Festlegung von Gewinnen und Verlusten für den Probanden nicht durchschaubar ist, ist eine rein rationale Kosten-Nutzen-Analyse nicht möglich. Der typische Befund hierbei ist, dass Personen einer Kontrollgruppe die gewinnbringendste Strategie im Verlauf des Tests erlernen, während Patienten mit einer Schädigung des VMPFC dies nicht tun.

„As the task progressed, normal and brain-damaged controls gradually shifted their selection toward the good decks (C and D), and away from the bad decks (A and B). By the last two blocks (trials 61-100), in fact, the subjects were choosing almost exclusively from the good decks. In sharp contrast, ventromedial prefrontal subjects failed to demonstrate this shift in behavior“ (Tranel, Bechara & Damasio, 1999, S. 1051)

Da aber, laut SMH, nicht nur der VMPFC für die Generierung somatischer Marker notwendig ist, sondern auch andere Strukturen wie die Amygdala, sagt die SMH auch eine Beeinträchtigung in der IGT von Patienten mit Amygdalaverletzungen vorher. Auch diese Hypothese konnte wiederholt bestätigt werden (Bechara, Damasio & Damasio, 2003; vgl. Bechara et al., 1999; Dunn, Dalgleish & Lawrence, 2006, S. 245; Martínez-Selva, Sánchez-Navarro, Bechara & Román, 2006). Neuere Ergebnisse deuten darauf hin, dass besonders eine Schädigung des rechten Präfrontalkortex zu einer Beeinträchtigung in der IGT führt (Clark, Manes, Antoun, Sahakian & Robbins, 2003).

Eine Kombination der neuropsychologischen und der physiologischen Befunde stellt einen weiteren empirischen Beleg der SMH dar. Es werden hierbei physiologische Parameter während der Bearbeitung der IGT erfasst. Meist wurde hierzu wiederum der Hautleitwiderstand herangezogen (vgl. Bechara et al., 1996). Da laut SMH die fehlende

somatische Reaktion auf imaginierte Handlung-Ergebnis-Szenarien kausal verantwortlich für das schlechte Abschneiden der VMPFC-geschädigten Patienten ist, sagt die SMH voraus, dass, während die Probanden in der IGT ihre Entscheidungen bedenken, sich bei gesunden Personen somatische Veränderungen feststellen lassen, welche Patienten mit Verletzungen des VMPFC oder der Amygdala nicht zeigen. Auch diese Hypothese konnte wiederholt bestätigt werden (vgl. Bechara, Damasio & Damasio, 2003; Bechara et al., 1999; Bechara et al., 1997; Bechara et al., 1996). Gesunde Versuchspersonen entwickeln im Laufe der Durchführung der IGT eine antizipatorische elektrodermale Reaktion, unmittelbar bevor sie eine Karte von einem 'schlechten Stapel' (also A oder B) ziehen (vgl. Tranel, Bechara & Damasio, 1999, S. 1052ff.). Sowohl Patienten mit Verletzungen des VMPFC als auch solche mit Verletzungen der Amygdala entwickeln keine antizipatorischen elektrodermalen Reaktionen vor einer 'riskanten' Entscheidung (vgl. Bechara et al., 1999, S. 5477).

„We have interpreted these results to indicate that the absence of anticipatory SCRs [elektrodermale Reaktion] in the ventromedial patients is a physiological correlate for their insensitivity to future outcomes. This interpretation is compatible with the idea that these subjects fail to activate biasing signals which would serve as value markers in the distinction between choices with good and bad future outcomes“ (Tranel, Bechara & Damasio, 1999, S. 1054).

3.3.2. Belege mittels bildgebender Verfahren

Die meisten Studien, welche bildgebende Verfahren nutzen, um die SMH zu prüfen, verwenden die IGT. In diesen Untersuchungen bearbeiten die Versuchspersonen meist eine Variante der IGT, während eine Positronen-Emissions-Tomografie (PET) oder eine funktionelle Magnet-Resonanz-Tomografie (fMRT) durchgeführt wird. Vorausgesagt wird dabei eine Aktivierung der Gebiete, welche von der SMH als essenziell für die Generierung und Verarbeitung somatischer Marker betrachtet werden. Diese Vorhersage bestätigend wurde wiederholt festgestellt, dass die Ergebnisse in der IGT mit dem cerebralen Blutfluss im rechten orbitofrontalen Kortex während der Bearbeitung korrelieren (vgl. Bolla, Eldreth, London, Kiehl, Mouratidis, Contoreggi, Matochik, Kurian, Cadet, Kimes, Funderburk & Ernst, 2003; Ernst, Bolla, Mouratidis, Contoreggi, Matochik, Kurian, Cadet, Kimes & London, 2002; Northoff, Grimm, Boeker, Schmidt, Bermpohl, Heinzl, Hell & Boesiger, 2006; Volz, Schubotz & von Cramon, 2006). Ein mögliches Problem für die SMH stellt die Tatsache dar, dass einige Studien eine Aktivierung des dorsolateralen Präfrontalkortex

während der Bearbeitung der IGT zeigten (vgl. Dunn, Dalgleish & Lawrence, 2006, S. 261).⁶³ Studien, welche die neuronale Grundlage von Emotionen unabhängig von Entscheidungsfindung mittels bildgebender Verfahren untersuchten, konnten hingegen übereinstimmend eine Aktivierung des medialen Präfrontalkortex bei der Verarbeitung emotionaler Information feststellen (vgl. Phan, Wager, Taylor & Liberzon, 2002, S. 344). Aus mehreren Gründen ist hier allerdings keine abschließende Bewertung möglich: (1) Eine Aktivierung des dorso-lateralen Präfrontalkortex bei der Bearbeitung der IGT, welcher oft mit dem Arbeitsgedächtnis in Zusammenhang gebracht wird, widerspricht nicht direkt der SMH. (2) Es sind wesentliche methodische Probleme bildgebender Verfahren noch ungelöst (z. B. Verzerrungsfakte im orbitofrontalen Bereich bei der fMRT) (vgl. Dunn, Dalgleish & Lawrence, 2006). (3) Die SMH wird derzeit mit dem Ziel weiterentwickelt, verschiedenen Gebieten des präfrontalen Kortex spezifische Aufgaben bei der Entscheidungsfindung zuzuschreiben (vgl. Bechara & Damasio, 2005; Volz, Schubotz & von Cramon, 2006).

3.3.3. Psychopharmakologische Daten

Die Vertreter der SMH postulieren eine spezifische Wirkung unterschiedlicher Neurotransmitter auf das Entscheidungsverhalten.

„DA [Dopamin] biases decisions covertly (perhaps through action in the striatum). On the other hand, 5-HT [Serotonin] biases decisions overtly (perhaps through action in the anterior cingulate and probably the adjacent SMA [supplementär-motorisches Areal])“ (Bechara & Damasio, 2005, S. 361).

Die Belege hierfür sind zurzeit noch nicht eindeutig. Zwar konnten Belege für den Einfluss von Dopamin und Serotonin auf das Entscheidungsverhalten erbracht werden (vgl. Bechara, Damasio & Damasio, 2001; Sevy, Hassoun, Bechara, Yechiam, Napolitano, Burdick, Delman & Malhotra, 2006), doch konnte kein Effekt von Noradrenalin festgestellt werden (vgl. O'Carroll & Papps, 2003). Zusätzliche Unterstützung der SMH bieten Befunde, welche Serotonin und Dopamin mit Entscheidungsverhalten in Verbindung bringen, ohne auf die SMH Bezug zu nehmen (vgl. Blair, Finger, Marsh, Morton, Mondillo, Buzas, Goldman, Drevets & Blair, 2007; Clark, Cools & Robbins, 2004; Denk, Walton, Jennings, Sharp, Rushworth & Bannerman, 2005; van der Plasse, La Fors, Meerkerk, Joosten, Uylings & Feenstra, 2007).

⁶³ Der dorsolaterale Kortex wird oft mit Funktionen des Arbeitsgedächtnisses in Verbindung gebracht. Seine Aktivierung ließe demnach eher auf nicht-emotionale Verarbeitungsmechanismen schließen.

Zusammenfassend lässt sich sagen, dass die SMH sich auf eine breite und robuste empirische Datenbasis stützen kann. Auch wenn durchaus verschiedene Probleme in der empirischen Forschung zur SMH aufgetaucht und einige Präzisierungen notwendig sind, kann die Theorie bisher nicht als empirisch falsifiziert gelten (vgl. Dunn, Dalgleish & Lawrence, 2006, S. 265). Es gab jedoch in jüngster Zeit zwei Herausforderungen an die SMH, welche ich im Folgenden besprechen möchte.

3.3.4. Zwei empirische Herausforderungen an die SMH

Beide empirischen Herausforderungen setzen an der IGT an, da ein Großteil der empirischen Belege der SMH mittel der IGT erbracht wurden. Während der erste Einwand gegen Befunde mittels IGT behauptet, dass das Verhalten gesunder Probanden bei der IGT nicht durch die langfristigen Gewinnaussichten, sondern ausschließlich durch die relativen Gewinn- und Verlust-Häufigkeiten bestimmt ist, besagt der zweite Einwand, dass das Verhalten bei der IGT durch bewusste Kosten-Nutzen-Analyse determiniert ist. Interessant hierbei ist, dass sich die beiden Einwände zumindest teilweise widersprechen.

3.3.4.1. Eine erste alternative Erklärung der Befunde mittels IGT

Die Kernthese der Kritik Ching-Hung Lins und Yao-Chu Chius ist, dass die 'Gewinn-Verlust-Verteilung' der unterschiedlichen Stapel bei der IGT nicht ausreichend kontrolliert wird (vgl. Chiu & Lin, 2007; Lin et al., 2007). Diese Kritik beruht darauf, dass in der ursprünglichen IGT eine Karte mit sich ausgleichenden Gewinnen und Verlusten als Verlust gerechnet wird. Betrachtet man diese Karten als eine eigenständige neutrale Kategorie, so ergibt sich, dass die Häufigkeiten der Gewinne und Verluste bei den einzelnen Stapeln so variieren, dass sie zur Erklärung der Ergebnisse herangezogen werden können. Demnach wären es nicht die größeren langfristigen Gewinne, welche die gesunden Probanden zur Wahl von Stapel C oder D bewegen, sondern das größere Gewinn-Verlust-Häufigkeitsverhältnis. Diese Alternativerklärung könnte verständlich machen, warum gesunde Kontrollpersonen mehr Karten vom Stapel B ziehen als gemäß SMH zu erwarten wäre (vgl. Lin et al., 2007). Die Autoren konnten empirische Ergebnisse anführen, welche diese Alternativerklärung unterstützen (vgl. Chiu & Lin, 2007; Lin et al., 2007).

Bevor allerdings die Schlagkraft dieser Kritik bewertet werden kann, müssen weitere unabhängige Überprüfungen der vorgeschlagenen Alternativerklärung abgewartet werden. Chiu und Lin bestreiten interessanterweise nicht nur (wie die Iowa-Gruppe), dass die Probanden die IGT kognitiv durchschauen können, sondern auch, dass sie bei der Bearbeitung der IGT

‘emotionale Voraussicht’ nutzen können. Die zweite Herausforderung behauptet nun genau, dass die Probanden bei der Bearbeitung der IGT aufgrund rationaler Überlegungen die voraussichtlich gewinnmaximierende Strategie wählen.

3.3.4.2. Die kognitive Durchschaubarkeit der IGT

Nachdem die Iowa-Gruppe 1997 Ergebnisse berichtete, aus welchen geschlossen wurde, dass Versuchspersonen bei der Bearbeitung der IGT, ohne explizites Wissen um eine vorteilhafte Spielstrategie, sich gewinnmaximierend verhalten können und diese unbewusste Verhaltenssteuerung mit der antizipatorischen elektrodermalen Reaktion kovariiert (vgl. Bechara et al., 1997), wurde dieser Befund 2004 von Tiago V. Maia und James L. McClelland bestritten (vgl. Maia & McClelland, 2004). Während die Iowa-Gruppe das explizierte Wissen der Probanden mittels zweier breiter, offener Fragen erfasste, welche nach den ersten 20 und dann nach jeder zehnten Karte gefragt wurden, verwendeten Maia und McClelland einen aufwendigeren Fragebogen.

„We have found that when participants behave advantageously in the IGT, (i) they have conscious access to the relative goodness and badness of the deck, and (ii) they have explicit, reportable knowledge that could provide the basis for such judgments and behavior“ (Maia & McClelland, 2004, S. 16078).

Als Alternativerklärung für die bekannten Daten greifen die Autoren auf das Konzept des *reversal learning* zurück, welches die Fähigkeit betrifft, sich veränderten Umweltkontingenzen anzupassen (vgl. Maia & McClelland, 2004, S. 16080). Die Antwort der Iowa-Gruppe auf diese Kritik besteht in drei Punkten: (1) Es wird darauf hingewiesen, dass die Annahme unbewusster Entscheidungsverzerrungen nicht die zentrale These der SMH darstellt, sondern dass die Rolle der Emotion bei der Entscheidungsfindung den Kern der Hypothese ausmacht; (2) die Iowa-Gruppe bestreitet, dass Alternativerklärungen die Daten zur Veränderung des Hautleitwiderstandes bei der Durchführung der IGT erklären können; (3) die Erklärung durch *reversal learning* wird als inkompatibel mit anderen neuropsychologischen Befunden betrachtet (vgl. Bechara et al., 2005). Es sind also auch bei der Bewertung dieser Kritik noch einige empirische Befunde abzuwarten.

3.4. Zusammenfassung

Als zusammenfassende Darstellung der SMH möchte ich die Beschreibung des Zustandekommens einer Entscheidung anfügen. Damit soll der neurobiologische Mechanismus,

welchen die SMH postuliert, synoptisch veranschaulicht werden. Person *P* befinde sich in einer Entscheidungssituation. Ihr stehen drei von ihr wahrgenommene Handlungsoptionen zur Verfügung: A, B und C. Nun geschieht, laut SMH, Folgendes: In den frühen sensorischen Kortizes des Gehirns von *P* treten Repräsentationen von A, B und C auf. Diese verursachen Repräsentationen der von *P* erwarteten Handlungsfolgen von A, B und C: A', B' und C'. Diese Repräsentationen bilden jeweils Handlungs-Ergebnis-Szenarien: (A-A'), (B-B') und (C-C'). Jedes dieser Szenarien wird an den VMPFC weitergeleitet. Der VMPFC löst für jedes der Szenarien einen Impuls aus. Dieser Impuls kann entweder somatische Veränderungen hervorrufen (*body-loop*) oder eine Simulation solcher Veränderungen in somatosensorischen Kortizes auslösen (*as-if-body-loop*). In beiden Fällen entsteht in den somatosensorischen Kortizes eine Repräsentation somatischer Veränderungen. Diese Repräsentationen somatischer Veränderungen werden somatische Marker genannt. Es entstehen also somatische Marker für jedes der drei Szenarien. Nennen wir diese: SM-A, SM-B und SM-C. Die somatischen Marker werden gemäß ihres 'Wertes' geordnet. Die somatischen Marker werden nun mit den Repräsentationen, welche sie verursacht haben, zu Einheiten verbunden: (A-A')-(SM-A), (B-B')-(SM-B) und (C-C')-(SM-C). Stark positive somatische Marker wirken als Anreiz, stark negative hingegen als Abschreckung. Dies geschieht, indem die Einheiten aus somatischen Markern und imaginierten Handlungsoptionen (zum Teil über dazwischen geschaltete kognitive Prozesse) unterschiedliche motorische Areale des Gehirns von *P* derart beeinflussen, dass diejenige Handlungsoption mit dem größten positiven somatischen Marker, *ceteris paribus*, die höchste Wahrscheinlichkeit erhält, sich im Verhalten von *P* zu manifestieren. Die *ceteris-paribus*-Klausel bezieht sich dabei auf andere mögliche Einflussfaktoren wie eine rationale Kosten-Nutzen-Analyse, welche *P* unabhängig vom Mechanismus der somatischen Marker durchführen könnte.

Zum Stand der empirischen Forschung bezüglich der SMH lässt sich sagen, dass die SMH innerhalb der biologischen Psychologie als gut bestätigt gilt, auch wenn sich an unterschiedlichen Stellen die Notwendigkeit zu Präzisierungen und Überarbeitungen ergeben hat. Darüber hinaus hat sich die SMH als extrem fruchtbar erwiesen und viele Forschungsarbeiten auf den unterschiedlichsten Gebieten angestoßen. Damit kann die SMH als eine aktuelle neurobiologische Theorie des menschlichen Entscheidungsverhaltens gelten. Eine solche Theorie scheint einer genaueren Analyse wert zu sein. Mit einer solchen Analyse will ich im nächsten Kapitel beginnen.

4. DER PSYCHOLOGISCHE GEHALT DER SMH

Nachdem die SMH im vorangegangenen Kapitel dargestellt wurde, will ich in diesem Kapitel mit der begrifflichen Analyse dieser Theorie beginnen. Insgesamt soll sich diese Analyse in zwei Schritten vollziehen, von welchen der erste in diesem Kapitel geleistet wird. Ziel dieses ersten Schrittes ist es, folgende These zu belegen:

(T1) Wer die SMH akzeptiert, ist auf die Behauptung folgender psychologischer Gesetzmäßigkeit festgelegt:

(PG) Wenn eine Person das erwartete Handlungsergebnis für eine bestimmte Handlung emotional positiver bewertet als die erwarteten Handlungsergebnisse für alle anderen Handlungen, so wird sie, *ceteris paribus*, die entsprechende Handlungsoption mit größerer Wahrscheinlichkeit auswählen als alle anderen Handlungsoptionen.

Für (T1) will ich in vier Schritten argumentieren, welche den Abschnitten dieses Kapitels entsprechen. Diese vier Schritte können durch folgende Thesen charakterisiert werden:

- (1) Psychologisch interessante neurowissenschaftliche Theorien müssen psychologische Prädikate enthalten. Diese können in solchen Theorien an drei unterschiedlichen Orten auftauchen (nämlich im Antecedens empirischer Gesetze, im Consequens solcher Gesetze und in Brückenprinzipien).
- (2) In der SMH treten psychologische Prädikate an allen drei Orten auf.
- (3) Es muss eine psychologische Gesetzmäßigkeit geben, welche von der SMH reduziert werden soll.
- (4) Diese Gesetzmäßigkeit ist (PG).

Im nächsten Kapitel soll eine begriffliche Analyse des Satzes (PG) folgen. Ich werde nicht versuchen, den Terminus „psychologisches Prädikat“ zu definieren, da es bei theoretischen Analysen solcher Begriffe im Allgemeinen sinnvoller ist, sich ihrer Bedeutung anhand von Beispielen zu nähern. Eine Definition kann, sofern sie zustande kommt, als das Ergebnis einer solchen Reflexion betrachtet werden. Als typische psychologische Prädikate betrachte ich solche, die auf Intentionen, propositionalen Gehalt oder Rationalität hindeuten, wie „denken“, „beabsichtigen“, „bewerten“, „entscheiden“ oder „urteilen“. Typischerweise werden solche

Prädikate von intensionalen Vorkommnissen⁶⁴ anderer sprachlicher Ausdrücke (z. B. dass-Sätzen) begleitet und lassen sich im eigentlichen Sinn nur auf Lebewesen, oder gar nur auf den Menschen, anwenden.

4.1. Psychologische Prädikate in der Biopsychologie

Dieser Abschnitt soll zeigen, dass zumindest die systematisch interessanten biopsychologischen Theorien psychologische Prädikate verwenden müssen.⁶⁵ Wie kann diese Behauptung begründet werden? Es ist hier zunächst festzuhalten, dass die Neurowissenschaften im Allgemeinen mit dem Anspruch auftreten, uns etwas über das zu sagen, was wir im Alltag „Geist“ oder „Psyche“ nennen. So lautet etwa der erste Satz des Standardlehrbuchs *Principles of Neural Science*: „*The goal of neural science is to understand the mind – how we perceive, move, think, and remember*“ (Kandel, Schwartz & Jessell, 2000, S. XXXV). Daraus folgt, dass eine derartige Wissenschaft sich irgendwelcher Begriffe bedienen muss, welche sich auf psychische Sachverhalte, Entitäten oder Eigenschaften beziehen. Wie sollten wir anderenfalls etwas über diese erfahren? Auf den ersten Blick scheinen diese Überlegungen trivial, dennoch sind sie in den letzten Jahrzehnten alles andere als unbestritten geblieben. Ich will deshalb im Folgenden die möglichen Positionen zu diesem Thema aufzeigen und prüfen.

4.1.1. Braucht die Neurowissenschaft psychologische Prädikate?

Auf die Frage, ob die Neurowissenschaft psychologische Prädikate verwenden muss, kann es drei unterschiedliche Antworten geben: „ja“, „nein“, und „Neurowissenschaft ist als Wissenschaft von Psychischem unmöglich“. Ich behaupte, dass die SMH die erste Antwortmöglichkeit vertreten muss, und will deshalb später näher auf diese eingehen.

Die dritte Antwort wird typischerweise von Vertretern eines Substanzdualismus verfochten. Eine solche Position wurde z. B. von René Descartes und in neuerer Zeit von Karl R. Popper und John C. Eccles (1984) vertreten. Die Argumentation verläuft dabei typischerweise folgendermaßen: Zuerst wird dafür argumentiert, dass der Geist immateriell sei, und dann wird gefolgert, dass der Geist niemals Gegenstand einer Naturwissenschaft sein kann. Der Substanzdualismus wird aber zurzeit von keinem einflussreichen Philosophen oder Neuro-

⁶⁴ s. Kapitel 5.2.2.1.

⁶⁵ Es ist klar, dass neurobiologische Grundlagenforschung oft ohne psychologische Prädikate auskommt; z. B. in der Zellphysiologie, der Biochemie oder der Genetik. Unter „systematisch interessanten Theorien“ möchte ich hier deshalb diejenigen verstehen, welche mit dem Anspruch auftreten, sich auf ein Gegenstandsgebiet zu beziehen, welches mit dem der Psychologie direkt in Kontakt steht. Es soll also das angesprochen sein, was im eigentlichen Sinn „biologische Psychologie“ oder „kognitive Neurowissenschaft“ genannt werden kann.

wissenschaftler vertreten (vgl. Heil, 2004, S. 177; Russell, 2008, S. 141). Und so scheint auch die SMH schlicht vorauszusetzen, dass die Erforschung psychischer Phänomene mit naturwissenschaftlichen Methoden gelingen kann. Es wäre schwer zu erkennen, wie man sonst z. B. die Rede von „*neural systems responsible for retrieval of knowledge about emotions*“ (Adolphs & Damasio, 2000, S. 195) oder die Behauptung „*To say that mind comes from brain is indisputable*“ (Damasio, 2006, S. 251) verstehen sollte.

Um meine Kritik später auch auf andere Theorien als die SMH verallgemeinern zu können, will ich die zweite Antwortmöglichkeit auf die aufgeworfene Frage kritisch diskutieren; obwohl die Vertreter der SMH diese Möglichkeit nicht ergreifen. Diese zweite Antwortmöglichkeit bestreitet die Notwendigkeit der Verwendung psychologischer Prädikate in der Neurowissenschaft. Wir hätten dann eine Wissenschaft vor uns, welche sich ausschließlich in einer physikalischen Sprache ausdrückt, aber nichtsdestoweniger Erklärungen bezüglich desjenigen Gegenstandsbereichs liefert, welchen wir normalerweise als „mental“, „geistig“ oder „psychisch“ beschreiben.⁶⁶ Eine solche Position wurde in den letzten Jahrzehnten als „eliminativer Materialismus“ oder „Eliminativismus“ bezeichnet. In jüngerer Zeit wurde der Eliminativismus von Quine (1980) und Feyerabend (1963) sowie insbesondere von Paul M. Churchland (1981) und Patricia S. Churchland (1986) vertreten. Der Eliminativismus behauptet, dass die Psychologie letztlich durch eine rein biologische Beschreibung und Erklärung menschlichen Verhaltens ersetzt werden wird.

„Eliminative Materialism is the thesis that our commonsense conception of psychological phenomena constitutes a radically false theory, a theory so fundamentally defective that both the principles and the ontology of that theory will eventually be displaced, rather than smoothly reduced, by completed neuroscience“ (Churchland, 1981, S. 67).

Die Elimination psychologischer Prädikate kann aber auf zweierlei Weise geschehen: Erstens könnten die Ergebnisse der Neurowissenschaft den meisten gängigen Verwendungen psychologischer Prädikate widersprechen. Oder die Neurowissenschaft könnte, zweitens, psychologische Prädikate zum Zwecke eines schnellen und ungehinderten Fortschritts ignorieren. Ich möchte im Folgenden gegen beide Möglichkeiten argumentieren.

⁶⁶ Wir könnten dann natürlich nicht im eigentlichen Sinne von „Biopsychologie“ reden, sondern nur von „verhaltenswissenschaftlich relevanter Neurowissenschaft“. (Es ist damit fraglich, ob diese Position tatsächlich die Möglichkeit einer neurowissenschaftlichen Erforschung psychologischer Phänomene annimmt.) Ich werde trotzdem auch in diesem Fall von Biopsychologie reden, da diese zukünftige Wissenschaft, laut Eliminativismus, im Kanon der Wissenschaften genau die systematische Stelle von Psychologie und Biopsychologie einnehmen soll.

Beginnen wir mit der zweiten Möglichkeit: Es sei also angenommen, dass keine reife biopsychologische Theorie psychologische Prädikate enthält, da sich im Verlauf der Forschung keine fruchtbare Verwendung für diese ergibt. Es sei auch zugestanden, dass diese reifen neurowissenschaftlichen Theorien extrem erfolgreich sind. Wir wären dann z. B. in der Lage, sämtliche Muskelkontraktionen und endokrinen Prozesse wesentlich weiter, exakter und einfacher vorherzusagen, als dies heute möglich ist. Eine entscheidende Frage ist dann, warum eine solche Wissenschaft behaupten sollte psychologische Phänomene zu erklären. Der Eliminativist wird hierauf rasch eine Antwort finden: „Ich kann sämtliches Verhalten eines menschlichen Körpers erklären und vorhersagen; und was kann von einer psychologischen Theorie darüber hinaus verlangt werden?“ Aber ist dies eine schlagende Antwort? Ich denke nicht (vgl. Schon, 2005, S. 227). In der beschriebenen Situation besteht, *ex hypothesi*, kein Widerspruch zwischen gängigen Verwendungen psychologischer Prädikate und neurowissenschaftlichen Sätzen. Es gibt aber unterschiedliche Möglichkeiten, den Gegenstandsbereich genuin psychologischer Erklärungen so zu definieren, dass es zu keiner Überschneidung mit dem Bereich neurowissenschaftlicher Erklärungen kommt. Wie bereits erwähnt, betrachtet etwa der Substanzdualismus die Gegenstandsbereiche dieser Erklärungen ohnehin als grundverschieden. Und daneben gibt es unterschiedliche Formen des nicht-reduktiven Materialismus, welche hier ebenfalls keine Überschneidung sehen. Während manche z. B. behaupten, dass psychologische Erklärungen keine kausalen Erklärungen sind (vgl. Schon, 2005), meinen andere, zum Teil aufgrund der Annahme der multiplen Realisierbarkeit mentaler Phänomene, dass die Erklärungen von Neurowissenschaft und Psychologie unabhängig voneinander sind, da sie sich auf unterschiedlichen „Ebenen“ bewegen (vgl. Botterill & Carruthers, 1999, S. 43ff.; Crane, 2001). Die Unfähigkeit der Neurowissenschaft, auch nur einen einzigen Satz zu behaupten, welcher sich eindeutig auf den menschlichen Geist bezieht, könnte geradezu als Beleg für solche Alternativpositionen angeführt werden. Der Eliminativist bräuchte also auch unter den angenommenen idealen Bedingungen noch überzeugende Argumente gegen all diese Möglichkeiten. Anderenfalls wäre völlig unklar, ob die hypothetische Wissenschaft uns tatsächlich etwas über den Gegenstandsbereich sagt, auf welchen sich psychologische Prädikate beziehen. Und es fällt äußerst schwer, zu erkennen, wie der Eliminativismus die angesprochenen Alternativen ausschließen will, ohne gegen die Verwendung psychologischer Prädikate selbst zu argumentieren. Dann wäre sein Ziel aber, einen Widerspruch zu zeigen. Auf diese Möglichkeit möchte ich jetzt eingehen.

Diese zweite Möglichkeit der Elimination psychologischer Prädikate geht von einem Widerspruch gängiger Verwendungen dieser Prädikate mit reifen neurowissenschaftlichen Theorien

aus. Diese Variante scheint bei Eliminativisten weiter verbreitet zu sein (vgl. z. B. Churchland, 1981). Es sei also wieder angenommen, dass wir im Besitz einer reifen neurowissenschaftlichen Theorie sind, und dass diese Theorie gängigen Verwendungsweisen psychologischer Prädikate widerspricht. Es ergeben sich hier mindestens fünf gravierende Probleme:

Erstens setzt diese Position voraus, dass psychologische Prädikate theoretischer Natur sind. Es gibt aber gute Gründe, dies anzuzweifeln. Ist die Beschreibung eines Gesichts als schmerzverzerrt tatsächlich theoretischer als die Angabe des Kontraktionsgrades jedes Gesichtsmuskels? Es scheint Fälle zu geben, in denen wir keine direktere Beschreibung geben können, als durch die Verwendung psychologischer Prädikate (vgl. Bennett & Hacker, 2003, S. 366ff.). Auch scheinen psychologische Prädikate sich in wesentlichen Teilen so stark von naturwissenschaftlichen zu unterscheiden, dass ihre Klassifikation als theoretische zweifelhaft ist: Sie tragen starke normative Konnotationen. Sie folgen nicht dem Referenzfixierungsmechanismus, welchen man für naturwissenschaftliche Termini (*natural kind terms*) zu erwarten hätte. Und sie sind, anders als naturwissenschaftliche Termini, extrem kontextsensitiv (vgl. Sehon, 2005, S. 55ff.). Wenn psychologische Prädikate aber keine theoretischen Termini darstellen, dann kann auch ihre Elimination nicht so funktionieren wie bei Begriffen wie „Phlogiston“. Doch auch wenn wir dies, zum Zwecke des Arguments, unbeachtet lassen, ergeben sich weitere Probleme.

Angenommen, es ergäbe sich der prognostizierte Widerspruch, woher weiß der Eliminativist, dass die Verwendung psychologischer Prädikate nicht flexibel und plastisch genug ist, um sich den neuen Erkenntnissen anzupassen? Die Wechselwirkung zwischen begrifflichem Wandel und empirischem Erkenntnisfortschritt ist ein viel diskutiertes Thema; und Voraussagen über diese Wechselwirkung in einem historischen Einzelfall scheinen eine solide Theorie diesbezüglich vorauszusetzen. Warum hat der wissenschaftliche Fortschritt Begriffe wie „Kraft“, „Bewegung“ oder „Element“ durch unendliche Veränderungen geführt, während andere wie „Kristallsphäre“ oder „Lebensgeister“ eliminiert wurden? Eine allgemeine Theorie, welche diese Fragen beantwortet, wurde bisher von den Vertretern des Eliminativismus noch nicht vorgelegt. Statt dessen werden uns Zukunftsvisionen und Versprechungen angeboten. „[I]t offers us the promise of a better world, a world in which mental terms will have disappeared from our language, and in which materialism will be victorious“ (Popper & Eccles, 1984, S. 97).

Ein drittes Problem ergibt sich, wenn wir genauer nachfragen, wie sich denn der angenommene Widerspruch ergeben sollte. Ein Widerspruch besteht im Allgemeinen zwischen

zwei Sätzen bzw. Satzmengen. Zu zeigen, dass zwei Sätze, welche grundsätzlich unterschiedliches Vokabular gebrauchen, sich widersprechen, ist in der Regel nicht trivial. Abgesehen von Kontradiktionen, scheint zumindest ein gemeinsames Prädikat vonnöten, damit zwei Sätze sich widersprechen können. Nach den Voraussetzungen des Eliminativismus können wir aber aus unserer hypothetischen neurowissenschaftlichen Theorie keine Folgerungen ziehen, welche ein psychologisches Prädikat enthalten. So scheint nur die Möglichkeit zu bleiben, aus Sätzen, welche psychologische Prädikate verwenden, andere Sätze abzuleiten, welche keine solchen Prädikate verwenden und mit unserer neurowissenschaftlichen Theorie in Widerspruch stehen. Da es aber dasjenige, was psychologische Prädikate anscheinend bezeichnen, laut Eliminativismus nicht gibt, ist diese Aufgabe nicht einfach. Denn empirische Gesetzmäßigkeiten lassen sich bezüglich inexistenter Gegenstände nicht aufstellen. Der Eliminativist ist also bei dieser Ableitung physikalischer Sätze aus solchen, welche psychologische Prädikate verwenden, auf diejenigen Sätze angewiesen, welche der Verwender psychologischer Prädikate zugesteht (vgl. Schon, 2005, S. 27ff.). Damit bieten sich aber dem Verwender psychologischer Prädikate so viele Möglichkeiten, seine Behauptungen für den Eliminativisten unübersetzbar zu machen, dass die Aussichten auf eine solche Ableitung verschwindend gering erscheinen. Der Verwender psychologischer Prädikate könnte ja bei jeder Ableitung, welche zum Widerspruch führt, Modifikationen vornehmen.⁶⁷

Das vierte Problem dieser Position kann als klassischer Einwand gegen den Eliminativismus gelten. Dieser Einwand macht geltend, dass der Eliminativismus sich selbst widerspricht. Denn wenn der Eliminativismus recht hat, so gibt es nichts, was wir zurecht „glauben“, „meinen“ oder „denken“ nennen könnten. Dann scheint aber auch der Eliminativist nicht von sich behaupten zu können, an die Wahrheit des Eliminativismus zu *glauben*. Ja, er kann noch nicht einmal an unser Vermögen zu *denken* oder rational zu *schlussfolgern* appellieren. Der Eliminativist *will* uns offensichtlich auch nicht von dem *überzeugen*, was wir „seine *Meinung*“ nennen. Und wenn der Eliminativist offensichtlich nicht *glauben* kann, was er selbst sagt, so stellt sich die Frage, wie erst wir ihn zu nehmen haben (vgl. Pauen, 2001, S. 101ff.). Patricia Churchland will dieses Problem umgehen, indem sie behauptet, dass sich der Eliminativismus nur in der Sprache der zukünftigen Neurowissenschaft konsistent wird ausdrücken lassen (vgl. Churchland, 1986, S. 397). Und Paul Churchland meint, dass die Selbst-

⁶⁷ Dieses Problem taucht bei der Elimination von Begriffen wie „Phlogiston“ oder „Kristallsphäre“ nicht auf, da hier der gemeinsame Gegenstandsbereich zweier konkurrierender Theorien unstrittig ist und eine klare gemeinsame Beobachtungsbasis besteht (sofern man kein Kuhn'sches Bild dieser Eliminationen vertritt (vgl. Kuhn, 1976)). Im Fall der Psychologie scheint es aber so zu sein, dass unklar ist, ob nicht schon die Beschreibung der Beobachtung psychologischer Sachverhalte psychologische Prädikate voraussetzt.

widersprüchlichkeit des Eliminativismus nur durch eine *petitio principii* gezeigt werden könne, da das Argument eine Bedeutungstheorie voraussetze, welche psychologische Prädikate verwende (vgl. Churchland, 1981). Angesichts des Mangels einer Bedeutungstheorie, welche frei von psychologischen Prädikaten ist, und der Unbekanntheit der Sprache einer zukünftigen Neurowissenschaft betrachte ich aber das Argument gegen den Eliminativismus als schlagend. Eine Theorie, welche ihre eigene Unmöglichkeit behauptet, muss falsch sein; und über eine Theorie, die nicht artikuliert werden kann, lässt sich nicht sprechen.

Ein fünftes Problem ergibt sich aus Überlegungen zum 'Gewissheitsgefälle': Damit ein Widerspruch zwischen Neurowissenschaft und der Verwendung psychologischer Prädikate uns zur Elimination dieser Prädikate berechtigt, ist es notwendig, dass die fragliche neurowissenschaftliche Theorie größere Gewissheit beanspruchen kann als unsere Überzeugung, dass es mentale Phänomene gibt (vgl. Müller, 2000). Was würden wir von einer Physik halten, die uns lehrte, dass es keine sichtbaren Gegenstände gebe, oder dass schwere Körper nach oben „fielen“? Ich denke, wir würden einige Bedenken gegen eine solche Theorie vorbringen. Und sollte die Theorie nicht einige sehr gute Erklärungen für unsere hartnäckigen angeblichen Illusionen haben, wäre dies wohl ein guter Grund die Theorie abzulehnen. Eine Theorie, welche mit dem Anspruch auftritt, die Inexistenz mentaler Phänomene zu zeigen, scheint nicht nur guter Gründe, sondern überwältigender Evidenz zu bedürfen. Denn damit wäre ein enormer Teil unserer alltäglichen Gewissheiten in Frage gestellt. Darüber hinaus sollte diese Theorie einige sehr überzeugende Erklärungen dafür bereit halten, dass die Menschheit, soweit wir von der Vergangenheit wissen, immer psychologische Prädikate verwendet hat. Angesichts der Tatsache, dass sich bisher für diese Position einige Probleme aufzeigen ließen, überwältigende Gründe aber ausblieben, scheint der Eliminativismus nicht der Mühe einer weiteren Beschäftigung wert zu sein. Damit will ich die Möglichkeit, dass die Biopsychologie ohne psychologische Prädikate auskommen könnte, beiseitelassen und also annehmen, dass biopsychologische Theorien psychologische Prädikate enthalten müssen.

4.1.2. Wo können psychologische Prädikate in biopsychologischen Theorien auftauchen?

Ich habe dafür argumentiert, dass die Vertreter neurowissenschaftlicher Theorien menschlichen Verhaltens akzeptieren müssen, dass die Sprache der Biologie und die der Psychologie innerhalb der kognitiven Neurowissenschaft in 'Kontakt' treten. Im Folgenden soll gezeigt werden, dass dieser Kontakt auf drei verschiedene Weisen hergestellt werden kann.

Die Wissenschaft kann als der Versuch betrachtet werden uns Erklärungen zu bieten. Es wird allgemein anerkannt, dass naturwissenschaftliche Erklärungen idealerweise Gesetzmäßigkeiten beinhalten (vgl. Poser, 2001; Rosenberg, 2005, S. 21ff.).⁶⁸ Gesetze haben die logische Form von Konditionalen (vgl. Nagel, 1979, S. 47). Diese bestehen aus zwei Teilen: dem Antecedens und dem Consequens. Das Consequens eines Gesetzes stellt eine Beschreibung dessen dar, was durch die zugehörige Theorie erklärt werden soll. Das Antecedens beschreibt im Allgemeinen die Ursache⁶⁹ dessen, was im Consequens beschrieben wird. Psychologische Prädikate können in diesen beiden Teilen Verwendung finden. Eine dritte Möglichkeit der Verwendung psychologischer Prädikate ergibt sich bei der Verbindung unterschiedlicher Theorien. Solche Verbindungen verschiedener Erklärungen sind das Ziel reduktionistischer Ansätze. Seit der klassischen Arbeit von Ernest Nagel (1979) wird in diesem Zusammenhang oft von „Brückenprinzipien“ gesprochen. Die Grundidee des Reduktionismus besteht darin, dass der Gehalt einer Theorie durch eine andere Theorie erklärt werden kann. Die Brückenprinzipien stellen dabei die Möglichkeit einer solchen Erklärung sicher, indem sie das Vokabular der verschiedenen Theorien miteinander in Beziehung setzen (vgl. Bechtel & Hamilton, 2007).⁷⁰

Wie könnte das Vorkommen psychologischer Prädikate an den drei möglichen Stellen in einer biopsychologischen Theorie aussehen? Tauchen psychologische Prädikate nur im Consequens der Gesetze auf, so haben wir es mit einer Theorie zu tun, mittels derer das Zustandekommen psychologischer Phänomene durch biologische Begriffe erklärt werden soll. Als ein klassisches Beispiel kann hier die Erklärung der visuellen Wahrnehmung unter Verwendung von Begriffen wie „Rezeptor“, „Aktionspotenzial“, „*corpus geniculatum laterale*“ oder „Hypersäulen“ dienen (vgl. Goldstein, 2002). Werden psychologische Prädikate nur im Antecedens eines Gesetzes verwendet, so haben wir es mit einer Theorie zu tun, welche biologische Phänomene durch Verweis auf psychologische Phänomene erklären will. Solche Theorien werden von Neurowissenschaftlern im Allgemeinen als unbefriedigend und vorläufig betrachtet. Als Beispiele ließen sich hier die Erklärung der Atrophie von Dendriten

⁶⁸ Ich möchte hier nicht auf die ausgedehnte Diskussion über das Wesen wissenschaftlicher Erklärungen eingehen. Der Begriff „Gesetz“ soll hier so schwach verstanden werden, dass kaum mehr gemeint ist als „eine Regelmäßigkeit mit empirischem Erklärungswert“. Hiermit sollten die wenigsten Konzeptionen wissenschaftlicher Erklärung Probleme haben.

⁶⁹ Die Diskussion des Kausalbegriffs im Wiener Kreis will ich hier beiseite lassen. Wer den Ursachenbegriff als metaphysisch ablehnt, kann im Folgenden an seiner Stelle „Antezedens-Bedingung eines empirisch gehaltvollen Gesetzes“ lesen.

⁷⁰ Seitdem das deduktiv-nomologische Erklärungsschema und die positivistische Auffassung von Reduktion als unhaltbar gelten, hat sich eine größere Diskussion bezüglich der Fragen entwickelt, was eine Erklärung und was eine Reduktion sei. Daher versuche ich, in beiderlei Hinsicht eine liberale Position einzunehmen, sodass meine Ausführungen nicht von speziellen Annahmen bezüglich dieser Punkte abhängen.

bestimmter Pyramidenzellen des *Hippocampus* mittels des Begriffs „Stress“ anführen (vgl. Magarinos, McEwen, Flugge & Fuchs, 1996). Im Fall der Reduktion sind zwei Theorien zu beachten: die reduzierte und die reduzierende. Innerhalb der Biopsychologie wird die reduzierte Theorie eine psychologische Gesetzmäßigkeit enthalten, während die reduzierende Theorie eine biologische Gesetzmäßigkeit enthält. Die Brückenprinzipien verbinden dabei das Vokabular der beiden Theorien so miteinander, dass sich die psychologische durch die biologische Gesetzmäßigkeit erklären lässt. Es ist bekanntermaßen umstritten, ob sich hierfür überzeugende Beispiele aus der Neurowissenschaft anführen lassen. John Bickle behauptet, es sei gelungen, psychologische Gesetzmäßigkeiten bezüglich der Gedächtniskonsolidierung auf biologische Gesetzmäßigkeiten bezüglich des Phänomens der Langzeitpotenzierung zu reduzieren (vgl. Bickle, 2003, S. 43ff.). In Bickles Version einer Reduktion wird die Funktion der Brückenprinzipien durch mengentheoretische Abbildungsbeziehungen übernommen.⁷¹ Diese Abbildungsbeziehungen bringen die psychologische Gesetzmäßigkeit mit der biologischen in 'Kontakt'. Dabei wird eine Rückwirkung der Reduktion auf die beiden Theorien zugelassen.

