

Emotionale Einstellungen.
Ein moralpsychologischer Lösungsansatz zu
Michael Smith's „Moral Problem“

Inaugural-Dissertation
zur
Erlangung der Doktorwürde
der
Philosophischen Fakultät
der
Rheinischen
Friedrich-Wilhelms-Universität
zu Bonn

vorgelegt von

SEBASTIAN HARRER

aus
Düsseldorf

Bonn 2006

Gedruckt mit Genehmigung der Philosophischen Fakultät der
Rheinischen Friedrich-Wilhelms-Universität zu Bonn.

Tag der mündlichen Prüfung: 19.06.2006

Zusammensetzung der Prüfungskommission:

Prof. Dr. Wolfram Högbe (Vorsitzender)

Prof. Dr. Christoph Horn (Betreuer und Gutachter)

Prof. Dr. Hans-Joachim Pieper (Gutachter)

Prof. Dr. Theo Kobusch (weiteres prüfungsberechtigtes Mitglied)

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn

http://hss.ulb.uni-bonn.de/diss_online

elektronisch publiziert.

Vorwort und Danksagung

Die vorliegende Arbeit ist unter der Betreuung von Herrn Prof. Dr. Christoph Horn am Institut für Philosophie der Rheinischen Friedrich-Wilhelms-Universität zu Bonn entstanden und wurde großzügig durch ein Promotionsstipendium der Friedrich-Ebert-Stiftung gefördert. Zuvörderst sei an dieser Stelle meinem Betreuer Herrn Prof. Dr. Horn für die stets engagierte Unterstützung sowie die stetige Bereitschaft zu fachlichem Ratschlag herzlich gedankt. Ohne Herrn Prof. Dr. Horn hätte diese Arbeit nicht entstehen können. Weiterhin bedanke ich mich bei der Friedrich-Ebert-Stiftung für die Zuerkennung des besagten Promotionsstipendiums.

Während der Anfertigung der vorliegenden Arbeit verbrachte ich einen neunmonatigen Forschungsaufenthalt an der University of New South Wales, Sydney, Australien, im Rahmen des „Professional Practicum Programmes“. Die dort verbrachte Zeit hat sich als enorme fachliche wie persönliche Bereicherung erwiesen, und ich danke meinen damaligen Kollegen (den *postgraduate students* und *PhD candidates*) sowie Dozenten (zuvörderst meinem Gastbetreuer Herrn A/Prof. Philip Cam und dem damaligen Institutsdirektor Herrn Prof. Paul Patton) für die ausgezeichnete Zusammenarbeit.

Meiner Familie, meiner Partnerin sowie meinen Freunden im In- und Ausland gebührt mein Dank für ihre persönliche Unterstützung während meiner Arbeit an der Dissertation.

Inhaltsverzeichnis

1. Einleitung und Überblick über die Argumentation	6
1.1. Problemlage und Umfeld des „moral problem“	6
1.2. Emotionstheorien.....	8
1.3. Emotionen, Werte und Wünsche	15
1.4. Moralische Überzeugungen als Emotionale Einstellungen	17
2. Was ist „Emotion“? Kritischer Abriss einer Theoriegeschichte und Entwicklung eines Thesenkatalogs	21
2.1. Einleitung.....	21
2.2. Die Tradition des Kognitivismus.....	22
2.3. Psycho-physische Emotionstheorien	31
2.4. Moderne kognitivistische Emotionstheorien in der empirischen Psychologie.....	46
2.5. Die „emotion-cognition“-Debatte in der Psychologie: Richard Lazarus und Robert Zajonc	54
2.6. Logisch-konzeptuelle Kritik des Kognitivismus	65
2.7. Neurologische Grundlagen von Emotionen	82
2.8. Rezente neuro-psychologische Emotionstheorien: Antonio Damasio	83
2.9. Neurologische Grundlagen von Emotion: Exkurs zu Schachter & Singer 1962, Hohmann 1966 und Damasio 1994.....	95
2.10. Ausblick: Emotionen und Akteure	101
2.11. „Emotion“. Ein Thesenkatalog	106
3. Emotionen, Werte und „desires“	115
3.1. Einleitung.....	115
3.2. Der Begriff des „Wertes“	118
3.3. „Valuing: desiring or believing?“	120
3.3.1. Werte und „desires“	120
3.3.2. Werte und Glaubenszustände	125
3.4. Wertrealismus	131
3.5. Axiologischer Anti-Realismus	136
3.6. Naturalismus und Supervenienz	138
3.7. Dispositionalismus.....	141
3.8. Emotionen als die Grundlage von Werten.....	144
3.9. Individuelle emotionale Reaktionen als eine Praxis des Wertens	151