4.2. Wo verwendet die SMH psychologische Prädikate?

Nachdem wir gesehen haben, dass es prinzipiell drei Orte gibt, an welchen psychologische Prädikate in biopsychologischen Theorien auftauchen können, stellt sich die Frage, wo solche Prädikate in der SMH verwendet werden. Diese Frage soll in diesem Abschnitt behandelt werden. Dabei wird sich herausstellen, dass die SMH psychologische Prädikate sowohl im Antecedens als auch im Consequens und in Brückenprinzipien verwendet.

4.2.1. Psychologische Prädikate im Consequens der Gesetze der SMH

Tauchen im Consequens der Gesetzmäßigkeit der SMH psychologische Prädikate auf? Um diese Frage zu beantworten, ist es notwendig, zuerst zu klären, worin dieses Consequens besteht. Was soll die SMH erklären? Hierzu möchte ich zunächst einige Zitate anführen, welche belegen sollen, dass das Erklärungsziel der SMH psychologische Phänomene sind und deshalb auch psychologische Prädikate zu ihrer Beschreibung verwendet werden.

„I struggled to understand the cognitive and neural machinery behind reasoning and decision making“ (Damasio, 2006, S. XXIV).

⁷¹ Für eine Kritik von Bickles Ansatz siehe Schon (2005, S. 33ff.).

„I propose that human reason depends on several brain systems, working in concert across many levels of neural organization [...]. The lower levels in the neural edifice of reason are the same ones that regulate the processing of emotions and feelings [...]“ (Damasio, 2006, S. XXIII).

„The brain systems necessary to experience the emotional feelings arising from social emotions constitute a critical part of the neural grounding of human values“ (Damasio, 2005, S. 55).

„The somatic marker hypothesis is based on the following main assumptions: (i) that human reasoning and decision making depend on many levels of neural operation“ (Bechara, Damasio & Damasio, 2000, S. 295).

„The central feature of the SMH is [...] that emotion-related signals assist cognitive processes even when they are non-conscious“ (Bechara et al., 2005, S. 159)

„The key idea of this hypothesis [SMH] is that decision-making is a process that is influenced by marker signals“ (Bechara & Damasio, 2005, S. 336).

Es ließen sich beliebig viele analoge Textstellen anführen (vgl. Bechara, Damasio & Damasio, 2003, S. 356; Damasio, 1996, S. 1413; Tranel, Bechara & Damasio, 1999, S. 1047). Ich denke, hiermit ist eine sichere Textbasis für die Schlussfolgerung gegeben, dass die SMH psychologische Phänomene erklären will und diese regelmäßig durch psychologische Prädikate beschreibt. Erklärt werden sollen Prozesse der Entscheidungsfindung (*decision-making*), der Überlegung (*reasoning*) und der Bewertung (*human values*). Da diese Prozesse *erklärt* werden sollen, kann gefolgert werden, dass sie in irgendeiner Form im Consequens eines Gesetzes der SMH benannt werden müssen. Weiter unten werde ich näher darauf eingehen, in welcher Form diese Prädikate hier auftauchen.

4.2.2. Psychologische Prädikate im Antecedens der Gesetze der SMH

Wie steht es um den Gebrauch psychologischer Prädikate im Antecedens der Gesetze der SMH? Wiederum will ich zunächst einige Textstellen wiedergeben, die zeigen sollen, dass eine sichere Textbasis für die Behauptung besteht, dass psychologische Prädikate auch im Antecedens von Gesetzen der SMH auftauchen:

„Emotion, feeling, and biological regulation all play a role in human reason“ (Damasio, 2006, S. XXIII).

„Feelings offer us a glimpse of what goes on in our flesh, as a momentary image of that flesh is juxtaposed to the image of other objects and situations; in so doing, feelings modify our comprehensive notion of those objects and situations. By dint of juxtaposition, body images give to other images a quality of goodness or badness, of pleasure or pain“ (Damasio, 2006, S. 159).

„You may think of it [somatic markers] as a system for automated qualification of predictions, which acts, whether you want it or not, to evaluate the extremely diverse scenarios of the anticipated future before you“ (Damasio, 2006, S. 174).

„The somatic marker hypothesis posits that deficits in emotional signaling (somatic states) lead to ... “ (Bar-On et al., 2003, S. 1790).

„The hypothesis attributes these [VMPFC impaired] patients’ inability to make advantageous decisions in real-life to a defect in emotional mechanism“ (Bechara & Damasio, 2005, S. 339).

„The somatic marker hypothesis provides a system-level neuroanatomical and cognitive framework for decision making and the influence on it by emotion“ (Bechara, Damasio & Damasio, 2000, S. 295).

Diesen Textbelegen ließen sich beliebig viele hinzufügen (vgl. z. B. Burns & Bechara, 2007, S. 265; Damasio, Tranel & Damasio, 1991, S. 220). Es wird in den angeführten Zitaten unverkennbar deutlich, dass die SMH psychischen Prozessen eine entscheidende kausale Rolle bei der Entscheidungsfindung zuschreibt. Es ist darüber hinaus offensichtlich, dass die SMH diese kausale Rolle von Gefühlen, Bewertungen und Erwartungen zur *Erklärung* von Entscheidungsprozessen nutzen will. In diesem Fall müssen aber auch im Antecedens der Gesetze der SMH Formulierungen auftauchen, welche sich auf die kausal relevanten emotionalen Prozesse beziehen. Es werden also sowohl im Consequens als auch im Antecedens der Gesetzmäßigkeit der SMH psychologische Prädikate verwendet.

4.2.3. Psychologische Prädikate in Brückenprinzipien

Ich kann jetzt auf einen Einwand eingehen, welchen der Leser vielleicht schon seit Längerem gegen meine Argumentation vorbringen will. Er lautet: „Die SMH muss sich in ihren Erklärungen in keiner Weise auf psychologische Phänomene beziehen, es genügt völlig, wenn sich die SMH auf die neurobiologischen Prozesse bezieht, welche diesen psychologischen Phänomenen zugrunde liegen“. Die Antwort auf diesen Einwand lautet: Es ist richtig, dass die

SMH sich nicht direkt auf psychologische Phänomene beziehen muss. Aber dennoch müssen ihre Vertreter annehmen, dass es Brückenprinzipien für den Übergang von einer biologischen Beschreibung zu einer psychologischen Beschreibung gibt. Damit beinhaltet die SMH eine biologische und eine psychologische Beschreibungsebene, welche sich wechselseitig in-einander überführen lassen. Was dies genau heißt, wird im Folgenden deutlich werden.

Es soll hier zunächst dargelegt werden, dass sich innerhalb der SMH Brückenprinzipien finden lassen, welche einen Übergang von biologischen zu psychologischen Beschreibungen erlauben, und umgekehrt. Der Terminus „Brückenprinzip“ soll dabei so weit verwendet werden, dass er in allen gängigen Konzeptionen von Reduktion (vgl. Sarkar, 1992) ein Äquivalent besitzt. Der Terminus ist der klassischen Konzeption des Reduktionismus entnommen. In strukturalistischen Konzeptionen findet er ein Äquivalent in Abbildungsbeziehungen (Bickle, 2003, S. 21ff.). In mechanistischen Ansätzen wäre ein Äquivalent in einer Erklärung des Zusammenwirkens der Teilprozesse eines Mechanismus, um die Gesamtfunktion des Mechanismus zu verstehen, zu sehen (vgl. Bechtel & Hamilton, 2007, S. 406f.). Ich will deshalb die Kernfunktion von Brückenprinzipien in der Vermittlung zwischen zwei Ebenen oder Theorien betrachten, wie auch immer diese Vermittlung charakterisiert werden sollte.

Um die These zu untermauern, dass sich solche Möglichkeiten der Vermittlung zwischen unterschiedlichem Vokabular innerhalb der SMH finden lassen, will ich wieder einige Textbelege anführen:

„There is abundant evidence to indicate that the ventromedial sector of the prefrontal cortices (VMPFC) is a region necessary for the triggering of social emotions“ (Damasio, 2005, S. 55).

„Thus the significance of this neural arrangement is that [...] somatic states [...] can gain access to cortical and subcortical neurons subserving cognition. Thus, depending on their strength, they have the capacity to modify and influence cognition, especially decision making“ (Burns & Bechara, 2007, S. 268).

„Decision making is ultimately dependent on the integrity of the VMPC [VMPFC]“ (Weller et al., 2007, S. 959).

„The ensemble of all these enacted responses in the body proper and in the brain constitutes an emotion. The ensemble of the signals as mapped in somatosensory regions of the brain itself provide the essential ingredients for what is ultimately perceived as a feeling“ (Bechara & Damasio, 2005, S. 339).

„[T]he images of the scenario are ‘judged’ and marked by the juxtaposed images of the somatic state“ (Damasio, 1996, S. 1415).

„Both the amygdala and OFC [orbitofrontal cortex] are known to receive rapid, multimodal sensory input; both are bidirectionally connected and are known to function together in evaluating the value of stimuli“ (Hsu, Bhatt, Adolphs, Tranel & Camerer, 2005, S. 1681).

„Ventral and medial sectors of the prefrontal cortex (PFC) have been implicated in guiding behavior on the basis of the motivational value of the choices available“ (Oya, Adolphs, Kawasaki, Bechara, Damasio & Howard, 2005, S. 8351).

„[T]he neural systems that support emotional and social intelligence appear to overlap with the neural systems subserving somatic state activation and personal judgment in decision-making“ (Bar-On et al., 2003, S. 1797).

„[S]omatic marker hypothesis [...] focuses [...] on the neurobiological mechanisms underlying feelings and their elicitation by ‘thoughts’“ (Bechara & Damasio, 2005, S. 359).

An all den angeführten Textstellen wird versucht, eine neurobiologische und eine psychologische Beschreibung zu verbinden. Im Englischen werden dafür oft folgende Ausdrücke verwendet: „ Φ is necessary for Ψ “, „ Ψ depends on Φ “, „ Φ is perceived as Ψ “, „ Φ functions in Ψ “, „ Φ is the neurobiological mechanism underlying Ψ “ und „ Φ subserves Ψ “, wobei hier Φ für ein neurobiologisches und Ψ für ein psychologisches Phänomen stehen soll. Obwohl es der Vagheit dieser Formulierungen deutlich anzumerken ist, dass sie versuchen, starke metaphysische Thesen bezüglich des Leib-Seele-Problems zu vermeiden, ist offensichtlich, dass die angeführten Formulierungen die Funktion von Brückenprinzipien übernehmen sollen. Sie ermöglichen den Übergang zwischen biologischen und psychologischen Beschreibungen. Auf der neurobiologischen Seite finden wir Termini wie „VMPFC“, „Amygdala“, „somatosensorische Kortizes“, „somatische Veränderungen“ und „motorischer Kortex“. Auf der psychologischen Seite stehen Begriffe wie „Entscheidungsfindung“, „Emotion“, „Gefühl“, „Bewertung“ und „Assoziation“. Die Brückenprinzipien können als Abbildung dieser beiden Begriffssysteme aufeinander verstanden werden. Im Folgenden will ich der Frage nachgehen, wie diese Abbildungsbeziehungen innerhalb der SMH genau aussehen.

4.3. Die reduktionistische Struktur der SMH

Wir haben gesehen, dass die SMH psychologische Prädikate sowohl im Consequens als auch im Antecedens ihrer Gesetze und in Brückenprinzipien verwendet. Damit scheint es eine gut

begründete Vermutung zu sein, dass die SMH eine reduktionistische Struktur aufweist. Diese Vermutung lässt sich wie folgt spezifizieren: Die SMH postuliert kausale Gesetzmäßigkeiten, welche sich sowohl auf psychologischer als auch auf biologischer Ebene beschreiben lassen. Sie zielt dabei darauf ab, die auf psychologischer Ebene beschriebene Gesetzmäßigkeit durch jene auf biologischer Ebene beschriebene Gesetzmäßigkeit zu erklären. Ich möchte diese Vermutung erhärten, in dem ich näher auf die Brückenprinzipien der SMH eingehe.

4.3.1. Die Brückenprinzipien der SMH

Für folgende fünf psychologische Begriffe möchte ich Brückenprinzipien angeben, welche durch die SMH impliziert werden: „Emotion“, „Gefühl“, „erwartetes Handlungsergebnis“, „etwas emotional bewerten“ und „Auswahl einer Handlungsoption“.⁷²

4.3.1.1. „Emotion“

Im Fall von Emotionen ist ein Brückenprinzip besonders leicht zu finden, da die Vertreter der SMH hier eine einfache Identitätsbeziehung postulieren.

„An emotion is defined as a collection of changes in body and brain states triggered by dedicated brain systems that responds to specific contents of one’s perceptions, actual or recalled, relative to a particular object or event“ (Bechara & Damasio, 2005, S. 339).

„I use emotion for a collection of changes occurring in both brain and body, usually prompted by a particular mental content“ (Damasio, 2006, S. 270).

Die SMH nimmt also an, dass Emotionen identisch sind mit einem Ensemble von viszeralen, endokrinen, vaskulären und neuronalen Prozessen (vgl. Tranel, Bechara & Damasio, 1999, S. 1049). In philosophischer Terminologie spricht man in solchen Fällen von Typen-Identität. Ein Typ psychischer Phänomene wird dabei mit einem Typ materieller Phänomene identifiziert.⁷³ Damit ist natürlich eine denkbar einfache Möglichkeit des Übergangs von psychologischen zu biologischen Beschreibungen gegeben: Ist bei einer Person eine Emotion vorhanden, so lassen sich die entsprechenden somatischen Prozesse bei dieser Person finden.

⁷² Das Auffinden solcher Brückenprinzipien ist ein traditionelles Anliegen der kognitiven Neurowissenschaft: „We will ultimately need bridging hypotheses to link the psychological and the neural, theories which specify the potential physiology of psychological processes“ (Lloyd, 1989, S. 8).

⁷³ Es ist fraglich, ob innerhalb der SMH mit dem Ausdruck „Emotion“ tatsächlich ein psychisches Phänomen bezeichnet werden soll. Sollte dies nicht der Fall sein, kann auch nicht von einer Identität psychischer und physischer Phänomene gesprochen werden. Für die weitere Analyse ist dieses Problem allerdings unerheblich, da der Ausdruck „Gefühl“ innerhalb der SMH eindeutig zur Bezeichnung eines psychologischen Phänomens verwendet wird. Das, was weiter unten „emotionale Bewertung“ genannt wird, bezeichnet tatsächlich die angebliche Bewertung von Vorstellungsinhalten durch *Gefühle*. „Emotional“ ist diese Bewertung, da Gefühle, laut SMH, in der Regel auf Emotionen beruhen.

4.3.1.2. „Gefühl“

Da die Vertreter der SMH explizit jeden Rekurs auf phänomenales Bewusstsein und Erleben aus dem Begriff der Emotion ausschließen, müssen diese Phänomene innerhalb der SMH mit anderen Termini beschrieben werden. Ein „Gefühl“ ist, laut SMH, das bewusste Erleben einer Emotion (vgl. Damasio, 2006, S. 133).

„The constituted somatic state is then signalled to cortical and subcortical somatosensory structures, whence it triggers either a covert process modifying appetitive/aversive behavior or a conscious perception in the form of a feeling“ (Tranel, Bechara & Damasio, 1999, S. 1048).

„[N]eural representations of feelings are thought to involve the insula, surrounding parietal cortices, and the cingulate“ (Bechara, 2002, S. 1673).

„Once somatic states from primary inducers are induced, signals from them are relayed to the brain. Representations of these signals can remain covert at the level of the brainstem, or can reach the parietal cortices (insular/SI, SII) and posterior cingulate cortices and be perceived as a feeling“ (Bar-On et al., 2003, S. 1791).

Die SMH postuliert also, dass Gefühle die subjektive Wahrnehmung der neuronalen Repräsentation von somatischen Prozessen, welche Emotionen darstellen, sind (vgl. Damasio, 2006, S. 143). Diese Repräsentationen sollen in parietalen und cingulären Kortizes aufzufinden sein. Im Fall von Gefühlen will die SMH anscheinend keine starken Identitätsthese vertreten. Es wird lediglich behauptet, dass gewisse neurobiologische Prozesse „als Gefühl wahrgenommen werden“. Diese Formulierung lässt Raum für eine eigenschaftsdualistische Auffassung. Es könnte demnach der Fall sein, dass die Eigenschaft des neuronalen Prozesses ‚als-Gefühl-wahrgenommen-zu-werden‘ sich nicht auf biologische Fakten reduzieren ließe. Nichtsdestoweniger scheint impliziert zu sein, dass die Eigenschaft ‚als-Gefühl-wahrgenommen-zu-werden‘ eine Eigenschaft der fraglichen neuronalen Repräsentationen ist. ‚Ein-Gefühl-zu-sein‘ ist demnach eine Eigenschaft eines neurobiologischen Prozesses. Hierdurch ergibt sich die gesuchte Übergangsmöglichkeit von biologischen zu psychologischen Beschreibungen. Fühlt eine Person ein Gefühl, so müssen auch die neurobiologischen Prozesse vorliegen, deren Eigenschaft das Gefühl ist. Und wo somatische Veränderungen, welche eine Emotion darstellen, in frühen somatosensorischen Kortizes repräsentiert werden, da muss auch ein Gefühl wahrgenommen werden. Es gibt also innerhalb der SMH klare Brückenprinzipien für den psychologischen Begriff „Gefühl“.

4.3.1.3. „Erwartetes Handlungsergebnis“

Das, was ich hier mit „erwartetes Handlungsergebnis“ ansprechen will, wird von den Vertretern der SMH oft als „*predicted future outcomes of certain scenarios*“ (vgl. Damasio, 2006, S. 174; Tranel, Bechara & Damasio, 1999, S. 1049) bezeichnet. Gemeint ist die Erwartung, welche eine Person bezüglich der voraussichtlichen Folgen des Ergreifens einer Handlungsoption hegt. Da Erwartungen für die Vertreter der SMH eine Klasse bewusster Kognitionen bzw. Gedanken darstellen, diese aber angeblich aus Vorstellungsbildern aufgebaut sind, so müssen Erwartungen ebenfalls wesentlich aus Vorstellungsbildern bestehen. Auch dieser Punkt sei mit einigen Zitaten belegt:

„The factual knowledge required for reasoning and decision making comes to the mind in the form of images. [...] As the planning process unfolded you were forming images of objects and movements [...]. Images of something that has not yet happened and that may in fact never come to pass are no different in nature from the images you hold of something that already has happened“ (Damasio, 2006, S. 96f.).

„[C]onscious, overtly cognitive operations depend on sensory images based on the activity of early sensory cortices“ (Bechara, Damasio & Damasio, 2000, S. 295).

„[W]hen a situation of a given class recurs, factual knowledge pertaining to the situation – possible options of acting, outcomes of such actions immediately and at long term – is evoked in sensory images based on the appropriate sensory cortices“ (Damasio, 1996, S. 1415).

„Imagine yourself as the owner of a large business, faced with the prospects of meeting or not with a possible client who can bring valuable business but also happens to be the arch-enemy of your best friend, and proceeding or not with a particular deal. The brain of a normal, intelligent, and educated adult reacts to the situation by rapidly creating scenarios of possible response options *and* related outcomes“ (Damasio, 2006, S. 170).

Die SMH postuliert also, dass, wenn eine Person gewisse Ereignisse als Folge ihres eigenen möglichen Handelns betrachtet, sie Vorstellungsbilder erlebt, welche die erwarteten Ereignisse abbilden. Diese Vorstellungsbilder gründen sich angeblich auf neuronale Repräsentationen der Abbilder in frühen sensorischen Kortizes. Laut SMH werden beim Überdenken verschiedener Handlungsoptionen die zugehörigen Repräsentationen in frühen

sensorischen Kortizes in synchronisierter Weise aktiviert⁷⁴ (vgl. Damasio, Tranel & Damasio, 1991, S. 220). Eine Person erwartet demnach genau dann gewisse Ereignisse als Folgen ihres möglichen Handelns, wenn sich in der Situation des Überdenkens der entsprechenden Handlungsoption neuronale Repräsentationen dieser Ereignisse in frühen sensorischen Kortizes manifestieren. Diese Formulierung lässt sich noch präzisieren: Eine Person erwartet Ergebnis E als Folge ihrer Handlung H genau dann, wenn eine neuronale Repräsentation von E die kausale Folge einer neuronalen Repräsentation von H ist und dieser kausale Prozess unter geeigneten Randbedingungen (Wahrnehmungs- und Situationsvariablen etc.) abläuft. Damit ist auch in diesem Fall ein Brückenprinzip gefunden. Es kann ein klarer Übergang von der psychologischen Beschreibung „Handlungsergebnis X wird erwartet“ zu der biologischen Beschreibung „Aktivitätsmuster Y verursacht Aktivitätsmuster Z in frühen sensorischen Kortizes“ gemacht werden.

4.3.1.4. „Etwas emotional bewerten“

Auch für die psychologische Beschreibung „etwas emotional bewerten“ lässt sich innerhalb der SMH ein neurobiologisches Äquivalent finden. Laut SMH wird ein Vorstellungsinhalt bewertet, indem er mit einem Gefühl verbunden wird. Übersetzt in die neurobiologische Sprache der SMH heißt dies: Ein Geschehen wird bewertet, indem die neuronale Repräsentation dieses Geschehens in den frühen sensorischen Kortizes mit der somatosensorischen Repräsentation einer somatischen Veränderung, welche eine Emotion darstellt, verbunden wird. Dabei wird für die Bezeichnung dieser Verbindung von den Vertretern der SMH meist der Terminus „Juxtaposition“ gewählt.

„Because the feeling is about the body, I gave the phenomenon the technical term somatic state [...]; and because it ‘marks’ an image, I called it a marker“ (Damasio, 2006, S. 173).

„[W]hen the somatosensory image which defines a certain emotional response is juxtaposed to the images which describe a related scenario of future outcome, and which triggered the emotional response via the ventromedial linkage, the somatosensory pattern marks the scenario as good or bad“ (Bechara, Damasio & Damasio, 2000, S. 297).

„[T]he somatic marker hypothesis, which posits that overt or covert somatic states juxtapose a value mark in the processing of a cognitive scenario, and that the qualifying value mark helps endorse or reject an option for acting“ (Bechara et al., 1996, S. 215).

⁷⁴ Die Synchronizität soll hier als Lösung des sogenannten ‘Bindungsproblems’ in Anschlag gebracht werden (vgl. Damasio, 2006, S. 95).

„[T]he images of the scenario are ‘judged’ and marked by the juxtaposed images of the somatic state“ (Damasio, 1996, S. 1415).

„[W]hen the somatosensory pattern image is juxtaposed both to the images prompting the somatic state and those depicting potential outcomes, the somatosensory pattern *marks* outcomes as good or bad“ (Tranel, Bechara & Damasio, 1999, S. 1048).

Laut SMH wird also ein Geschehen als „gut“ oder „schlecht“ bewertet oder markiert, wenn seine neuronale Repräsentation mit der somatosensorischen Repräsentation einer Emotion (d. i. ein Gefühl im Sinne der SMH) in Juxtaposition steht. Damit haben wir eine deutliche Übergangsmöglichkeit von psychologischen zu biologischen Beschreibungen gefunden.

4.3.1.5. „Auswahl einer Handlungsoption“

Immer wieder wird das zentrale Thema der SMH als „*decision-making*“ bezeichnet (vgl. z. B. Bechara, Damasio & Damasio, 2003; Tranel, Bechara & Damasio, 1999). Insofern kann der gesamte Prozess, welcher den Gegenstand der SMH darstellt, als „Auswahl einer Handlungsoption“ bezeichnet werden. Ich möchte hier aber ein Brückenprinzip für einen enger gefassten Begriff suchen. Die Frage soll dabei lauten: Wie kann das Resultat des Mechanismus beschrieben werden, welchen die SMH postuliert? Das Resultat des postulierten Mechanismus ist die Wahl eines Verhaltens. „*It is perhaps accurate to say that the purpose of reasoning is deciding and that the essence of deciding is selecting a response option, that is, choosing a nonverbal action, a word, a sentence, or some combination thereof*“ (Damasio, 2006, S. 165). Die Vertreter der SMH spezifizieren die Bedeutung von „*decision-making*“ auch als die Selektion eines Antwortverhaltens: „*decision making, [is] defined as the ability to select an advantageous response from an array of available options*“ (Bechara, Damasio, Tranel & Anderson, 1998, S. 428). Die SMH nimmt an, dass eine Handlungsoption genau dann gewählt wurde, wenn eine entsprechende Aktivität in motorischen Arealen des zentralen Nervensystems auftaucht.

„[S]omatic states signals influence activity in regions concerned with motor responses and behavioral action (e.g. striatum and anterior cingulate/supplementary motor area (SMA)). They interfere with response selection and thus render the occurrence of a given behavior more likely or less likely. [...] Thus changes in neurotransmitter release induced by somatic state signals modulate the synaptic activities of telencephalic neurons subserving behavior and cognition, thereby providing a mechanism for somatic states to exert a bias-

ing effect on behaviors (e.g., selection of a response over another)“ (Bechara & Damasio, 2005, S. 343).

„A [...] reason why the prefrontal cortices are ideally suited for participation in reasoning and deciding is that they are directly connected to every avenue of motor and chemical response available to the brain. The dorsolateral and upper medial sectors can activate the primary motor cortex (M1), the supplementary motor area (M2), and the third motor area (M3). The subcortical motor machinery of the basal ganglia is equally accessible to the prefrontal cortices. Last but not least [...] the ventromedial prefrontal cortices send signals to autonomic nervous system effectors“ (Damasio, 2006, S. 183).

„Thus over the course of pondering a decision, positive and negative somatic markers that are strong are reinforced, while weak ones are eliminated. [...] Ultimately, a winner takes all; an overall, more dominant, somatic state emerges [...], which then provides signals to the brain that modulate activity in neural structures involved in biasing decisions. This ‘winner takes all’ view is consistent with the conception by Strack and Deutsch of competition between motor schemata (Strack & Deutsch, 2004)“ (Burns & Bechara, 2007, S. 265).

In der Arbeit von Strack und Deutsch, auf welche die Vertreter der SMH hinweisen, wird die Ansicht vertreten, dass es verschiedene Verhaltensschemata gäbe, welche unterschiedlich stark aktiviert werden könnten. „*An important assumption is that more than one behavioral schemata can receive activation at a time. However, activation must exceed a given threshold to result in overt behavior*“ (Strack & Deutsch, 2004, S. 229). Es wird angeblich jenes Verhalten ausgeführt, dessen Schema-Aktivität zuerst den Schwellenwert überschreitet. Demnach wäre die Wahl einer Handlungsoption dann abgeschlossen, wenn das Verhaltensschema einer Handlungsoption seine Schwellenaktivität erreicht hat. Die Aktivität der Verhaltensschemata muss, nach den angeführten Zitaten, eine neuronale Aktivität in motorischen Arealen zugrunde liegen. Demzufolge stellt Folgendes ein Brückenprinzip dar: Eine Person wählt genau dann eine Handlungsoption aus, wenn die neuronale Repräsentation des zugehörigen Verhaltensschemas in den motorischen Arealen seine Aktivitätsschwelle erreicht.

Hiermit sind für die fünf angeführten psychologischen Prädikate Brückenprinzipien gefunden, welche den Übergang zu einer neurobiologischen Beschreibung ermöglichen.

4.3.2. Es muss einen genuin psychologischen Gehalt der SMH geben

Wir haben gesehen, dass die SMH psychologische Prädikate sowohl im Consequens als auch im Antecedens von Gesetzen und innerhalb von Brückenprinzipien verwendet. Die Vermutung, dass die SMH eine reduktive Struktur aufweist, in welcher psychologische Gesetzmäßigkeiten durch neurobiologische Gesetzmäßigkeiten erklärt werden sollen (indem erstere auf letztere reduziert werde), kann damit als gut begründet gelten.

Trifft diese Vermutung zu, so muss sich aber ein Inhalt der SMH isolieren lassen, welcher in psychologischen Begriffen allein beschrieben werden kann. Die letzte Möglichkeit, diese Schlussfolgerung zu blockieren, besteht darin, zu behaupten, die SMH verwende in der Formulierung ihrer postulierten Gesetzmäßigkeiten ein unentwirrbares Amalgam psychologischer und neurobiologischer Begriffe. Die Brückenprinzipien wären in diesem Fall lediglich Behauptungen bezüglich der Äquivalenz bestimmter Beschreibungen dieses Amalgams. Gegen diese letzte Möglichkeit zu leugnen, dass die SMH einen genuin psychologischen Gehalt besitzen muss, will ich nun argumentieren.

Der Kern meines Arguments wird sein, dass sich psychologische und neurobiologische Beschreibungen grundlegend unterscheiden. Aufgrund dieses Unterschieds bewegen sich diese Beschreibungen gleichsam auf zwei verschiedenen 'Ebenen', welche sich nicht beliebig 'mischen' lassen. Der entscheidende Unterschied psychologischer und neurobiologischer Prädikate besteht darin, dass psychologische Prädikate sich auf einen lebenden Organismus *als ganzen* beziehen, während neurobiologische Prädikate sich auf *Teile* eines Organismus beziehen.

„The brain is not a logically appropriate subject for psychological predicates. [...] Mereology is the logic of part/whole relations. The neuroscientists' mistake of ascribing to constituent parts of an animal attributes that logically apply only to the whole animal we shall call 'the mereological fallacy' in neuroscience“ (Bennett & Hacker, 2003, S. 72f.).⁷⁵

Diese Unterscheidung zwischen einer Beschreibungsebene auf dem Niveau einer ganzen Person und einer Beschreibungsebene auf dem Niveau von Teilen einer Person ist für das Verhältnis von Neurobiologie und Psychologie ungemein wichtig (vgl. Bechtel, 1994; Bechtel & Hamilton, 2007; Block, 1995; Dennett, 1986, S. 90ff.).

⁷⁵ Insbesondere Max R. Bennett und Peter M. S. Hacker haben ausführlich dargelegt, wie die Vertreter der SMH, in unterschiedlichen Zusammenhängen, diesen mereologischen Fehler begehen (vgl. Bennett & Hacker, 2003, S. 199ff.; Bennett & Hacker, 2005; Hacker, 2004).

Eine entscheidende Beobachtung ist, dass die Menschheit psychologische Prädikate anwenden konnte, lange bevor sie von den Erkenntnissen der Neurowissenschaft erfahren hat. Dies lässt darauf schließen, dass es für die richtige Anwendung psychologischer Begriffe Kriterien geben muss, welche unabhängig sind von neurobiologischen Fakten. Anderenfalls hätten wir, bevor wir die relevanten neurobiologischen Fakten erführen, überhaupt keine Kriterien für die richtige Anwendung psychologischer Begriffe. Dies ist aber ganz eindeutig nicht der Fall. Wir wissen im Allgemeinen, wann wir von jemandem sagen können, er wisse etwas, er wünsche etwas, er glaube etwas.

„Das Denken in den Begriffen physiologischer Vorgänge ist für die Klarstellung der begrifflichen Probleme in der Psychologie höchst gefährlich. Das Denken in physiologischen Hypothesen spiegelt uns manchmal falsche Schwierigkeiten, manchmal falsche Lösungen vor. Die beste Kur dagegen ist der Gedanke, daß ich garnicht weiß, ob die Menschen, die ich kenne, wirklich ein Nervensystem haben“ (BPP I, § 1063, Wittgenstein, 1984, S. 192).

Wir können psychologische Begriffe in Abwesenheit aller Kenntnisse über das zentrale Nervensystem anwenden. Und diese Anwendungen können dennoch richtig oder falsch, berechtigt oder unberechtigt sein. Sogar die *Beurteilung* der Richtigkeit und Falschheit der Anwendung psychologischer Begriffe hängt nicht von unserem biologischen Wissen ab. „Die Begriffe der Psychologie sind eben Begriffe des Alltags“ (BPP II, § 62, Wittgenstein, 1984, S. 230). Die Kriterien der Anwendung neurobiologischer Begriffe dagegen sind aufs Engste mit unserem Wissen bezüglich des Nervensystems verbunden. Ob ein bestimmter neurobiologischer Begriff gerechtfertigterweise auf einen bestimmten Gegenstand angewendet wird, bemisst sich nach Tatsachen, welche uns im Alltag überhaupt nicht zugänglich sind. Dies ist kein Zufall. So wie wir in der Chemie das alltägliche Phänomen der Auflösung von Kochsalz in Wasser durch den Rückgriff auf unser naturwissenschaftliches Wissen bezüglich der molekularen Struktur von Kochsalz und Wasser erklären wollen; so will die Neurowissenschaft alltägliche psychologische Phänomene durch den Rückgriff auf naturwissenschaftliche Theorien erklären. Es kann nicht verwundern, dass die Beschreibung dessen, was erklärt werden soll, sich nicht nach Kriterien bemessen kann, welche erst durch die gewünschte Erklärung zugänglich werden. Anderenfalls könnten wir überhaupt nicht nach einer Erklärung verlangen, welche wir noch nicht haben, denn die Beschreibung des zu Erklärenden würde die Erklärung voraussetzen. Wir müssen z. B. Fälle der Auflösung von Kochsalz in Wasser identifizieren können, um nach einer Erklärung dafür zu verlangen. Und so müssen wir auch

in der Lage sein, psychische Phänomene unabhängig von neurobiologischem Wissen identifizieren zu können, um nach einer neurobiologischen Erklärung für diese Phänomene zu fragen. Hierzu bedürfen wir bestimmter Kriterien. „Ein ‘innerer Vorgang’ bedarf äußerer Kriterien“ (PU, § 580, Wittgenstein, 1997b, S. 455).

Ein Einwand gegen diese These könnte lauten, dass wir, wenn wir hinreichend sichere neurobiologische Theorien besitzen, neurobiologische Kriterien für die Verwendung psychologischer Prädikate einführen könnten. Wir könnten z. B. eine Aktivität der Amygdala als Kriterium für das Vorliegen von Angst verwenden. Dann könnten wir etwa Folgendes sagen: „Wir wissen, dass Person *P* Angst hat, denn wir haben bei ihr eine starke Aktivität ihrer Amygdala gefunden“. Um diesen möglichen Einwand zu entkräften, reicht es hin, sich zu vergegenwärtigen, dass dieses neue Kriterium (Amygdala-Aktivität) seinerseits von anderen Kriterien abhängt. Denn um dieses Kriterium für die Anwendung des Begriffs „Angst“ gerechtfertigterweise zu verwenden, ist eine gut bestätigte Theorie erforderlich, welche Aussagen über die Korrelation von Angst und Amygdala-Aktivität macht. Um eine solche Theorie zu überprüfen, bedürfen wir aber anderer Kriterien für die Anwendung des Begriffs „Angst“. Sollten diese Kriterien nicht letztlich an unseren alltagssprachlichen Kriterien für diesen Begriff orientiert sein, so gibt es keine Möglichkeit, radikale Abweichungen zwischen der neurobiologischen und der alltäglichen Zuschreibung von Angst zu verhindern. Ein Zusammenfallen dieser Zuschreibungen wäre in der Tat reiner Zufall. Dann wäre es aber fraglich, ob sich diese Zuschreibungen überhaupt auf den gleichen Gegenstand beziehen. Scott R. Sehon illustriert den gleichen Punkt am Beispiel einer Überzeugung:

„In particular contexts we will associate various behaviors with a belief that *p*, if a neurophysiologist were to claim that beliefs that *p* actually had rather different sorts of behavioural connections, we would have no reason to believe that the neurophysiologist was talking about the same thing as we were“ (Sehon, 2005, S. 68).

Unsere Zuschreibungen psychologischer Prädikate sind vom beobachtbaren Verhalten eines ganzen Organismus abhängig, die Zuschreibung neurobiologischer Prädikate hingegen nicht (vgl. Bennett & Hacker, 2003, S. 81ff.).⁷⁶ Damit ergibt sich eine einfache Möglichkeit, psychologische und neurobiologische Beschreibungen zu entwirren. Psychologische Beschreibungen beziehen sich auf Personen *als ganze*, neurobiologische Beschreibungen be-

⁷⁶ Im Fall neurobiologischer Prädikate dreht sich das Abhängigkeitsverhältnis um. Wenn es Kriterien für die Zuschreibung neurobiologischer Prädikate gibt, welche auf dem beobachtbaren Verhalten des gesamten Organismus beruhen, so nur deshalb, weil dieser Zusammenhang zuvor empirisch entdeckt wurde.

ziehen sich auf *Teile* von Lebewesen. Die Phänomene, welche die SMH erklären will, können also in eine Beschreibung gebracht werden, welche keine neurobiologischen Termini enthält. Dies kann schlicht dadurch geschehen, dass die Phänomene auf der 'Ebene' des gesamten Organismus beschrieben werden (vgl. Bennett & Hacker, 2003, S. 388ff.). Vernachlässigt man sowohl auf Seiten des Antecedens als auch des Consequens alle Prädikate, deren Zuschreibung auf neurobiologischem Wissen beruht, und behält ausschließlich jene bei, deren Zuschreibung letztlich auf beobachtbarem Verhalten beruht, so ergibt sich eine genuin psychologische Gesetzmäßigkeit. Da die SMH, wie oben gezeigt wurde, sowohl im Antecedens als auch im Consequens psychologische Prädikate verwendet, ist gewährleistet, dass beiden Teilen der Gesetzmäßigkeit eine psychologische Beschreibung entspricht. Sollte sich für einen Teil der Gesetzmäßigkeit der SMH keine Beschreibung in psychologischen Termini finden lassen, so stellt der resultierende genuin psychologische Zusammenhang zwar keine kausal hinreichende, aber dennoch eine kausal notwendige Verbindung dar. Hiermit ist gezeigt, dass die SMH einen genuin psychologischen Gehalt haben muss, welchen zumindest die kausale Notwendigkeit bestimmter, mittels psychologischer Prädikate zu beschreibender Phänomene für andere behauptet.

4.4. Welches ist der psychologische Gehalt der SMH?

Um den genuin psychologischen Gehalt der SMH zu bestimmen, will ich zunächst eine klare Formulierung der neurobiologischen Gesetzmäßigkeit angeben, welche die SMH postuliert. Danach soll diese neurobiologische Gesetzmäßigkeit, mittels der oben gefundenen Brückenprinzipien, in eine psychologische Gesetzmäßigkeit 'übersetzt' werden. Diese Gesetzmäßigkeit stellt den gesuchten Gehalt dar. Wie bereits angekündigt, wird sich der Satz (PG) als diese Gesetzmäßigkeit erweisen.

4.4.1. Das neurobiologische Gesetz der SMH

Welche Gesetzmäßigkeit postuliert die SMH? Um diese Frage zu klären, ist es hilfreich, das Consequens und das Antecedens der postulierten Gesetzmäßigkeit nacheinander zu betrachten.

Die SMH will die Ursachen eines Entscheidungsverhaltens angeben. Also muss im Consequens des postulierten Kausalgesetzes die Beschreibung einer Entscheidung auftauchen. Wie oben dargelegt,⁷⁷ betrachtet die SMH eine Entscheidung dann als getroffen, wenn ein

⁷⁷ s. Kapitel 4.3.1.5.

bestimmtes motorisches Schema seinen Schwellenwert erreicht (vgl. Burns & Bechara, 2007, S. 265). Dies gilt es also kausal zu erklären. Daher muss eine Beschreibung des Erreichens des Schwellenwerts eines motorischen Schemas im Consequens des gesuchten Gesetzes auftauchen.

Welches sind, gemäß der SMH, die Ursachen des genannten Vorgangs? Insbesondere sollen die somatischen Marker kausal relevant dafür sein, welches motorische Schema seinen Schwellenwert erreicht (vgl. Bechara & Damasio, 2005, S. 343). Somatische Marker sind aber nichts anderes als Repräsentationen somatischer Veränderungen in somatosensorischen Kortizes (vgl. Tranel, Bechara & Damasio, 1999, S. 1048). Um wirksam zu werden, müssen die somatischen Marker allerdings in Juxtaposition mit den Repräsentationen des zugehörigen Handlungs-Ergebnis-Szenarios gebracht werden (vgl. Bechara, Damasio & Damasio, 2000, S. 297). Diese Repräsentationen befinden sich in frühen sensorischen Kortizes und lösen die zugehörigen somatischen Marker mittels einer Assoziation im VMPFC aus. Im Antecedens des gesuchten Gesetzes muss also eine Bezugnahme auf Repräsentationen somatischer Veränderungen, auf Repräsentationen von Handlungs-Ergebnis-Szenarien und auf die Juxtaposition dieser beiden auftauchen. Das gesuchte neurobiologische Gesetz lässt sich nun angeben:

- (NBG) Wenn die neuronale Repräsentation eines Handlungs-Ergebnis-Szenarios (in den frühen sensorischen Kortizes) mit einer positiveren Repräsentation somatischer Veränderungen (in somatosensorischen Kortizes) in Juxtaposition steht als alle anderen Repräsentationen von Handlungs-Ergebnis-Szenarien, so wird, *ceteris paribus*, das motorische Schema, welches die Handlung dieses Szenarios repräsentiert, mit größerer Wahrscheinlichkeit seinen Schwellenwert erreichen als alle anderen (damit unverträglichen) motorischen Schemata.

Dieses Konditional kann als der neurobiologische Kerngehalt der SMH bezeichnet werden, es ist die zentrale kausale Gesetzmäßigkeit, welche durch die SMH postuliert wird.

Ich will kurz auf einen möglichen Einwand eingehen. Dieser könnte lauten: Das angeführte Konditional kann nicht der zentrale neurobiologische Gehalt der SMH sein, weil es die Begriffe „Repräsentation“, „Juxtaposition“ und „Schema“ enthält, welche evidenterweise nicht neurobiologischer Natur sind. Ich halte diesen Einwand für überaus gerechtfertigt, will aber behaupten, dass sich dahinter ein grundlegendes Problem der gesamten kognitiven Neurowissenschaft verbirgt. Der Begriff „Repräsentation“ ist in neurowissenschaftlichen Veröffentlichungen allgegenwärtig, obwohl dies eindeutig kein biologischer Begriff ist. Tatsächlich

handelt es sich um einen semantischen Begriff, da eine Relation zwischen Zeichen bzw. Bild und Bezeichnetem bzw. Abgebildetem impliziert wird. Und es ist bis heute nicht gelungen, eine naturalistische Definition semantischer Relationen zu geben. Die SMH, und mit ihr die meisten anderen Theorien der kognitiven Neurowissenschaft, scheint aber keinesfalls ohne den Begriff der Repräsentation oder ein Äquivalent auszukommen. Die Zuordnung neuronaler Prozesse zu kognitiven Inhalten würde ohne solche Begriffe noch nicht einmal beginnen können. Denn kognitive Inhalte besitzen wesentlich semantischen Gehalt. Ich werde auf dieses Thema zurückkommen.⁷⁸ Im Moment sei nur darauf hingewiesen, dass eine Formulierung der SMH ohne den Gebrauch semantischer Begriffe unmöglich ist, auch in ihrer *rein biologischen* Variante.

4.4.2. Anwendung der Brückenprinzipien

Um nun aus dem Satz (NBG) eine psychologische Gesetzmäßigkeit zu erhalten, reicht es hin, die oben angeführten Brückenprinzipien anzuwenden. Das Antecedens von (NBG) hat die Form: „Eine neuronale Repräsentation von X steht in Juxtaposition zu der Repräsentation somatischer Veränderungen in somatosensorischen Kortizes“. Die angesprochenen somatischen Veränderungen stellen laut den Brückenprinzipien eine Emotion dar.⁷⁹ Damit ergibt sich die Form: „Eine neuronale Repräsentation von X steht in Juxtaposition zu der Repräsentation einer Emotion“. Dies ist aber, gemäß den Brückenprinzipien, die neurobiologische Entsprechung für „X wird emotional bewertet“.⁸⁰ An der Stelle von „Repräsentation von X“ steht in (NBG): „Repräsentation eines Handlungs-Ergebnis-Szenarios (in den frühen somatosensorischen Kortizes)“. Dies ist aber nichts anderes als die Repräsentation der Handlung verbunden mit der davon verursachten Repräsentation der Folgen. Nach den Brückenprinzipien ist dies die Entsprechung für: „Das erwartete Handlungsergebnis für eine bestimmte Handlung“.⁸¹ Damit ergibt sich für das Antecedens des Satzes (NBG): „Wenn das erwartete Handlungsergebnis für eine bestimmte Handlung emotional positiver bewertet wird als alle erwarteten Handlungsergebnisse für andere Handlungen, ...“.

Das Consequens von (NBG) lautet: „so wird, *ceteris paribus*, das motorische Schema, welches die Handlung dieses Szenarios repräsentiert, mit größerer Wahrscheinlichkeit seinen Schwellenwert eher erreichen als alle anderen (damit unverträglichen) motorischen Schemata“. Ein motorisches Schema ist ein Verhaltensschema. Gemäß den Brückenprinzipien

⁷⁸ s. Kapitel 5.4.2.3.

⁷⁹ s. Kapitel 4.3.1.1.

⁸⁰ s. Kapitel 4.3.1.4.

⁸¹ s. Kapitel 4.3.1.3.

ist das Erreichen des Schwellenwerts eines Verhaltensschemas äquivalent zur Auswahl einer Handlungsoption.⁸² Damit ergibt sich für das Consequens: „so wird, *ceteris paribus*, die entsprechende Handlungsoption mit größerer Wahrscheinlichkeit ausgewählt als alle anderen Handlungsoptionen“. Da die psychologische Beschreibung von ganzen Organismen handelt, können wir diese Beschreibungen einer Person zuordnen. Es ergibt sich:

(PG) Wenn eine Person das erwartete Handlungsergebnis für eine bestimmte Handlung emotional positiver bewertet als die erwarteten Handlungsergebnisse für alle anderen Handlungen, so wird sie, *ceteris paribus*, die entsprechende Handlungsoption mit größerer Wahrscheinlichkeit auswählen als alle anderen Handlungsoptionen.

Hiermit ist gezeigt, dass die SMH den Satz (PG) impliziert. Da (PG) lediglich eine Umformulierung des Gesetzes (NBG) ist, muss (PG) eine kausale Gesetzmäßigkeit wiedergeben.⁸³ Damit kann (T1) als gesichert betrachtet werden. Zur übersichtlichen Zusammenfassung des Kapitels will ich eine schematische Darstellung der Struktur der SMH geben (s. Abbildung 6). Zusammenfassend lässt sich Folgendes sagen: Die SMH besteht im Wesentlichen aus zwei Gesetzen, einem neurobiologischen und einem psychologischen. Das psychologische Gesetz soll, durch die SMH, auf das neurobiologische Gesetz reduziert werden. Hierzu verwendet die SMH Brückenprinzipien, welche das psychologische und das biologische Vokabular miteinander verbinden. Das psychologische Gesetz, welches die SMH postuliert, besteht in (PG).

⁸² s. Kapitel 4.3.1.5.

⁸³ Ich werde auf diesen Punkt zurückkommen und ihn näher erläutern (s. Kapitel 5.1).

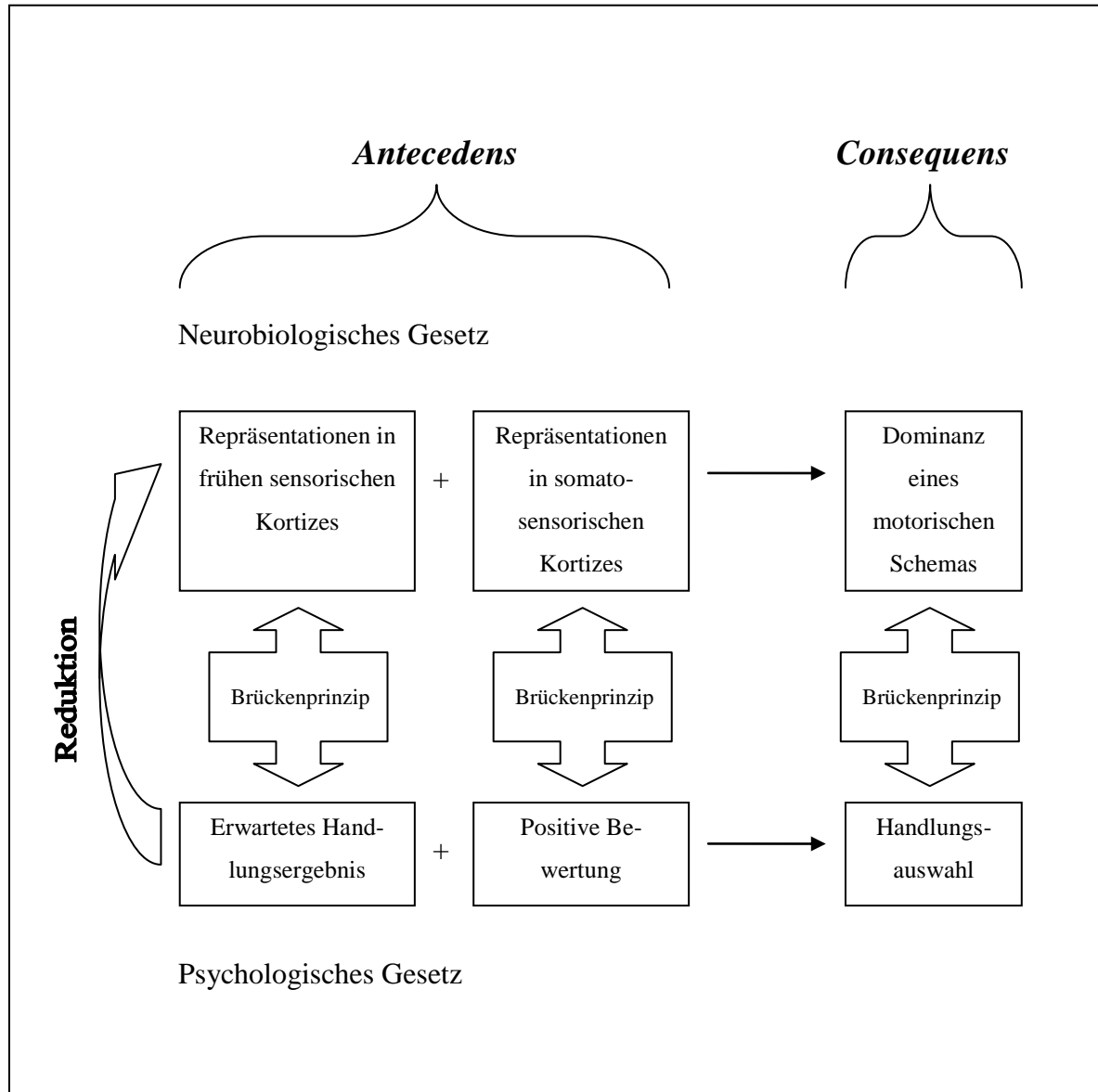


Abbildung 6: Schematische Darstellung der Struktur der SMH

Es werden zwei Gesetzmäßigkeiten postuliert. Die psychologische Gesetzmäßigkeit soll auf die neurobiologische reduziert werden. Hierzu werden Brückenprinzipien benötigt.

5. EMOTIONALE BEWERTUNG UND HANDLUNGSKLÄRUNG

Im vorangegangenen Kapitel wurde dargelegt, dass der Satz (PG)⁸⁴ als Kern des psychologischen Gehalts der SMH betrachtet werden muss. Ziel dieses Kapitels ist es nun, zu zeigen, dass (PG) nicht als psychologische Gesetzmäßigkeit mit empirischem Gehalt betrachtet werden kann. Beim Satz (PG) handelt es sich vielmehr um die Wiedergabe einer begrifflichen Relation, wie sie in Kapitel 2 besprochen wurde. Aber Sätze, welche begriffliche Relationen wiedergeben, können nicht mittels eines Rückgriffs auf empirische Gesetzmäßigkeiten reduktiv erklärt werden. Dieses Beweisziel lässt sich als These formulieren:

(T2) (PG) gibt einen begrifflichen Zusammenhang wieder und kann damit nicht als Gesetzmäßigkeit betrachtet werden, welche einer reduktiven Erklärung zugänglich ist.

(T2) folgt aus folgenden zwei Prämissen:

(T2.1) Wenn der Satz (PG) empirischen Gehalt hat, dann gibt er eine kausale Gesetzmäßigkeit wieder.

(T2.2) (PG) gibt keine kausale Gesetzmäßigkeit wieder.

Ich will nur kurz auf (T2.1) eingehen, um dann ausführlich für (T2.2) zu argumentieren.

5.1. Laut SMH erklärt (PG) Handlungen kausal

Dass die SMH davon ausgehen muss, dass (PG) eine kausale Gesetzmäßigkeit darstellt, scheint offensichtlich. Hierauf lassen sich deutliche Hinweise in den Texten der Vertreter der SMH finden:

„Emotion, feeling, and biological regulation all play a role in human reason“ (Damasio, 2006, S. XXIII).

„The central feature of the SMH is [...] that emotion-related signals assist cognitive processes“ (Bechara et al., 2005, S. 159)

„The key idea of this hypothesis [SMH] is that decision-making is a process that is influenced by marker signals“ (Bechara & Damasio, 2005, S. 336).

Dies scheint auch durchaus folgerichtig, denn sofern (PG) auf (NBG)⁸⁵ reduziert werden soll und (NBG) eine kausale Gesetzmäßigkeit darstellt,⁸⁶ muss auch (PG) eine solche wieder-

⁸⁴ s. S. 97 oben.

⁸⁵ s. S. 120 oben.

geben. Repräsentationen in frühen sensorischen Kortizes und im somatosensorischen Kortex sollen, laut SMH, das Entscheidungsverhalten kausal beeinflussen. Die emotionale Bewertung von Handlungsergebnissen besteht aber, der SMH zufolge, genau in der Juxtaposition dieser Repräsentationen. Der Übergang von (NBG) zu (PG) ist also schlicht eine Veränderung der Beschreibungsebene. Wie sollte aber die Beschreibungsebene dieser Vorgänge einen Einfluss auf die kausale Natur ihrer Beziehungen haben? Damasio spricht deshalb auch von „*brain caused` actions`*“ (Damasio, 2006, S. 89) und „*a mechanism for somatic states to exert a biasing effect on behaviors (e.g., selection of a response over another)*“ (Bechara & Damasio, 2005, S. 343).

Doch lassen sich hier auch Gründe anführen, welche gänzlich unabhängig von der SMH sind. Es wäre schlicht inkonsistent zu versuchen, eine begriffliche Relation durch die Angabe eines Naturgesetzes zu erklären. Welches Naturgesetz könnte uns reduktiv erklären, warum alle Rüden Hunde sind? Hier ein Naturgesetz anzuführen heißt, zu behaupten, dass der fragliche Zusammenhang ein empirischer ist. Die einzige Möglichkeit, zu bestreiten, dass (PG) laut SMH ein kausales Gesetz angibt, besteht darin, zu behaupten, dass die Reduktion auf (NBG) diese Kausalität ausschließt. Diese Möglichkeit scheint insbesondere bei der Annahme einer supervenienten Kausalbeziehung plausibel zu sein (vgl. Kim, 2007). Doch im gegenwärtigen Zusammenhang entsteht hieraus kein Problem. Denn reduktionistische neurowissenschaftliche Theorien, welche beanspruchen, menschliches Handeln erklären zu können, müssen eine der folgenden Positionen beziehen: (1) Typen-Identität:⁸⁶ Dann wird keine superveniente Kausalbeziehung behauptet, sondern eine Identität der Kausalbeziehungen. (2) Token-Identität: Dann wird eine superveniente Kausalbeziehung behauptet. Aber damit trägt auch die jeweilige Theorie die Beweislast, die Möglichkeit einer solche Beziehung zu belegen. (3) Epiphänomenalismus: Mit dieser Position wäre explizit das fragliche Kausalverhältnis geleugnet. Allerdings ist dann die Frage zu stellen, wie die jeweilige Theorie sich vom Eliminativismus unterscheiden lässt. Darüber hinaus stellt der Epiphänomenalismus eine dualistische Position dar (vgl. Pauen, 2001, S. 64ff.). Damit ist aber fraglich, wie eine reduktive Erklärung innerhalb dieses Rahmens überhaupt aussehen könnte.

⁸⁶ Ich betrachte es als unstrittig, dass die SMH den Satz (NBG) als eine kausale Gesetzmäßigkeit betrachtet. Dunn et al. (2006) gründen einen Großteil ihrer Kritik der SMH genau auf die Behauptung, dass die kausale Natur dieser Beziehung nicht hinreichend belegt sei.

⁸⁷ Die Vertreter der SMH postulieren zumindest für Emotionen eine Typen-Identität (s. Kapitel 4.3.1.1). Viele der anderen Thesen der SMH sind unter der Annahme einer allgemeinen Identitätsbeziehung zwischen neuronalen Prozessen und psychischen Prozessen am plausibelsten. Eine solche Interpretation scheitert aber an der allgemeinen Zurückhaltung neurowissenschaftlicher Theorien bezüglich ontologischer Fragen, welcher sich auch die Vertreter der SMH anschließen. Diese Zurückhaltung bleibt insofern unverständlich, als damit eine systematische Unklarheit der Theorien provoziert wird.

Auch unabhängig von der Tatsache, dass die Vertreter der SMH in (PG) eine kausale Gesetzmäßigkeit erblicken, ist also festzuhalten, dass reduktionistische neurowissenschaftliche Theorien im Allgemeinen in den von ihnen auf biologische Gesetze reduzierten psychologischen Gesetzen Aussagen über kausale Relationen erblicken müssen. Anderenfalls müssen ausführliche ontologische Erörterungen von der jeweiligen Theorie erwartet werden. Es findet sich also auch an diesem Punkt kein Grund, an der Verallgemeinerbarkeit der vorgetragenen Kritik auf andere Theorien der kognitiven Neurowissenschaft zu zweifeln.

5.2. Die SMH muss Handlungserklärungen geben können

Ziel dieses Abschnitts ist es zu belegen, dass (PG) keine kausale Gesetzmäßigkeit wiedergibt. Dies entspricht der Prämisse (T2.2). Mein Beleg soll sich in drei Schritten vollziehen: Erstens will ich darlegen, dass (PG), sofern der Satz ein kausales Gesetz sein soll, auf die Erklärung von Handlungen zielt. Zweitens will ich zeigen, dass (PG) Rahmenbedingungen für eine Handlungserklärung aus Gründen (d. i. einer intentionalen Erklärung) angibt. Und drittens will ich dafür argumentieren, dass Rahmenbedingungen intentionaler Erklärungen Handlungen nicht kausal erklären können.

5.2.1. Das Verhalten, welches auf Entscheidung beruht, ist eine Handlung

Ein Verhalten, welches auf einer Entscheidung beruht, kann als paradigmatischer Fall einer Handlung betrachtet werden (vgl. Jungermann, Pfister & Fischer, 1998, S. 3). Kann es Verhaltensweisen geben, die auf Entscheidungen beruhen, über welche aber die Person keine Kontrolle hat, zu welchen sie keine Gründe angeben kann, welche sie nicht sich selbst zuschreibt und für welche sie nicht verantwortlich ist? Diese Merkmale scheinen dem Begriff der Entscheidung zu widersprechen. Und genau diese Merkmale werden im Allgemeinen zur Abgrenzung des Begriffs der „Handlung“ verwendet (vgl. Greve, 1994, S. 10ff.; Schaub, 1993, S. 15ff.). Dieser Zusammenhang lässt sich an Beispielen verdeutlichen: Dem Patellar-sehnenreflex liegt keine Entscheidung zugrunde, dem Aktienkauf hingegen durchaus. Das Stottern beruht nicht auf Entscheidungen, doch das Heiraten unbedingt. Der Patellarsehnenreflex und Stottern sind keine Handlungen, der Aktienkauf und das Heiraten sind hingegen paradigmatische Handlungen. Dieser Zusammenhang hängt nicht von besonderen ‚Prozessen‘ ab, welche dem Verhalten vielleicht vorausgehen: Das Setzen des Blinkers zum Überholen auf der Autobahn beruht auf einer Entscheidung, selbst wenn dieser Vorgang völlig auto-

matisiert und ohne 'bewusste Kognitionen' vor sich geht. Und es handelt sich hier auch um eine Handlung.

Zum selben Ergebnis gelangt man, wenn man Elizabeth Anscombes Klärung des Handlungsbegriffs zugrunde legt: Zwar können wir auf die Frage „Warum hast du das getan?“ antworten „Aus keinem besondern Grund“, und unser Tun könnte immer noch eine Entscheidung sein, z. B. wenn ich mich für einen von zwei gleich langen Wegen zu einem Ziel entscheide (vgl. Anscombe, 2000, S. 25f., § 17). Aber es ist kaum möglich, ein Verhalten als auf Entscheidung beruhend zu beschreiben, wenn die Antwort auf die Warum-Frage lautet, dass die Person nicht (bzw. nur aus Beobachtung) wusste, dass sie das fragliche Verhalten zeigte oder wenn das Verhalten unfreiwillig war (vgl. Anscombe, 2000, S. 11ff., § 6f.). Ein Verhalten, das gerechtfertigterweise als auf Entscheidung beruhend bezeichnet wird, muss also als eine Handlung verstanden werden. In anderen Worten: Verhalten, welches auf Entscheidungen beruht, stellt stets eine Handlung dar.

5.2.2. Erklärungen von Verhaltensweisen, welche auf Entscheidungen beruhen, sind Handlungserklärungen

Die Tatsache, dass Verhaltensweisen, welche auf Entscheidungen beruhen, Handlungen sind, hat entscheidende Auswirkungen auf die Erklärung solchen Verhaltens. Denn sofern ein Verhalten *als entscheidungsbasiert* erklärt werden soll, muss es *als Handlung* erklärt werden. Dies ist zwar an sich trivial, da es zunächst nicht mehr sagt als dass, wenn die Neurobiologie Handlungen *qua Handlungen* erklären will, sie Handlungserklärungen bieten können muss. Jedoch sind damit besondere Bedingungen für die neurobiologische Erklärung menschlichen Verhaltens verbunden.

Es ist dabei entscheidend, zu beachten, dass Erklärungen intensionale Kontexte sind. Hieraus folgt, dass Handlungserklärungen ihren Gegenstand adäquat beschreiben müssen. Handlungserklärungen können allein aufgrund der Art und Weise, wie sie ihren Gegenstand beschreiben, falsch sein. Zunächst will ich daher darlegen, was es heißt, dass Erklärungen intensionale Kontexte darstellen.

5.2.2.1. Erklärungen sind intensionale Kontexte

Ein intensionaler Kontext ist ein Satz oder Teilsatz, in welchem die Ersetzung bestimmter Termini durch extensionsgleiche Termini den Wahrheitswert des Gesamtsatzes verändern kann (vgl. Bussmann, Trauth, Kazzazi & Bussmann, 1996, S. 234; Carnap, 1972, § 11). Der Wahrheitswert von Sätzen, welche intensionale Kontexte darstellen, hängt also nicht allein

von der Referenz der verwendeten Ausdrücke ab, sondern auch von dem semantischen Gehalt der Ausdrücke. Es ist leicht einzusehen, dass Erklärungen intensionale Kontexte darstellen. Beispielsweise kann der Satz „Das dritte Inferenzmaximum befindet sich an der Stelle xy, *weil* das Licht durch ein Gitter mit $g=833\text{nm}$ gefallen ist“ wahr sein, obwohl der Satz „Das dritte Inferenzmaximum befindet sich an Stelle xy, *weil* das Licht durch Konrads Gitter gefallen ist“ falsch ist und zwar auch dann, wenn Konrads Gitter eben genau jenes mit $g=833\text{nm}$ (also extensionsgleich) ist, denn die erklärende Variable ist hier die Gitterkonstante und nicht der Gitterbesitzer. Es handelt sich hier also um einen intensionalen Kontext. „*Explanations are intensional. So one thing explains another only 'under a description'*“ (Schueler, 2003, S. 10). Diese Tatsache wurde bereits in diversen Kontexten diskutiert (vgl. Achinstein, 1983, S. 195; de Muijnck, 2003, S. 130; s. auch Keil, 2000, S. 152). Bei näherer Betrachtung kann sie auch kaum überraschen, da es ganz offensichtlich ist, dass wir in jeder Erklärung das „*weil*“ jedweden Sinnes berauben können, indem wir die Referenten bestimmter Termini auf hinreichend bizarre Weise herausgreifen. Dies unterscheidet Erklärungen von der Angabe von Kausalverhältnissen: „X verursachte Y“ stellt keinen intensionalen Kontext dar, da hier die Termini X und Y durch beliebige extensionsgleiche Termini *salva veritate* ersetzt werden können. „X geschah, weil Y geschah“ ist hingegen durchaus ein intensionaler Kontext.

Wenn Handlungserklärungen intensionale Kontexte darstellen, dann ist die Art und Weise, in welcher in solchen Erklärungen entsprechende Referenten herausgegriffen werden, nicht gleichgültig. Dies will ich folgendermaßen ausdrücken: Handlungserklärungen müssen Handlungen *qua Handlungen* erklären. Eine rein 'extensional korrekte' Erklärung erfüllt diese Bedingung nicht. Eine Erklärung wie „Carla kaufte die Aktien, um jenes Geschehen herbeizuführen, welches 215 Jahre nach der Französischen Revolution genau 656,345 km östlich von Paris geschah“ könnte durchaus extensional korrekt sein (wenn z. B. an genannter Raum-Zeit-Stelle die von Carla intendierten Übernahmeverträge unterschrieben wurden). Aber eine solche Erklärung ist falsch, da sie den fraglichen Sachverhalt nicht auf die richtige Weise beschreibt. Diese angebliche Erklärung erklärt nichts. Es wird sich unten zeigen, dass eine intensional adäquate Erklärung menschlicher Handlungen auf *Gründe* rekurren muss.

Für die SMH bedeutet dies, dass es eine Adäquatheitsbedingung ihres Erfolgs darstellt, ob sie eine Erklärung von Entscheidungsverhalten *qua Handlungen* entwickeln kann. Im Folgenden wird sich zeigen, dass die SMH hieran scheitert.

5.2.2.2. Die SMH und das Wunsch-Überzeugungs-Modell

Welche Strukturen und Bedingungen der SMH haben sich bisher ergeben? Erstens haben wir gesehen, dass die Vertreter der SMH auf die Wahrheit und den empirischen Gehalt des Satzes (PG) verpflichtet sind. Zweitens habe ich dargelegt, dass die Vertreter der SMH in (PG) die Angabe eines Kausalverhältnisses erblicken müssen. Und drittens hat sich herausgestellt, dass die SMH eine Erklärung von Entscheidungsverhalten *qua Handlungen* bieten muss, wenn sie beansprucht, Entscheidungsverhalten zu erklären und nicht nur Prozesse erklären will, welche zufällig identisch (extensionsgleich) mit Entscheidungsverhalten sind. Ich denke nicht, dass die Vertreter der SMH in diesen drei Punkten ein Problem für ihre Theorie erblicken. Tatsächlich bin ich überzeugt, dass die Vertreter der SMH allen meiner bisherigen Ausführungen zustimmen würden. Es ist sogar möglich, alles bisher Gesagte zur Unterstützung der SMH anzuführen. Diese Möglichkeit möchte ich kurz aufzeigen, um sie anschließend zu kritisieren.

Zwei weitverbreitete Behauptungen lauten, dass Handlungserklärungen erstens Kausalserklärungen darstellen (diese These wird oft als „Kausalismus“ bezeichnet), und dass diese Erklärungen zweitens paradigmatischerweise in der Angabe eines Wunsch-Überzeugungs-Paares bestehen. Betrachten wir zunächst die zweite These: das Wunsch-Überzeugungs-Modell der Handlungserklärung. In jüngerer Zeit wurde diese These, welche auf David Hume zurückgeht, insbesondere von Donald Davidson vertreten. Davidson verwendet anstelle des Begriffs „Wunsch“ den Begriff „Proeinstellung“, um „Wünsche, Begehren, Impulse, Reize und eine große Vielfalt von moralischen Ansichten, ästhetischen Grundsätzen, ökonomischen Vorurteilen, gesellschaftlichen Konventionen [und] von öffentlichen und privaten Zielen und Werten“ (Davidson, 1998d, S. 20) zu bezeichnen. Nach Davidson sind Handlungserklärungen Erklärungen aus Gründen. Die kanonische Form einer Handlungserklärung ist eine Erklärung aus *primären Gründen*. Dabei gilt:

„*G* ist nur dann ein primärer Grund, weshalb eine handelnde Person die Handlung *H* unter der Beschreibung *b* ausführt, wenn sich *G* zusammensetzt aus einer Proeinstellung des Handelnden zu Handlungen mit einer bestimmten Eigenschaft sowie der Überzeugung des Handelnden, daß *H* unter der Beschreibung *b* diese Eigenschaft hat“ (Davidson, 1998d, S. 22)

Dieses Wunsch-Überzeugungs-Modell der Handlungserklärung ist bis heute vielfach vertreten und ausgearbeitet worden (vgl. Goldman, 2005; Mele, 1992). In der psychologischen Forschung wurden weitgehend analoge Modelle unter dem Titel „Erwartungs-Wert-Theorien“

vorgelegt (vgl. Ajzen & Fishbein, 1980; s. auch Greve, 1994, S. 131). Viele Vertreter unterschiedlicher Wunsch-Überzeugungs-Modelle verfechten gleichzeitig die These, dass Handlungserklärungen mittels Wunsch-Überzeugungs-Paaren kausale Erklärungen darstellen. So schreibt etwa Davidson: „Die Ursache einer Handlung ist ihr primärer Grund“ (Davidson, 1998d, S. 20). Ich werde unten näher auf dieses Thema eingehen, will aber im Moment nur verdeutlichen, welche Konsequenzen die genannten Thesen für die Interpretation der SMH haben.

Es liegt nahe, den Ausdruck „emotional positiv bewerten“ in (PG) als die Bezeichnung einer Proeinstellung zu lesen. Entsprechend können wir den Ausdruck „erwartetes Handlungsergebnis für eine bestimmte Handlungsoption“ so verstehen, dass damit die Überzeugung des Handelnden ausgedrückt wird, dass die fragliche Handlung die positive Eigenschaft, auf welche sich die Proeinstellung bezieht, auch tatsächlich hat. Wir haben es demnach bei (PG) mit der Formulierung eines Wunsch-Überzeugungs-Paares bzw. eines Erwartungs-Wert-Paares zu tun. Setzt man das Wunsch-Überzeugungs-Modell voraus, so wäre hiermit eine Erklärung von Entscheidungen *qua Handlungen* gelungen. Wenn wir nun zusätzlich akzeptieren, dass Wunsch-Überzeugungs-Erklärungen kausaler Natur sind, so können alle drei der zu Beginn dieses Abschnitts erwähnten Anforderungen an die SMH erfüllt werden. Erstens wäre (PG) wahr und empirisch gehaltvoll, zweitens gäbe (PG) ein Kausalverhältnis wieder und drittens leistete (PG) eine Erklärung von Entscheidungsverhalten *qua Handlung*. Zusätzlich wäre in diesem Fall eine Reduktion der Gesetzmäßigkeit (PG) auf eine neurobiologische Gesetzmäßigkeit gelungen (sofern sich die SMH empirisch bestätigen lässt).⁸⁸ Dieses Ergebnis erscheint auf den ersten Blick bestechend. Entscheidend ist allerdings, dass man nun deutlich sehen kann, dass die Wahrheit der SMH die Richtigkeit des Wunsch-Überzeugungs-Modells und des Kausalismus impliziert. Damit lässt sich aber aus der Falschheit des Wunsch-Überzeugungs-Modells oder des Kausalismus *per modus tollens* auf die Falschheit der SMH schließen. Genau dies möchte ich im Folgenden versuchen.

5.3. Emotionale Bewertungen sind Rahmenbedingungen von Handlungserklärungen

Der folgende Abschnitt dient dazu, darzulegen, dass emotionale Bewertungen von Handlungs-Ergebnis-Szenarien nicht Gründe, sondern Rahmenbedingungen von Handlungs-

⁸⁸ Innerhalb der Davison'schen Konzeption wäre dies, aufgrund des anomalen Monismus, prinzipiell ausgeschlossen. Hier will ich jedoch diese Problematik vernachlässigen.

erklärungen darstellen. Hierzu ist es notwendig, den Begriff der „Handlungserklärung“ näher zu untersuchen. Daran anschließend will ich einige Betrachtungen zu Emotionen und emotionalen Bewertungen anstellen.

5.3.1. Handlungserklärungen

Wir haben bereits gesehen, dass Handlungserklärungen intensionale Kontexte darstellen. Aber wie lassen sie sich darüber hinaus charakterisieren? Wann sind Handlungserklärungen intensional adäquat? Auf diese Fragen versucht der folgende Abschnitt zu antworten.

5.3.1.1. Handlungserklärungen sind Erklärungen aus Gründen

Die Erklärung einer Handlung *qua Handlung* ist eine Erklärung aus Gründen (vgl. Dickenson, 2007, S. 1). Dies leitet sich vor allem daraus ab, dass wir unter einer Handlung ein Tun verstehen, welches eine bestimmte Art von Warum-Frage zulässt (vgl. Anscombe, 2000).⁸⁹ Dasjenige, was in der Antwort auf diese Frage genannt wird, bezeichnen wir als den Grund, aus welchem die Person jene Handlung vollzogen hat. Es ist also ein wesentliches Merkmal von Handlungen, dass sie in der Regel aus Gründen vollzogen werden (vgl. Bishop, 1989, S. 99). „[I]n understanding some event as an action one is necessarily understanding it as having some point or purpose, at least under some description“ (Schueler, 2003, S. 8). Hier könnte ein Vertreter der SMH einwenden, dass es genau einen attraktiven Zug der SMH darstellt, dass sie eine unbewusste Wirkung somatischer Marker zulässt und unbewusste Vorgänge nicht als Elemente einer Erklärung aus Gründen auftauchen können. Dieser Einwand ist insofern falsch, als das Vorliegen eines Grundes keine bewusste Überlegung voraussetzt.

„Nehmen wir an, ich blättere beim Lesen eines Textes um und werde von einem Kind gefragt, warum ich das tue. Ich sage: `Weil ich sehen will, was hier auf dieser neuen Seite steht`. Ist das gelogen, weil ich ganz sicher diesen Gedanken auch in meiner Vorstellung nicht artikuliert habe? Nein. Meine Auskunft bringt tatsächlich den wirklichen Grund für mein Blättern zum Vorschein“ (Müller, 1998, S. 135).

Es mag Fälle geben, in welchen wir sagen, dass wir den Grund unserer Handlungen nicht kennen, oder dass es keinen besonderen Grund gibt (vgl. Anscombe, 2000, S. 25f., § 17). Doch diese Fälle sind notwendigerweise Ausnahmen. Würden sie zum Regelfall, so würde unser Begriff der Handlung in sich zusammenfallen.

⁸⁹ Anscombes Abgrenzung von Handlungen (gegenüber anderem Verhalten) basiert auf dem Begriff der Absicht, doch Absichten sind wesentlich mit Gründen verbunden. Davidson folgt Anscombe in beiden Punkten, sodass das im Text Gesagte auch auf ihn zutrifft (vgl. Davidson, 1998a, S. 131; Davidson, 1998c, S. 77).

Wenn nun eine Handlung ein Tun aus Gründen ist, dann scheint klar, dass wir bei der Erklärung von Handlungen *qua Handlungen* auf Gründe rekurrieren müssen. Wir nennen genau das einen Grund, was eine angemessene Antwort auf die Frage darstellt, warum jemand etwas getan hat. Und wir sagen, dass wir verstehen, warum eine Person eine Handlung vollzieht, wenn wir ihre Gründe kennen (vgl. Bittner, 2001, S. 91; Stoutland, 1998, S. 44ff.). „*The explanation of an action succeeds to the extent that it enables us to see how the agent might have taken certain features of the action as good reason to do it*“ (Dancy, 2000, S. 95). Wir erklären also Handlungen durch die Angabe von Gründen. Anders ausgedrückt: Es ist ein Kriterium der intensionalen Adäquatheit einer Handlungserklärung, dass sie Gründe nennt.

Natürlich können wir dasselbe Geschehen, welches wir als Handlung erklären können, auch als physikalisches Ereignis erklären, aber in diesem Fall haben wir es eben *nicht als Handlung* erklärt. Zumindest haben wir das nicht, solange wir keine reduktive Theorie der Handlungserklärung besitzen, welche uns erklärt, wie wir methodisch gesichert von einer physikalischen Erklärung zu einer Handlungserklärung kommen, die in nichts über die physikalische Erklärung hinausgeht. Genau weil Handlungserklärungen Erklärungen aus Gründen sind, war es für den Kausalismus immer eine entscheidende Herausforderung, zu zeigen, wie Gründe Ursachen sein können. Ein schlichter Hinweis auf einen Kausalnexus, der zu einer Handlung führt, reicht hier nicht aus.

5.3.1.2. Normative und motivierende Gründe

In der Literatur zum Charakter von Handlungserklärungen hat sich die Unterscheidung zwischen normativen und motivierenden Gründen etabliert. Manchmal werden normative Gründe auch als rationalisierende Gründe und motivierende Gründe als erklärende Gründe bezeichnet (vgl. Davidson, 1998d, S. 19). Normative Gründe sind solche, welche eine Handlung rechtfertigen. Sie sind jene Gründe, aus welchen eine Handlung vollzogen werden *sollte*; sie geben an, was *für* eine Handlung spricht. Motivierende Gründe hingegen sind diejenigen Gründe, aus welchen eine Person tatsächlich handelt. Motivierende Gründe können Handlungen somit erklären (vgl. Dancy, 2000, S. 1ff.). Es ist offensichtlich, dass jede gute Handlungstheorie die Beziehung zwischen diesen beiden Typen von Gründen klären muss. Denn rationales Handeln ist Handeln aus guten Gründen, also aus normativen Gründen. Eine Handlungstheorie, welche die Beziehung zwischen den beiden Typen von Gründen im Dunkeln lässt, kann also nicht erklären, wie Menschen aus guten Gründen, d. h. rational, handeln können. In diesem Fall wäre die entsprechende Handlungstheorie bestenfalls das Fragment eines Erklärungsansatzes (vgl. Davidson, 1998d, S. 19; Davis, 2005). In jüngster Zeit hat

Jonathan Dancy zwei Bedingungen formuliert, welchen jede Theorie der Handlungserklärung gerecht werden muss. Ich will diese Bedingungen jeweils nennen und kurz für sie argumentieren.⁹⁰

5.3.1.2.1. Die Erklärungsbedingung

Die erste Bedingung lautet:

„[A]ny theory about the relation between normative and motivating reasons [must] show that and how any normative reason is capable of contributing to the explanation of an action that is done for that reason“ (Dancy, 2000, S. 101).

Um diese Bedingung zu rechtfertigen, sei das Gegenteil angenommen, d. i., dass es möglich ist, eine Theorie der Handlungserklärung zu formulieren, nach welcher normative Gründe⁹¹ keine Rolle für die Erklärung von Handlungen spielen. Dann wäre es möglich, eine Handlung aus Gründen zu erklären, ohne auf normative Gründe Bezug zu nehmen. Doch hier taucht ein Problem auf. Inwieweit könnten wir in einem solchen Fall noch von Gründen sprechen? Von Gründen sprechen wir nur dort, wo wir ein gewisses Verständnis dafür entwickeln können, warum jemand gehandelt hat. Der Grund einer Handlung lässt diese Handlung in irgendeiner Hinsicht als sinnvoll erscheinen (vgl. Dancy, 2000, S. 97). Ein Beispiel: „Warum hast du diese Aktien gekauft?“, „Weil Kohlenstoff die Ordnungszahl 6 hat“. Hier wissen wir nicht, ob wir es mit einem Grund zu tun haben. Wenn wir der Sache weiter nachgehen und dabei immer nur ähnlich absurde Antworten bekommen, werden wir entweder vermuten, dass wir zum Besten gehalten werden, oder dass wir es hier überhaupt nicht mit einem Grund zu tun haben.⁹² Anders ist es, wenn sich herausstellt, dass die Antwort soviel heißt, wie: „Ich habe die Gründe XY zu der Annahme, dass die Ordnungszahl von Kohlenstoff einen Anstieg des Aktienkurses hervorrufen wird“. In diesem Fall haben wir es mit einem Grund zu tun.⁹³

⁹⁰ Für eine ausführliche Darstellung und Argumentation sei auf Dancy (2000) verwiesen.

⁹¹ Mit dem Ausdruck „normative Gründe“ sind hier immer objektive gute Gründe gemeint. Eine zusätzliche Komplikation entsteht dadurch, dass die Beurteilung der objektiven Güte dieser Gründe bei jeder einzelnen Handlungserklärung natürlich beim erklärenden Subjekt liegt. Doch sollten hieraus keine voreiligen skeptischen Schlüsse gezogen werden.

⁹² Vielleicht ist der Handelnde ein Neurobiologe und weiß, dass diese Eigenschaft des Kohlenstoffs ein wichtiger Kausalfaktor beim Zustandekommen seiner Handlung war.

⁹³ Diese Überlegungen sind eng verbunden mit dem sogenannten *Principle of Charity*, wie es von Davidson und Quine vertreten wurde (vgl. Davidson, 1998b, S. 312; Quine, 1980, S. 115). Wenn wir diesem Prinzip bei der Zuschreibung sprachlicher Äußerungen, mentaler Phänomene und sinnvollen Tuns folgen müssen, so müssen wir diese Zuschreibungen auf genau jene Weise vornehmen, dass die jeweilige Person maximal rational erscheint – nach *unseren eigenen* Maßstäben. Damit ist aber auch klar, dass keine Theorie sinnvollen Tuns (Handelns) ohne Rekurs auf *unsere eigenen Rationalitätsmaßstäbe* auskommt.

Anscombe (2000, S. 70ff., § 37) spricht in diesem Zusammenhang von *desirability-characterisation* und bezeichnet damit jene Beschreibung eines gewollten Gegenstands oder Geschehens, welche auf die Frage danach, warum der Gegenstand oder das Geschehen gewollt wird, eine abschließende Antwort gibt. Damit widerspricht Anscombe explizit der Ansicht, dass die Worte „wollen“ und „beabsichtigen“ auf beliebige Gegenstände anwendbar sind. Damit wir eine Person als „wollend“ oder „beabsichtigend“ beschreiben können, muss das Gewollte oder Beabsichtigte bzw. dessen letztes Ziel als aus der Perspektive des Erklärenden (also aus *unserer* Perspektive) begehrenswert betrachtet werden können.

„But is not anything wantable, or at least any perhaps attainable thing? It will be instructive to anyone who thinks this to approach someone and say: ‘I want a saucer of mud’ [...] He is likely to be asked what for; to which let him reply that he does not want it for anything, he just wants it. It is likely that the other will then perceive that a philosophical example is all that is in question, and will pursue the matter no further; but supposing he did not realize this, and yet did not dismiss our man as a dull babbling loon, would he not try to find out in what aspect the object desired is desirable? Does it serve as a symbol? Is there something delightful about it? Does the man want to have something to call his own, and no more? Now if the reply is: ‘Philosophers have taught that anything can be an object of desire; so there can be no need for me to characterise these objects as somehow desirable; it merely so happens that I want them’ then this is fair nonsense“ (Anscombe, 2000, S. 70f., § 57).

Es lässt sich natürlich fragen, wie es möglich ist, dass wir Menschen im Allgemeinen unsere *desirability-characterisations* wechselseitig anerkennen und verstehen. Doch diese Frage würde uns weit vom gegenwärtigen Thema wegführen.⁹⁴ Im gegenwärtigen Kontext reicht es hingegen völlig aus zu erkennen, dass Handlungserklärungen aus motivierenden Gründen nicht ohne Rekurs auf normative Gründe auskommen.

Um Missverständnisse zu vermeiden, sei darauf hingewiesen, dass mit dem hier Gesagten nicht behauptet wird, Menschen könnten sich nicht sehr unterschiedliche Ziele setzen. Unsere Zuschreibung von Handlungen und Gründen setzt aber eine grundlegende Gemeinsamkeit voraus, auf welcher sämtliche Unterschiede in unseren Zielen, Präferenzen und Neigungen erst bestehen können. Ich kann vielleicht nicht verstehen, wie jemand Freude am Briefmarkensammeln haben kann, doch ich weiß, was ein Hobby ist. Das Unverständnis für die

⁹⁴ Für eine Antwort auf diese Frage wäre der Begriff der menschlichen Lebensform entscheidend (vgl. Foot, 2001; Thompson, 2008).

Ziele unserer Mitmenschen ist in Wirklichkeit oft lediglich ein Unverständnis für die *Rangfolge* ihrer Wertungen und Ziele oder die besondere Ausformung eines Ziels.

5.3.1.2.2. Die Normativitätsbedingung

Die zweite Bedingung lautet:

„[A] motivating reason, that in the light of which one acts, must be the sort of thing that is capable of being among the reasons in favour of so acting; it must, in this sense, be possible to act for a good reason“ (Dancy, 2000, S. 103).

Dieselbe Bedingung einer befriedigenden Handlungserklärung wurde auch von anderen Autoren vorgebracht (Stoutland, 2005; Wiland, 2002, S. 451; Williams, 1995, S. 38f.). So schreibt etwa Ulrike Heuer:

„[W]hen an agent acts for a (specific) reason that very reason is also the explanation (or at least part of the explanation) of why she did what she did. Normative or justificatory reasons and explanatory reasons are the same reasons in such a case, and not different kinds of reasons altogether. [...] [J]ustificatory reasons explain if a person acts for them“ (Heuer, 2004, S. 45).

Anders ausgedrückt: Normative Gründe müssen eine entscheidende Rolle in Handlungserklärungen spielen können. Der einfache Grund dieser These ist, dass wir genau denjenigen als rational bezeichnen, der aus guten (normativen) Gründen handelt. Wäre die Normativitätsbedingung nicht erfüllt, so bestritten wir damit die Möglichkeit der menschlichen Rationalität. Ein solches Bestreiten der menschlichen Rationalität ist zwar in der psychologischen Literatur nicht selten,⁹⁵ es ist aber leicht einzusehen, dass diese These inkonsistent ist. Denn wie sollten wir zu diesem Ergebnis kommen? Entweder durch ein rationales Verfahren, dies ist aber laut der in Frage stehenden These ausgeschlossen, oder durch ein irrationales Verfahren, dann sollten wir diesem Ergebnis aber keinen Glauben und kein Gewicht schenken.⁹⁶ Es ist also eine Bedingung der Möglichkeit aller wissenschaftlicher Unternehmungen überhaupt (tatsächlich unseres gesamten gewöhnlichen Lebensvollzugs), dass Menschen aus guten Gründen handeln können. Natürlich ist hierdurch nicht ein unerfreulich hohes Maß an tatsächlich vorliegender Irrationalität ausgeschlossen, aber diese Irrationalität kann nicht das gesamte

⁹⁵ Eine Sammlung von Beispielen findet sich bei Botterill & Carruthers (1999, S. 105ff.).

⁹⁶ Im Wesentlichen ist dies dieselbe Kritik, wie sie oben gegen den Eliminativismus vorgebracht wurde. Sie trifft alle Arten von Skeptizismus und ist seit der Antike bekannt. Unglücklicherweise scheint sie heutzutage nicht Teil der Allgemeinbildung zu sein. Für eine – meines Erachtens erfolglose – zeitgenössische Gegenkritik siehe Stein (1996).

menschliche Handeln ausmachen. Wenn wir nun aber tatsächlich aus guten (normativen) Gründen handeln, dann sind diese auch unsere motivierenden Gründe. Und da Handlungserklärungen Erklärungen aus (motivierenden) Gründen sind, erklären in diesem Fall normative Gründe unsere Handlungen. Und genau dies besagt die Normativitätsbedingung.

5.3.1.3. Gründe sind Sachverhalte

Das soeben Gesagte hat weitreichende Folgen. Denn wir sind oft geneigt zu glauben, dass motivierende Gründe psychische Zustände seien, z. B. Wünsche und Überzeugungen. Es ist aber eine äußerst unplausible These, zu behaupten, dass normative Gründe typischerweise psychische Zustände sind. Der normative Grund dafür, dass ich Tante Erna eine Dankeskarte schreibe ist, dass ich ein Geschenk von Tante Erna bekommen habe. Mein Glauben, dass ich ein Geschenk bekommen habe, stellt *keinerlei* normativen Grund für mein Handeln dar. Mein Grund zu behaupten, dass „5!“ Gleich „120“ ist, besteht nicht in meinem *Glauben*, sondern darin, dass „ $5 \times 4 \times 3 \times 2$ “ *tatsächlich* gleich „120“ ist. „Wünschen und Überzeugungen mag durchaus eine bedeutsame Rolle zukommen, etwa als notwendige Bedingungen dafür, dass Sachverhalte Gründe sind, aus denen gehandelt wird. Aber sie *konstituieren* nicht diese Gründe“ (Stoutland, 2005, S. 4). Ein motivierender Grund muss etwas sein, von dem wir verstehen können, wie es dem Handelnden als ein normativer Grund erscheinen konnte. Normative Gründe bestehen aber typischerweise in Sachverhalten: „dass Tante Erna mir ein Geschenk gemacht hat“, „dass $5 \times 4 \times 3 \times 2$ gleich 120 ist“ oder „dass er mich betrogen hat“ (vgl. Bittner, 2001; Schueler, 2003; Stoutland, 1998). „*The crucial point here is that believing that p is never (or hardly ever) a good reason for Φ -ing. It is what is believed, that p, that is the good reason for Φ -ing, if there is one*“ (Dancy, 2000, S. 107).⁹⁷ Eine adäquate Handlungstheorie muss also Sachverhalten, welche typischerweise nicht psychische Zustände des Handelnden sind, eine entscheidende Rolle bei der Erklärung von Handlungen zuweisen können.

Schon hier zeichnen sich gravierende Probleme für eine kausalistische Handlungstheorie ab. Diese werden insbesondere an Beispielen deutlich, in welchen mathematische Sachverhalte Gründe unseres Tuns sind. Denn sofern wir Mathematik auf rationale Weise betreiben können, müssen mathematische Sachverhalte nicht nur unter den normativen, sondern auch unter den motivierenden Gründen unserer Handlungen sein können. Mathematische Sachver-

⁹⁷ Die Ausnahmefälle, in welchen psychische Zustände gute Gründe darstellen, sind z. B. solche, in welchen pathologische Phänomene (Angst, Halluzination etc.) Gründe für den Besuch eines Psychiaters darstellen (vgl. Dancy, 2000, S. 125).

halte scheinen aber kaum die richtige 'Art von Gegenständen' zu sein, um kausale Relationen einzugehen. Mathematische Sachverhalte können daher nicht das Notieren der richtigen Lösung einer Aufgabe *verursachen*. Aber der normative Grund des Notierens ist eben jener mathematische Sachverhalt.

5.3.1.4. Gründe und Rahmenbedingungen

Es scheint sich aber unmittelbar ein Problem für die soeben skizzierte Position zu ergeben: Wir sagen oft Dinge wie „sie rannte, weil sie *dachte*, der Bus würde bald abfahren“, „sie holte den Hammer, weil sie ein Bild aufhängen *wollte*“ oder „sie ging diesen Weg, weil sie nach Trier *wollte* und *dachte*, es sei der kürzeste Weg“. In solchen Fällen nehmen wir Bezug auf kognitive und volitionale 'psychische Zustände',⁹⁸ um Handlungen zu erklären. Objektive Sachverhalte scheinen dabei keine wesentliche Rolle zu spielen. Durch diese Überlegungen kann man leicht zu der Überzeugung kommen, dass Handlungserklärungen wie „Sie schrieb die Karte, weil Tante Erna ihr etwas geschickt hatte“ in Wahrheit elliptisch sind. Demnach würde hier in Wirklichkeit gesagt: „Sie dachte, dass Tante Erna ihr etwas geschenkt hatte, wollte sich bedanken und war überzeugt dies durch das Schreiben einer Karte erreichen zu können, deshalb schrieb sie die Karte“.

Die Falschheit dieser Auffassung, dass Handlungserklärungen aus objektiven Sachverhalten (normativen Gründen) als elliptisch zu betrachten sind, folgt daraus, dass in diesem Fall die nicht-elliptische Version der Erklärung weder die Normativitäts- noch die Erklärungsbedingung erfüllen könnte. Um dies zu umgehen, müsste man den normativen Grund (Sachverhalt) an einer entscheidenden Stelle in die Handlungserklärung aufnehmen. Man könnte dies entweder versuchen, indem man den normativen Grund zur Ursache des motivierenden Grundes erklärt, oder indem man den normativen Grund als den Inhalt derjenigen psychischen Zustände bestimmt, welche den motivierenden Grund ausmachen. Die erste Alternative scheitert daran, dass abstrakte und zukünftige Sachverhalte keine effizienten Ursachen psychischer Zustände sein können, hingegen unbedingt als Gründe unseres Handelns gelten können müssen. Denn sowohl mathematische Tatsachen als auch die Verwirklichung bestimmter Ziele können normative Gründe unseres Handelns sein. Die zweite Alternative scheitert daran, dass die Formulierung „Inhalt psychischer Zustände“ zweideutig ist. Mit dem „Inhalt psychischer Zustände“ können erstens Propositionen gemeint sein. So glaube ich,

⁹⁸ Bei der Verwendung des Ausdrucks „psychischer Zustand“ ist äußerste Vorsicht geboten (vgl. PU § 308, Wittgenstein, 1997b, S. 377f.). Viele der begrifflichen Verwirrungen der 'wissenschaftlichen' Psychologie und des Kausalismus haben hier ihren Ursprung.

wenn ich glaube, dass Schnee weiß ist, die Proposition „Schnee ist weiß“. Propositionen sind aber genauso wenig normative Gründe wie psychische Zustände, sie sind abstrakte Objekte (z. B. Funktionen von möglichen Welten auf Wahrheitswerte oder – je nach Theorie – etwas Ähnliches) (vgl. Dancy, 2000, S. 112). Zweitens kann mit der Formulierung „Inhalt psychischer Zustände“ die Referenz der Proposition, welche Inhalt des psychischen Zustands ist, gemeint sein. In diesem Fall könnte der „Inhalt psychischer Zustände“ tatsächlich ein Sachverhalt und also auch ein normativer Grund sein (sofern nicht mir Frege der Wahrheitswert als die Referenz von Propositionen angenommen wird). Aber es stellt sich hier wieder die Frage, inwieweit die Normativitätsbedingung erfüllt ist. Die Referenz des psychischen Zustandes, d. i. der Sachverhalt, spielt nämlich in dieser Konzeption keine entscheidende Erklärungsrolle; sondern es ist die Tatsache, dass die entsprechende Proposition geglaubt wird, welche allein die entscheidende Erklärungsarbeit leistet. Der Grund, aus dem heraus gehandelt wurde, ist dann auch nicht der Sachverhalt, auf den der psychische Zustand referiert, sondern der psychische Zustand selbst. Damit sind wir aber wieder bei der Behauptung, dass Menschen nicht aus guten Gründen handeln können.

„Perhaps, however the thought is merely that we can *call* ‘good’ any motivating reason that represents a state of affairs that is a good reason. [...] The purpose of the normative constraint was to demand that motivating reasons be not excluded from the normative realm. We do not do much to meet that demand by finding a simple correlation between motivating reasons, conceived as stubbornly non-normative, and those states of affairs that do play a normative role“ (Dancy, 2000, S. 119).

Also sind Handlungserklärungen aus Gründen, welche Sachverhalte sind, nicht elliptischer Natur. Aber was ist dann die Rolle von Wünschen und Überzeugungen in Handlungserklärungen? Denn dass sie eine Rolle bei der Erklärung von Handlungen spielen, steht außer Frage.

Die Antwort hierauf lautet: *Psychische Zustände des Handelnden sind Rahmenbedingungen*⁹⁹ *von Erklärungen aus Gründen, sie sind selbst nicht Teil der Gründe*. Um die Richtigkeit und Bedeutung dieser Behauptung zu erkennen, ist es hilfreich, sich zunächst vor Augen zu halten, wann wir typischerweise im Kontext von Handlungserklärungen auf psychische Zustände verweisen. Es gibt drei unterschiedliche Arten solcher Kontexte. Erstens nehmen wir

⁹⁹ In der Literatur werden für dasjenige, was ich hier „Rahmenbedingungen“ nenne, unterschiedliche Termini verwendet. Stoutland spricht von „*background conditions for an explanation*“ (Stoutland, 1998, S. 60), während Dancy von „*enabling conditions for an explanation*“ (Dancy, 2000, S. 127) spricht.

auf Wünsche (Proeinstellungen) Bezug, wenn uns das Ziel einer Handlung nicht unmittelbar als wünschenswert erscheint. Ein Beispiel: „Warum ruft sie diese schreckliche Person an?“ „Weil sie sich Geld von ihr leihen *will*“. Hier haben wir einen Sachverhalt, welcher einen Grund für das Handeln darstellen kann, nämlich Geldnot. Doch auf Anhieb verstehen wir nicht, dass dieser Grund mit der Handlung verbunden ist. Die Auskunft über den psychischen Zustand des Handelnden macht uns diese Verbindung klar. Zweitens nehmen wir auf Überzeugungen Bezug, wenn uns eine Handlung nicht zielführend erscheint. Auch hier ein Beispiel: „Warum schlägt sie mit dem Hammer auf den Motor?“ „Weil sie *weiß*, dass sie den Motor damit wieder in Gang bekommt“. Auch hier gibt es einen Grund, welcher in einem Sachverhalt besteht, nämlich das Stillstehen des Motors. Die Bezugnahme auf den psychischen Zustand hilft uns, eine Verbindung zwischen diesem Grund und der Handlung zu sehen. Am wichtigsten ist jedoch der dritte Fall. Hier nennen wir psychische Zustände, weil es keinen Sachverhalt gibt. Dies ist der Fall, wenn der Handelnde einem Irrtum unterliegt. Wieder ein Beispiel: „Warum rief sie ihn an?“ „Weil sie *dachte*, er sei zu Hause“. Es ist wichtig zu sehen, dass alle diese drei Kontexte Ausnahmesituationen darstellen. Im Normalfall reicht das Nennen der Gründe hin, um eine Handlung zu erklären: Der Schaden des Autos kann erklären, warum Franz es zur Reparatur bringt. Der Ruf der Freundin kann erklären, warum Peter ins Wohnzimmer geht. Die große Rendite kann erklären, warum Heinz sein Geld dort anlegt. Dass Klaus Ingrid ins Schach gesetzt hat, kann erklären, warum Ingrid Klaus' Läufer geschlagen hat. Hier möchte man vielleicht einwenden, dass aber bei jeder dieser Erklärungen stillschweigend ein Motiv angenommen wurde. Dass hier stillschweigende Voraussetzungen am Werk sind, ist zwar richtig, aber nicht spezifisch für die psychischen Zustände des Handelnden. Dass Klaus Ingrid ins Schach gesetzt hat, erklärt Ingrids Zug nur unter der Voraussetzung der Regeln des Schachs. Hieraus folgt aber nicht, dass in Wirklichkeit alle Regeln des Schachs Teil der richtigen Erklärung von Ingrids Handlung sein müssen. Der Ruf seiner Freundin erklärt Peters Kommen nur unter der Voraussetzung, dass Peter nicht vor genau 4 Minuten und 32 Sekunden einen Hörsturz erlitten hat. Trotzdem muss diese Tatsache nicht in einer vollständigen Erklärung von Peters Verhalten auftauchen. Anderenfalls könnten wir keine Erklärung jemals als vollständig betrachten, denn es lässt sich für jedes Ereignis eine unendliche Liste notwendiger Bedingungen anführen (insbesondere solcher Bedingungen, welche irgendwelche besonderen Umstände ausschließen, wie die Explosion der Erde oder notwendige Umstände, welche weit in der Vergangenheit liegen, wie der Urknall). Dies bedeutet: Um zu zeigen, dass X in der Erklärung von Y auftauchen muss, reicht es nicht hin, zu belegen, dass X eine notwendige Bedingung für Y ist. Und damit ist das Argument,

dass psychische Zustände in der Erklärung von Handlungen auftauchen müssen, da sie eine notwendige Bedingung dieser Handlungen sind, hinfällig.

Was wir wissen wollen, wenn wir nach psychischen Zuständen im Zusammenhang von Handlungserklärungen fragen, ist nicht so sehr, was die Gründe waren, aus welchen heraus eine Handlung vollzogen wurde, sondern vielmehr, wie dieser oder jener Sachverhalt einer Person als Grund zum Handeln *erscheinen* konnte. Nach dem Sachverhalt zu fragen und danach zu fragen, wie dieser Sachverhalt als Grund erscheinen konnte, sind zwei unterschiedliche Dinge. Ich will nur die Antwort auf die erste Frage als eine „Handlungserklärung aus Gründen“ bezeichnen, da nur diese der Erklärungs- und Normativitätsbedingung genügt. Die Antwort auf die zweite Frage will ich eine „Klärung der Rahmenbedingungen“ nennen. Der Erklärungswert des Nennens von Rahmenbedingungen ist wesentlich abhängig von einer Handlungserklärung aus Gründen (Sachverhalten). Denn Rahmenbedingungen können nur dort Erklärungswert für eine Erklärung von Handlungen *qua Handlungen* haben, wo es eine solche Handlungserklärung gibt. Dieselben Tatsachen, welche als Rahmenbedingungen in Handlungserklärungen figurieren, können natürlich auch eigenständigen Erklärungswert besitzen, aber dann sind die zugehörigen Erklärungen keine Handlungserklärungen. Weiter unten werden ich näher auf diesen Themenbereich eingehen. Hier soll zunächst gefragt werden, wie das soeben Ausgeführte sich auf die SMH beziehen lässt.

5.3.2. Emotionale Bewertungen

Ich will kurz die Ergebnisse zusammenfassen, welche bis hierher erzielt wurden: Der Satz (PG) stellt eine Implikation der SMH dar und behauptet, dass emotionale Bewertungen von Handlungs-Ergebnis-Szenarien kausale Faktoren sind, welche das Entscheidungsverhalten beeinflussen. Verhaltensweisen, welche auf Entscheidungen beruhen, fallen unter den Begriff der Handlung. Deshalb müssen die Erklärungen solcher Verhaltensweisen Erklärungen von Handlungen sein. Erklärungen sind aber intensionale Kontexte, sodass die Erklärung von entscheidungsbasiertem Verhalten nicht nur eine Erklärung von Handlungen sein muss, sondern eine Erklärung von entscheidungsbasiertem Verhalten *qua Handlung*. Die Erklärung von Handlungen *qua Handlungen* sind Erklärungen aus Gründen. Bei Gründen lassen sich normative und motivierende Gründe unterscheiden. Da die menschliche Rationalität nicht rational in Frage gestellt werden kann, müssen normative Gründe unter den motivierenden Gründen sein können. Psychische Zustände kommen nicht als motivierende Gründe in Frage,

da sie keine normativen Gründe darstellen können.¹⁰⁰ Gründe sind Sachverhalte. Da aber psychische Zustände ohne Zweifel für unser Verständnis menschlicher Handlungen wichtig sind, muss ihnen ein Platz in der Erklärung von Handlungen gegeben werden. Dieser Platz liegt in den Rahmenbedingungen von Handlungserklärungen. Da emotionale Bewertungen ohne Zweifel psychische Zustände darstellen, folgt aus dem soeben Gesagten, dass emotionale Bewertungen Rahmenbedingungen von Handlungserklärungen darstellen. Diese These möchte ich nun unabhängig von den vorangegangenen Überlegungen erhärten. Diese Erhärtung wird zugleich als Beleg des soeben Ausgeführten gelten können.

5.3.2.1. Emotionen und Werturteile

Ich will hier folgende These vertreten: Emotionen konstituieren sich zum Teil durch Werturteile und Begründungsmuster. Ich will also Emotionen als wesentlich rational betrachten. Hiermit stimme ich Damasio darin zu, dass Emotionen ein wesentlicher Bestandteil der menschlichen Rationalität sind. Doch dieser Zusammenhang ist nicht empirischer Natur und beruht nicht auf Kausalrelationen.

Wie ich oben erläutert habe, ist Damasios Theorie der Emotionen im Wesentlichen eine Weiterentwicklung der James-Lange-Theorie. Damasio betrachtet, wie James, Emotionen im Wesentlichen als körperliche Reaktionen. Er schreibt: „*I see the essence of emotions as the collection of changes in body state*“ (Damasio, 2006). Viel entscheidender für eine Kritik dieser Position als die Unspezifität somatischer Reaktionen (vgl. Schachter & Singer, 1962) erscheint mir, dass Emotionen offensichtlich Intentionalität aufweisen. Emotionen sind typischerweise auf etwas gerichtet, sie haben ein Objekt (vgl. Kenny, 1963, S. 187ff.). „*It is uncontroversial that these experiences typically have an intentional content*“ (Tye, 2008, S. 23). Schon 1965 wies George Pitcher darauf hin, dass die Intentionalität von Emotionen ein entscheidendes Problem für alle Theorien aufwirft, welche Emotionen als ´nicht-kognitive Prozesse` betrachten. Darüber hinaus zeigte Pitcher, dass das kognitive Element, welches angenommen werden muss, um die Intentionalität von Emotionen zu erklären, den Emotionen nicht äußerlich (d. i. bloß kontingent mit ihnen verbunden) sein kann (vgl. Pitcher, 1965). Wir individuieren und identifizieren nämlich Emotionen anhand ihrer Intentionalität. Meine Wut auf Peter ist eine andere Wut als die auf Klaus, weil Peter und Klaus verschieden sind. Ich kann zwei Gründe haben, auf Klaus wütend zu sein, aber ich kann nicht zwei Wut-Emotionen auf Klaus haben. Eine Kausalrelation reicht hier schon deshalb nicht hin, weil das intentionale

¹⁰⁰ Dies gilt mit den oben genannten Ausnahmen.

Objekt von Emotionen nicht existieren muss. Ich kann mich vor einem Einbrecher fürchten, obwohl keiner da ist. Das intentionale Objekt einer Emotion kann also nicht mit ihrer Ursache identisch sein. Weitere Gründe, eine begriffliche Verbindung zwischen Emotionen und 'kognitiven Elementen' anzunehmen, bestehen darin, dass Emotionen Überzeugungen implizieren. Wenn ich auf Klaus wütend bin, weil er meine Geldbörse gestohlen hat, dann kann diese Wut nicht als Identische weiterbestehen, wenn ich herausfinde, dass meine Geldbörse nicht gestohlen wurde. Ich mag dann immer noch erregt sein, aber ich habe nicht mehr dieselbe Emotion. In diesen Zusammenhang gehört auch die Tatsache, dass wir Emotionen durch Kognitionen 'beeinflussen' können.

„Auf die Äußerung 'Ich kann nicht ohne Furcht daran denken...' antwortet man etwa: 'Es ist kein Grund zur Furcht, denn ...' Das ist jedenfalls *ein* Mittel, Furcht zu beseitigen, im Gegensatz zum Schmerz“ (BPP II, § 161, Wittgenstein, 1984, S. 249).

„Es ist als könnte man sagen: Furcht, oder gar Hoffnung, können geradezu aus Gedanken bestehen, aber doch nicht der Schmerz“ (BPP I, § 153, Wittgenstein, 1984, S. 247).

Emotionen müssen also ein 'kognitives Element' enthalten (vgl. Brandtstädter, 1987, S. 7). In jüngerer Zeit wurden jedoch unterschiedliche Vorschläge gemacht, wie dieses 'kognitive Element' der Emotionen zu charakterisieren sei.¹⁰¹ In der psychologischen Forschung wurden insbesondere im Zuge der sogenannten kognitiven Wende unterschiedliche Konzeptionen der kognitiven Grundlagen von Emotionen ausgearbeitet (vgl. Gessner, 2004, S. 42ff.). Eine der wohl einflussreichsten Theorien dieser Art wurde von Richard S. Lazarus vorgelegt. Dieser Theorie zufolge sind die entscheidenden kognitiven Elemente von Emotionen einerseits persönliche Ziele und andererseits Bewertungen (*appraisals*¹⁰²) (Lazarus & Lazarus, 1994, S. 139).¹⁰³

„To survive and flourish [we] need to decide whether events are or are not significant for [our] well-being, and in what way. Without personal significance, there is no emotion. Appraisal is an evaluative judgment about this significance. It is the main process of rea-

¹⁰¹ Für Literaturhinweise hierzu sei auf Solomon (2003, S. 3) verwiesen.

¹⁰² In deutschen Übersetzungen wird *appraisal* oft mit „Einschätzung“ wiedergegeben. Dies scheint mir nicht das zu treffen, was Lazarus sagen will. Er erklärt diesen Terminus etwa durch Wendungen wie „*personal meaning*“, „*evaluations*“ und „*evaluative judgment*“ (Lazarus & Lazarus, 1994, S. 290, 203), welche mehr enthalten als das Wort „Einschätzungen“.

¹⁰³ Ich will hier nicht auf die bekannte Lazarus-Zajonc-Debatte eingehen, da diese Debatte hauptsächlich auf einem zu engen Begriff von Kognitionen zu beruhen scheint. Zajonc scheint anzunehmen, dass Kognitionen Informationsverarbeitung in einem anspruchsvollen Sinne voraussetzen (vgl. Power & Dalgleish, 1997, S. 112). Es erscheint mir aber mehr als fragwürdig, ob ein Streit um den *terminus technicus* „Kognition“ uns an dieser Stelle weiterhilft.

soning on which emotions depend, the heart of the emotion process“ (Lazarus & Lazarus, 1994, S. 143)

Diese Position findet ihr philosophisches Gegenstück in der These Roberts C. Solomons: „[E]motions are evaluative judgments“ (Solomon, 2003, S. 3). Anders als Lazarus unterscheidet Solomon aber in seinen Untersuchungen genauer zwischen empirischen und begrifflichen Relationen. Eine kreative Ausarbeitung seiner These findet sich im Werk Martha C. Nussbaums. Nach Nussbaum stellen Emotionen Werturteile dar. Aber es sind nur bestimmte Urteile, welche wir als Emotionen erleben. Diese Urteile haben drei besondere Merkmale (vgl. Nussbaum, 2001, S. 4):

- (1) Sie beziehen sich typischerweise auf Dinge, welche wir für wertvoll oder wichtig achten.
- (2) Sie beziehen sich typischerweise auf Dinge, welche unser eigenes Wohlergehen, unser Gedeihen oder unsere Ziele und Vorhaben betreffen.
- (3) Sie beziehen sich typischerweise auf Dinge, welche nicht vollständig unter unserer Kontrolle stehen.

Da sich diese Werturteile auf das eigene Wohlergehen (in einem weiten Sinn) richten, spricht Nussbaum hier von *eudaimonistischen* Urteilen. „[An] eudaimonistic judgment must ultimately be formed in order for the emotion to occur“ (Nussbaum, 2001, S. 55). Mit dem Verweis auf den Begriff der *Eudaimonia* ist auch impliziert, dass das Wort Wohlergehen hier nicht als Lustgewinn zu verstehen ist. Vielmehr soll alles das angesprochen werden, was zu einem gelungenen Dasein gehört (vgl. Müller, 1982, S. 138f.). Wichtig ist hier insbesondere, dass sich durch das *eudaimonistische* Element Emotionen als zentriert auf eine Person verstehen lassen. Das Werturteil, welches von einer Emotion impliziert wird, bezieht sich auf das Wohl der jeweiligen Person und ist damit in einem relevanten Sinn nicht abstrakt oder allgemein. Das an dritter Stelle genannte Merkmal der Urteile, welche konstitutiv für Emotionen sind, lässt uns verstehen, wie Emotionen eine so entscheidende Rolle in unserem Leben spielen können. Da Emotionen sich typischerweise auf Dinge beziehen, welche sich nicht vollständig unter unserer Kontrolle befinden, stellen sie eine entscheidende Verbindung zu unserer Umwelt her. Emotionen verbinden uns mit den Aspekten und Dingen unserer Umwelt, um welche wir uns, um des Gelingens unseres Daseins willen, *sorgen* müssen.¹⁰⁴

Ich will kurz auf vier Kritikpunkte eingehen, welche gegen eine solche `kognitive` Emotionstheorie vorgebracht wurden:

¹⁰⁴ Lazarus spricht in diesem Zusammenhang von „relational meaning“ (Lazarus, 2006, S. 10).

(a) Manche Autoren kritisieren den dargestellten Ansatz aufgrund seiner aprioristischen Methode (vgl. Spicer, 2004). Dabei wird angenommen, dass es keinen Unterschied zwischen empirischen und begrifflichen Relationen gebe und daher jede begriffliche Analyse wertlos sei. Wie zu Beginn dieser Arbeit dargestellt, teile ich diese Ansicht nicht und kann auch nicht erkennen, wie Quines Arbeiten – welche in diesem Zusammenhang meist angeführt werden – hier entscheidende Argumente liefern können.

(b) Ronald de Sousa erachtet es als entscheidendes Problem kognitiver Emotionstheorien, dass es nicht möglich ist, Emotionen hypothetisch zu fühlen, während man dagegen Überzeugungen und Urteile hypothetisch annehmen kann (vgl. de Sousa, 1987, S. 156). Hierauf ist zu sagen, dass es einen Unterschied gibt zwischen dem Erwägen eines Urteils, was hypothetisch sein kann, und dem Fällen des Urteils. Wenn ich ein Urteil fälle, so muss ich von der Wahrheit des Geurteilten überzeugt sein, und dies ist auch bei Emotionen der Fall. Es besteht hier also kein Unterschied (vgl. Nussbaum, 2001, S. 78).

(c) Andere Autoren kritisieren, dass Emotionen ein anderes phänomenales Erleben beinhalten als Urteile oder Überzeugungen (vgl. Goldie, 2000, S. 40). Das angesprochene phänomenale Erleben wird dabei meist als Gefühl bezeichnet. Ich betrachte diesen Einwand als grundsätzlich richtig. Emotionen fühlen wir, Urteile nicht. Hier sind aber einige Klärungen nötig. Erstens will ich, anders als Nussbaum und Solomon, nicht die These vertreten, Emotionen seien auf Urteile reduzierbar. Wir haben unterschiedliche Kriterien zur Zuschreibung von Emotionen und allen diesen Kriterien sollte Bedeutung eingeräumt werden. Es ist aber wichtig zu sehen, dass Urteile ein wesentlicher 'Bestandteil' von Emotionen sind. Zweitens muss zwischen Emotionen als Gemütsbewegungen und Emotionen als Gemütsdispositionen unterschieden werden (vgl. Müller, 1998, S. 133).¹⁰⁵ Emotionen als Gemütsbewegungen (z. B. der Ärger, der mich dazu bringt, etwas Unvernünftiges zu sagen) beinhalten typischerweise das phänomenale Erleben, welches wir ein Gefühl nennen. Emotionen als Gemütsdispositionen beinhalten dies nicht. Beispielsweise beinhaltet mein jahrzehntelanger Ärger über die Vereitelung einer großen Chance kein spezifisches phänomenales Erleben. Auch wenn dieser Ärger sich in einem solchen Gefühl zeigen mag (z. B., wenn ich meinen alten Feind wiedersehe). Dass unsere Worte zur Bezeichnung von Emotionen hier systematisch zwei-

¹⁰⁵ Es ist hier wichtig, Emotionen als Gemütsdispositionen nicht mit Dispositionen zu einer Emotion zu verwechseln. Erstere besitzt ein konkretes Objekt, letztere ist eine Neigung, bei unterschiedlichen Anlässen etwas als ein geeignetes Objekt anzuerkennen. Deshalb ist diese Unterscheidung auch nicht mit der gängigen *state/trait* Unterscheidung identisch. Die angesprochene Unterscheidung wird in der Psychologie im Allgemeinen nicht beachtet. Hierdurch kommt es, insbesondere in der Persönlichkeitspsychologie, immer wieder zu einem Mangel der Unterscheidung zwischen Eigenschaften, welche Dispositionen sind, und Dispositionen zu Eigenschaften.

deutig sind, deutet darauf hin, dass das phänomenale Erleben von Gefühlen nicht so zentral für unseren Emotionsbegriff ist, wie oft angenommen wird. Das, was beide Emotionsarten verbindet, liegt offensichtlich im kognitiven Gehalt von Emotionen.

(d) Ein vierter Kritikpunkt lautet: Emotionen können bestehen, obwohl die Person die zugehörigen Urteile nicht teilt. So ist es möglich, dass eine Angst bestehen bleibt, auch wenn die Person überzeugt ist, dass keine Gefahr droht, oder dass jemand sich aufgrund des Nicht-Erhaltens einer Auszeichnung gekränkt fühlt, obwohl er glaubhaft versichert, er halte jene Auszeichnung für völlig irrelevant (vgl. DeLancey, 2002, S. 31ff.). Um diesem Kritikpunkt auszuräumen, ist es wichtig, zu sehen, dass die Möglichkeit von Inkonsistenz nicht *gegen*, sondern *für* einen kognitiven Gehalt spricht. Die Situation, in der ich Angst vor dem Fliegen habe, obwohl ich diese Angst für völlig unbegründet halte, sind deshalb so interessant, weil sich hierin eine Inkonsistenz meiner Bewertungen ausdrückt. Eine Inkonsistenz setzt aber einen kognitiven Gehalt voraus. Denn wie sollte meine Überzeugung meiner Emotion widersprechen, wenn diese kein Urteil beinhaltet (vgl. Nussbaum, 2001, S. 36)? Dass Emotionen mit Überzeugungen konfliktieren können – wie andere Urteile auch – spricht also in Wirklichkeit *für* und nicht *gegen* meine These. Doch die inhärent rationale Struktur von Emotionen geht über die Implikation von Urteilen hinaus: Emotionen beinhalten Begründungsmuster. Was dies heißt, soll nun näher untersucht werden.

5.3.2.2. Emotionen und Gründe

Ich habe dargelegt, dass Emotionen Werturteile implizieren. Dies heißt, dass Emotionen ihren Gegenstand gemäß einem solchen Urteil präsentieren. „Die Furcht präsentiert ihn als bedrohlich, der Stolz als Leistung oder Verdienst, die Bewunderung als großartig, die Liebe als liebenswert“ (Müller, 1998, S. 134). Emotionen implizieren aber auch die Anerkennung ihres Gegenstandes als Handlungsgrund.

„Gemeinsam ist demnach allen Emotionen folgende quasi-rationale Struktur: 1. Das *emotionale Urteil* schreibt dem Gegenstand der Emotion eine bestimmte Bedeutung zu. 2. Nach einem für die jeweilige Emotion charakteristischen *Begründungsmuster* wird diese Bedeutung zum *Motiv oder Grund*, sich soundso zu verhalten“ (Müller, 1998, S. 135).

Demnach impliziert z. B. Furcht einerseits das Urteil, dass das gefürchtete Objekt bedrohlich ist und andererseits die Anerkennung dieser Bedrohlichkeit als einen Handlungsgrund. Emotionen beinhalten also ein Muster praktischen Folgerns. Eine solche Folgerung könnte beispielsweise folgendermaßen aussehen:

Tante Erna hat mir ein Geschenk geschickt.

Also: Ich schreibe Tante Erna eine Dankeskarte.

Die Emotion, welche dieser Folgerung entspricht, nennen wir Dankbarkeit. D. h., Dankbarkeit besteht zum Teil im Ziehen dieser praktischen Folgerung.¹⁰⁶ Durch den Rückgriff auf Begründungsmuster ergibt sich also eine Verbindung zwischen Emotionen und Handlungstendenzen. Dies ist ein Vorzug der vorgestellten Theorie, da eine Explikation dieser Verbindung von jeder Theorie der Emotionen erwartet werden muss (vgl. Frijda, 2004). Bei dem eben genannten Beispiel fällt sofort auf, dass die Emotion nicht in der praktischen Folgerung auftaucht. Die Emotion der Dankbarkeit ist kein Glied des praktischen Schlusses, in welchem sie sich manifestiert. Daher kann die Emotion auch nicht in der Handlungserklärung genannt werden, welche auf dem praktischen Schluss basiert. Dies könnte verwirren, ist aber vollkommen folgerichtig. Wie ich oben gezeigt habe, bestehen Gründe typischerweise in Sachverhalten und nicht in psychischen Zuständen der Handelnden. Diese psychischen Zustände sind Rahmenbedingungen. Dies impliziert, dass auch Emotionen typischerweise Rahmenbedingungen und nicht Gründe darstellen. Dies heißt: Das Nennen von Emotionen beantwortet in der Regel nicht die Frage „Was war sein Grund zu handeln? / Warum hat er das getan?“, sondern die Frage „Wie konnte ihm das als ein Grund erscheinen? / Wie kam er von diesen Gründen zu dieser Handlung?“.

Ich will auch hier wieder meine These verdeutlichen, indem ich kurz auf zwei mögliche Einwände eingehe: (a) Ein erster Einwand könnte lauten, dass die genannte praktische Folgerung keine Folgerung ist, da sie nicht formal gültig ist. Hierauf lässt sich antworten, dass die *formale* Gültigkeit des praktischen Syllogismus generell kontrovers diskutiert wurde (vgl. Anscombe, 2000, S. 57ff.; von Wright, 1984, S. 93ff.). In diesem Zusammenhang wurde ein hilfreicher Vorschlag von Brandom vorgelegt:

„Wir müssen nicht alle korrekten Inferenzen als solche behandeln, die aufgrund ihrer Form korrekt sind und implizite oder unterdrückte Prämissen liefern, die das erforderliche logische Vokabular enthalten. Statt dessen können wir Inferenzen wie die von ‘Pittsburgh liegt westlich von Philadelphia’ auf ‘Philadelphia liegt östlich von Pittsburgh’ [...] als *material* gute Inferenzen behandeln, d. h. als Inferenzen, die aufgrund des Gehalts ihres

¹⁰⁶ Das praktische Schlussfolgern kann als jene Leistung unserer Vernunft betrachtet werden, welche unsere Handlungen an Maßstäben der Rationalität ausrichtet. Typische Beispiele sind zweckrationale Überlegungen, wie etwa die Überlegung, welche vom Erkennen eines Schadens meines Fahrrads zum Einkauf bestimmter Ersatzteile führt. Was ein praktischer Schluss letztlich ist, wird kontrovers diskutiert (vgl. Anscombe, 2000, S. 57ff., §§ 33ff.).

nichtlogischen Vokabulars korrekt sind. Ich schlage vor, sich beim Nachdenken über praktische Inferenzen dieser nichtformalistischen Strategie anzuschließen“ (Brandom, 2001a, S. 112f.).

Hier wird eine wichtige Parallele zwischen einigen theoretischen und praktischen Folgerungen angesprochen: Folgerungen wie „Bello ist ein Rüde, also ist Bello ein Hund“ sind nicht *formal* gültig. Sie beziehen ihre Gültigkeit aus ihrem semantisch-grammatischen Gehalt. Der Schluss könnte *formale* Gültigkeit beanspruchen, wenn wir eine Prämisse hinzufügen, z. B. „Alle Rüden sind Hunde“. Wenn wir dies aber tun, so gleichen wir die Inferenz einem Schluss folgenden Musters an: „Dieses Buch gehört Klaus. Jedes Buch, das Klaus gehört, ist rot. Also: Dieses Buch ist rot“. Es ist aber offensichtlich, dass die beiden Folgerungen „Bello ist ein Rüde, also ist er ein Hund“ und „Dieses Buch gehört Klaus, also ist es rot“ sich grundlegend unterscheiden. Während der erste Schluss allein aufgrund seines Inhalts richtig ist,¹⁰⁷ stellt der zweite bestenfalls ein Enthymem dar. Folgerungen als *formal* gültig zu betrachten, ist also nicht die einzige Möglichkeit, ihnen gerechtfertigterweise Gültigkeit zuzuschreiben.

(b) Ein zweiter Einwand könnte lauten, dass ein praktischer Syllogismus eine konative und eine kognitive Prämisse benötigt, und die Emotion das konative Element liefert. Hierauf habe ich bereits oben reagiert, indem ich auf Dancys Erklärungs- und Normativitätsbedingung hingewiesen habe. Psychische Zustände können nicht als Gründe betrachtet werden, da sie keine guten Gründe sein können. Gründe und gute Gründe müssen aber identisch sein können, wenn die menschliche Rationalität nicht bestritten werden soll. Psychische Zustände – wie Emotionen – sind Rahmenbedingungen. Emotionen beinhalten Begründungsmuster. Das Begründungsmuster einer Schlussfolgerung aber ist nicht Teil der Folgerung. Im Schluss „ $p \wedge (p \supset q) \vdash q$ “ fehlt nicht die Angabe des *modus ponens* Musters. Anderenfalls ergäbe sich Lewis Corrolls bekanntes Problem Achills mit der Schildkröte, nämlich das des infiniten Regresses hinsichtlich jeder Schlussfolgerung (vgl. Carroll, 1895). Und so wie hier nicht die Angabe des Begründungsmusters fehlt, so fehlt sie auch nicht im Fall von Emotionen. Natürlich kann die Nennung des Begründungsmusters uns zu entscheidenden Erkenntnissen verhelfen, wie etwa wenn wir die Gültigkeit eines Schlusses nicht einzusehen vermögen, bis uns klar wird, dass es sich um einen äußerst komplizierten *modus ponens* handelt. Aber hieraus folgt nichts bezüglich der Vollständigkeit von Begründungen.

¹⁰⁷ s. Kapitel 2.

5.3.2.3. Bewertungen

Ich möchte das Ergebnis der bisherigen Überlegungen zusammenfassen: Ich habe dafür argumentiert, dass Emotionen auf begrifflicher Ebene zwei kognitive Elemente beinhalten: Werturteile und Begründungsmuster. Die Begründungsmuster, welche Emotionen beinhalten, sind abhängig von den Werturteilen, die sie implizieren. Die Furcht präsentiert einen Gegenstand als bedrohlich und beinhaltet hierdurch ein Begründungsmuster folgender Art: „Ich bin konfrontiert mit X“, also „Ich werde fliehen“. Da Emotionen Werturteile implizieren, stellt die Formulierung „emotionale Bewertung“ in einem gewissen Sinne einen Pleonasmus dar. Handlungserklärungen sind Erklärungen aus Gründen. Emotionen – ebenso wie andere psychische Zustände – spielen in dieser Art von Erklärung die Rolle von Rahmenbedingungen. Sie tauchen nicht direkt in einer Erklärung aus Gründen auf, sondern werden angegeben, um den Zusammenhang zwischen Gründen und Handlungen weiter zu erläutern (z. B., indem sie das vorliegende Begründungsmuster nennen). Dies heißt nicht, dass eine Erklärung aus Gründen ohne solche weiteren Erläuterungen unvollständig wäre.

Wie hängen diese Ergebnisse nun mit der SMH zusammen? Die SMH beansprucht, Handlungen zu erklären, welche auf Entscheidungen beruhen. Da sie Handlungen *qua Handlungen* erklären will, muss sie Erklärungen aus Gründen bieten. Die SMH will diese Erklärungen aus Gründen auf neurobiologische Erklärungen reduzieren. Das zu reduzierende Erklärungsschema besteht in (PG). (PG) muss damit ein kausales Erklärungsschema darstellen, das Erklärungen aus Gründen bietet. Als Antecedensbedingung einer entsprechenden Erklärung gibt (PG) emotionale Bewertungen an. Es ist nun aber klar, dass emotionale Bewertungen nicht die Rolle von Gründen übernehmen können. Sie sind Rahmenbedingungen für Erklärungen aus Gründen. Damit kann (PG) keine Erklärungen aus Gründen liefern. Hiermit hat sich die SMH als höchst problematisch erwiesen, denn es bleiben nur zwei Strategien zu ihrer Verteidigung.

Erstens könnten die Vertreter der SMH den Anspruch aufgeben, Handlungserklärungen zu bieten. Dann ist aber nicht klar, inwieweit wir es hier mit einer psychologischen Theorie zu tun haben. Denn wenn Menschen aus Gründen handeln können – und das müssen sie können, wenn Wissenschaft einen Sinn haben soll¹⁰⁸ –, dann lassen sich menschliche Handlungen durch Gründe erklären. Dann gibt es aber potenziell zwei Erklärungen für eine Handlung, welche auf einer Entscheidung beruht. Erstens eine Erklärung der Handlung *qua Handlung* und zweitens eine biologische Erklärung der Handlung, wie sie die SMH uns anbietet. Damit

¹⁰⁸ s. Kapitel 5.3.1.2.

wäre eine Reduktion nicht gelungen. Wir können dann natürlich die biologische Erklärung für interessant erachten, aber es ist nicht klar, inwiefern es sich hier um eine genuin psychologische Erklärung handelt. Diese Antwortmöglichkeit kommt einer so starken Schwächung der SMH gleich, dass ich sie im Folgenden vernachlässigen möchte.

Zweitens könnten die Vertreter der SMH den Anspruch darauf aufrechterhalten, eine wesentliche Verbindung zwischen Handlungserklärung und Neurobiologie aufgezeigt zu haben. Zwar gäbe (PG) kein Schema für Erklärungen aus Gründen an, aber eben ein Schema zur weiteren Erläuterung solcher Erklärungen. Diese Position ist insbesondere deshalb plausibel, da die SMH auf die These verpflichtet ist, dass (PG) ein Kausalverhältnis angibt. Damit ließe sich argumentieren, dass es zwar sein mag, dass unsere Handlungserklärungen aus Gründen nicht direkt auf psychische Zustände Bezug nehmen müssen, dies aber nicht entscheidend sei, da die kausal relevante Rolle eben doch von den Rahmenbedingungen gespielt würde. Wir hätten es dann natürlich nicht mit adäquaten Handlungserklärungen zu tun.¹⁰⁹ Doch die genannte Position wird demjenigen, der die Psychologie als Naturwissenschaft betrachtet, dennoch attraktiv erscheinen. Deshalb möchte ich im Rest dieses Kapitels gegen diese zweite Antwortmöglichkeit argumentieren und zeigen, dass angebliche psychische Zustände oder Prozesse wie Absichten auch unabhängig von der bisher vorgetragenen Argumentation nicht als Ursachen von Handlungen betrachtet werden können.

5.4. Rahmenbedingungen und Kausalerklärungen

Wenn die SMH eine Erklärung von Handlungen, welche auf Entscheidungen beruhen, *qua Handlungen* geben soll, dann muss (PG) ein Kausalverhältnis zwischen Rahmenbedingungen und Handlungen angeben. Hiergegen will ich in diesem Abschnitt argumentieren. Damit wird der dritte und letzte Schritt der Begründung der Prämisse (T2.2) vollzogen. Meine These lautet hier also: Rahmenbedingungen von Handlungserklärungen können Handlungen nicht kausal erklären.¹¹⁰

5.4.1. Der Kausalismus

Wie oben bereits erwähnt, ist es seit Davidsons einflussreichem Aufsatz *Actions, Reasons, and Causes* von 1963 eine weitverbreitete Auffassung, dass Handlungserklärungen kausale Erklärungen darstellen. Da aber, laut dem Kausalismus, Handlungserklärungen Erklärungen durch Hinweis auf psychische Zustände oder Prozesse sind, solche Zustände oder Prozesse

¹⁰⁹ s. Kapitel 5.3.1.3.

¹¹⁰ s. S. 126 oben.

aber tatsächlich lediglich die Rahmenbedingungen von Handlungserklärungen darstellen, kann die Kernthese des Kausalismus folgendermaßen formuliert werden: Handlungen werden durch die Rahmenbedingungen der jeweiligen Handlungserklärungen verursacht. Wobei diese Formulierung dem Kausalismus selbst natürlich nicht möglich ist, da er den Unterschied zwischen Handlungserklärungen und der Klärung von Rahmenbedingungen übersieht.¹¹¹ Im Folgenden will ich die Unhaltbarkeit des Kausalismus darlegen.

5.4.1.1. Davidson's challenge

Als Kernargument für den Kausalismus gilt allgemein ein Argument, welches in dem soeben genannten Aufsatz von Davidson erstmals vorgebracht wurde und oft als *Davidson's challenge* bezeichnet wird (vgl. Dickenson, 2007; Mele, 2003, S. 38ff.; Risjord, 2005; Schueler, 2003, S. 8ff.; Tanney, 1995). Dieses Argument kann in der Form einer Frage formuliert werden: Wenn eine Person zwei normative Gründe für eine Handlung anerkennt, aber die Handlung nur aus einem der beiden Gründe vollzieht, was zeichnet dann den motivierenden Grund aus? Die Antwort des Kausalisten hierauf ist klar: Es ist genau derjenige Grund der motivierende, welcher die Ursache der Handlung ist. Die angebliche '*challenge*' für den intentionalistischen Handlungstheoretiker besteht nun darin, erklären zu können, wie nur einer von zwei anerkannten normativen Gründen zu einem motivierenden Grund werden kann, ohne kausale Begriffe zu verwenden.

„[E]s kann sein, daß man einen Grund für eine Handlung hat und diese Handlung auch ausführt, ohne daß dieser Grund derjenige ist, weshalb man die Handlung vollzogen hat. Wesentlich für die Beziehung zwischen einem Grund und einer durch ihn erklärten Handlung ist die Vorstellung, daß der Handelnde die Handlung ausgeführt hat *weil* er diesen Grund hatte. [...] Wenn kausale Erklärungen [...] 'für das angestrebte Verständnis' menschlichen Handelns 'völlig belanglos sind', fehlt uns eine Analyse des 'weil' in Sätzen wie 'Er tat es, weil ...' (mit anschließender Angabe des Grundes)“ (Davidson, 1998d, S. 28ff.).

Obwohl dieses Argument vielfach für überzeugend gehalten wurde, besitzt es entscheidende Schwächen. Ich möchte hier vier anführen:

¹¹¹ Ich werde im folgenden Abschnitt die Terminologie des Kausalismus übernehmen und dort von Gründen sprechen, wo wir es tatsächlich mit Rahmenbedingungen zu tun haben. Dies ist ein Zugeständnis an den Kausalismus, um diesem maximale Rationalität zu unterstellen (s. die Anmerkungen zum *Principle of Charity* oben (S. 73 und S. 133)).

(a) *Davidson's challenge* ist ein Schluss auf die beste Erklärung (vgl. Dickenson, 2007). Es wird dabei auf das Problem hingewiesen, wie man 'wirksame' Gründe von 'unwirksamen' Gründen unterscheiden kann, um dann einen Kausalnexus als die beste Lösung dieses Problems anzubieten. Hilft aber ein Kausalnexus hier tatsächlich weiter?

„To be told by Davidson that efficacious reasons are causes and non-efficacious reasons are not causes is fine and good, but this does not yet tell us anything about what we really want to know. How do we make sense of cases in which an agent has reasons that are non-efficacious (not causes)?“ (Dickenson, 2007, S. 15).

Wie Jason Dickenson überzeugend darlegt, muss der Kausalist hier einen Begriff motivationaler Stärke einführen. Der Begriff motivationaler Stärke ist aber dem Intentionalisten ebenso zugänglich wie dem Kausalist; denn ob das Phänomen der Motivation kausaler Natur ist, steht ja genau in Frage. Tatsächlich ist also völlig unklar, wer hier die beste Erklärung bieten kann. Aber damit verliert ein Schluss auf die beste Erklärung sämtliche Kraft.

(b) Julia Tanney (1995) weist darauf hin, dass die von uns akzeptierte Erklärung hinsichtlich der Frage, wie die vom Handelnden anerkannten normativen Gründe nicht die motivierenden Gründe seines Handelns sein können, keinen Bezug auf Verursachung nehmen muss. Wenn die Situation tatsächlich auftritt, in der eine Person einen Grund anerkennt, die entsprechende Handlung vollzieht, aber nicht aus dem entsprechenden Grund gehandelt hat, so verlangen wir eine ausgefeiltere Interpretation, eine Geschichte über das Abwägen konkurrierender Gründe, das Offenlegen einer Wertschätzung oder Ähnliches, aber nicht eine differenzierte Kausalanalyse (vgl. Tanney, 1995). Tatsächlich ist es mehr als fragwürdig, ob wir durch irgendeine neurobiologische Kausalverbindung Zweifel an der Wahrheit der Ausführungen unseres aufrichtigen Gesprächspartners aufkommen ließen, wenn uns dieser sagt, er habe sein Erspartes auf die Bank gebracht, weil er einen Einbruch fürchte, aber nicht weil er auf die Zinsen aus war, obwohl er beides als völlig überzeugende Gründe anerkennt und beide Gründe ihm bewusst waren. Würden wir hier an der Aufrichtigkeit unseres Gesprächspartners zweifeln aufgrund der Ergebnisse eines neurobiologischen Messverfahrens? Ich denke, es wäre in diesem Fall angemessener, die neurobiologische Theorie zu überdenken, denn wir hätten es mit einer möglichen Falsifikationsinstanz zu tun.¹¹²

¹¹² Wie sollten wir zu einer entsprechenden neurobiologischen Theorie kommen, ohne Korrelationen zu finden zwischen – letztlich auf Alltagssprachlichen Begriffen beruhenden – unzweifelhaften Feststellungen bezüglich des Vorliegens bestimmter Gründe-Konstellationen und neurobiologischer Tatsachen? Unser Alltagsbegriff

(c) Das wohl bekannteste Problem des Kausalismus besteht in den sogenannten 'abweichenden Kausalketten'. Kausalisten behaupten, dass vom Handelnden anerkannte potenzielle Gründe genau dann motivierende Gründe sind, wenn diese Gründe die Ursachen der entsprechenden Handlung sind. Es hat sich aber als relativ leicht herausgestellt, Fälle zu konstruieren, in welchen Personen gewisse Wunsch-Überzeugungs-Paare besitzen, die vom Kausalisten als Gründe akzeptiert werden müssen, und in denen diese Wunsch-Überzeugungs-Paare die Handlung auch *verursachen*, wir aber nicht sagen, dass die Person aus diesem *Grund* gehandelt hat (vgl. Keil, 2000, S. 72ff.). Das bekannteste Beispiel ist folgendes: Ein Bergsteiger will sich seines Kameraden entledigen. Dieser hängt an einem Seil und der Bergsteiger weiß, dass er seinen Kameraden durch das Loslassen des Seiles töten kann. Durch diese Situation ist der Bergsteiger aber so nervös, dass seine Hände anfangen zu zittern und er seinen Kameraden loslässt (vgl. Davidson, 1998e, S. 121f.). Dieser Fall zeigt, dass eine kausale Relation allein nicht ausreicht, um zu bestimmen, wann eine Handlung aus einem bestimmten Grund heraus vollzogen wurde. Der Kausalismus benötigt deshalb eine Spezifikation. Ein Wunsch-Überzeugungs-Paar muss eine Handlung nicht nur verursachen, sondern auch *auf die richtige Weise* verursachen. Dies hat sich als eine nicht lösbare Aufgabe entpuppt. Davidson schreibt (in einem späteren Aufsatz): „Ich halte es für aussichtslos, im einzelnen angeben zu wollen, in welcher Weise Handlungen durch Einstellungen bewirkt werden müssen, damit die Handlung durch sie rationalisiert wird“ (Davidson, 1998e, S. 121). Bis heute konnte der Kausalismus keine befriedigende Antwort auf diese Frage geben (vgl. Sehon, 1997).

(d) Über den vorigen Punkt hinaus ergibt sich aber aus der Möglichkeit abweichender Kausalketten für *Davidson's challenge* ein weiteres Problem. Davidsons Argument basiert darauf, keine teleologischen Lösungen zuzulassen für das Problem, wie anerkannte normative Gründe nicht unter den motivierenden Gründen sein können. Denn es ist natürlich ein Leichtes, anerkannte normative Gründe und motivierende Gründe zu unterscheiden, wenn Formulierungen wie „um zu“ erlaubt sind. Es wird argumentiert, dass teleologische Begriffe unklar seien, und Handlungserklärungen ausschließlich mit den Begriffen „Proeinstellung/Wunsch“, „Überzeugung“ und „Kausalität“ auszukommen haben (vgl. Owens, 1998). Dies ist eine für den Kausalismus entscheidende Annahme (vgl. Mele, 2003, S. 38). Denn

eines Grundes wird nicht daran gemessen, ob er mit bestimmten neurobiologischen Ergebnissen übereinstimmt, sondern umgekehrt. (In diesem Zusammenhang wird oft von „Inhaltsvalidität“ der Messinstrumente gesprochen.) Damit ist hier aber ein völliges Divergieren alltäglicher Auffassungen und wissenschaftlicher Ergebnisse gar nicht möglich (s. Kapitel 4.3.2).

anderenfalls ist nicht einzusehen, warum der Intentionalist sein „um zu“-Vokabular nicht gebrauchen darf. Es hat sich aber in den letzten Jahrzehnten gezeigt, dass der Kausalismus, um auch nur einen Ansatzpunkt für die Lösung des Problems abweichender Kausalketten zu gewinnen, Begriffe einsetzen muss, welche über das Vokabular von Wünschen, Überzeugungen und Kausalität hinausgeht (vgl. Owens, 1998). Damit wird genau die Beschränkung, welche es dem Intentionalisten schwer gemacht hat, auf *Davidson's challenge* zu antworten, von kausalistischer Seite selbst nicht mehr anerkannt.

Zusammenfassend lässt sich sagen, dass heute eine nicht-kausalistische Position wesentlich bessere Chancen hat, das Problem, welches durch Davidson aufgeworfen wurde, zu lösen als eine kausalistische Position. *Davidson's challenge* ist zu einer Herausforderung an den Kausalismus selbst geworden. Doch darüber hinaus hat der Kausalismus hierdurch sein wichtigstes Argument eingebüßt. Ohne dieses Argument scheint die These, dass Handlungserklärungen Kausalerklärungen seien, keine große Attraktivität zu besitzen. Doch um nicht vorschnell eine weitverbreitete Position zu verwerfen, will ich einige weitere Argumente für den Kausalismus näher betrachten.

5.4.1.2. Weitere Argumente für den Kausalismus

Es ist bemerkenswert, dass der Kausalismus seit den achtziger Jahren keine wesentlichen neuen Argumente hervorbringen konnte (vgl. Hursthouse, 2000, S. 96, 105). Eine hilfreiche Auflistung der Argumente für den Kausalismus findet sich bei Raimo Tuomela (1977, S. 221ff.). Ich werde mich an dieser Liste orientieren und Schwächen der einzelnen Argumente aufzeigen.

(a) Der Funktionalismus war in der zweiten Hälfte des zwanzigsten Jahrhunderts die dominante Auffassung bezüglich psychischer Phänomene (vgl. Botterill & Carruthers, 1999, S. 8ff.; Levin, 2004). Dabei wird behauptet, dass psychische Zustände durch ihre kausale Rolle definiert werden können. Eine offensichtliche Möglichkeit, den Begriff „Absicht“ oder „motivierender Grund“ funktionalistisch zu definieren, besteht darin, genau denjenigen mentalen Zustand „Absicht“ zu nennen, der typischerweise eine entsprechende Handlung *verursacht*.¹¹³ Solange der Funktionalismus als eine überzeugende Position bezüglich psychischer Zustände betrachtet wird, lässt sich aus ihm also ein Argument für den Kausalismus gewinnen (vgl. Tuomela, 1977, S. 221f.). Dieses Argument ist aber deshalb nicht trag-

¹¹³ Der Funktionalismus stimmt mit der hier vertretenen These insofern überein, als auch gemäß dem Funktionalismus Behauptungen bezüglich der Verbindung von Absichten/Gründen und Handlungen empirisch gehaltslos sind. Absichten bzw. Gründe werden ja genau anhand dieser Verbindungen definiert (vgl. Tuomela, 1977, S. 221).

fähig, weil der Funktionalismus keine überzeugende Position darstellt. Ich möchte hier nur auf einige wenige Probleme des Funktionalismus hinweisen. Erstens kann der Funktionalismus – wie Putnam gezeigt hat – nicht die Referenz mentaler Zustände fixieren, diese Zustände werden aber gerade anhand ihrer Referenz individuiert (vgl. Shagrir, 2005, S. 235). Zweitens bedarf der Funktionalismus weitgehender Anpassungen, wenn man eine Form des – heute gut etablierten – Externalismus bezüglich psychischer Inhalte vertreten will (vgl. Burge, 1979). Drittens hat sich gezeigt, dass der Funktionalismus kaum den Vorwurf ausräumen kann, eine Spielart des Epiphänomenalismus darzustellen (vgl. Kim, 1993, S. 358ff., 1998, S. 55, 2007). Damit steht aber genau die kausale Relevanz mentaler Zustände infrage. Viertens wurde seit Langem bemerkt, dass der Funktionalismus die Normativität mentalen Inhalts nicht erklären kann (vgl. Boghossian, 2003b). „[C]ausal-role functionalism does not have the resources to account for the normativity that is central to intentionality“ (Okrent, 2007, S. xi).

(b) Tuomelas zweites Argument lautet: Bestimmte Handlungen, wie das Folgern mittels *modus ponens*, können nicht intentionalistisch durch Gründe erklärt werden, da sich sonst ein infinites Regress ergibt. Beispielsweise können wir bei der Handlung des praktischen Folgerns nicht fragen, aus welchen Gründen sie vollzogen wurde, da die Antwort wieder eine praktische Folgerung nennen müsste, bei welcher dieselbe Frage auftauchen würde. Ein praktischer Syllogismus beruht nicht auf einem anderen praktischen Syllogismus, welcher ersteren als geboten oder ratsam erweist. Dies ist ohne Zweifel wahr, denn sonst wäre ein infinites Regress unvermeidlich; doch ziehen die Kausalisten hieraus falsche Schlüsse. Der Kausalist betrachtet die Kausalerklärung als die einzige mögliche Erklärung für eine Folgerung, welche nicht nach dem Muster des praktischen Syllogismus erklärt werden kann (vgl. Tuomela, 1977, S. 224). Aber diese Position wirft mehr Probleme auf, als sie löst. Denn ihr zufolge wären praktische Schlussfolgerungen nicht zielgerichtet. Gleiches gälte für das theoretische Schließen. Es könnte sich in beiden Fällen nicht um ein rationales Tun handeln, sofern darunter etwas verstanden wird, was sich von Prozessen in der unbelebten Natur abgrenzen lässt. Rational ist eine Handlung für den Kausalisten ja genau dann, wenn sie durch das richtige Wunsch-Überzeugungs-Paar verursacht wurde. Durch Müller (1992) wurde überzeugend dargelegt, dass genau die angesprochenen Fälle gute Argumente für die Möglichkeit einer Finalität ohne Intentionalität liefern.

„Unreasoned teleology has to be accepted if syllogistic teleology is. This is perhaps most obvious when we consider the example of a practical reflection which is not itself performed as a practical conclusion of an antecedent reflection. Such a practical reflection

occurs *for the sake of* the purpose which it is about, no less than its practical conclusion (some action) is going to occur *for the sake of* this purpose. Yet, *ex hypothesi*, this action comes under syllogistic teleology, the reflection does not“ (Müller, 1992, S. 179).

Tuomelas Argument muss daher letztlich als ein Argument gegen eine mechanistische Weltanschauung und damit auch gegen den Kausalismus verstanden werden. Dieses Ergebnis ist natürlich für eine Position, welche teleologische Erklärungen nur anerkennt, sofern sie sich in Kausalerklärungen übersetzen lassen, höchst problematisch. Aber dies ist kein Problem des Intentionalismus, sondern des Kausalismus.

(c) Ein drittes Argument besteht in der Kritik, dass das intentionalistische Erklärungsschema nicht streng allgemeingültig sei (vgl. Tuomela, 1977, S. 225). Tatsächlich ist aber auch dies ein Vorzug der intentionalistischen Erklärung. Denn es ist eine bekannte Tatsache, dass die psychologische Forschung bis heute nicht in der Lage war, streng allgemeingültige Gesetze aufzustellen (vgl. Keil, 2000, S. 116; McCrae & Costa, 1995). Darüber hinaus wurden insbesondere von Davidson selbst Argumente dafür vorgebracht, warum solche streng allgemeingültigen Gesetze für die Psychologie prinzipiell nicht aufgestellt werden können. Er kommt, durch Überlegungen im Zusammenhang mit dem *Principle of Charity*, zu dem Schluss: „Es gibt gar keine strikten Gesetze, auf deren Basis wir geistige Phänomene prognostizieren und erklären können“ (Davidson, 1998b, S. 315). Wie auch immer man hier einzelne Argumente bewerten mag, scheint es doch mehr als zweifelhaft, ob sich aus dem Faktum, dass intentionalistische Erklärungen nicht streng allgemeingültig sind, ein Argument für den Kausalismus gewinnen lässt.

(d) Als ein weiteres Argument für den Kausalismus wird oft angeführt, dass er Erklärungen liefere, welche ´tiefer` in die Natur eindringen. „A non-causal theory might give [...] ´verbal` answers [...]. Our causal theory, on the contrary, looks for the ontically and dynamically real solution to the problem“ (Tuomela, 1977, S. 226). Für ein naturalistisch gesinntes Publikum mag sich dies plausibel anhören, es ist aber schwer zu erkennen, ob sich mehr hinter dieser Bemerkung verbirgt als die Kundgabe einer Weltanschauung. Dafür, was eine ´nur verbale` und was eine ´echte` Lösung eines Problems ist, gibt es offensichtlich keine Kriterien. Es sei denn, es werden nur kausale Erklärungen als ´echte` Erklärungen anerkannt. Aber damit wird die *petitio principii* völlig offensichtlich.

(e) Ein fünftes Argument wird oft darin erblickt, dass Handlungserklärungen kontrafaktische Behauptungen zulassen. Die Rechtfertigung kontrafaktischer Behauptungen wiederum wird seit David Lewis' klassischer Arbeit (1973) oft als Kennzeichen kausaler Relationen be-

trachtet (vgl. Tuomela, 1977, S. 227). Die Schwäche dieses Argumentes liegt in seiner logischen Ungültigkeit. Es handelt sich schlicht um den klassischen Fehler der Bejahung des Consequens. Dass ein Satz, der eine kausale Relation wiedergibt, einen Satz impliziert, der eine kontrafaktische Behauptung darstellt, berechtigt nicht zum Schluss von einer kontrafaktischen Behauptung auf ein Kausalverhältnis (vgl. Stoutland, 1998, S. 58).

(f) Zuletzt möchte ich auf das Argument eingehen, dass nur eine kausale Theorie die Möglichkeit biete, unterschiedliche wissenschaftliche Gebiete zu verbinden und so Erklärungen durch Vereinheitlichung zu erzeugen (vgl. Tuomela, 1977, S. 227). Tatsächlich ist es der Fall, dass die vorliegende Arbeit genau gegen eine solche Erklärung durch Vereinheitlichung gerichtet ist. Eine nicht-kausale Theorie von Handlungen bietet nicht die Möglichkeit, auf eine andere (z. B. neurobiologische) kausale Theorie reduziert zu werden. Die Frage ist nur, ob dies eine Schwäche der intentionalistischen Position darstellt. Als Erstes ist darauf zu sagen, dass man nicht verbinden sollte, was nicht zusammen gehört. Und die Möglichkeit einer *illegitimen* Verbindung stellt *kein* Argument dar. Die Legitimität dieser Verbindung steht aber hier genau in Frage. Weshalb sie nicht ohne *petitio principii* als Prämisse eines Arguments verwendet werden kann. Zweitens wird in dem Argument übersehen, dass Verbindungen zu Naturwissenschaften nicht die einzigen respektablen Verbindungen darstellen. Ich betrachte es als schlagkräftiges Argument der hier vertretenden Position, mit einer tragfähigen (aristotelischen) Ethik vereinbar zu sein (vgl. Foot, 2001; Hursthouse, 1999). Und es lässt sich schwer erkennen, wie der Kausalismus die 'Vereinheitlichung' mit einer tragfähigen Ethik erreichen sollte. Auch auf dem Gebiet der 'Vereinigung' mit verschiedenen Geistes- und Sozialwissenschaften schneidet die intentionalistische Position besser ab. Dass sich eine Theorie hingegen in ein überkommenes positivistisches Weltbild fügt, ist *kein* Vorzug (vgl. Putnam, 2002).

Es hat sich also herausgestellt, dass keines der Argumente für den Kausalismus überzeugen kann. Doch was spricht positiv dafür, dass emotionale Bewertungen Handlungen nicht kausal erklären können?

5.4.2. Argumente für den Intentionalismus

Meines Erachtens gibt es vier wichtige Tatsachen, aus welchen sich Argumente für den Intentionalismus gewinnen lassen: Erstens besteht zwischen Gründen (im Sinne von Rahmenbedingungen) und Handlungen eine begriffliche Abhängigkeit; zweitens lässt die Abhängigkeit mentaler Inhalte von Kontextbedingungen, welche nicht in der Person liegen, eine

kausale Verbindung nicht zu; drittens führen die Begriffe „Absicht“ und „Handlung“ starke normative Implikationen mit sich, was deren Naturalisierung unmöglich macht; und viertens besitzen psychische Zustände bzw. Prozesse (also Rahmenbedingungen) und Handlungen ‘ontologische Eigenschaften’, durch welche sie für kausale Relationen ungeeignet sind.

5.4.2.1. Das Argument aus der begrifflichen Abhängigkeit

Das Argument, dass Gründe nicht die Ursachen von Handlungen sein können, da hier eine begriffliche Abhängigkeit besteht, wird oft als das *Logical Connection Argument* bezeichnet (vgl. Brandtstädter, 2007, S. 59). Das Argument wurde in sehr unterschiedlichen Formen vorgebracht und so handelt es sich eigentlich nicht um *ein* Argument, sondern eher um einen Typ von Argumenten, welcher einer genaueren Ausformulierung bedarf. Während Frederick Stoutland und Abraham I. Melden das Argument als ein semantisches betrachten (vgl. Melden, 1961, S. 89ff.; Stoutland, 1970), hält Georg H. von Wright es im Wesentlichen für ein epistemologisches Argument (vgl. Peijnenburg 1996, S. 72; von Wright, 1984, S. 92). Gegen das *Logical Connection Argument* wurde verschiedene Kritik vorgebracht. Von dieser Kritik enthält wohl der schon genannte Artikel Davidsons (Davidson, 1998d) die einflussreichste und gewichtigste. Ich möchte zunächst das Argument in einer mir tragfähig erscheinenden Form darstellen, um dann auf Kritik einzugehen.

Das Argument beginnt mit Überlegungen zu unserem Begriff der Absicht. Denn für den Kausalist ist Gründe Paare von Proeinstellungen und Überzeugungen (vgl. Davidson, 1998d, S. 20); diese wiederum sind nichts anderes als Absichten (vgl. Davidson, 1998a, S. 131). Nun lässt sich folgende Frage stellen: Wenn jemand sich und anderen Absichten zuschreibt, ohne dabei zu erwarten, dass die jeweilige Person die entsprechende Handlung auch ausführt, sofern dem nichts im Wege steht, können wir dann von dieser Person behaupten, sie hätte verstanden, was mit dem Wort „Absicht“ gemeint ist? Ich erachte es für offensichtlich, dass die Antwort hierauf „Nein“ lauten muss. Wenn der A dem B Absichten zuschreibt, aber keinen Anlass und keine Berechtigung dazu sieht, in Umständen, in welchen einer Verwirklichung der Absicht durch B nichts entgegensteht, Person B aber trotzdem nicht handelt, eine Erklärung zu verlangen, so können wir nicht mehr davon sprechen, dass Person A das Wort „Absichten“ richtig verwendet hat. Dies heißt aber, dass folgender Satz eine begriffliche Relation wiedergibt:¹¹⁴ Wenn eine Person eine Absicht hegt, so wird sie, sofern keine widrigen Umstände eintreten, ihre Absicht ausführen. Da dieser Zusammenhang in unseren

¹¹⁴ s. Kapitel 2.

Begriffen enthalten ist, besitzt er keinen empirischen Gehalt. Es verhält sich hier ähnlich wie mit dem Satz, „Wer eine Trauungszeremonie im Standesamt als Braut erlebt, wird, sofern keine widrigen Umstände eintreten (z. B. Formfehler, Tod beim Trauungsakt, Unzurechnungsfähigkeit etc.), danach verheiratet sein“. Auch dieser Satz besitzt keinen empirischen Gehalt, obwohl er von zeitlich auseinanderliegenden Ereignissen handelt, seine Wahrheit nicht der formalen Logik verdankt und Ausnahmen zulässt. Dies erkennt man unter anderem daran, dass die Vermutung, Trauungen könnten den Ehestatus kausal erklären, keine sinnvolle Forschungshypothese darstellt. Ebenso stellt es keine sinnvolle Forschungshypothese dar, zu vermuten, Absichten könnten Handlungen kausal erklären. Der Satz „Wenn eine Person eine Absicht hegt, so wird sie, sofern keine widrigen Umstände eintreten, nach dieser Absicht handeln“ gibt einen begrifflichen und keinen empirischen Zusammenhang wieder. Genauer handelt es sich hier um grammatische Sätze (im Gegensatz zu analytischen Sätzen).¹¹⁵ Dass dieser grammatische Zusammenhang tatsächlich besteht, zeigt sich daran, dass wir Handlungen anhand ihrer Gründe bzw. Absichten individuieren; aber auch daran, dass Gründe bzw. Absichten nicht unabhängig von der zugehörigen Handlung ‘auftauchen’. Absichten sind immer Absichten *zu etwas*. Wenn aber Absichten bzw. Gründe Handlungen kausal erklären sollen, so muss es einen empirisch gehaltvollen Allsatz geben, welcher im Antecedens von Absichten bzw. Gründen und im Consequens von den entsprechenden Handlungen spricht. Nun liegt die Beweislast, zu zeigen, dass es einen solchen Satz gibt, aber auf der Seite des Kausalismus.

„We typically require of a causal relation that it does not already follow from logical or conceptual relations, so that antecedents and consequents can be independently verified. Action explanation that refer to the agent’s desires, wishes, beliefs, expectancies, and so on, do not seem to easily satisfy this requirement“ (Brandtstädter, 2007, S. 58).

Es wurden verschiedene Versuche unternommen, ein empirisch gehaltvolles Gesetz zu finden, welches Absichten bzw. Gründe mit Handlungen verbindet (vgl. Churchland, 1970; Tuomela, 1977). In den letzten Jahrzehnten wurde aber deutlich, dass die kausalistische Handlungstheorie bei der Formulierung einer entsprechenden, empirisch gehaltvollen Gesetzmäßigkeit auf extreme Schwierigkeiten stößt (vgl. Greve, 1994, S. 116ff.; Keil, 2000, S. 53ff.). Das Hauptproblem besteht dabei darin, Antecedensbedingungen anzugeben, welche unabhängig von der Handlung überprüft werden können. Typischerweise enthalten solche Gesetze als

¹¹⁵ s. Kapitel 2.3.

Antecedensbedingung z. B., dass der Akteur zu der Handlung in diesem Moment fähig sein muss. Lässt man diese Bedingung fallen, so lässt sich das Gesetz zu leicht falsifizieren, um ernsthaftes Interesse beanspruchen zu können. Doch wie sollte die Fähigkeit überprüft werden, wenn nicht durch die Handlung selbst (vgl. Brandtstädter, 2007, S. 58)? Hier wäre die Kenntnis einer strukturellen Personeneigenschaft notwendig, anhand derer wir die Fähigkeit unabhängig überprüfen könnten (z. B. eine biologische Eigenschaft). Doch derzeit besitzen wir keine Kenntnis einer Eigenschaft, welche so hohe Erwartung erfüllen kann (z. B. kennen wir keine Eigenschaft, welche uns erlaubt, vorherzusagen, ob sich ein Akteur an eine bestimmte Information erinnern wird). Ähnliches lässt sich von Wünschen und Überzeugungen sagen. Ob ein Akteur eine Handlung als einem bestimmten Ziel dienlich erachtet, bemisst sich unter anderem daran, ob er diese Handlung vollzieht, wenn er jenes Ziel erreichen will. Und welche Ziele eine Person hat, bemisst sich unter anderem daran, was die Person tut, wenn sie überzeugt ist, dass diese oder jene Handlung dieses oder jenes Ergebnis zeitigen wird. Als besonders problematisch hat sich hierbei die Tatsache erwiesen, dass es notwendig ist, die Rationalität des Akteurs als eine Antecedensbedingung in die zu formulierende Gesetzmäßigkeit aufzunehmen.¹¹⁶ Dass dies notwendig ist, zeigt sich daran, dass Erklärungen aus Absichten bzw. Gründen ihren Erklärungswert einbüßen, wenn der Akteur irrational handelt. Nun müssen wir aber auch hier wieder ein unabhängiges Kriterium zur Feststellung der Rationalität von Akteuren verlangen (vgl. Keil, 2000, S. 63ff.). Anderenfalls schrumpft der empirische Gehalt der angeblich kausalen Erklärungen unserer Handlungen immer weiter zusammen. Unter Rationalität ist in diesem Zusammenhang das Orientieren der eigenen Handlungen an Gründen gemeint. Aber wie sollten wir diese Rationalität überprüfen, wenn nicht mittels der Handlungen selbst? Von demjenigen, der stets beteuert, all sein Handeln an Gründen auszurichten, aber gleichzeitig maximal irrational handelt, können wir nur sagen, dass er entweder sich selbst oder uns betrügt, psychisch krank ist oder schlicht nicht weiß, was „handeln“ heißt. Wir lassen hier nicht die verbalen Äußerungen, sondern letztlich die Handlungen selbst darüber entscheiden, wer seine Handlungen an Gründen orientiert. Damit ist aber auch klar, dass eine unabhängige Überprüfbarkeit der Rationalität von Akteuren nicht in Sicht ist (vgl. Greve, 1994, S. 118ff.). Eine formallogische Ausarbeitung dieser Argumentation findet sich bei Adriana J. M. Peijnenburg (1996, S. 67ff.). Zusammenfassend lässt sich sagen, dass es bis heute nicht gelungen ist, ein Gesetz

¹¹⁶ Churchland (1970) umgeht die Notwendigkeit der Aufnahme der Rationalität in die Antecedensbedingungen, indem er eine allgemeine *ceteris paribus* Klausel einführt. Dies macht aber den Vorwurf der empirischen Gehaltlosigkeit nur noch schwerwiegender (vgl. Keil, 2000, S. 66f.).

zu formulieren, welches Absichten oder Gründe mit Handlungen verbindet und den einschlägigen Ansprüchen an kausale Gesetzmäßigkeiten genügt. Entweder ein solches Gesetz kann unmittelbar falsifiziert werden oder es ist fraglich, ob es empirischen Gehalt besitzt.

Dieses sogenannte *Logical Connection Argument* genießt heute keine sehr gute Reputation. Warum? Der wichtigste Kritikpunkt besteht darin, dass wir es mit einem Argument bezüglich kausaler Erklärungen und nicht bezüglich kausaler Relationen zu tun haben. Davidson schreibt hierzu: „Die Wahrheit einer Kausalaussage hängt davon ab, *welche* Ereignisse beschrieben werden; ihr Status als analytische oder synthetische Aussage hängt davon ab, *wie* die Ereignisse beschrieben werden“ (Davidson, 1998d, S. 35). Dies lässt sich folgendermaßen wiedergeben: Erklärungen stellen intensionale Kontexte dar, Sätze, welche von Verursachung reden, hingegen extensionale (vgl. Bradley, 1979; Deutscher, 1976). Der Satz „Die Ursache von A hat A verursacht“ ist empirisch gehaltlos, aber nichtsdestoweniger besteht das angegebene Kausalverhältnis tatsächlich. Es trifft also zu, dass das Bestehen einer begrifflichen Beziehung an sich kein Grund ist, an einer Kausalrelation zu zweifeln. Es ist möglich und tatsächlich oft der Fall, dass ein Ereignis die Ursache eines zweiten ist, und sich trotzdem Beschreibungen finden lassen, unter welchen ersteres Ereignis das zweite nicht kausal *erklären* kann. („Ereignis A geschah, weil die Ursache von Ereignis A auftrat“ ist keine Erklärung.) Zweifelhaft ist allerdings, ob diese Kritik das Argument im Kern trifft. Denn wenn zugestanden wird, dass kausale *Erklärungen* empirischen Gehalt besitzen müssen, so macht das Argument es unmöglich, eine kausale *Erklärung* von Handlungen *qua Handlungen* zu geben, welche auf Absichten *als Absichten* Bezug nehmen (vgl. Schueler, 2003, S. 10ff.). Dies schließt tatsächlich nicht aus, dass Absichten die Ursachen von Handlungen sind, aber es ist dennoch nicht möglich, ein empirisch gehaltvolles Gesetz zu finden, welches Handlungen *als Handlungen* mit Absichten *als Absichten* verbindet. Bekannterweise ist genau dies auch Davidsons Lösung. Während er Absichten und Gründen *als solchen* eine Funktion zugesteht, welche er Rationalisierung nennt, sind dieselben Absichten und Gründe nicht *als solche*, sondern als materielle Phänomene kausal effizient (vgl. Davidson, 1998b). Dies wird als Davidsons „anomaler Monismus“ bezeichnet, doch eine Vielfalt alternativer nicht-reduktionistischer Positionen wie etwa der Funktionalismus vertreten ähnliche Thesen. Damit kann zwar der Kausalismus vor den Folgen des *Logical Connections Arguments* bewahrt werden, aber nur um den Preis, auf eine Form des nicht-reduktionistischen Materialismus festgelegt zu sein. Abgesehen von der Tatsache, dass damit eine Reduktion, wie sie die SMH verlangt, unmöglich wird, wurden in den letzten Jahrzehnten immer mehr Schwächen des nicht-reduktionistischen Materialismus deutlich (vgl. Kim, 1998, S. 120, 2007). Im gegen-

wärtigen Zusammenhang scheint das *Logical Connections Argument* trotz der genannten Schwäche also immer noch schlagkräftig, selbst wenn es nur auf Versuche zutrifft, deduktiv-nomologische Handlungserklärungen unter Bezug auf Absichten bzw. Gründen aufzustellen. Eine andere Kritik übt Tuomela am *Logical Connections Argument*. Er betrachtet den Schluss von einer epistemischen Abhängigkeit auf die empirische Gehaltlosigkeit als unberechtigt.

„[It] sounds like confusing meaning with evidence, even if only the *logical* rather than something like a *factual* possibility of verification seems to be under consideration. I think that, for instance, two theoretical scientific propositions could be conceptually independent even if they were not independently testable in any circumstances“ (Tuomela, 1977, S. 185).

Aber hier stellt sich die Frage, wann wir dazu bereit sind, den Fall, welchen Tuomela schildert, zu akzeptieren. Denn wenn unsere gesamte Theorie aus der Behauptung über Kausalverhältnisse besteht, welche sich nicht unabhängig voneinander überprüfen lassen, so werden wir gut beraten sein, uns nach einer besseren Theorie umzusehen.¹¹⁷ Behauptungen über Zusammenhänge, welche nicht unabhängig überprüfbar sind, akzeptieren wir nur dort als empirisch gehaltvoll, wo es eine Rahmentheorie gibt (z. B. im Fall von Messtheorien). Die Rahmentheorie garantiert dabei den empirischen Gehalt. Ist aber eine solche Rahmentheorie in unserem Falle verfügbar? Ich denke nicht. Auch diese Kritik scheint also nicht den Kern des Arguments zu treffen.

5.4.2.2. Das Argument aus dem Externalismus

Ich habe oben betont, dass eine Handlung *qua Handlung* zu erklären heißt, sie aus Gründen zu erklären. Dies wird auch von Kausalisten allgemein anerkannt; nur betrachtet der Kausalismus Rahmenbedingungen dabei als Gründe (vgl. Davidson, 1998d; Klein, 1996). Der Kausalismus identifiziert darüber hinaus Erklärungen aus Gründen mit Kausalerklärungen. Dabei sollen mentale Zustände bzw. Prozesse die Ursachen von Handlungen sein. Im Fall der SMH sollen emotionale Bewertungen von erwarteten Handlungsergebnissen die Ursachen für das Ausführen entsprechender Handlungen sein. Dabei sind diese mentalen Zustände bzw. Prozesse mit einem bestimmten Zustand des zentralen Nervensystems korreliert. Philosophen sprechen hier oft davon, dass bestimmte mentale Zustände physisch „realisiert“ seien. Nun ist

¹¹⁷ Der Gehalt unserer Sätze ist tatsächlich nicht völlig unabhängig von unseren epistemischen Möglichkeiten, wie Tuomela zu implizieren scheint. Dies erkennt man daran, dass Behauptungen wie, dass ein Geist in meiner Kaffeetasse lebt, der aber leider durch keine Kriterien oder Evidenzen nachzuweisen ist, sinnlose Sätze sind. Dasjenige, was meine Behauptung zu beschreiben scheint, ist nicht das, was wir einen „Geist in meiner Kaffeetasse“ nennen würden.

es aber entscheidend zu erkennen, dass es hier nicht der mentale Zustand als solcher ist, der die eigentliche 'Erklärungsarbeit' übernimmt, sondern der *Inhalt* des jeweiligen mentalen Zustandes (vgl. Klein, 1996, S. 159f.). Meine positive emotionale Bewertung frischer Brötchen kann meinen Gang zum Bäcker nur erklären (wenn überhaupt), weil sie die Brötchen *zum Inhalt* hat. Mein Wunsch nach einem neuen Auto kann meinen Gang zum Bäcker hingegen nur unter Zuhilfenahme sehr bizarrer Annahmen erklären. Allgemein formuliert: „*Cognitive science [...] attempts to explain behavior as causally resulting from the interplay of content-individuated states*“ (Owens, 1993, S. 243). Und hier bildet weder der Kausalismus im Allgemeinen noch die SMH im Besonderen eine Ausnahme. Somatische Marker erklären Entscheidungen nur insofern, als sie bestimmte *Inhalte* 'markieren'. Verhalten, welches auf Entscheidungen beruht, soll durch intentionale psychische Zustände kausal erklärt werden. Dies bedeutet, dass der intentionale Gehalt eines psychischen Zustandes als kausal effizient betrachtet werden muss.

Nun ist es aber offensichtlich, dass kausal relevante Eigenschaften von Gegenständen intrinsische Eigenschaften dieser Gegenstände sind; wobei mit „intrinsisch“ hier gesagt werden soll, dass diese Eigenschaften nicht von Objekten oder Prozessen außerhalb des jeweiligen Gegenstandes begrifflich abhängen.¹¹⁸ Denn anderenfalls müsste man annehmen, dass Gegenstände, welche alle intrinsischen Eigenschaften teilen, sich in ihren kausalen Eigenschaften unterscheiden könnten. Dies dürfte aber wohl niemand ernsthaft behaupten (vgl. Kim, 2002, S. 173). Dass mein Fahrrad mir gehört, ändert seine kausal relevanten Eigenschaften nicht im Mindesten. Ein intrinsisch identisches Fahrrad würde in jeder möglichen Situation gleich reagieren und gleiche Wirkungen zeitigen, auch wenn es nicht meines wäre. Dies bedeutet aber, dass kausal relevante Eigenschaften anhand intrinsischer Eigenschaften individuiert werden können müssen. „*[W]e expect causal properties involved in behavior production to be nonrelational, or intrinsic, properties of the organism*“ (Kim, 2002, S. 173). Wenn nun die Eigenschaft bestimmter mentaler Phänomene, einen bestimmten Inhalt zu besitzen, kausal relevant sein soll, so muss sie anhand intrinsischer Eigenschaften individuiert werden können. Da aber mentale Phänomene genau anhand ihres Inhalts individuiert werden, muss die Eigenschaft, einen bestimmten Inhalt zu besitzen, selbst eine intrinsische Eigenschaft darstellen. Der Externalismus bezüglich mentaler Inhalte stellt nun genau diese These in Frage. Er be-

¹¹⁸ Sarah Sawyer drückt denselben Gedanken folgendermaßen aus: „*A property is individualistic if there is some specification of it that makes no reference to the environment, and anti-individualistic if there is no specification of it that does not*“ (Sawyer, 2007, S. 32). So ist beispielsweise meine Eigenschaft, ein Bruder zu sein, nicht-intrinsisch (*anti-individualistic*), hingegen ist meine Haarfarbe eine intrinsische Eigenschaft (*individualistic*).

hauptet, dass die Eigenschaft, einen bestimmten Inhalt zu besitzen, nicht allein anhand intrinsischer Eigenschaften individuiert werden kann. Daher ist der Kausalismus *prima facie* nicht mit einem Externalismus bezüglich mentaler Inhalte vereinbar (vgl. Hursthouse, 2000, S. 88f.). Die meisten zeitgenössischen Forscher vertreten jedoch irgendeine der unterschiedlichen Versionen des Externalismus (vgl. Jacob, 1993, S. 131). Welche Gründe lassen sich für den Externalismus vorbringen?

Ich habe bereits zu Beginn dieser Arbeit Putnams Zwillingserden-Gedankenexperiment dargestellt.¹¹⁹ Ich habe auch bereits dort darauf hingewiesen, dass die essenzialistischen Annahmen, auf welchen viele der Argumente Putnams (1979) und Kripkes (1981) für den Externalismus beruhen, höchst unplausible Konsequenzen nach sich ziehen, und daher unhaltbar sind.¹²⁰ Deshalb möchte ich hier einen anderen Argumentationsweg einschlagen. Hierzu ist zunächst festzuhalten, dass der Inhalt mentaler Phänomene sehr eng mit den entsprechenden Wahrheitsbedingungen verbunden ist.¹²¹ Nach dem Inhalt eines Glaubens, also einer Überzeugung, kann man fragen, indem man danach fragt, was der Fall sein müsste, damit dieser Glaube wahr ist. Nach dem Inhalt eines Wunsches lässt sich fragen, indem man fragt, was der Fall sein müsste, damit dieser Wunsch erfüllt wäre. Außerdem können die Inhalte mentaler Phänomene in eine logische Beziehung zueinander treten; und auch hierfür scheinen ihre Wahrheitsbedingungen entscheidend zu sein (vgl. Frances, 1999, S. 53). Anders ausgedrückt: Wenn wir mentalen Phänomenen intentionalen Inhalt zuschreiben, so schreiben wir ihnen etwas zu, das – zumindest zum größten Teil – die Eigenschaften des Frege'schen Sinns oder der Carnap'schen Intension besitzt. Damit ist auch der Anspruch verbunden, dass eben jene Intension (Sinn) auch die Extension (Bedeutung) festlegt; denn nur so können auch die Wahrheitsbedingungen festgelegt werden. Dies bedeutet aber, dass sich von einem Unterschied in den Extensionen bestimmter mentaler Phänomene unmittelbar auf einen Unterschied ihres Inhalts schließen lässt.¹²² Die Extension von Sätzen bzw. Propositionen wird im Allgemeinen als ihr Wahrheitswert betrachtet. Hieraus ergibt sich: „[I]f it is possible for the ascriptions ‘S believes that P’ and ‘S believes that P*’ to differ in truth-value, then they attribute different states“ (Owens, 1993, S. 244). Sollten sich nun intrinsisch identische Personen

¹¹⁹ s. Kapitel 2.2.1.3.2.

¹²⁰ s. Kapitel 2.2.2.1.2.

¹²¹ Ich will hier Überlegungen bezüglich der Frage ignorieren, ob und inwieweit die Rede von Wahrheitsbedingungen ersetzt werden kann durch die Rede von Behauptbarkeitsbedingungen. Ich denke, es ist leicht zu sehen, dass das folgende Argument sich auch unter Rückgriff auf Behauptbarkeitsbedingungen konstruieren lässt.

¹²² Wie diese Position mit der dieser Arbeit zugrundeliegenden gebrauchsbasierten Bedeutungstheorie vereinbar ist, habe ich in Kapitel 2.3 zu zeigen versucht.

denken lassen, deren mentale Inhalte sich in ihrem Wahrheitswert unterscheiden, kann geschlossen werden, dass die Eigenschaft, einen bestimmten Inhalt zu besitzen, keine intrinsische Eigenschaft ist. Dann aber wäre kaum noch verständlich, wie Erklärungen mittels solcher Eigenschaften kausale Erklärungen sein können.

Einer der prominentesten Fälle, der zeigen soll, dass die mentalen Inhalte intrinsisch identischer Personen sich in ihrem Wahrheitswert unterscheiden können, ist der Fall des arthritischen Alf (vgl. Burge, 1979).¹²³ Gegeben seien zwei Situationen, die eine aktuell, die andere sei eine kontrafaktische Situation. In beiden Situationen gibt es eine Person des Namens Alf. Zu Zwecken der Klarheit wollen wir den kontrafaktischen Alf mit „Alf*“ bezeichnen. Alf und Alf* sind intrinsisch identische Personen und haben außerdem dieselben intrinsischen Veränderungen seit ihrer Geburt erlebt. Die Umwelten, in denen Alf und Alf* leben, sind weitgehend identisch. Doch es besteht auch ein Unterschied: Die sozialen Gemeinschaften, in welchen unsere Protagonisten leben, unterscheiden sich in den Regeln ihres Gebrauchs des Wortes „Arthritis“. Während das Wort in Alfs Umwelt nur für Entzündungen in Gelenken gebraucht wird, ist in Alf*s Gemeinschaft auch die Anwendung des Wortes auf bestimmte Beschwerden in Knochen zulässig. Da sich weder Alf noch Alf* für Medizin interessieren, sind ihnen solche Subtilitäten des richtigen Gebrauchs des Wortes Arthritis nicht bekannt. Beiden haben außerdem eine große Menge an richtigen Überzeugungen bezüglich Arthritis (z. B. welcher ihrer Verwandten an Arthritis leidet, wo die beste Behandlung von Arthritis zu bekommen ist etc.). Nun wachen beide eines Morgens mit Schmerzen in ihrem Oberschenkel auf und sagen zu sich selbst: „Ich werde wohl an Arthritis in meinem Oberschenkel leiden“. Beim Arzt wird Alf darüber aufgeklärt, dass er etwas falsche Vorstellungen bezüglich Arthritis habe, während Alf* mitgeteilt wird, dass er völlig richtig liege. An dem besagten Morgen haben wir also eine Situation, in welcher intrinsisch identische Personen ein mentales Phänomen aufweisen, dessen Inhalt beide beschreiben würden mit: „Ich werde wohl an Arthritis in meinem Oberschenkel leiden“. Genau dieser Inhalt ist aber bei einer Person wahr, bei der anderen hingegen falsch. Da aber die jeweiligen Umstände (bis auf die Unterschiede in den Regeln des Sprachgebrauchs) gleich sind, liegen hier unterschiedliche Extensionen des mentalen Inhalts vor und wie oben gezeigt, lässt sich daraus auf einen Unterschied des Inhalts schließen. Dies ist nur eines von vielen Beispielen, welche in der Literatur angeführt werden, um zu belegen, dass der Inhalt mentaler Phänomene nicht allein anhand intrinsischer Eigenschaften bestimmt werden kann (vgl. Frances, 1999, S. 67f.).

¹²³ Dieses Beispiel hat gegenüber den Beispielen von Putnam den Vorteil, dass es keine essenzialistischen Annahmen voraussetzt.

„What do the thought experiments show? They show that the intentional content of ordinary propositional attitudes, as indicated by obliquely occurring expressions in that-clauses, cannot be accounted for in terms of physical, phenomenal, causal-functional, computational, or syntactical states or processes that are specified nonintentionally and are defined purely on the individual in isolation from his physical and social environment“ (Burge, 1982, S. 286).

Wenn der Kausalist diesem Ergebnis zustimmen will, so ist er uns eine sehr gute Antwort auf die Frage schuldig, wie wir uns den Kausalnexus zwischen bestimmten mentalen Inhalten und bestimmten Handlungen vorzustellen haben. „*How can extrinsic, relational properties be causally efficacious in behavior production*“ (Kim, 2002, S. 173)? Einige kausalistische Autoren experimentieren mit der Idee, die Zuschreibung kausal relevanter Eigenschaften genauso an extrinsische Faktoren zu binden, wie dies offensichtlich bei der Zuschreibung mentaler Inhalte geschieht (vgl. Klein, 1996; Saidel, 1994). Mit einem solchen Manöver wird aber genau dasjenige realistische Verständnis kausaler Relationen untergraben, welches dem Kausalismus seine eigentliche Brisanz verleiht (vgl. Butler, 1996). Denn hierdurch werden Aussagen über kausale Verbindungen zu intensionalen Kontexten, aber genau dieses Missverständnis wurde ja von kausalistischer Seite den Vertretern des *Logical Connection Argument* vorgeworfen. Fodor schlägt einen anderen Ausweg für den Kausalismus vor: „*I propose to resolve the difficulty by distinguishing between the content of a belief and its truth conditions*“ (Fodor, 1982, S. 111). Damit schuldet uns der Kausalist aber eine sehr gute Erklärung dafür, wie Zustände, deren Identität nichts mit ihren Wahrheitsbedingungen zu tun hat, in logische und begriffliche Beziehungen zueinander treten können. Es war ja genau das Anliegen fast der gesamten (nicht-konnektionistischen) Kognitionswissenschaft, physikalische Systeme so zu analysieren, dass deutlich wird, wie physikalische Prozesse die Verarbeitung von Information darstellen können und dies heißt auch, wie physikalische Prozesse als Implementierung wahrheitsfunktionaler Zusammenhänge verstanden werden können. Es stellt sich damit überhaupt die Frage, warum Phänomene, welche in keine begriffliche oder logische Beziehung zu anderen Phänomenen treten können, als Phänomene mit einem bestimmten Inhalt betrachtet werden sollten. Die wohl verbreitetste Strategie, um auf das Argument aus dem Externalismus zu reagieren, besteht darin, zwar anzuerkennen, dass unsere alltäglichen Begriffe mentaler Inhalte diese von extrinsischen Eigenschaften abhängig machen, dieser Zug unserer normalen Sprache aber durch die wissenschaftliche Psychologie bereinigt werden muss (vgl. Segal, 2007; Wedgwood, 2006). Demnach stellt es ein Ziel der wissen-

schaftlichen Psychologie dar, ausschließlich mit der Zuschreibung von Inhalten auszukommen, die keinen Bezug auf extrinsische Eigenschaften nehmen. Solche Inhalte werden auch als *narrow content* (im Gegensatz zu *broad content*) bezeichnet (vgl. Beckermann, 2001, S. 371ff.; Brown, 2007). Damit ist der Kausalist aber auf zwei starke Thesen verpflichtet: erstens, dass es einen gangbaren Weg gibt, den Begriff des *narrow content* zu spezifizieren, und zweitens, dass der gesamte Erklärungswert mentaler Inhalte ausschließlich unter Rückgriff auf *narrow content* erhalten werden kann. Beide Thesen sind äußerst fragwürdig. Es ist bis heute nicht gelungen, eine haltbare Spezifikation des Begriffs *narrow content* zu finden (vgl. Sawyer, 2007). Dies kann nicht verwundern, wenn man bedenkt, dass wir von nichts sagen können, es trage einen Inhalt (eine Bedeutung), sofern es keine Möglichkeit für Fehler gibt. Zumindest bei Aussagesätzen ist klar, dass sie ihren Inhalt der Tatsache verdanken, dass sie richtig und falsch sein können. Richtigkeit und Falschheit sind aber selbst keine intrinsischen Eigenschaften des Organismus, welcher den Inhalt der Aussage ausspricht oder denkt. Kurz: Der Begriff des *narrow content* scheint Fehler auszuschließen. Damit steht aber in Frage, inwieweit wir hier überhaupt von Inhalt reden sollten. Der zweite Punkt ist nicht weniger problematisch. Wenn wir Inhalte zuschreiben, sind wir ohne Zweifel oft in der Lage, diesen Inhalt so zu beschreiben, dass er uns als *narrow content* erscheint. Im genannten Beispiel können wir sagen, dass die Gedanken Alfs und Alf*s denselben Inhalt haben, weil wir diesen Inhalt in beiden Fällen beschreiben können als „Ich leide wohl an einer Krankheit mit folgenden Eigenschaften: a, b, c, ...“.¹²⁴ Wenn wir diese Inhaltszuschreibung verwenden, können wir die genannten Konsequenzen umgehen. Nun lässt sich aber natürlich ein ähnliches Beispiel mit dem Wort „Krankheit“ anführen. Wenn jedoch alle Inhaltszuschreibungen ausschließlich auf *narrow content* beruhen sollen, so muss ausgeschlossen werden, dass die Beschreibung des jeweiligen *narrow content* auf der Beschreibung eines anderen *broad content* beruht. Es ist alles andere als klar, dass man ein grundlegendes Vokabular finden kann, welches für diese Aufgabe geeignet ist (vgl. Frances, 1999; Jacob, 1993), insbesondere, wenn man bedenkt, dass Sprache in einem sozialen Kontext gelernt wird. Es scheint vielmehr so zu sein, dass wir ein vielfältiges Vokabular teilen und mithilfe dieses Vokabulars die Fehler bestimmter Sprecher so weitgehend beschreiben können, dass diese Beschreibung von denjenigen Umweltfaktoren unabhängig wird, auf welchen die entsprechenden Fehler beruhen. Wir sind also in der Lage, mentale Inhalte so zu beschreiben, dass wir alle eine gemeinsame

¹²⁴ Dies würde bedeuten, dass Alf nicht die Sprache der ihn umgebenden Sprachgemeinschaft spricht, sondern jene der Sprachgemeinschaft, in welcher Alf* lebt.

Grundlage an Sprachverständnis und Rationalität teilen, aber hieraus folgt keineswegs, dass diese Grundlage sich in Begriffen von *narrow content* beschreiben lässt.

Zusammenfassend lässt sich sagen, dass bis heute keine ernsthaften Gegenargumente gegen folgende Bemerkungen Anscombes vorgebracht wurden:

„[W]e cannot ascribe a belief like that about the bank’s opening hours, to someone not living in a world of banks and clocks. [...] And the same goes for intentions, decisions, and thoughts. [...] There might come a point at which there is no further puzzle about *how* each link in the causal chain produces the next one. And still nothing has been said about intentions, beliefs, thoughts or decisions“ (Anscombe, 2005, S. 99f.).

Wenn dies korrekt ist, wir aber in unseren Erklärungen menschlicher Handlungen *als Handlungen* eben genau auf solche Phänomene wie Absichten oder Entscheidungen eingehen und wir anderenfalls nicht verstehen können, wie wir Wissenschaft treiben und aus Gründen handeln können, so ist klar, dass Handlungserklärungen keine Kausalerklärungen sein können.

Die Tatsache, dass die Eigenschaft eines psychischen Zustandes, propositionalen Gehalt zu besitzen, keine intrinsische Eigenschaft des `Trägers` des jeweiligen Zustands ist, macht es also unmöglich, dass solche gehaltvollen psychischen Zustände die Ursachen von Handlungen sind. Psychische Zustände sind daher nicht nur keine Gründe unseres Handelns (und tauchen deshalb nicht in Handlungserklärungen im eigentlichen Sinne auf), sondern können nicht einmal als Rahmenbedingungen von Handlungserklärungen eine Ursache der jeweiligen Handlung darstellen. Doch die Intentionalität der fraglichen psychischen Zustände bringt noch weitere Eigenschaften mit sich, welche sich nicht mit einer kausalistischen Auffassung menschlichen Handelns vereinbaren lassen. Das Tragen propositionalen Gehalts bringt normative Implikationen mit sich, welche sich effizient kausal nicht erklären lassen.

5.4.2.3. Das Argument aus der Normativität

Handlungserklärungen nehmen oft implizit oder explizit auf Normen oder Regeln Bezug. Ein Beispiel: „Warum hat Klaus an der Kreuzung angehalten?“ – „Weil das andere Auto Vorfahrt hatte“. Der Kausalismus kann aber normativen Tatsachen (wie etwa der Tatsache, dass jemand ein Recht oder eine Pflicht hat) offensichtlich keinen echten Erklärungswert zubilligen. Denn wenn normativen Fakten Erklärungswert zugestanden wird, warum dann nicht auch Zielen, welche in der Zukunft liegen? Sowohl Normen wie auch Ziele scheinen *prima facie* nicht in die physische und kausale Ordnung der Dinge zu passen, welche der Kausalist

als allein wissenschaftlich respektabel akzeptiert. Sollte der Kausalist Normen oder Zielen echten Erklärungswert zugestehen, so ist er uns eine sehr gute Erklärung schuldig hinsichtlich der Frage, wie und warum genau jene Dinge, welche er zulässt, in eine kausale Erklärung passen. Kurz: Der Kausalist ist darauf verpflichtet, in seinen Erklärungen entweder ohne Rückgriff auf Normen oder Ziele auszukommen oder diese auf eine Weise einzubinden, durch welche verständlich wird, wie diese Entitäten eine Rolle in kausalen Erklärungen spielen können; d. h., der Kausalist muss Normativität bzw. Finalität entweder naturalisieren oder auf sie verzichten (vgl. Owens, 1994; Owens, 1998). Beide Möglichkeiten erscheinen aber in Anbetracht der jüngeren Forschung alles andere als attraktiv.

Der gängigste Weg, Normen in kausale Erklärungen einzuführen, besteht in der vorherigen Einführung von Zielen. Diese Ziele wiederum werden dann meist zuerst auf den psychologischen Prozess der Zielsetzung und schließlich auf Wünsche und Bedürfnisse zurückgeführt (vgl. Audi, 1993, S. 282ff.). So könnte etwa eine kausalistische Erklärung für das gerade genannte Beispiel lauten: Klaus hat den Wunsch, dass sein Wagen unversehrt bleibt. Dieser Wunsch verursacht Klaus' Zielsetzung sich an die Verkehrsregeln zu halten (welche er in seinem Gedächtnis zuvor als 'Information gespeichert' hat).¹²⁵ Diese Zielsetzungen wiederum verursachen Klaus' Anhalten. Innerhalb der SMH würde diese Erklärung folgende Form annehmen: Klaus hat durch gewisse Umweltkontingenzen die Disposition erworben, auf bestimmte Situationen mit *secondary emotions* zu reagieren (vgl. Damasio, 2006, S. 134ff.). Zudem hat Klaus die Disposition erworben, in Entscheidungssituationen unterschiedliche, von ihm erwartete, Handlungs-Ergebnis-Szenarien (weitgehend unbewusst) zu verarbeiten, wobei die Ergebnisse dieses Verarbeitungsprozesses (als *secondary inducers*) Veränderungen in somatosensorischen Kortizes verursachen (vgl. Bechara & Damasio, 2005, S. 339ff.). Diese Prozesse führen schließlich zu Klaus' Anhalten. Die Veränderungen im somatosensorischen Kortex fungieren dabei als Maß für die Diskrepanz zwischen erwartetem Ist-Zustand und Soll-Zustand. Damit werden Zielsetzungen letztlich durch biologische Prozesse zur Aufrechterhaltung der Homöostase im menschlichen Körper erklärt. Und hier liegt demnach auch die Wurzel der Normativität: „*It is a demonstrable fact that what we usually call good and evil is aligned with categories of actions related to particular ranges of homeostatic regulation*“ (Damasio, 2005, S. 48). Auch innerhalb der SMH werden also Normativität und Finalität auf

¹²⁵ Die Vielfalt der unterschiedlichen Ausformulierungen solcher Erklärungen innerhalb der kognitiven Psychologie ist natürlich grenzenlos. „TOTE-Einheiten“, „kognitive Schemata“ oder „Erwartung-Wert-Modelle“ sind nur einige wenige der Vokabeln, welche in solchen Zusammenhängen fallen (vgl. Barone, Maddux & Snyder, 1997, S. 277ff.).

Wünsche und Bedürfnisse zurückgeführt und damit naturalisiert. Es ist offensichtlich, dass diese Kette von Reduktionen (Normativität auf Finalität, Finalität auf den psychologischen Prozess der Zielsetzung, Zielsetzung auf Wünsche bzw. Bedürfnisse, Wünsche bzw. Bedürfnisse schließlich auf biologische Prozesse) dem Kausalisten nur dann helfen kann, wenn erstens diese Reduktion gelingt (im Sinne von logischer Widerspruchsfreiheit und empirischer Bestätigung), und zweitens sich Wünsche und Bedürfnisse ohne Bezug auf Normativität bestimmen lassen. Es gibt gute Gründe, die Erfüllung dieser beiden Bedingungen für unmöglich zu halten. Da aber als Reaktion auf eine Kritik der ersten Bedingung meist unzulässigerweise¹²⁶ auf zukünftigen wissenschaftlichen Fortschritt verwiesen wird, will ich mich hier auf die zweite Bedingung beschränken.

Seit Kripkes (1987) klassischer Arbeit zu Wittgensteins Überlegungen bezüglich des Regel-folgens hat sich die Ansicht verbreitet, dass eine Theorie intentionaler Inhalte nicht ohne Rückgriff auf normatives Vokabular auskommen wird (vgl. Brandom, 2001b, S. 589ff.; Millar, 2004). Es haben sich sehr unterschiedliche Ausgestaltungen dieser These entwickelt, aber für den derzeitigen Kontext sollte es genügen, einige Überlegungen anzuführen, welche die grundlegende These plausibel erscheinen lassen. Was kann es also heißen, dass normatives Vokabular für das Verständnis der Zuschreibung intentionaler Inhalte notwendig ist? Paul Boghossian gibt eine Antwort:

„I want to say that content attributions are normative just in case oughts are constitutive of content attribution. And I want to say that oughts are constitutive of content attributions just in case it's a condition on understanding a content attribution that one understands that if it is true, then the corresponding ought claim is also true“ (Boghossian, 2003b, S. 38).

In einem nächsten Schritt lässt sich nun leicht erkennen, dass genau diese Bedingungen für die Zuschreibung von Überzeugungen tatsächlich gelten: Für jeden Satz *p* gilt, man *soll* von *p* nur dann überzeugt sein, wenn *p*.

„I would maintain that the holding of this norm is one of the defining features of the notion of belief: it's what captures the idea that it is constitutive of belief to aim at truth. The truth is what you ought to believe, whether or not you know how to go about it, and whether or not you know if you have attained it“ (Boghossian, 2003b, S. 38f.).

¹²⁶ Es geht hier ja genau um die (logische) Möglichkeit und nicht um die Wirklichkeit der angestrebten Erklärung.

Der Begriff der Überzeugung bzw. des Glaubens setzt also normative Tatsachen voraus. Wir können anderen Menschen nur dann Überzeugungen zuschreiben, wenn wir wissen, was „sollen“ heißt. Wer Überzeugungen zuschreiben will, muss also die Verwendung normativen Vokabulars akzeptieren.

In einem letzten Schritt kann nun gefragt werden, ob wir den Begriff des Wunsches oder des Bedürfnisses haben könnten, ohne einen Begriff der Überzeugung zu besitzen. Sollte dies nicht möglich sein, so können wir auch Wünsche nicht zuschreiben, ohne ein Verständnis normativer Fakten zu besitzen. *„Can I coherently think of a creature as wanting various things to be true, without thinking of it as having any views whatsoever about how they actually are? This does seem bizarre“* (Boghossian, 2003b, S. 42). Wenn wir diesen bizarren Gedanken ablehnen, so bedeutet dies, dass wir keinen Begriff von Wünschen bzw. Bedürfnissen haben können, ohne bereits normatives Vokabular zu verwenden. Damit kann das Projekt, Normativität bzw. Finalität durch Bezug auf Wünsche oder Bedürfnisse wissenschaftlich respektabel zu machen, als gescheitert betrachtet werden. Gleichzeitig ist aber auch die Möglichkeit, auf normatives bzw. teleologisches Vokabular ganz zu verzichten, hiermit blockiert. Denn wenn die Zuschreibung intentionaler Inhalte normatives Vokabular voraussetzt, stünde dieses Vokabular insgesamt dem Kausalisten nicht mehr zur Verfügung. Dies käme aber einer eliminativistischen Position gleich, welche sich schon im vorigen Kapitel als unhaltbar erwiesen hat.¹²⁷

Natürlich bleibt die Möglichkeit bestehen, normatives oder teleologisches Vokabular auf andere Weise für den Kausalismus so respektabel zu machen, dass seine Verwendung keine Gefahr mehr darstellt. Die Chancen hierfür stehen aber meines Erachtens nicht sehr gut (vgl. Brandom, 2001b), insbesondere angesichts der Tatsache, dass die Ablehnung normativen Vokabulars ohnehin ihre Überzeugungskraft einem antiquierten Weltbild des logischen Positivismus verdankt und wir in der Wissenschaft wohl kaum ohne Normen wie die der Konsistenz und Einfachheit auskommen werden (vgl. Putnam, 2002, S. 28ff.; Siegel, 1996). Darüber hinaus sollte nicht unterschätzt werden, wie eminent wichtig normatives Vokabular für Handlungserklärungen ist, nicht nur als Grundlage für die Zuschreibung intentionaler Inhalte. So weist etwa Dancy darauf hin, dass der Kausalismus ohne Rückgriff auf Normativität nicht verständlich machen kann, warum bestimmte Kombinationen von Wünschen und Überzeugungen offensichtlich mehr Erklärungswert besitzen als andere:

¹²⁷ s. Kapitel 4.1.1.

„Let us consider the difference between two supposedly causal explanations:

A: $S \Phi$'s because S believes that p and S desires to Φ if p

B: $S \Phi$'s because S believes that p and S desires to Φ if not- p

One of these explanations seems to have some kind of advantage over the other. We can see how it might be that the 'process' specified by A should occur. We cannot see how it might be that the 'process' specified by B should occur. But what is the ground for this different attitude? Is it that we have learnt by experience that things work the A-way and not the B-way, by and large? I doubt it“ (Dancy, 2004, S. 31).

Darüber hinaus gibt es gute Gründe, viele Begriffe, welche für Handlungserklärungen unerlässlich sind, als teleologische Begriffe zu betrachten. Teleologischen Begriffen ist ein bestimmter Wertmaßstab inhärent (vgl. Müller, 2003, S. 329). So kann beispielsweise gefragt werden: Kann man wirklich verstehen, was ein Grund ist, wenn man nicht versteht, was ein guter Grund ist? Es scheint offensichtlich, dass dies unmöglich ist. „Grund“ scheint also ein teleologischer Begriff zu sein und damit Normativität bereits vorauszusetzen (vgl. Müller, 2003, S. 347). In anderen Worten: Normativität und Finalität tauchen bei Handlungserklärungen überall auf, und der Kausalist muss jedes dieser Vorkommnisse entweder auf naturalistischer Grundlage erklären oder dessen Entbehrlichkeit zeigen – kein leichtes Unterfangen.

Zusammenfassend lässt sich sagen, dass der Kausalist hier vor der Wahl zwischen zwei für ihn gleichermaßen unattraktiven Positionen steht: Entweder er akzeptiert normatives bzw. teleologisches Vokabular und die entsprechenden Erklärungen als wissenschaftlich respektabel. Dann ist unklar, auf welcher Grundlage intentionalistische Handlungserklärungen abgelehnt werden und nach einer kausalen Erklärung verlangt wird. Oder er versucht sich an der Naturalisierung von propositionalem Gehalt, Normativität und Finalität. „*And there is every reason to think that this is a hopeless program*“ (Owens, 1994, S. 928).

Der Kausalist, welcher mit all diesen Problemen konfrontiert ist, wird sich vielleicht fragen, woher es kommen mag, dass seine Auffassung an so vielen Stellen in Schwierigkeiten gerät und ob nicht vielleicht all diese Probleme eine gemeinsame Wurzel haben. Und tatsächlich lässt sich solche eine gemeinsame Wurzel finden: Das Bedürfnis, auf etwas zu zeigen. Der Kausalist stellt sich Psychisches als ein 'Ding' vor, auf welches man entweder körperlich oder in Gedanken 'zeigen' kann. Diese Vorstellung wiederum entspringt der Tendenz, bei jedem Substantiv nach etwas zu suchen, was diesem entspricht (vgl. hierzu auch BB, Wittgenstein,

1997a, S. 21). Dies deutet darauf hin, dass sich eine ontologische Betrachtung der fraglichen 'Gegenstände' lohnt.

5.4.2.4. Das Argument aus der Ontologie

Die These dieses Abschnitts lautet: Handlungen, Gründe und Absichten besitzen 'ontologische Eigenschaften', welche ausschließen, dass Handlungserklärungen Erklärungen kausaler Natur sind. Mit 'ontologischen Eigenschaften' meine ich hier nicht mehr als 'die Art dieser Dinge zu sein'. Die Art, wie ein Ding ist, bestimmt sich nach unserem Gebrauch der Begriffe, mit welchem wir über dieses Ding reden. „Das Wesen ist in der Grammatik ausgesprochen“ (PU § 371, Wittgenstein, 1997b, S. 398).¹²⁸ Die 'ontologischen Eigenschaften' eines 'Dinges' bemessen sich also wesentlich nach unseren Kriterien für seine Identifikation und Individuation. Eine solche Eigenschaft ist z. B. die mögliche zeitliche Struktur einer Entität (z. B. lässt sich die Existenz des *modus ponens* nicht zeitlich datieren). Dies ist deshalb interessant, da es offensichtlich 'ontologische Eigenschaften' gibt, welche kausale Relationen ausschließen (z. B. die Eigenschaft, eine Zahl zu sein). Um in kausalen Relationen zu anderen Entitäten zu stehen, scheint es für eine Entität erforderlich zu sein, in einem näher zu bestimmenden Sinn physikalischer Natur zu sein (vgl. Davidson, 1998b, 1998f).¹²⁹ Hier wird oft vom *Prinzip der kausalen Geschlossenheit* (des Physischen) gesprochen, welches angeblich für eine wissenschaftliche Weltansicht unverzichtbar sei (vgl. Walter & Heckmann, 2003, S. 133ff.). Die Frage ist nun, ob Handlungen, Absichten und Gründe in dieser Art geeignet sind, kausale Relationen einzugehen. Es gibt einige Fakten, welche hieran Zweifel aufkommen lassen sollten.

Ich will zuerst auf Handlungen eingehen, denn hier wird man wohl kaum dem Vorwurf des Dualismus ausgesetzt sein, wenn man bestreitet, dass Handlungen geeignete Entitäten sind, um bestimmte kausale Beziehungen einzugehen. Der einfachste Weg zu sehen, dass Handlungen in einem bestimmten Sinn ungeeignet für kausale Erklärungen sind, besteht darin, zu sehen, dass Handlungen nicht in physikalischen Prozessen bestehen müssen. Auch Unterlassungen können Handlungen sein; und von einer adäquaten Theorie der Erklärung von Handlungen dürfen wir erwarten, dass sie die Erklärung von Unterlassungen nicht von vornherein ausschließt (vgl. Hornsby, 2004, S. 4ff.). Auch in den Begriffen der SMH lässt sich dieses Problem formulieren. Was die SMH maximal erklären kann, ist Folgendes: „[A]ctivity

¹²⁸ s. Kapitel 2.

¹²⁹ Dies gilt nicht unter der Voraussetzung eines interaktionistischen Dualismus. Da der interaktionistische Dualismus aber heute nicht mehr ernsthaft vertreten wird, lasse ich diese Position hier außer Betracht.

in regions concerned with motor responses and behavioural actions (e.g. striatum and anterior cingulate/supplementary motor area (SMA))“ (Bechara & Damasio, 2005, S. 343). Aktivitäten in motorischen Kortizes konstituieren aber keine Handlung und für die Bewegungen, die daraus folgen, gilt das Gleiche. Menschen können sich auch dazu entscheiden, nichts zu tun; und in diesen Fällen kann das Nichts-Tun auch eine Handlung sein (vgl. Zhu, 2004). Aber wie soll hier eine kausale Erklärung aussehen? Was wurde hier verursacht? Das Nichts-Tun bzw. die alternative Bewegung? Aber hierfür können wir doch andere hinreichende Erklärungen finden. Wurde die Nicht-Bewegung verursacht? Diese scheint ein zu leichtes Opfer für Ockhams Rasiermesser zu sein. Fand überhaupt keine Handlung statt, und muss deshalb auch nichts kausal erklärt werden? Dies widerspricht unserer Alltagserfahrung. „Warum schreibst du ihr nicht?“ – „Weil ich nicht will, dass sie denkt, sie sei mir wichtig“. Solche Phänomene können wir im Alltag erklären und wir sollten wohl von einer wissenschaftlichen Psychologie erwarten dürfen, dass sie dies ebenso kann. Nun mag man dieses Problem mittels einer Analyse kontrafaktischer Bedingungssätze lösen wollen. So können wir etwa problemlos das Nicht-Platzen eines Luftballons durch die Dicke seiner Gummihaut erklären. Aber es besteht hier ein Unterschied darin, dass das Nicht-Platzen des Luftballons von uns nicht als ein realer Sachverhalt betrachtet wird; wenn ich aber aufhöre, eine bestimmte Person in bestimmten Situationen zu grüßen, so haben wir es hier mit einem sehr realen Sachverhalt zu tun. Für mein Nicht-Grüßen werde ich sozial sanktioniert werden und ich werde mich dafür rechtfertigen müssen. Was die Analyse mittels kontrafaktischer Bedingungssätze ergeben könnte, wäre also nicht die Erklärung einer Handlung, sondern bestenfalls die Erklärung des Ausbleibens einer Handlung. In unserem Fall wollen wir aber nicht nur das Nicht-Grüßen, sondern auch das Beleidigen erklären. Im Alltag sind wir hierzu auch problemlos in der Lage und diese Fähigkeit erwarten wir auch von einer wissenschaftlichen Psychologie.

Weitere Probleme ergeben sich durch die zeitliche Struktur von Handlungen. Zunächst ist festzustellen, dass Handlungen nicht immer genau zeitlich datierbar sind. Wann hat man beispielsweise ein Auto gekauft? Wenn man festentschlossen zum Händler geht? Wenn man den Vertrag unterschreibt? Wenn man über das Auto verfügt? Wenn man die erste Rate bezahlt hat? Wenn man die letzte Rate bezahlt hat? Wenn die Handlung des Autokaufs aber durch die entsprechende Absicht verursacht worden ist, so sollte doch die kausale Wirkung zeitlich datierbar sein. Man wird hier vielleicht sagen, der Autokauf bestehe aus vielen Einzelhandlungen und diese seien die kausalen Folgen bestimmter psychischer Prozesse, nicht der Autokauf insgesamt. Aber damit stehen wir ohne eine einheitliche Erklärung für den Autokauf da,

und eine solche Erklärung kann es doch ganz offensichtlich geben (z. B. „weil die Bahn so teuer geworden ist“). Doch die Schwierigkeiten liegen hier tiefer. Dies lässt sich beispielsweise an der Tatsache erkennen, dass eine Handlung unbeschadet ihrer Einheit unterbrochen werden kann. Wenn ich z. B. bei der Reparatur meines Fahrrades unterbrochen werde, so habe ich es deshalb nicht zweimal repariert. Aber wie viele kausale Prozesse haben sich hier abgespielt? Wurde dieselbe Handlung zweimal verursacht? Wurden zwei unterschiedliche Handlungen verursacht? Wie viele Absichten bzw. Gründe waren hier ursächlich beteiligt?¹³⁰ Der Kausalist wird hier wohl geneigt sein, Konzepte wie TOTE-Einheiten (oder neuere Äquivalente solcher kybernetischen Begriffe) ins Spiel zu bringen (vgl. Miller, Galanter & Pribram, 1960; Schaub, 1993, S. 20ff.). Es stellt sich bei einer solchen Strategie aber sofort die Frage, wie es dabei um die Einheit der Handlung bestellt ist. Die angesprochenen Modelle bieten keine einheitliche Handlungserklärung, sondern vielmehr ein detailliertes Bild hierarchisch geordneter Mechanismen. Haben wir es hier noch mit einer Handlungserklärung zu tun? Viele der ‘O’s’¹³¹ solcher TOTE-Einheiten können wohl kaum als Handlungen betrachtet werden. Und wie kann zwischen ‘O’s’ unterschieden werden, welche Handlungen darstellen, und solchen, die es nicht tun? Hier gilt: „Das Explanandum der empirischen Bemühungen wird vielmehr, wenn man es nicht mehr als Handlung fassen darf, konturenlos bis zur Unkenntlichkeit“ (Greve, 1994, S. 141). Zudem wird in den angesprochenen Theorien die Fähigkeit des Menschen, sich Ziele zu setzen, einfach vorausgesetzt, dies ist aber durchaus problematisch, wenn der Anspruch besteht, Zielhaftigkeit zu naturalisieren. Allgemein lässt sich das Problem folgendermaßen formulieren: Wir individuieren Handlungen anhand ihrer zugehörigen Absichten, Ziele oder Gründe (vgl. Anscombe, 2000, S. 40f., 45ff., §§ 23, 26f.).¹³² Das hat zur Folge, dass Handlungen keine ‘übersichtlich’ zeitliche Struktur haben. Dies wiederum stellt ein Problem für den Kausalismus dar, denn bei kausalen Wechselwirkungen scheint die Frage „Wann genau?“ durchaus berechtigt zu sein. Handlungen sind schlicht keine physikalischen Vorkommnisse in einem gewöhnlichen Sinn. Und dies wirft kein gutes Licht auf eine kausale Theorie der Handlungserklärung, denn das beste Verständnis

¹³⁰ Im Rahmen der SMH würden diese Fragen folgende Form annehmen: Wie viele somatische Marker wirkten? Wann genau lief die Wechselwirkung ab? Muss derselbe somatische Marker nochmals erzeugt werden, wenn eine Handlung fortgesetzt wird? Und wenn ja, warum dann nicht, wenn die Handlung nicht unterbrochen wurde?

¹³¹ Die Abkürzung TOTE steht für *test – operate – test – exit* (vgl. Miller, Galanter & Pribram, 1960).

¹³² Hier tauchen natürlich auch Probleme auf, aber diese beziehen sich im Allgemeinen nicht darauf, dass das grundlegende Kriterium für die Individuation von Handlungen Absichten oder Gründe sind, sondern entzünden sich an der Frage, wie viele Handlungen in einzelnen Fällen ausgeführt werden (vgl. Enç, 2003, S. 85ff.; Ruben, 1999). Dabei ergibt sich unmittelbar ein weiteres Problem für den Kausalismus. Denn wenn eine Körperbewegung mehrere Handlungen ‘verwirklichen’ kann, muss wohl die Frage berechtigt sein, wie viele Ursachen hier im Spiel waren. Der Kausalist scheint also diese Möglichkeit nicht zulassen zu dürfen.

dessen, was der Kausalismus unter Kausalität versteht, besitzen wir wohl auf dem Gebiet des Physikalischen. Doch wie sieht es auf der Seite des Explanans aus? Haben Absichten, Erwartungen oder Wünsche (also Rahmenbedingungen) weniger problematische ontologische Eigenschaften?

Hier sind die Aussichten ganz offensichtlich keinen Deut besser. Wann genau hat sich meine Absicht, Psychologie zu studieren, gebildet? Wann und wie oft hat sie ihre Wirkung entfaltet? Diese Fragen scheinen keine sinnvolle Antwort zuzulassen. Ähnlich steht es mit dem, was Davidson „primäre Gründe“ nennt. Wann fing meine Proeinstellung gegenüber dem Psychologiestudium an? Wann war sie kausal relevant? Wann genau hat meine Überzeugung eingesetzt, dass ich das Ziel, Psychologie zu studieren, erreichen kann, indem ich mich für das Studium bewerbe? In welcher Form besteht diese Überzeugung, sofern ich nicht aktuell an sie denke oder sie für mein Verhalten relevant ist? Gibt es einen materiellen Zustand in meinem Gehirn, welcher die Proposition „Bewerbungen sind für das Erlangen von Studienplätzen instrumentell“ repräsentiert (vgl. Bunge & Souza, 2007)? Befindet sich dieser Zustand in einer Art ‘Überzeugungs-Kiste’ in meinem Geist (vgl. Fodor, 1998, S. 8)? Dies sind gängige Vorschläge. Aber wieviele Überzeugungen sind so in mir repräsentiert? Ich behaupte, dass ich, schon bevor ich mir jemals die Frage vorlegte, davon überzeugt war, dass die deutsche Bundeskanzlerin mindestens einen Schreibtisch besitzt oder dass die Addition von 146789 und 37682 eine eindeutige Lösung hat. Besaß ich also entsprechende Repräsentationen?

Der einzige Weg, wie man zu solch absurden Behauptungen kommen kann, besteht darin, den Menschen als eine Rechenmaschine zu betrachten. Und tatsächlich hat die sogenannte Computer-Metapher in der kognitiven Psychologie bekanntlich eine ‘große Karriere’ gemacht. Wir sagen von unseren Computern, dass auf ihren Festplatten Information gespeichert sei oder dass sie Berechnungen ausführen.¹³³ Wir reden manchmal so, als ob unsere Computer mit ‘inneren Bildern’ hantierten. Da scheint ein Verständnis psychischer Phänomene als innerer, privater, materiell realisierter Prozesse nahe zu liegen. Doch lässt sich der Mensch tatsächlich in Analogie zu einem Computer verstehen? Für diese These sollte der Kausalist wenigstens unabhängige Gründe anführen können. Doch diese sind rar (vgl. Descombes, 2001, S. 108ff.). In der Tat scheint es zweifelhaft, ob dem Ausdruck „innere Repräsentation“

¹³³ Es sei hier angemerkt, dass ich bereits die Behauptung, dass Computer Berechnungen ausführen, als höchst problematisch betrachte. Wir sagen nicht, dass ein Abakus Berechnungen ausführe, warum sollten wir es beim Computer tun? Berechnet ein geworfener Stein die Funktion einer Parabel? – Nein. – Und dennoch können wir sie von ihm ablesen.

überhaupt eine Bedeutung gegeben werden kann. Wie hätte man sich solche Repräsentationen vorzustellen? Lehrbücher der kognitiven Psychologie sagen hierzu etwa Folgendes:

„Internal or mental representations serve similar purposes to external representations. They provide a means of manipulation, exploring and describing things in their absence. Mental representations are created by processes of perception and are exchanged between cognitive processes, serving as the currency of thought“ (Andrade & May, 2004, S. 86).

Aber hier fragt sich natürlich, wer denn hier all das Manipulieren, Explorieren und Beschreiben eigentlich *tut*. Ein intelligenter Geist? Aber der sollte doch erklärt werden. Ein bloßer Mechanismus? Aber zu was sollten diesem solche Repräsentationen dienen (vgl. Bennett & Hacker, 2003, S. 68ff.; Kenny, 1991)? Was hier passiert, ist, dass mentale Phänomene vorschnell als Vorgänge und Zustände verstanden werden; womit dann natürlich sofort äußerst schwierige Fragen hinsichtlich der Natur dieser Vorgänge auftauchen. Wittgenstein kritisierte dieses Verständnis schon im Zusammenhang mit dem Behaviourismus:

„Wie kommt es nur zum philosophischen Problem der seelischen Vorgänge und Zustände und des Behaviourismus? – Der erste Schritt ist der ganz unauffällige. Wir reden von Vorgängen und Zuständen, und lassen ihre Natur unentschieden! Wir werden vielleicht einmal mehr über sie wissen – meinen wir. Aber eben dadurch haben wir uns auf eine bestimmte Betrachtungsweise festgelegt. Denn wir haben einen bestimmten Begriff davon, was es heißt: einen Vorgang näher kennen zu lernen. (Der entscheidende Schritt im Taschenspielerkunststück ist getan, und gerade er schien uns unschuldig.)“ (PU § 308, Wittgenstein, 1997b, S. 377f.).

Zu all diesen Problemen kommt hinzu, dass der Kausalismus auf eine völlig ‘überfrachtete’ Ontologie festgelegt ist. Denn ganz offensichtlich sind Absichten, Proeinstellungen und Überzeugungen nicht hinreichend, um Handlungen zu verursachen. Nicht jede Absicht führt unmittelbar zur Handlung. Damit ist aber auch klar, dass hier ein zusätzliches Element benötigt wird: Ein Auslöser, der festlegt, wann die richtige Zeit gekommen ist, um die Absicht zu verwirklichen. Während ich hier sitze und diese Arbeit schreibe, kann jeder meiner Tastenanschläge als Handlung betrachtet werden. Dann muss aber jeder dieser Anschläge nicht nur durch eine Absicht verursacht worden sein, sondern auch durch einen entsprechenden Auslöser zur Verwirklichung gebracht worden sein. Wir hätten es hier also mir unübersehbar vielen Ereignissen und Prozessen zu tun. Dies ist zwar an sich kein Problem, aber es sollten doch Zweifel aufkommen, ob wir hier nicht unseren Begriff der Absicht bzw. der Über-

zeugung enorm überstrapazieren. Reden wir hier noch von dem, was wir mit diesen Begriffen meinen? Dass jeder meiner Anschläge verursacht worden ist, kann der Intentionalist durchaus zugeben. In Frage steht aber, ob wir es hier mit Absichten oder psychischen Phänomenen (Rahmenbedingungen von Handlungserklärungen) zu tun haben (vgl. Hursthouse, 2000). Abschließend lässt sich hier nur Rosalind Hursthouse zustimmen, wenn sie schreibt:

„Now I maintain that the really fundamental fault here is the existence claim – the thought of all these beliefs, desires, intentions or whatever as already there, existing in the mind – which, given causal realism, is required by the causalist claim. What is the argument for their existence? It might be thought to be no more than causal realism plus the argument for the causal thesis. But actually there isn't, in the literature, an argument for the causalist thesis. The only one that is standardly given, i.e. that the 'because' in 'He Φ -ed because he wanted x / believes that Φ -ing was a means to x ' is not equivalent to 'and', shows only that a reason for action explains the action. This leaves open whether the sort of explanation is event-causality or whether the 'because' is one of Aristotle's other three 'because's', for instance, the teleological one“ (Hursthouse, 2000, S. 95f.).

5.5. Zusammenfassung

Ich will kurz die Argumentation dieses Kapitels zusammenfassen. Ich habe für die These argumentiert, dass der Satz (PG) kein empirisches Verhältnis wiedergibt, welches einer Reduktion im Sinne der SMH zugänglich wäre. Hierzu habe ich dargelegt, dass (PG) gemäß der SMH ein kausales Verhältnis wiedergeben muss. Da Verhalten, welches auf Entscheidungen beruht, eine Handlung darstellt und genau dieses Verhalten durch die SMH erklärt werden soll, sind die Vertreter der SMH damit auf eine kausalistische Theorie der Handlungserklärung festgelegt. Ich habe versucht, die Falschheit einer solcher kausalistischen Theorie der Handlungserklärung in zwei Schritten zu zeigen.

Der erste Schritt begann mit der Feststellung, dass Handlungserklärungen Erklärungen aus Gründen sind. Danach wurde dargelegt, dass die logische Möglichkeit menschlicher Rationalität verlangt, dass Gründe gute Gründe sein können müssen. Da aber, wie anschließend ebenfalls gezeigt werden konnte, gute Gründe sehr selten in psychischen Phänomenen bestehen, sondern in der Regel Sachverhalte darstellen, folgt hieraus, dass psychische Phänomene in aller Regel nicht selbst in Handlungserklärungen auftauchen, sondern lediglich deren Rahmenbedingungen bilden. Dies schließt emotionale Bewertungen, wie sie durch die SMH zur Handlungserklärung herangezogen werden, ein. Allgemein spielen

Emotionen in Handlungserklärungen nicht die Rolle der Erklärung selbst, sondern bilden die Begründungsmuster, welchen solche Erklärungen folgen. Da gute Gründe auch in *abstrakten* Sachverhalten bestehen können, ist eine kausalistische Position hier von vornherein ausgeschlossen.

In einem zweiten Schritt habe ich versucht zu zeigen, dass selbst für die abgeschwächte These, dass der Satz (PG) zwar keine Grundlage für Handlungserklärungen bieten kann, aber dennoch als Rahmenbedingung von Handlungserklärungen ein kausales Verhältnis wiedergibt, schwerwiegende Probleme entstehen. Hier hätten wir es mit einer klassischen kausalistischen Position zu tun, gegen welche seit geraumer Zeit vielfältige Argumente bekannt sind. Deshalb wurde das einflussreichste Argument für den Kausalismus (*Davidson's challenge*) kritisch besprochen, um anschließend die wichtigsten Argumente gegen eine solche Position kurz wiederzugeben. Das Ergebnis dieses Kapitels lässt sich folgendermaßen formulieren: Anders als die SMH voraussetzt, gibt (PG) kein kausales Verhältnis wieder und ist damit keiner reduktionistischen Erklärung zugänglich. Tatsächlich stellt (PG) einen begrifflichen Zusammenhang dar, welcher uns angibt, wie wir den Begriff „emotionale Bewertung“ im Zusammenhang von Rahmenbedingungen für die Erklärung von Handlungen aus Gründen verwenden sollten.

6. ERGEBNISSE UND DISKUSSION

Das folgende Kapitel wird zunächst einen Überblick über die erarbeiteten Ergebnisse bieten und anschließend einige mögliche Einwände gegen die vorgetragene Argumentation besprechen. Des Weiteren werde ich versuchen darzulegen, welche Teile der SMH als gehaltvolle empirische Hypothesen betrachtet werden können, um schließlich die Möglichkeit zu diskutieren, die gewonnenen Ergebnisse zu verallgemeinern.

6.1. Ergebnisse

Die vorgetragene Argumentation ging von der Feststellung aus, dass es einen grundlegenden Unterschied zwischen begrifflichen und empirischen Relationen gibt. Es ist eine Aufgabe der Psychologie sowie der kognitiven Neurowissenschaft, zwischen diesen Beziehungen zu unterscheiden. Anderenfalls können Behauptungen als das Ergebnis empirischer Forschung präsentiert werden, die entweder die Grenze sinnvollen Sprachgebrauchs überschreiten oder Wahrheiten sind, welche auf unserem Sprachgebrauch und nicht auf der Beschaffenheit des Gegenstandsgebiets der jeweiligen Wissenschaft beruhen (vgl. Bennett & Hacker, 2003). Hieran anschließend wurde die SMH dargestellt und der Stand der einschlägigen empirischen Forschung wiedergegeben. Dabei erwies sich die SMH als eine breit rezipierte, aktuelle und fruchtbare neurobiologische Theorie, welche beansprucht menschliches Entscheidungsverhalten zu erklären. Darauf aufbauend wurde versucht, einen genuin psychologischen Gehalt der SMH herauszuarbeiten. Dies gelang vor allem durch die Trennung zweier Beschreibungsebenen: einer neurobiologischen Beschreibungsebene, welche sich auf subintentionalem und subpersonalem Niveau bewegt und einer psychologischen Ebene, welche auf der Ebene des gesamten Organismus anzusiedeln ist, und damit auch auf intentionalem und personalem Niveau. Es stellt sich heraus, dass es die genuin psychologische Behauptung der SMH ist, dass die emotionale Bewertung erwarteter Handlungsergebnisse die Auswahl von Handlungsoptionen bestimmt. Diese These wurde als (PG) bezeichnet. Es wurde des Weiteren gezeigt, dass die SMH annehmen muss, dass (PG) eine kausale Relation wiedergibt. Schließlich wurde dargelegt, warum eben diese Annahme falsch ist. Hierbei habe ich mich wesentlich darauf gestützt, dass Handlungserklärungen Erklärungen aus Gründen sind und dass Gründe immer gute Gründe sein können müssen. Auf dieser Grundlage ließ sich zeigen, dass emotionale Bewertungen, ebenso wie die meisten anderen psychischen Zustände, nicht dazu geeignet sind, Handlungen zu erklären. Sie spielen hierbei vielmehr die Rolle von Rahmen-

bedingungen, welche eine tatsächliche Handlungserklärung erst möglich machen. Darüber hinaus wurde gezeigt, dass psychische Zustände, auch wenn man sie als Rahmenbedingungen versteht, Handlungen nicht kausal erklären können. Als direktes Ergebnis dieser Arbeit lässt sich also festhalten: *Die SMH ist falsch.*

Dieses unmittelbare Ergebnis ist natürlich zunächst auf die SMH in ihrer derzeitigen und hier besprochenen Form beschränkt. Ich bin jedoch der Überzeugung, dass die hier vorgetragenen Überlegungen sehr allgemeiner Natur sind, sodass es schwierig ist, zu erkennen, wie die angesprochenen Schwierigkeiten von anderen neurowissenschaftlichen Theorien umgangen werden können. Ich werde auf die Verallgemeinerbarkeit der Ergebnisse zurückkommen. Doch zunächst sollen einige möglichen Einwände gegen meine Argumentation besprochen werden.

6.2. Mögliche Einwände

Meines Erachtens lassen sich drei Einwände gegen die vorgetragene Position vorbringen. Erstens lässt sich daran zweifeln, ob ‘theoretische Spekulationen’ und ‘psychologisches Alltagsverständnis’ dazu geeignet sind, empirische Theorien zu widerlegen. Zweitens lässt sich einwenden, dass die vorliegende Arbeit die Möglichkeit teleologischer Erklärungen übersehen habe. Und drittens gilt es den Vorwurf zu entkräften, die vorliegende Arbeit vertrete eine Form des Epiphänomenalismus.

6.2.1. Modus tollens vs. modus ponens

Bekanntlich sagt uns die Logik, wenn wir mit inkonsistenten Überzeugungen konfrontiert sind, nicht welche der beteiligten Überzeugungen wir aufzugeben haben (vgl. Russell, 2008, S. 22f.). Die Logik sagt uns lediglich, dass wir *irgendwelche* der beteiligten Überzeugungen aufgeben müssen. Auf dieser Grundlage könnte ein Einwand gegen hier vorgetragene Argumentation vorgebracht werden. Ein Opponent könnte fragen: „Ist es nicht vernünftiger, unsere alltagspsychologischen Überzeugungen bezüglich der adäquaten Erklärung menschlicher Handlungen fallen zu lassen als eine wissenschaftliche Theorie zu bezweifeln, welche zumindest die Vorteile hat, auf ‘harten Daten’ zu beruhen und wissenschaftlich kritisierbar zu sein?“ Tatsächlich ist die sogenannte ‘naive Psychologie’ bei vielen wissenschaftlichen Psychologen nicht hoch angesehen. Was sollte uns also dazu bewegen, bei der festgestellten Unverträglichkeit von SMH und nicht-kausalistischer Handlungstheorie uns für die nicht-kausalistische Handlungstheorie zu entscheiden? Warum sollten wir die Tatsache, dass die SMH den Kausalismus impliziert, nicht als ein Argument für den Kausalismus und nicht

gegen die SMH auffassen? Russell spricht in diesem Zusammenhang von: „*the cheap (but effective), ‘your modus ponens is my modus tollens’ defense*“ (Russell, 2008, S. 13).

Die Antwort hierauf lautet, dass es sich hier nicht um eine naive alltagspsychologische *Theorie* handelt, sondern um die *Bedeutung der Worte*, mit welchen wir uns verständigen. Was das Wort „Handlung“ bedeutet, wird zum Teil dadurch bestimmt, dass wir bei Handlungen sinnvoll nach Gründen fragen können (vgl. Anscombe, 2000). Für die Bedeutung des Wortes „Grund“ wiederum ist die Tatsache teilkonstitutiv, dass wir wissen, wann wir von „guten Gründen“ reden können. Der Vorschlag, der SMH den Vorzug gegenüber einer nicht-kausalistischen Handlungstheorie zu geben, läuft also darauf hinaus, für eine Veränderung unseres Sprachgebrauchs zu plädieren. Eine solche Forderung unterliegt natürlich keinen prinzipiellen Einwänden; nur muss man sich darüber klar sein, dass, was immer man mit Theorien innerhalb des neuen Sprachgebrauchs erklären kann, es kann sich dabei nicht um Handlungen nach unserem derzeitigen Verständnis handeln. Wie wir das Wort „Handlung“ gebrauchen, bestimmt dessen Bedeutung. Deshalb bedeutet dieses Wort vor und nach der Veränderung unseres Sprachgebrauchs nicht dasselbe; und hiermit ist klar, dass die neue Theorie keine Handlungen, sondern höchstens Handlungen*^{NEU} erklären kann (vgl. Bennett & Hacker, 2003, S. 75). Es ist also nicht eine im Alltag verwurzelte psychologische Theorie, welche vorzuziehen ist, sondern es sind die im Alltag verwurzelten Begriffe, die hier Priorität genießen. Dass diese Begriffe hier Vorrang haben, liegt darin begründet, dass sie die Grundlage dafür sind, überhaupt diejenigen Fragen zu stellen, welche die Wissenschaft beantworten will.

6.2.2. Begrifflicher oder teleologischer Zusammenhang?

Ein zweiter Einwand gegen die vorgelegte Argumentation könnte lauten, dass ich voraussetzen scheine, dass Handlungserklärungen entweder kausaler oder begrifflicher Natur sind. Aber es könnte noch eine dritte Möglichkeit geben: Handlungserklärungen könnten *teleologische Erklärungen* sein. Eine solche teleologische Erklärung würde auf mehr basieren als nur einem begrifflichen Verhältnis, wäre aber dennoch keine effizient-kausale Erklärung. In wieweit eine solche Erklärung *empirischen* Gehalt hätte, wäre in diesem Fall näher zu prüfen. Meine defensive Antwort auf diesen Einwand lautet schlicht, dass ich diese Frage offen lassen will. Ich beanspruche, gezeigt zu haben, dass die SMH keine adäquate Erklärung menschlichen Handelns bieten kann. Ob eine adäquate Erklärung von Handlungen ausschließlich auf begrifflichen Relationen beruht oder auch teleologische Beziehungen einschließen muss, ist deshalb für den von mir erhobenen Anspruch nicht relevant.

Meine inhaltliche Antwort soll lediglich aus einigen Bemerkungen und Hinweisen bestehen: Seit Beginn der Neuzeit und dem Siegeszug der Naturwissenschaften hat sich die europäische Auffassung von Kausalität auf eine der vier Ursachentypen des Aristoteles eingeeengt: die effiziente Ursache. Dies wurde allgemein als Fortschritt betrachtet, und auch die Psychologie betrachtet im Allgemeinen teleologische Erklärungen als nicht wissenschaftlich respektabel. Die entscheidende Frage, welche es hier zu stellen gilt, ist: Verbirgt sich hinter dieser Auffassung tatsächlich mehr als ein Dogma? In jüngster Zeit häufen sich die Hinweise, dass dies nicht der Fall ist (vgl. Dancy, 2000, 2004; Müller, 1992; Schon, 2005; Thompson, 2008). Natürlich mag so mancher moderne Forscher vor einer aristotelischen Naturteleologie zurückschrecken, doch hier sind keine emotionalen Reaktionen, sondern Argumente gefragt. Und bis hierzu schlagende Argumente vorgetragen werden, sollte die Psychologie mit der Möglichkeit rechnen, teleologische Erklärungen anerkennen zu müssen.

6.2.3. Epiphänomenalismus

Die vorliegende Arbeit behauptet, dass psychische Zustände und Prozesse nicht kausal relevant für unsere Handlungen sind. Damit liegt die Vermutung nahe, dass hier eine Form des Epiphänomenalismus vertreten wird. Da der Epiphänomenalismus im Allgemeinen als eine äußerst unattraktive Position bewertet wird, kommt diese Vermutung einer Kritik gleich. Der Unterschied zwischen der hier vertretenen Position und dem Epiphänomenalismus liegt in einer radikal anderen ontologischen und wissenschaftstheoretischen Grundhaltung. Der Epiphänomenalismus scheint deshalb eine unattraktive Position zu sein, weil er psychische Zustände und Prozesse überflüssig zu machen scheint. Laut Epiphänomenalismus kann jedes Geschehen adäquat kausal erklärt werden, nur ergibt sich bei manchen Prozessen ein phänomenaler Bewusstseinszustand als 'Nebenprodukt' (vgl. Pauen, 2001, S. 72f.). Die hier vertretene Position hingegen behauptet, dass nicht jedes Geschehen adäquat kausal erklärt werden kann; namentlich Handlungen können dies nicht. Insbesondere kann nicht jedes Geschehen unter der für uns relevanten Beschreibung (intensional adäquat) kausal erklärt werden. Dass eine rein extensional korrekte Erklärung für die Wissenschaft stets ausreiche, ist ganz offensichtlich falsch. Es ist darüber hinaus falsch, sich psychische Phänomene als (private) 'Dinge', im Sinn von Gegenständen, welche den zugehörigen Substantiven entsprächen, vorzustellen. Psychische Prozesse und Zustände spielen in unserem Leben eine durchaus relevante Rolle, unter anderem als Rahmenbedingungen für Handlungserklärungen. Aber hieraus lässt sich nicht auf ihre kausale Wirksamkeit schließen. In anderen Worten: Die vorgetragene Position schließt nicht von der kausalen Ineffizienz eines Phänomens auf dessen

wissenschaftliche Irrelevanz. Und dies reicht aus, um dem Vorwurf des Epiphänomenalismus zu begegnen.

6.3. Was bleibt von der SMH?

Das Ergebnis dieser Arbeit ist, dass die SMH falsch ist. Dies scheint ein ausschließlich negatives Ergebnis zu sein. Lässt sich kein positiver Erkenntnisfortschritt erzielen? Auf diese Frage sind zwei unterschiedliche Antworten zu geben:

Erstens ist die Widerlegung einer Theorie durchaus kein rein negatives oder destruktives Ergebnis, denn die Wissenschaft entwickelt sich nun einmal durch das Vorbringen von Vorschlägen und deren Kritik. Auch wenn die Rolle der Widerlegung in einigen wissenschaftstheoretischen Ansätzen – wie dem kritischen Rationalismus – deutlich überbewertet wird, so steht doch fest, dass ihr eine gewichtige Rolle innerhalb der Wissenschaft gebührt.

Zweitens lassen sich aus den hier gewonnenen Ergebnissen durchaus produktive Schlüsse ziehen. Sofern meine Argumentation gegen eine kausalistische Handlungstheorie korrekt ist, ist auch klar, dass wir in der Psychologie nicht zu adäquaten Handlungserklärungen kommen, indem wir versuchen, Gesetze des Handelns zu finden und diese anschließend auf neurobiologische Gesetze zu reduzieren. Es ist wichtig, diese Kluft zwischen intentionalen und nicht-intentionalen Erklärungen anzuerkennen. Die Neurobiologie sollte sich um nicht-intentionale Erklärungen bemühen, da dies die den Naturwissenschaften angemessene Art von Erklärungen ist. Das heißt aber auch, dass solche Wissenschaften charakteristisch andere Fragen stellen, verglichen mit Fragen auf einer psychologischen und intentionalen Beschreibungsebene. Wir haben hier also zwei Typen von Fragen zu unterscheiden. Ein Typ von Frage zielt auf eine kausale Erklärung körperlicher Veränderungen, der andere Typ zielt auf das Nennen von Gründen. Anscombe sagt hierzu: *„There might come a point at which there is no further puzzle about how each link in the causal chain produces the next one. And still nothing has been said about intentions, beliefs, thoughts or decisions“* (Anscombe, 2005, S. 100). Und ganz ähnlich Sehon: *„Behavior that is apparently purposeless troubles us; we long for an explanation. And our brain scanner’s report about which brain state is causing the behavior will not help; we want to know the reason for the behavior“* (Sehon, 2005, S. 226). Wir sollten also von den kognitiven Neurowissenschaften keine Erklärung von Handlungen *qua Handlungen* verlangen, sondern eine kausale Erklärung des Zustandekommens der Veränderungen unterschiedlicher Teile des menschlichen Körpers. Es ist schlicht der Anspruch der SMH, sich auf einer intentionalen Beschreibungsebene zu bewegen, welcher zu den von mir vor-

getragenen Argumenten Anlass bot. Die einfachste Möglichkeit, jenen Teil der SMH zu isolieren, welchen wir als empirisch gehaltvolle Hypothese betrachten können, besteht deshalb darin, alle Behauptungen der SMH, die sich auf einer intentionalen Beschreibungsebene befinden, fallen zu lassen. Was würde nach einer solchen Revision übrig bleiben? Um diese Frage zu beantworten, müssen wir die Kausalbeziehungen, welche die SMH behauptet, ohne die Verwendung psychologischer Prädikate oder anderer intentionaler Ausdrücke formulieren. Dies könnte etwa folgendermaßen aussehen: Bestimmte Muskelkontraktionen des Menschen werden durch neuronale Aktivität im primären motorischen Kortex sowie in supplementär motorischen Arealen und den Basalganglien verursacht. In manchen Fällen wird diese Aktivität wiederum vom Feuern bestimmter Neurone in somatosensorischen Kortizes¹³⁴ (mit-)verursacht. Das Feuern dieser Neurone wird wiederum durch neuronale Aktivität im VMPFC verursacht, wobei hier zwei unterschiedliche Wege der Kausalwirkung unterschieden werden müssen. Die Aktivität im VMPFC kann die Aktivität in den somatosensorischen Kortizes entweder vermittelt über eine Veränderung der Viszera und anderer Reaktionen des autonomen Nervensystems herbeiführen (*body-loop*), oder durch die Einwirkung auf Kerne des Hirnstamms, welche dann ihrerseits direkt auf die somatosensorischen Kortizes wirken (*as-if-body-loop*). Die neuronale Aktivität im VMPFC kann durch bestimmte Aktivitätsmuster verschiedener primärer sensorischer Kortizes verursacht werden.

Die Annahme eines solchen Zusammenhangs kann als der Ausgangspunkt einer empirisch gehaltvollen neurobiologischen Theorie betrachtet werden; auch wenn es hier an der wünschenswerten Präzision fehlt. Im gegenwärtigen Zusammenhang ist nur wichtig, dass in der angesprochenen Theorie nichts gesagt wird, was unser psychologisches Interesse beanspruchen kann. Es ist sehr gut möglich, dass eine Weiterentwicklung der SMH zu einer sehr überzeugenden Theorie bezüglich der Verursachung verschiedener Muskelkontraktionen, Extremitätsbewegungen oder Stimmbandschwingungen kommt. Doch wahrscheinlich wird es solche Fortschritte nur geben, wenn sich Theorien von der Art der SMH darauf besinnen, dass sie ihre Aufgabe ohne die Verwendung psychologischen bzw. intentionalen Vokabulars erfüllen müssen.

Was die Entwicklungsmöglichkeiten der SMH in diese Richtung anbelangt, sollte jedoch auch auf ein Problem hingewiesen werden: Die Beschreibung einer körperlichen Veränderung als Entscheidungsverhalten beruht auf intentionalem Vokabular. Das Verhalten, welches die SMH erklären will, wird aber genau durch seine Klassifizierung als Entscheidungsverhalten

¹³⁴ Es sei hier nochmals auf den eigenwilligen Gebrauch des Terminus „somatosensorische Kortizes“ innerhalb der SMH hingewiesen (s. Fußnote 56 auf S. 81).

bestimmt. Es ist somit fraglich, ob die SMH, wenn sie den illegitimen Gebrauch solchen Vokabulars aufgibt, noch die Ressourcen besitzt, den Gegenstand ihrer Erklärungen zu identifizieren. Aus der Perspektive der vorliegenden Arbeit wäre es also eine entscheidende Zukunftsaufgabe der SMH, eine Möglichkeit zu finden, ihren Untersuchungsgegenstand auf der Grundlage einer nicht-intentionalen Beschreibung herauszugreifen. Sollte diese Aufgabe gelingen, ist eine produktive Weiterentwicklung der SMH durchaus möglich. Doch sollte sie dabei immer als eine biologische und keinesfalls als eine psychologische Hypothese verstanden werden. Das menschliche Nervensystem bietet die Möglichkeit, unzählige kausale Relationen zu entdecken und zu erforschen, doch die *Gründe* für das *Handeln* von *Personen* werden wir darin niemals finden.

6.4. Verallgemeinerbarkeit der Ergebnisse

Es wird dem Leser nicht entgangen sein, dass ich stets bemüht war, meine Argumentation möglichst allgemein zu halten und auch Möglichkeiten auszuschließen, welche den Vertretern der SMH durchaus fernliegen.¹³⁵ Dies erlaubt es durchaus optimistisch bezüglich der Frage zu sein, inwiefern die Ergebnisse dieser Arbeit sich auf andere neurobiologische Theorien menschlichen Handelns übertragen lassen. Um die Frage näher zu erörtern, möchte ich einige neuere Entwicklungen der empirischen Forschung kommentieren.

Bekanntlich befindet sich die kognitive Neurowissenschaft immer noch in einer Phase rapider Expansion.¹³⁶ Die vielen neuen Erkenntnisse, welche in den letzten Jahrzehnten über das menschliche Nervensystem erzielt wurden, führten bei vielen Forschern zu der Hoffnung, auch zielgerichtetes menschliches Verhalten bald neurobiologisch adäquat erklären zu können. Da Entscheidungsverhalten als ein paradigmatischer Typ zielgerichteten menschlichen Verhaltens angesehen werden kann und darüber hinaus leicht experimentell untersuchbar scheint, haben sich besonders große Erwartungen an die Erklärung von entscheidungsbasiertem Verhalten geknüpft. Die SMH ist nur eine der vielen Ansätze, welche auf die Erfüllung dieser Erwartungen zielen. In den letzten Jahren nahm sich ein gesamtes neues Forschungsgebiet dieser Aufgabe an, welches oft als *neuroeconomics* bezeichnet wird. „[N]euroeconomics is now striving to better understand decision-making behavior“ (Sanfey, 2007, S. 151). Dieses neue Forschungsgebiet versucht insbesondere, mathematische Modelle

¹³⁵ Dies war beispielsweise der Grund für die Besprechung des Eliminativismus in Kapitel 4.1.1.

¹³⁶ Gerd Nollmann merkt hierzu an: „Deshalb fließen auch die Forschungsgelder reichlich – nicht zuletzt, weil die Hirnforscher ‘harte’ kausale ‘Tiefen’-Erklärungen auf experimenteller und quantitativer Basis anbieten“ (Nollmann, 2006, S. 62). Wie verfehlt und trügerisch dieser Drang nach ‘harten’, ‘tiefen’ und ‘quantitativ belegten’ Erklärungen ist, wurde oben dargelegt.

aus den Wirtschaftswissenschaften mit der neurobiologischen Forschung zu verbinden. Dabei wird das Gehirn als „*decision-making organ*“ (Gintis, 2007, S. 2f.) betrachtet. Und es wird die Hoffnung gehegt, die mathematisch modellierten Parameter bestimmter Entscheidungstheorien mit bestimmten Variablen im zentralen Nervensystem identifizieren zu können.

„[S]upporting evidence from contemporary neuroscience suggests that maximization is not simply an ‘as-if’ story. In fact, the brain’s neural circuitry makes choices by internally representing the payoffs of various alternatives as neural firing rates, choosing such a maximal rate“ (Gintis, 2007, S. 5).¹³⁷

Es soll hier also ein Maximalwert, welcher einem erwarteten Handlungsergebnis zugeordnet wird, eine bestimmte Entscheidung kausal erklären können. Die Natur dieses Maximalwertes bleibt dabei meist unklar. Verallgemeinernd wird in diesem Zusammenhang oft von *expected utility* gesprochen, wobei diese Größe definiert wird als das Produkt der Wahrscheinlichkeit einer ‘Belohnung’ und dem ‘Wert’ der Belohnung; meist wird darüber hinaus die höchst unplausible Annahme vertreten, dass sich der ‘Wert’ einer Belohnung mittels der Bezugnahme auf eine Geldmenge darstellen lässt (vgl. z. B. Rolls, 2005, S. 340ff.). Auf dieser Grundlage sucht man dann neuronale Prozesse, welche hoch mit solchen ‘Werten’, ‘Wahrscheinlichkeiten’, ‘Belohnungen’ und ‘*expected utility*’ korrelieren. Es werden gleichsam die Referenten jedes einzelnen Terms des mathematischen Modells gesucht. Es ist klar, dass hier nicht weniger nach einer mechanistischen, kausalen Erklärung menschlichen Handelns gesucht wird als innerhalb der SMH. Wobei sich die Theorien dabei nur oberflächlich unterscheiden, je nachdem, welche mathematische Modellierung einer Entscheidungstheorie zugrunde gelegt wird, welche neuronalen Prozesse als die Referenten der Variablen des Modells identifiziert werden und welche allgemeinen Annahmen über die Natur dieser Prozesse gemacht werden (z. B. ob diese als emotionale oder als ‘rein kognitive’ Prozesse verstanden werden).

Dass hier versucht wird, zielgerichtetes Handeln kausal zu erklären, wird schon aus den Titeln der in diesem Gebiet zurzeit veröffentlichten Artikel und Bücher deutlich: *Toward a Mechanistic Understanding of Human Decision Making* (O’Doherty & Bossaerts, 2008), *Neuroeconomics - Opening the "Black Box" Behind the Economic Behavior* (Yu & Zhou, 2007), *Social Neuroeconomics - The Neural Circuitry of Social Preferences* (Fehr & Camerer, 2007), *Taking Action - Cognitive Neuroscience Perspectives on Intentional Acts* (Johnson-Frey,

¹³⁷ Es ist natürlich völlig unklar, wie sich jemand anhand der Feuerraten von Neuronen entscheiden kann, sofern er von diesen Feuerraten nicht das Geringste weiß. In solchen Fällen verwenden wir weder das Wort „entscheiden“ noch „auswählen“ (vgl. Bennett & Hacker, 2003, S. 70). Herbert Gintis überschreitet hier also die Grenzen des sinnvollen Sprachgebrauchs (s. Kapitel 2.3).

2003), *Exploring the Brain Basis of Joint Action - Co-Ordination of Actions, Goals and Intentions* (Newman-Norlund, Noordzij, Meulenbroek & Bekkering, 2007) oder *Neuroscience of Rule-Guided Behavior* (Bunge & Wallis, 2007). Diese Liste ließe sich beliebig lange fortsetzen. Der Leser dieser Schriften wird nicht enttäuscht werden, wenn er angebliche neurobiologische Kausalerklärungen zielgerichteter Handlungen erwartet. Es werden also zwei Dinge deutlich: Erstens zeigt die Neurowissenschaft in jüngster Zeit ein starkes Interesse an Verhaltensweisen, welche als Handlungen bezeichnet werden müssen. Zweitens verstehen sich alle Erklärungen, welche die Neurowissenschaft hierzu vorgebracht hat, als Kausalerklärungen. Damit ist aber genau der Grundzug der in dieser Arbeit am Beispiel der SMH besprochenen Schwierigkeiten benannt. Die Neurowissenschaft bietet Kausalerklärungen für Phänomene an, deren adäquate Erklärung nicht kausaler Natur ist.

Hieraus lässt sich erkennen: Wenn die hier vertretene These, dass Handlungen nicht adäquat durch Kausalerklärungen erklärt werden können, korrekt ist, so müssen neurobiologische Erklärungen menschlichen Handelns ihrem Gegenstand unangemessen sein. Um die Übertragbarkeit der hier gewonnenen Ergebnisse auf eine andere Theorie wahrscheinlich zu machen, genügt es also zu zeigen, dass die Theorie Handlungen kausal erklären will. Natürlich muss jede Theorie als neuer Einzelfall betrachtet und analysiert werden, doch die allgemeine Struktur der Problematik ist höchst einfach; und es ist nicht einzusehen, wie neurobiologische Theorien diese Schwierigkeit umgehen könnten. Jede naturwissenschaftliche Theorie menschlichen Handelns steht vor der Aufgabe, ihre Position bezüglich der Natur von Handlungserklärungen darzulegen, zu verteidigen und mit der von ihr angebotenen Handlungserklärung in Einklang zu bringen. Die neurowissenschaftlichen Erklärungen menschlicher Handlungen sind dieser Aufgabe bisher nicht hinreichend nachgekommen. Es bleibt hier also noch viel Forschung zu leisten. Doch die Aufgabe der vorliegenden Arbeit war es, darauf hinzuweisen, dass diese lange bekannten Probleme sich nicht dadurch lösen lassen, dass die Forschung von einem kognitivistischen Paradigma zu einem neurowissenschaftlichen übergeht.

LITERATUR

- Achinstein, P. (1983). *The nature of explanation*. New York: Oxford University Press.
- Adolphs, R. & Damasio, A. R. (2000). Neurobiology of emotion at a systems level. In J. C. Borod (Hrsg.), *The neuropsychology of emotion* (S. 194-213). New York: Oxford University Press.
- Ajzen, I. & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Andrade, J. & May, J. (2004). *Cognitive psychology* (2. Aufl.). London: Bios Scientific Publishers.
- Anscombe, G. E. M. (2000). *Intention* (2. Aufl.). Cambridge, MA: Harvard University Press.
- Anscombe, G. E. M. (2005). The causation of action. In M. Geach & L. Gormally (Hrsg.), *Human life, action and ethics: Essays by G.E.M. Anscombe* (S. 89-108). Exeter: Imprint Academic.
- Aristoteles. (2000). *Metaphysik: Schriften zur ersten Philosophie*. Stuttgart: Reclam.
- Audi, R. (1993). *Action, intention, and reason*. Ithaca: Cornell University Press.
- Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, 4, 829-839.
- Bar-Elli, G. (1996). *The sense of reference: Intentionality in Frege*. Berlin: de Gruyter.
- Bar-On, R., Tranel, D., Denburg, N. L. & Bechara, A. (2003). Exploring the neurological substrate of emotional and social intelligence. *Brain*, 126, 1790-1800.
- Bard, P. (1934). Emotion: I. The neuro-humoral basis of emotional reactions. In C. Murchison (Hrsg.), *Handbook of general experimental psychology* (S. 264-311). Worcester, MA: Clark University Press.
- Barone, D. F., Maddux, J. E. & Snyder, C. R. (1997). *Social cognitive psychology: History and current domains*. New York: Plenum Press.
- Bechara, A. (2002). The neurology of social cognition. *Brain*, 125, 1673-1675.
- Bechara, A. (2004). A neural view of the regulation of complex cognitive functions by emotion. In P. Philippot & R. S. Feldman (Hrsg.), *The regulation of emotion* (S. 3-32). Mahwah, NJ: Lawrence Erlbaum.
- Bechara, A. & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, 52, 336-372.
- Bechara, A., Damasio, A. R., Damasio, H. & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, 50, 7-15.
- Bechara, A., Damasio, H. & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, 10, 295-307.
- Bechara, A., Damasio, H. & Damasio, A. R. (2001). Manipulation of dopamine and serotonin causes different effects on covert and overt decision-making. *Society for Neuroscience Abstracts*, 27.
- Bechara, A., Damasio, H. & Damasio, A. R. (2003). Role of the amygdala in decision-making. In P. Shinnick-Gallagher, A. Pitkanen, A. Shekhar & L. Cahill (Hrsg.), *The amygdala in brain function: Basic and clinical approaches* (S. 356-369). New York: New York Academy of Sciences.
- Bechara, A., Damasio, H., Damasio, A. R. & Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *The Journal of Neuroscience*, 19, 5473-5481.

- Bechara, A., Damasio, H., Tranel, D. & Anderson, S. W. (1998). Dissociation of working memory from decision making within the human prefrontal cortex. *The Journal of Neuroscience*, 18, 428-437.
- Bechara, A., Damasio, H., Tranel, D. & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, 275, 1293-1295.
- Bechara, A., Damasio, H., Tranel, D. & Damasio, A. R. (2005). The Iowa Gambling Task and the somatic marker hypothesis: Some questions and answers. *Trends in Cognitive Sciences*, 9, 159-162.
- Bechara, A., Tranel, D., Damasio, H. & Damasio, A. R. (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex*, 6, 215-225.
- Bechtel, W. (1994). Levels of description and explanation in cognitive science. *Minds and Machines*, 4, 1-25.
- Bechtel, W. & Hamilton, A. (2007). Reductionism, integration, and the unity of the sciences: Natural, behavioral, and social sciences and the humanities. In T. A. F. Kuipers (Hrsg.), *General philosophy of science: Focal issues* (S. 377-430). Amsterdam: Elsevier.
- Beckermann, A. (2001). *Analytische Einführung in die Philosophie des Geistes* (2. Aufl.). Berlin: de Gruyter.
- Beer, J. S. (2007). The importance of emotion: Social cognition interactions for social functioning - Insights from orbitofrontal cortex. In E. Harmon-Jones & P. Winkielman (Hrsg.), *Integrating biological and psychological explanations of social behavior* (S. 15-31). New York: Guilford Press.
- Bennett, M. R. & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience*. Malden, MA: Blackwell.
- Bennett, M. R. & Hacker, P. M. S. (2005). Emotion and cortical-subcortical function: Conceptual developments. *Progress in Neurobiology*, 75, 29-52.
- Bickle, J. (2003). *Philosophy and neuroscience: A ruthlessly reductive account*. Dordrecht: Kluwer.
- Bishop, J. (1989). *Natural agency: An essay on the causal theory of action*. Cambridge: Cambridge University Press.
- Bittner, R. (2001). *Doing things for reasons*. Oxford: Oxford University Press.
- Blair, K. S., Finger, E., Marsh, A. A., Morton, J., Mondillo, K., Buzas, B., Goldman, D., Drevets, W. C. & Blair, R. J. (2007). The role of 5-HTTLPR in choosing the lesser of two evils, the better of two goods: Examining the impact of 5-HTTLPR genotype and tryptophan depletion in object choice. *Psychopharmacology*, 196, 29-38.
- Blair, R. J. (2001). Neurocognitive models of aggression, the antisocial personality disorders, and psychopathy. *Journal of Neurology, Neurosurgery, and Psychiatry*, 71, 727-731.
- Block, N. (1995). The mind as the software of the brain. In D. N. Osherson, L. Gleitman, S. Kosslyn, E. Smith & S. Sternberg (Hrsg.), *An invitation to cognitive science, Volume 3: Thinking* (2. Aufl., S. 377-425). Cambridge, MA: MIT Press.
- Boghossian, P. A. (1994). Inferential role semantics and the analytic/synthetic distinction. *Philosophical Studies*, 73, 109-122.
- Boghossian, P. A. (1996). Analyticity reconsidered. *Noûs*, 30, 360-391.
- Boghossian, P. A. (1997). Analyticity. In C. Wright & B. Hale (Hrsg.), *A companion to the philosophy of language* (S. 331-368). Oxford: Blackwell.
- Boghossian, P. A. (2003a). Epistemic analyticity: A defense. *Grazer Philosophische Studien*, 66, 15-35.
- Boghossian, P. A. (2003b). The normativity of content. *Philosophical Issues*, 13, 31-45.

- Bolla, K. I., Eldreth, D. A., London, E. D., Kiehl, K. A., Mouratidis, M., Contoreggi, C., Matochik, J. A., Kurian, V., Cadet, J. L., Kimes, A. S., Funderburk, F. R. & Ernst, M. (2003). Orbitofrontal cortex dysfunction in abstinent cocaine abusers performing a decision-making task. *Neuroimage*, 19, 1085-1094.
- Botterill, G. & Carruthers, P. (1999). *The philosophy of psychology*. Cambridge: Cambridge University Press.
- Bowman, C. H., Evans, C. E. Y. & Turnbull, O. H. (2005). Artificial time constraints on the Iowa Gambling Task: The effects on behavioural performance and subjective experience. *Brain and Cognition*, 57, 21-25.
- Bradley, M. C. (1979). Two logical connection arguments and some principles about causal connection. *Erkenntnis*, 14, 1-23.
- Brandom, R. B. (1994). *Making it explicit: Reasoning, representing, and discursive commitment*. Cambridge, MA: Harvard University Press.
- Brandom, R. B. (2001a). *Begründen und Begreifen: Eine Einführung in den Inferentialismus*. Frankfurt a. M.: Suhrkamp.
- Brandom, R. B. (2001b). Modality, normativity, and intentionality. *Philosophy and Phenomenological Research*, 63, 587-609.
- Brandom, R. B. (2007). Inferentialism and some of its challenges. *Philosophy and Phenomenological Research*, 74, 651-676.
- Brandtstädter, J. (1987). "A rose has no teeth": Zum Problem der Unterscheidung zwischen Begriffsverwirrung und überraschenden empirischen Befunden in der Psychologie. In J. Brandtstädter (Hrsg.), *Struktur und Erfahrung in der psychologischen Forschung* (S. 1-13). Berlin: de Gruyter.
- Brandtstädter, J. (2007). Causality, intentionality, and the causation of intentions: The problematic boundary. In M. G. Ash & T. Sturm (Hrsg.), *Psychology's territories: Historical and contemporary perspectives* (S. 51-66). Mahwah, NJ: Lawrence Erlbaum.
- Brandtstädter, J. & Sturm, T. (2004). Apriorität, Erfahrung und das Projekt der Psychologie. *Zeitschrift für Sozialpsychologie*, 35, 15-32.
- Brown, C. (2007). *Narrow mental content*. In E. N. Zalta (Hrsg.), *The stanford encyclopedia of philosophy* (summer 2008 edition). Abgerufen am 01.05.2008, vom World Wide Web unter: <<http://plato.stanford.edu/entries/content-narrow/>>
- Brown, S. & Schäfer, E. A. (1888). An investigation into the functions of the occipital and temporal lobes of the monkey's brain. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 179, 303-327.
- Bunge, S. A. & Souza, M. J. (2007). Neural representations used to specify action. In S. A. Bunge & J. D. Wallis (Hrsg.), *Neuroscience of rule-guided behavior* (S. 45-66). Oxford: Oxford University Press.
- Bunge, S. A. & Wallis, J. D. (Hrsg.). (2007). *Neuroscience of rule-guided behavior*. New York: Oxford University Press.
- Burge, T. (1979). Individualism and the mental. In P. A. French, T. E. Uehling & H. K. Wettstein (Hrsg.), *Midwest studies in philosophy, Volume 4: Studies in metaphysics* (S. 73-121). Minneapolis: University of Minnesota Press.
- Burge, T. (1982). Two thought experiments reviewed: Comments on J. A. Fodor's paper "Cognitive science and the twin-Earth problem". *Notre Dame Journal of Formal Logic*, 23, 284-293.
- Burge, T. (1986). Individualism and psychology. *The Philosophical Review*, 95, 3-45.
- Burns, K. & Bechara, A. (2007). Decision making and free will: A neuroscience perspective. *Behavioral Sciences & the Law*, 25, 263-280.
- Bussmann, H., Trauth, G., Kazzazi, K. & Bussmann, H. (Hrsg.). (1996). *Routledge dictionary of language and linguistics* (3. Aufl.). London: Routledge.

- Butler, K. (1996). Content, causal powers, and context. *Philosophy of Science*, 63, 105-114.
- Campos, M. (2003). Analyticity and incorrigibility. *Philosophy and Phenomenological Research*, 66, 689-708.
- Cannon, W. B. (1927). The James-Lange theory of emotion: A critical examination and an alternative theory. *American Journal of Psychology*, 39, 106-124.
- Carnap, R. (1963). W. V. Quine on logical truth. In P. A. Schilpp (Hrsg.), *The philosophy of Rudolf Carnap* (S. 915-922). Chicago: Open Court Press.
- Carnap, R. (1972). *Bedeutung und Notwendigkeit: Eine Studie zur Semantik und modalen Logik*. Wien: Springer.
- Carroll, L. (1895). What the tortoise said to Achilles. *Mind*, 4, 278-280.
- Chalmers, D. J. (2006). The foundations of two-dimensional semantics. In M. García-Carpintero & J. Macià (Hrsg.), *Two-dimensional semantics* (S. 55-140). Oxford: Oxford University Press.
- Chiu, Y.-C. & Lin, C.-H. (2007). Is deck C an advantageous deck in the Iowa Gambling Task? *Behavioral and Brain Functions*, 3, 1-11.
- Churchland, P. M. (1970). The logical character of action-explanations. *The Philosophical Review*, 79, 214-236.
- Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *Journal of Philosophy*, 78, 67-90.
- Churchland, P. S. (1986). *Neurophilosophy: Toward a unified science of the mind-brain*. Cambridge, MA: MIT Press.
- Clark, L., Cools, R. & Robbins, T. W. (2004). The neuropsychology of ventral prefrontal cortex: Decision-making and reversal learning. *Brain and Cognition*, 55, 41-53.
- Clark, L., Manes, F., Antoun, N., Sahakian, B. J. & Robbins, T. W. (2003). The contributions of lesion laterality and lesion volume to decision-making impairment following frontal lobe damage. *Neuropsychologia*, 41, 1474-1483.
- Colombetti, G. (2008). The somatic marker hypotheses, and what the Iowa Gambling Task does and does not show. *The British Journal for the Philosophy of Science*, 59, 51-71.
- Colombetti, G. & Thompson, E. (2007). The feeling body: Toward an enactive approach to emotion. In W. F. Overton, U. Müller & J. Newman (Hrsg.), *Body in mind, mind in body: Developmental perspectives on embodiment and consciousness* (S. 45-68). Mahwah, NJ: Lawrence Erlbaum.
- Copleston, F. C. (1999). *A history of philosophy. Volume 4: The Rationalists - Descartes to Leibniz*. London: Burns & Oates.
- Cornford, F. M. (2000). *Plato's theory of knowledge: The Theaetetus and the Sophist of Plato*. London: Routledge.
- Craig, A. D. (2004). Human feelings: why are some more aware than others? *Trends in Cognitive Sciences*, 8, 239-241.
- Crane, T. (2001). Autonomy of psychology. In R. A. Wilson & F. C. Keil (Hrsg.), *The MIT encyclopedia of the cognitive sciences* (S. 64-65). Cambridge, MA: MIT Press.
- Creath, R. (2004). Quine on the intelligibility and relevance of analyticity. In R. F. Gibson (Hrsg.), *The Cambridge companion to Quine* (S. 47-64). Cambridge: Cambridge University Press.
- D'Agostino, M., Gabbay, D. M., Hähnle, R. & Posegga, J. (Hrsg.). (1999). *Handbook of tableau methods*. Dordrecht: Kluwer.
- Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 351, 1413-1420.

- Damasio, A. R. (2004). Emotions and feelings: A neurobiological perspective. In A. S. R. Manstead, N. Frijda & A. Fischer (Hrsg.), *Feelings and emotions: The Amsterdam symposium* (S. 49-57). New York: Cambridge University Press.
- Damasio, A. R. (2005). The neurobiological grounding of human values. In J. P. Changeux, A. R. Damasio, W. Singer & Y. Christen (Hrsg.), *The neurobiology of human values* (S. 47-56). Berlin: Springer Verlag.
- Damasio, A. R. (2006). *Descartes' error: Emotion, reason and the human brain* (12. Aufl.). London: Vintage.
- Damasio, A. R. (2007). Neuroscience and ethics: Intersections. *The American Journal of Bioethics*, 7, 3-7.
- Damasio, A. R., Tranel, D. & Damasio, H. (1990). Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. *Behavioural Brain Research*, 41, 81-94.
- Damasio, A. R., Tranel, D. & Damasio, H. (1991). Somatic markers and the guidance of behaviour: Theory and preliminary testing. In H. S. Levin, H. M. Eisenberg & A. L. Benton (Hrsg.), *Frontal lobe function and dysfunction* (S. 217-229). New York: Oxford University Press.
- Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M. & Damasio, A. R. (1994). The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science*, 264, 1102-1105.
- Dancy, J. (2000). *Practical reality*. Oxford: Oxford University Press.
- Dancy, J. (2004). Two ways of explaining actions. In J. Hyman & S. H. (Hrsg.), *Agency and action* (S. 25-42). Cambridge: Cambridge University Press.
- Darwin, C. (1965). *The expression of the emotions in man and animals*. Chicago: University of Chicago Press.
- Davidson, D. (1998a). Beabsichtigen. In D. Davidson (Hrsg.), *Handlung und Ereignis* (2. Aufl., S. 125-152). Frankfurt a. M.: Suhrkamp.
- Davidson, D. (1998b). Geistige Ereignisse. In D. Davidson (Hrsg.), *Handlung und Ereignis* (2. Aufl., S. 291-317). Frankfurt a. M.: Suhrkamp.
- Davidson, D. (1998c). Handeln. In D. Davidson (Hrsg.), *Handlung und Ereignis* (2. Aufl., S. 73-98). Frankfurt a. M.: Suhrkamp.
- Davidson, D. (1998d). Handlungen, Gründe und Ursachen. In D. Davidson (Hrsg.), *Handlung und Ereignis* (2. Aufl., S. 19-42). Frankfurt a. M.: Suhrkamp.
- Davidson, D. (1998e). Handlungsfreiheit. In D. Davidson (Hrsg.), *Handlung und Ereignis* (2. Aufl., S. 99-124). Frankfurt a. M.: Suhrkamp.
- Davidson, D. (1998f). Kausale Beziehungen. In D. Davidson (Hrsg.), *Handlung und Ereignis* (2. Aufl., S. 214-232). Frankfurt a. M.: Suhrkamp.
- Davidson, R. J. (2003). Seven sins in the study of emotion: correctives from affective neuroscience. *Brain and Cognition*, 52, 129-132.
- Davis, W. A. (2005). Reasons and psychological causes. *Philosophical Studies*, 122, 51-101.
- de Muijnck, W. (2003). *Dependencies, connections and other relations: A theory of mental causation*. Dordrecht: Kluwer.
- de Sousa, R. (1987). *The rationality of emotion*. Cambridge, MA: MIT Press.
- DeLancey, C. (2002). *Passionate engines: What emotions reveal about mind and artificial intelligence*. New York: Oxford University Press.
- Denburg, N. L., Cole, C. A., Hernandez, M., Yamada, T. H., Tranel, D., Bechara, A. & Wallace, R. B. (2007). The orbitofrontal cortex, real-world decision-making, and normal aging. *Annals of the New York Academy of Sciences*, 1121, 480-498.
- Denburg, N. L., Tranel, D. & Bechara, A. (2005). The ability to decide advantageously declines prematurely in some normal older persons. *Neuropsychologia*, 43, 1099-1106.

- Denk, F., Walton, M. E., Jennings, K. A., Sharp, T., Rushworth, M. F. & Bannerman, D. M. (2005). Differential involvement of serotonin and dopamine systems in cost-benefit decisions about delay or effort. *Psychopharmacology*, 179, 587-596.
- Dennett, D. C. (1986). *Content and consciousness* (2. Aufl.). London: Routledge.
- Descartes, V. (2001). *The mind's provisions: A critique of cognitivism*. Princeton: Princeton University Press.
- Deutscher, M. (1976). Conceptual connection and causal relation. *Australasian Journal of Philosophy*, 54, 3-13.
- Dickenson, J. (2007). Reasons, causes, and contrasts. *Pacific Philosophical Quarterly*, 88, 1-23.
- Dunn, B. D., Dalgleish, T. & Lawrence, A. D. (2006). The somatic marker hypothesis: A critical evaluation. *Neuroscience and Biobehavioral Reviews*, 30, 239-271.
- Ebert, P. A. (2005). Transmission of warrant-failure and the notion of epistemic analyticity. *Australasian Journal of Philosophy*, 83, 505-521.
- Enç, B. (2003). *How we act: Causes, reasons, and intentions*. Oxford: Oxford University Press.
- Ernst, M., Bolla, K. I., Mouratidis, M., Contoreggi, C., Matochik, J. A., Kurian, V., Cadet, J.-L., Kimes, A., S. & London, E., D. (2002). Decision-making in a risk-taking task: A PET study. *Neuropsychopharmacology*, 26, 682-691.
- Ertz, T.-P. (in Vorbereitung). *Regel und Witz: Wittgensteinsche Perspektiven auf Mathematik, Sprache und Moral*. Berlin: de Gruyter.
- Eslinger, P. J. & Damasio, A. R. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation: Patient EVR. *Neurology*, 35, 1731-1741.
- Fehr, E. & Camerer, C. F. (2007). Social neuroeconomics: The neural circuitry of social preferences. *Trends in Cognitive Sciences*, 11, 419-427.
- Fein, G., McGillivray, S. & Finn, P. (2007). Older adults make less advantageous decisions than younger adults: Cognitive and psychological correlates. *Journal of the International Neuropsychological Society*, 13, 480-489.
- Feyerabend, P. (1983). *Wider den Methodenzwang*. Frankfurt a. M.: Suhrkamp.
- Feyerabend, P. K. (1963). Mental events and the brain. *Journal of Philosophy*, 60, 295-296.
- Field, H. (2006). Recent debates about the a priori. In T. S. Gendler & J. Hawthorne (Hrsg.), *Oxford studies in epistemology, Volume 1* (S. 69-88). Oxford: Oxford University Press.
- Fodor, J. A. (1982). Cognitive science and the twin-Earth problem. *Notre Dame Journal of Formal Logic*, 23, 98-118.
- Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. Oxford: Oxford University Press.
- Fodor, J. A. & Lepore, E. (1991). Why meaning (probably) isn't conceptual role. *Mind and Language*, 6, 328-343.
- Foot, P. (2001). *Natural goodness*. Oxford: Oxford University Press.
- Frances, B. (1999). On the explanatory deficiencies of linguistic content. *Philosophical Studies*, 93, 45-75.
- Frege, G. (1986). *Die Grundlagen der Arithmetik: Eine logisch mathematische Untersuchung über den Begriff der Zahl*. Hamburg: Meiner.
- Frege, G. (1990). *Kleine Schriften* (2. Aufl.). Hildesheim: G. Olms.
- Frege, G. (1998). *Begriffsschrift und andere Aufsätze*. (2. Aufl.). Hildesheim: G. Olms.
- Frijda, N. H. (2004). Emotions and action. In A. S. R. Manstead, N. H. Frijda & A. Fischer (Hrsg.), *Feelings and emotions: The Amsterdam symposium* (S. 158-173). Cambridge: Cambridge University Press.

- George, A. (2000). On washing the fur without wetting it: Quine, Carnap, and analyticity. *Mind*, 109, 1-24.
- Gerrans, P. (2007). Mental time travel, somatic markers and "myopia for the future". *Synthese*, 159, 459-474.
- Gessner, W. (2004). *Die kognitive Emergenz von Emotionen*. Paderborn: Mentis.
- Gintis, H. (2007). A framework for the unification of the behavioral sciences. *Behavioral and Brain Sciences*, 30.
- Glock, H.-J. (1992). Wittgenstein vs. Quine on logical necessity. In S. Teghrarian, A. Serafini & J. Hintikka (Hrsg.), *Wittgenstein and contemporary philosophy* (S. 154-186). Wakefield, NH: Longwood Academic.
- Glock, H.-J. (2000). *Wittgenstein-Lexikon*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Glock, H.-J. (2008). Necessity and language: In defence of conventionalism. *Philosophical Investigations*, 31, 24-47.
- Gluck, A. L. (2007). *Damasio's error and Descartes' truth: An inquiry into consciousness, epistemology, and metaphysics*. Scranton: University of Scranton Press.
- Goldie, P. (2000). *The emotions: A philosophical exploration*. Oxford: Oxford University Press.
- Goldman, A. H. (2005). Reason internalism. *Philosophy and Phenomenological Research*, 71, 505-533.
- Goldstein, E. B. (2002). *Sensation and perception* (6. Aufl.). Pacific Grove: Wadsworth-Thompson.
- Gonzalez, R., Bechara, A. & Martin, E. M. (2007). Executive functions among individuals with methamphetamine or alcohol as drugs of choice: Preliminary observations. *Journal of Clinical and Experimental Neuropsychology*, 29, 155-159.
- Grafton, S. T. & Ivry, R. B. (2004). The representation of action. In M. S. Gazzaniga (Hrsg.), *The cognitive neurosciences* (3. Aufl., S. 451-451). Cambridge, MA: MIT Press.
- Gregory, P. A. (2003). Two dogmas - all bark and no bite?: Carnap and Quine on analyticity. *Philosophy and Phenomenological Research*, 67, 633-648.
- Greve, W. (1994). *Handlungsklärung: Die psychologische Erklärung menschlicher Handlungen*. Bern: Hans Huber.
- Grice, H. P. & Strawson, P. F. (1956). In defense of a dogma. *The Philosophical Review*, 65, 141-158.
- Haaland, V. O. & Landro, N. I. (2007). Decision making as measured with the Iowa Gambling Task in patients with borderline personality disorder. *Journal of the International Neuropsychological Society*, 13, 699-703.
- Hacker, P. M. S. (1996). *Wittgenstein's place in twentieth-century analytic philosophy*. Oxford: Blackwell.
- Hacker, P. M. S. (2004). The conceptual framework for the investigation of emotions. *International Review of Psychiatry*, 16, 199-208.
- Hanks, P. W. (2008). A dilemma about necessity. *Erkenntnis*, 68, 129-148.
- Harlow, J. M. (1868). Recovery from a passage of an iron bar through the head. *Publications of the Massachusetts Medical Society*, 2, 327-347.
- Harlow, J. M. (1999). Passage of an iron rod through the head. *The Journal of Neuropsychiatry & Clinical Neuroscience*, 11, 281-283.
- Harman, G. (1996). Analyticity regained? *Noûs*, 30, 392-400.
- Hascher, T. & Baillod, J. (2004). Soziale Integration in der Schulklasse als Prädiktor für Wohlbefinden. In T. Hascher (Hrsg.), *Schule positiv erleben: Ergebnisse und Erkenntnisse zum Wohlbefinden von Schülerinnen und Schülern* (S. 133-158). Bern: Haupt.
- Heil, J. (2004). *Philosophy of mind: A contemporary introduction*. London: Routledge.

- Heitner, R. (2006). From a phono-logical point of view: neutralizing Quine's argument against analyticity. *Synthese*, 150, 15-39.
- Heuer, U. (2004). Reasons for actions and desires. *Philosophical Studies*, 121, 43-63.
- Holzkamp, K. (1986). Die Verkennung von Handlungsbegründungen als empirische Zusammenhangsannahmen in sozialpsychologischen Theorien: Methodologische Fehlorientierung infolge von Begriffsverwirrung. *Zeitschrift für Sozialpsychologie*, 17, 216-238.
- Hornsby, J. (2004). Agency and action. In J. Hyman & H. Steward (Hrsg.), *Agency and action* (S. 1-23). Cambridge: Cambridge University Press.
- Hsu, M., Bhatt, M., Adolphs, R., Tranel, D. & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science*, 310, 1680-1683.
- Hursthouse, R. (1999). *On virtue ethics*. Oxford: Oxford University Press.
- Hursthouse, R. (2000). Intention. In R. Teichmann (Hrsg.), *Logic, cause & action: Essays in honour of Elizabeth Anscombe* (S. 83-105). Cambridge: Cambridge University Press.
- Jacob, P. (1993). Externalism and the explanatory relevance of broad content. *Mind & Language*, 8, 131-156.
- Johnson-Frey, S. H. (Hrsg.). (2003). *Taking action: Cognitive neuroscience perspectives on intentional acts*. Cambridge, MA: MIT Press.
- Johnson, C. A., Xiao, L., Palmer, P., Sun, P., Wang, Q., Wei, Y., Jia, Y., Grenard, J. L., Stacy, A. W. & Bechara, A. (2008). Affective decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in 10th grade Chinese adolescent binge drinkers. *Neuropsychologia*, 46, 714-726.
- Jungermann, H., Pfister, H.-R. & Fischer, K. (1998). *Die Psychologie der Entscheidung: Eine Einführung*. Heidelberg: Spektrum Akademischer Verlag.
- Kalhat, J. (2008). Has the later Wittgenstein accounted for necessity? *Philosophical Investigations*, 31, 1-23.
- Kandel, E. R. & Kupfermann, I. (1996). Emotionale Zustände. In E. R. Kandel, J. H. Schwartz & T. M. Jessel (Hrsg.), *Neurowissenschaften: Eine Einführung* (S. 607-624). Heidelberg: Spektrum Akademischer Verlag.
- Kandel, E. R., Schwartz, J. H. & Jessel, T. M. (Hrsg.). (1996). *Neurowissenschaften: Eine Einführung*. Heidelberg: Spektrum Akademischer Verlag.
- Kandel, E. R., Schwartz, J. H. & Jessell, T. M. (2000). *Principles of neural science* (4. Aufl.). New York: McGraw-Hill.
- Kant, I. (1966). *Kritik der reinen Vernunft*. Stuttgart: Reclam.
- Kaplan, D. (1989). Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives. In J. Almog, J. Perry & H. K. Wettstein (Hrsg.), *Themes from Kaplan* (S. 481-563). New York: Oxford University Press.
- Keil, G. (2000). *Handeln und Verursachen*. Frankfurt a. M.: Klostermann.
- Kenny, A. J. P. (1963). *Action, emotion and will*. London: Routledge.
- Kenny, A. J. P. (1991). The homunculus fallacy. In J. Hyman (Hrsg.), *Investigating psychology: Sciences of the mind after Wittgenstein* (S. 155-165). London: Routledge.
- Killgore, W. D. S. & Yurgelun-Todd, D. A. (2007). Neural correlates of emotional intelligence in adolescent children. *Cognitive, Affective & Behavioral Neuroscience*, 7, 140-151.
- Kim, J. (1993). *Supervenience and mind: Selected philosophical essays*. Cambridge: Cambridge University Press.
- Kim, J. (1998). *Mind in a physical world: An essay on the mind-body problem and mental causation*. Cambridge, MA: MIT Press.

- Kim, J. (2002). The many problems of mental causation. In D. J. Chalmers (Hrsg.), *Philosophy of mind: Classical and contemporary readings* (S. 170-179). Oxford: Oxford University Press.
- Kim, J. (2007). Causation and mental causation. In J. Cohen & B. P. McLaughlin (Hrsg.), *Contemporary debates in philosophy of mind* (S. 227-242). Malden, MA: Blackwell.
- Klein, M. (1996). Externalism, content, and causation. *Proceedings of the Aristotelian Society*, 96, 159-176.
- Klüver, H. & Bucy, P. C. (1937). "Psychic blindness" and other symptoms following bilateral temporal lobectomy in rhesus monkeys. *American Journal of Physiology*, 119, 352-353.
- Kolb, B. & Whishaw, I. Q. (1996). *Fundamentals of human neuropsychology* (4. Aufl.). New York: W.H. Freeman and Company.
- Kripke, S. A. (1981). *Name und Notwendigkeit*. Frankfurt a. M.: Suhrkamp.
- Kripke, S. A. (1987). *Wittgenstein über Regeln und Privatsprache: Eine elementare Darstellung*. Frankfurt a. M.: Suhrkamp.
- Kuhn, T. S. (1976). *Die Struktur wissenschaftlicher Revolutionen* (2. Aufl.). Frankfurt a. M.: Suhrkamp.
- Lange, C. G. (1887). *Über Gemüthsbewegungen: Eine psycho-physiologische Studie*. Leipzig: Thomas.
- Lazarus, R. S. (2006). Emotions and interpersonal relationships: Toward a person-centered conceptualization of emotions and coping. *Journal of Personality*, 74, 9-46.
- Lazarus, R. S. & Lazarus, B. N. (1994). *Passion and reason: Making sense of our emotions*. New York: Oxford University Press.
- LeDoux, J. (1998). *The emotional brain: The mysterious underpinnings of emotional life*. London: Weidenfeld & Nicolson.
- Lee, Y., Kim, Y. T., Seo, E., Park, O., Jeong, S. H., Kim, S. H. & Lee, S. J. (2007). Dissociation of emotional decision-making from cognitive decision-making in chronic schizophrenia. *Psychiatry Research*, 152, 113-120.
- Leibniz, G. W. (1996). *Philosophische Werke 3: Neue Abhandlungen über den menschlichen Verstand*. Hamburg: Meiner.
- Levin, J. (2004). *Functionalism*. In E. N. Zalta (Hrsg.), *The stanford encyclopedia of philosophy* (summer 2008 edition). Abgerufen am 01.06.2008, vom World Wide Web unter: <http://plato.stanford.edu/entries/functionalism/>
- Lewis, D. (1973). Causation. *Journal of Philosophy*, 70, 556-567.
- Lin, C.-H., Chiu, Y.-C., Lee, P.-L. & Hsieh, J.-C. (2007). Is deck B a disadvantageous deck in the Iowa Gambling Task? *Behavioral and Brain Functions*, 3, 1-16.
- Lloyd, D. E. (1989). *Simple minds*. Cambridge, MA: MIT Press.
- Locke, J. (1988). *Versuch über den menschlichen Verstand, Band 2*. Hamburg: Meiner.
- Loewer, B. (2007). Mental causation, or something near enough. In J. Cohen & B. P. McLaughlin (Hrsg.), *Contemporary debates in philosophy of mind* (S. 243-264). Malden, MA: Blackwell.
- Lormand, E. (2001). Frame problem. In R. A. Wilson & F. C. Keil (Hrsg.), *The MIT encyclopedia of the cognitive sciences* (S. 326). Cambridge, MA: MIT Press.
- MacLean, P. (1949). Psychosomatic disease and the 'visceral brain': Recent developments bearing on the Papez theory of emotion. *Psychosomatic Medicine*, 11, 338-353.
- Macmillan, M. (2000). *An odd kind of fame: Stories of Phineas Gage*. Cambridge, MA: MIT Press.
- Magarinos, A. M., McEwen, B. S., Flugge, G. & Fuchs, E. (1996). Chronic psychosocial stress causes apical dendritic atrophy of hippocampal CA3 pyramidal neurons in subordinate tree shrews. *Journal of Neuroscience*, 16, 3534-3540.

- Maia, T. V. & McClelland, J. L. (2004). A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa gambling task. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 16075-16080.
- Maia, T. V. & McClelland, J. L. (2005). The somatic marker hypothesis: Still many questions but no answers - Response to Bechara et al. *Trends in Cognitive Sciences*, 9, 162-164.
- Margolis, E. & Laurence, S. (2001). Boghossian on analyticity. *Analysis*, 61, 293-302.
- Martin, G. N. (2006). *Human neuropsychology* (2. Aufl.). Harlow: Prentice Hall Europe.
- Martínez-Selva, J. M., Sánchez-Navarro, J. P., Bechara, A. & Román, F. (2006). Brain mechanisms involved in decision-making. *Revista de Neurologia*, 42, 411-418.
- McCrae, R. R. & Costa, P. T. (1995). Trait explanations in personality psychology. *European Journal of Personality*, 9, 231-252.
- Megill, J. & Cogburn, J. (2005). Easy's getting harder all the time: The computational theory and affective states. *Ratio*, 18, 306-316.
- Melden, A. I. (1961). *Free action*. London: Routledge.
- Mele, A. R. (1992). *Springs of action: Understanding intentional behavior*. New York: Oxford University Press.
- Mele, A. R. (2003). *Motivation and agency*. New York: Oxford University Press.
- Millar, A. (2004). *Understanding people: Normativity and rationalizing explanation*. Oxford: Clarendon Press.
- Miller, G. A., Galanter, E. & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart & Winston.
- Morris, M. (2007). *An introduction to the philosophy of language*. Cambridge: Cambridge University Press.
- Mouly, v. G. J. (1970). *The science of educational research* (2. Aufl.). New York: Van Nostrand Reinhold.
- Müller, A. W. (1982). *Praktisches Folgern und Selbstgestaltung nach Aristoteles*. Freiburg: Verlag Karl Alber.
- Müller, A. W. (1992). Mental teleology. *Proceedings of the Aristotelian Society*, 92, 161-183.
- Müller, A. W. (1998). *Was taugt die Tugend?: Elemente einer Ethik des guten Lebens*. Stuttgart: Kohlhammer.
- Müller, A. W. (2000). Tolernaz als Tugend: Die Rationalität des Ethischen Diskurses im Dilemma. In D. Birnbacher & G. Patzig (Hrsg.), *Bioethik als Tabu?: Toleranz und ihre Grenzen* (S. 23-39). Münster: LIT.
- Müller, A. W. (2003). Handeln. *Zeitschrift für philosophische Forschung*, 57, 325-347.
- Nagel, E. (1979). *The structure of science: Problems in the logic of scientific explanation* (2. Aufl.). Indianapolis: Hackett.
- Naqvi, N. H., Rudrauf, D., Damasio, H. & Bechara, A. (2007). Damage to the insula disrupts addiction to cigarette smoking. *Science*, 315, 531-534.
- Newman-Norlund, R. D., Noordzij, M. L., Meulenbroek, R. G. J. & Bekkering, H. (2007). Exploring the brain basis of joint action: Co-ordination of actions, goals and intentions. *Social Neuroscience*, 2, 48-65.
- Nimtz, C. (2005). *Two-dimensional semantics: The basics*. Universität Bielefeld. Abgerufen am 13.02.2008, vom World Wide Web unter: <[http://www.uni-bielefeld.de/\(en\)/philosophie/personen/nimtz/NIMTZ_041005_2DSEMANTICS_LONG.pdf](http://www.uni-bielefeld.de/(en)/philosophie/personen/nimtz/NIMTZ_041005_2DSEMANTICS_LONG.pdf)>
- Noel, X., Bechara, A., Dan, B., Hanak, C. & Verbanck, P. (2007a). Response inhibition deficit is involved in poor decision making under risk in nonamnesic individuals with alcoholism. *Neuropsychology*, 21, 778-786.

- Noel, X., Van der Linden, M., d'Acremont, M., Bechara, A., Dan, B., Hanak, C. & Verbanck, P. (2007b). Alcohol cues increase cognitive impulsivity in individuals with alcoholism. *Psychopharmacology*, 192, 291-298.
- Nollmann, G. (2006). Das neuronale Korrelat und Max Webers Konzept der soziologischen Kausalerklärung: Warum die Neurowissenschaft keine Konkurrentin der Soziologie ist. In J. Reichertz & N. Zaboura (Hrsg.), *Akteur Gehirn: Oder das vermeintliche Ende des handelnden Subjekts - Eine Kontroverse* (S. 61-77). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Northoff, G., Grimm, S., Boeker, H., Schmidt, C., Bermpohl, F., Heinzel, A., Hell, D. & Boesiger, P. (2006). Affective judgment and beneficial decision making: ventromedial prefrontal activity correlates with performance in the Iowa Gambling Task. *Human brain mapping*, 27, 572-587.
- Nussbaum, M. C. (2001). *Upheavals of thought: The intelligence of emotions*. New York: Cambridge University Press.
- O'Carroll, R. E. & Papps, B. P. (2003). Decision making in humans: the effect of manipulating the central noradrenergic system. *Journal of Neurology Neurosurgery and Psychiatry*, 74, 376-378.
- O'Doherty, J. P. & Bossaerts, P. (2008). Toward a mechanistic understanding of human decision making: Contributions of functional neuroimaging. *Current Directions in Psychological Science*, 17, 119-123.
- O'Grady, P. (1999). Carnap and two dogmas of empiricism. *Philosophy and Phenomenological Research*, 59, 1015-1027.
- O'Neill, M. (2001). Explaining 'the hardness of the logical must': Wittgenstein on grammar, arbitrariness and logical necessity. *Philosophical Investigations*, 24, 1-29.
- Okrent, M. (2007). *Rational animals: The teleological roots of intentionality*. Athens: Ohio University Press.
- Owens, J. (1993). Content, causation, and psychophysical supervenience. *Philosophy of Science*, 60, 242-261.
- Owens, J. (1994). Psychological externalism and psychological explanation. *Philosophy and Phenomenological Research*, 54, 921-928.
- Owens, J. (1998). Psychological explanation and causal deviancy. *Synthese*, 115, 143-169.
- Oya, H., Adolphs, R., Kawasaki, H., Bechara, A., Damasio, A. & Howard, M. A., III. (2005). Electrophysiological correlates of reward prediction error recorded in the human prefrontal cortex. *Proceedings of the National Academy of Sciences*, 102, 8351-8356.
- Papez, J. (1937). A proposed mechanism of emotion. *Archives of Neurology and Psychiatry*, 38, 725-743.
- Pauen, M. (2001). *Grundprobleme der Philosophie des Geistes: Eine Einführung*. Frankfurt a. M.: Fischer.
- Pears, D. & Kenny, A. (1995). Von Mill bis Wittgenstein. In A. Kenny (Hrsg.), *Illustrierte Geschichte der westlichen Philosophie*. Frankfurt a. M.: Campus Verlag.
- Peijnenburg, A. J. M. (1996). *Acting against one's best judgement: An enquiry into practical reasoning, dispositions and weakness of will*. Groningen: University of Groningen.
- Peregrin, J. (2006). Meaning as an inferential role. *Erkenntnis*, 64, 1-35.
- Peregrin, J. (2007). *Viable inferentialism*. Academy of Sciences & Charles University, Prague. Abgerufen am 27.02.2008, vom World Wide Web unter: <<http://jarda.peregrin.cz/mybibl/PDFTxt/503.pdf>>
- Phan, L. K., Wager, T., Taylor, S. F. & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage*, 16, 331-348.
- Philie, P. (2005). Meaning-scepticism and analyticity. *Dialectica*, 59, 357-365.

- Pinel, J. P. J. (1997). *Biopsychologie: Eine Einführung*. Heidelberg: Spektrum Akademischer Verlag.
- Pitcher, G. (1965). Emotion. *Mind*, 74, 326-346.
- Platon. (1957). *Sämtliche Werke, Band 1*. Hamburg: Rowohlt.
- Platon. (1994). *Sämtliche Werke, Band 2*. Hamburg: Rowohlt.
- Popper, K. R. (1973). *Logik der Forschung* (5. Aufl.). Tübingen: J.C.B. Mohr (Paul Siebeck).
- Popper, K. R. & Eccles, J. C. (1984). *The self and its brain: An argument for interactionism*. London: Routledge.
- Poser, H. (2001). *Wissenschaftstheorie: Eine philosophische Einführung*. Stuttgart: Reclam.
- Power, M. J. & Dalgleish, T. (1997). *Cognition and emotion: From order to disorder*. Hove: Psychology Press.
- Putnam, H. (1962). The analytic and the synthetic. In H. Feigl & G. Maxwell (Hrsg.), *Minnesota studies in the philosophy of science, Volume 3* (S. 358-397). Minneapolis: University of Minnesota Press.
- Putnam, H. (1979). *Die Bedeutung von "Bedeutung"*. Frankfurt a. M.: Klostermann.
- Putnam, H. (1985). Meaning and reference. In A. P. Martinich (Hrsg.), *The philosophy of language* (4. Aufl., S. 288-295). Oxford: Oxford University Press.
- Putnam, H. (1987). Analyticity and apriority: Beyond Wittgenstein and Quine. In P. K. Moser (Hrsg.), *A priori knowledge* (S. 85-111). Oxford: Oxford University Press.
- Putnam, H. (2002). *The collapse of the fact/value dichotomy and other essays*. Cambridge, MA: Harvard University Press.
- Quine, W. V. O. (1964). Two dogmas of empiricism. In W. V. O. Quine (Hrsg.), *From a logical point of view* (2. Aufl., S. 20-46). Cambridge, MA.: Harvard University Press.
- Quine, W. V. O. (1976). Truth by convention. In W. V. O. Quine (Hrsg.), *The ways of paradox, and other essays* (S. 77-106). Cambridge, MA.: Harvard University Press.
- Quine, W. V. O. (1980). *Wort und Gegenstand (Word and object)*. Stuttgart: Reclam.
- Risjord, M. (2005). Reasons, causes, and action explanation. *Philosophy of the Social Sciences*, 35, 294-306.
- Rolls, E. T. (2005). *Emotion explained*. Oxford: Oxford University Press.
- Ros, A. (1989). *Begründung und Begriff: Wandlungen des Verständnisses begrifflicher Argumentationen, Band 1: Antike, Spätantike und Mittelalter*. Hamburg: Meiner.
- Ros, A. (1990). *Begründung und Begriff: Wandlungen des Verständnisses begrifflicher Argumentationen, Band 2: Neuzeit*. Hamburg: Meiner.
- Rosenberg, A. (2005). *Philosophy of science: A contemporary introduction* (2. Aufl.). London: Routledge.
- Rosenzweig, M. R., Leiman, A. L. & Breedlove, S. M. (1999). *Biological psychology: An introduction to behavioral, cognitive, and clinical neuroscience* (2. Aufl.). Sunderland, MA: Sinauer Associates.
- Ruben, D.-H. (1999). Act individuation: The Cambridge theory. *Analysis*, 59, 276-283.
- Russell, G. (2008). *Truth in virtue of meaning: A defence of the analytic/synthetic distinction*. Oxford: Oxford University Press.
- Saidel, E. (1994). Content and causal powers. *Philosophy of Science*, 61, 658-665.
- Sanfey, A. G. (2007). Decision neuroscience: New directions in studies of judgment and decision making. *Current Directions in Psychological Science*, 16, 151-155.
- Sankey, H. (1999). Incommensurability: An overview. *Divinatio*, 10, 1-13.
- Sarkar, S. (1992). Models of reduction and categories of reductionism. *Synthese*, 91, 167-194.
- Sawyer, S. (2007). There is no viable notion of narrow content. In J. Cohen & B. P. McLaughlin (Hrsg.), *Contemporary debates in philosophy of mind* (S. 20-34). Malden, MA: Blackwell.

- Schachter, S. & Singer, J. E. (1962). Cognitive, social and physiological determinants of emotional state. *Psychological Review*, 69, 379-407.
- Schaub, H. (1993). *Modellierung der Handlungsorganisation*. Bern: Hans Huber.
- Schmitt, W. A., Brinkley, C. A. & Newman, J. P. (1999). Testing Damasio's somatic marker hypothesis with psychopathic individuals: risk takers or risk averse? *Journal of Abnormal Psychology*, 108, 538-543.
- Schueler, G. F. (2003). *Reasons and purposes: Human rationality and the teleological explanation of action*. Oxford: Oxford University Press.
- Segal, G. (2007). Cognitive content and propositional attitude attributions. In J. Cohen & B. P. McLaughlin (Hrsg.), *Contemporary debates in philosophy of mind* (S. 5-19). Malden, MA: Blackwell.
- Sehon, S. R. (1997). Deviant causal chains and the irreducibility of teleological explanation. *Pacific Philosophical Quarterly*, 78, 195-213.
- Sehon, S. R. (2005). *Teleological realism: Mind, agency, and explanation*. Cambridge, MA: MIT Press.
- Sevy, S., Hassoun, Y., Bechara, A., Yechiam, E., Napolitano, B., Burdick, K., Delman, H. & Malhotra, A. (2006). Emotion-based decision-making in healthy subjects: Short-term effects of reducing dopamine levels. *Psychopharmacology*, 188, 228-235.
- Shagrir, O. (2005). The rise and fall of computational functionalism. In Y. Ben-Menahem (Hrsg.), *Hilary Putnam* (S. 220-250). Cambridge: Cambridge University Press.
- Shirley, E. S. (1973). Putnam on analyticity. *Philosophical Studies*, 24, 268-271.
- Siegel, H. (1996). Naturalism and the abandonment of normativity. In W. T. O'Donohue & R. F. Kitchener (Hrsg.), *The philosophy of psychology* (S. 4-18). London: Sage Publications.
- Smedslund, J. (1988). *Psycho-logic*. Berlin: Springer.
- Smedslund, J. (1997). *The structure of psychological common sense*. Mahwah, NJ: Lawrence Erlbaum.
- Soames, S. (2006). The philosophical significance of the Kripkean necessary aposteriori. *Philosophical Issues*, 16, 288-309.
- Solomon, R. C. (2003). Emotions, thoughts and feelings: What is a 'cognitive theory' of the emotions and does it neglect affectivity? In A. Hatzimoysis (Hrsg.), *Philosophy and the emotions* (S. 1-18). Cambridge: Cambridge University Press.
- Spicer, F. (2004). Emotional behaviour and the scope of belief-desire explanation. In D. Evans & P. Cruse (Hrsg.), *Emotion, evolution and rationality* (S. 51-68). Oxford: Oxford University Press.
- Stamm, I. (2007). *Zwischen Neurobiologie und Sozialethik: Zum soziologischen Gehalt der Gefühle bei Jürgen Habermas und Antonio Damasio*. Oldenburg: BIS-Verlag der Carl von Ossietzky Universität Oldenburg.
- Stein, E. (1996). *Without good reason: The rationality debate in philosophy and cognitive science*. Oxford: Oxford University Press.
- Stoutland, F. (1970). The logical connection argument. In N. Rescher. (Hrsg.), *Studies in the theory of knowledge* (S. 117-129). Oxford: Oxford University Press.
- Stoutland, F. (1998). The real reasons. In J. Bransen & S. E. Cuypers (Hrsg.), *Human action, deliberation, and causation* (S. 43-66). Dordrecht: Kluwer.
- Stoutland, F. (2005). *Zur Verteidigung einer nichtpsychologischen Theorie der Handlungsgründe*. e-Journal Philosophie der Psychologie. Abgerufen am 02.03.2008, vom World Wide Web unter: <<http://www.jp.philo.at/texte/StoutlandF1.pdf>>
- Strack, F. & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8, 220-247.

- Suzuki, A., Hirota, A., Takasawa, N. & Shigemasa, K. (2003). Application of the somatic marker hypothesis to individual differences in decision making. *Biological Psychology*, 65, 81-88.
- Tanney, J. (1995). Why reasons may not be causes. *Mind & Language*, 10, 105-128.
- Tchanturia, K., Liao, P. C., Uher, R., Lawrence, N., Treasure, J. & Campbell, I. C. (2007). An investigation of decision making in anorexia nervosa using the Iowa Gambling Task and skin conductance measurements. *Journal of the International Neuropsychological Society*, 13, 635-641.
- Thomas, v. A. (2008). *Summa theologiae*. The latin library. Abgerufen am 01.02.2008, vom World Wide Web unter: <<http://www.thelatinlibrary.com/aquinas/q1.15.shtml>>
- Thompson, M. (2008). *Life and action: Elementary structures of practice and practical thought*. Cambridge, MA: Harvard University Press.
- Tranel, D. (1994). "Acquired sociopathy": the development of sociopathic behavior following focal brain damage. In D. C. Fowles, P. Sutker & S. H. Goodman (Hrsg.), *Progress in Experimental Personality and Psychopathology Research* (S. 285-311). New York: Springer.
- Tranel, D., Bechara, A. & Damasio, A. (1999). Decision making and the Somatic Marker Hypothesis. In M. S. Gazzaniga (Hrsg.), *The new cognitive neurosciences* (S. 1047-1061). Cambridge, MA: MIT Press.
- Tranel, D., Damasio, H. & Damasio, A. R. (1995). Double dissociation between overt and covert face recognition. *Journal of Cognitive Neuroscience*, 7, 425-432.
- Tuomela, R. (1977). *Human action and its explanation: A study on the philosophical foundations of psychology*. Dordrecht: D. Reidel.
- Tye, M. (2008). The experience of emotion: An intentionalist theory. *Revue Internationale de Philosophie*, 62, 25-50.
- van der Plasse, G., La Fors, S. S., Meerkerk, D. T., Joosten, R. N., Uylings, H. B. & Feenstra, M. G. (2007). Medial prefrontal serotonin in the rat is involved in goal-directed behaviour when affect guides decision making. *Psychopharmacology*, 195, 435-449.
- van Honk, J., Hermans, E. J., Putman, P., Montagne, B. & Schutter, D. J. (2002). Defective somatic markers in sub-clinical psychopathy. *Neuroreport*, 13, 1025-1027.
- Verdejo-Garcia, A., Bechara, A., Recknor, E. C. & Perez-Garcia, M. (2007). Negative emotion-driven impulsivity predicts substance dependence problems. *Drug and Alcohol Dependence*, 91, 213-219.
- Volkman-Schluck, K.-H. (1984). *Nicolaus Cusanus: Die Philosophie im Übergang vom Mittelalter zur Neuzeit* (3. Aufl.). Frankfurt a. M.: Klostermann.
- Volz, K. G., Schubotz, R. I. & von Cramon, D. Y. (2006). Decision-making and the frontal lobes: Neuroimaging. *Current Opinion in Neurology*, 19, 401-406.
- von Wright, G. H. (1984). *Erklären und Verstehen* (2. Aufl.). Frankfurt a.M.: Athenäum Verlag.
- Walter, S. & Heckmann, H.-D. (2003). Introduction. In S. Walter & H.-D. Heckmann (Hrsg.), *Physicalism and mental causation: The metaphysics of mind and action* (S. 133-136). Exeter: Imprint Academic.
- Wedgwood, R. (2006). The Internal and external components of cognition. In R. Stainton (Hrsg.), *Contemporary debates in cognitive science* (S. 307-325). Malden, MA: Blackwell.
- Weller, J. A., Levin, I. P., Shiv, B. & Bechara, A. (2007). Neural correlates of adaptive decision making for risky gains and losses. *Psychological Science*, 18, 958-964.
- Wiland, E. (2002). Theories of practical reason. *Metaphilosophy*, 33, 450-467.
- William, J. (1884). What is an emotion? *Mind*, 9, 185-205.

- Williams, B. (1995). Internal reasons and the obscurity of blame. In B. Williams (Hrsg.), *Making sense of humanity: And other philosophical papers* (S. 35-45). Cambridge: Cambridge University Press.
- Wilson, N. L. (1959). Substances without substrata. *Review of Metaphysics*, 12, 521-539.
- Wittgenstein, L. (1970). *Über Gewißheit*. Frankfurt a. M.: Suhrkamp.
- Wittgenstein, L. (1978). *Wittgensteins Vorlesungen über die Grundlagen der Mathematik*. Frankfurt a. M.: Suhrkamp.
- Wittgenstein, L. (1984). *Bemerkungen über die Philosophie der Psychologie; Letzte Schriften über die Philosophie der Psychologie*. Frankfurt a. M.: Suhrkamp.
- Wittgenstein, L. (1997a). *Das Blaue Buch: Eine Philosophische Betrachtung (Das Braune Buch)* (7. Aufl.). Frankfurt a. M.: Suhrkamp.
- Wittgenstein, L. (1997b). *Tractatus logico-philosophicus (u.a.)* (11. Aufl.). Frankfurt a. M.: Suhrkamp.
- Wright, C. (2007). Rule-following without reason: Wittgenstein's quietism and the constitutive question. *Ratio*, 10, 481-502.
- Yu, R. J. & Zhou, X. L. (2007). Neuroeconomics: Opening the "black box" behind the economic behavior. *Chinese Science Bulletin*, 52, 1153-1161.
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35, 151-175.
- Zemach, E. M. (1976). Putnam's theory on the reference of substance terms. *The Journal of Philosophy*, 73, 116-127.
- Zhu, J. (2004). Passive action and causalism. *Philosophical Studies*, 119, 295-314.