3.10. Die Intentionalität emotionalen Wertens	155
3.11. Transpersonale Geltung von emotionalen Wertungen	159
3.12. Verantwortung für Emotionen und emotionale Wertungen	163
4. Handeln auf Grund von emotionalen Einstellungen: metaethische Situierung des Modells.....	182
4.1. Einleitung.....	182
4.2. Michael Smith's „moral problem“	183
4.3. Internalismus und Externalismus.....	188
4.4. Noch einmal: die Humeanische Motivationstheorie	203
4.5. Motivierende Gründe.....	205
4.5.1. Der Humeanismus	207
4.5.2. Der Kognitivismus.....	208
4.6. Normative Gründe	211
4.6.1. Die Quellen normativer Gründe	212
4.6.2. Wünsche (desires) als Quellen normativer Gründe	213
4.6.3. Werte als Quellen normativer Gründe.....	217
4.7. Wie ist das „moral problem“ zu lösen? Oder: das „reconciliation problem“	218
4.8. Michael Smith's Lösungsansatz	222
4.9. Die Lösung des „moral problem“ aus moralpsychologischer Perspektive.....	227
5. Bibliografie	235

1. Einleitung und Überblick über die Argumentation

1.1. Problemlage und Umfeld des „moral problem“

Der von A.J. Ayer (1936) entwickelte Emotivismus war begründet auf den positivistischen Bestrebungen, eine Urteilstheorie zu entwickeln, in der nur zwei Typen von Sätzen als bedeutungsvoll gelten sollen: analytische Urteile und diejenigen synthetischen Urteile, die sich empirisch verifizieren lassen. Moralische Urteile fallen unter keine der beiden Klassen. Anstatt daraus zu folgern, dass sie keine Bedeutung haben, schlägt Ayer nun vor, eine alternative Theorie der Bedeutung moralischer Urteile zu entwickeln. Gemäß dieser Theorie des ethischen Emotivismus haben moralische Aussagen eine „emotive Bedeutung“ (Ayer 1936, Kap. 6, vgl. auch Stevenson 1937 und 1944). Eine Aussage wie „Stehlen ist (moralisch) schlecht.“ ist weder wahr noch falsch, weil sie sich nicht auf einen Zustand der Welt bezieht; wir drücken mit dieser Aussage vielmehr unsere Befürwortung oder Ablehnung eines Sachverhalts aus. Daher ist der Emotivismus auch unter der Bezeichnung der „Buh-/ Hurra-Theorie“ der Moral bekannt geworden: die oben angeführte Aussage ließe sich umformen in „Stehlen, buh!“.

Im Rückblick gelang es dem Emotivismus, ein eklatantes Problem jeder Moraltheorie zu lösen, jedoch nur, indem er in einem anderen Punkt eine stark kontraintuitive Position bezog. Zunächst zum entscheidenden Vorteil einer emotivistischen Theorie der Moral: Nach Davidson besteht eine Handlung aus zwei Teilen: einem Motiv (*pro-attitude*) und einem Glaubenszustand (vgl. Davidson 1963 und 1970). Jegliche kognitivistische Theorie der Moral wird ein moralisches Urteil als Glaubenszustand identifizieren, da es ein Urteil über einen Sachverhalt darstellt und wahrheitsfähig ist. Daraus folgt nun das sog. „moral problem“ (Smith 1994)¹, dass trotz des Habens einer moralischen Überzeugung – in einer *ceteris-paribus*-Situation – dennoch keine entsprechende Handlung erfolge, einfach weil ein moralisches Urteil nicht notwendig eine motivationale Einstellung („*pro-attitude*“) einschließe.

Kognitivistische Theorien handeln sich leicht das Problem des Externalismus ein, d.h. sie können nicht erklären, wie ein moralisches Urteil (mit seiner typischen „Geist-auf-Welt“

¹ Die Problemlage des von Smith so apostrophierten „moral problem“ wurde von Lewis Carroll schon 1895 für die Logik ausgearbeitet im Paradoxon von Achilles und der Schildkröte, vgl. Carroll 1895, dazu: Blackburn 1995; zur Parallele von theoretischer und praktischer Rationalität anhand des Carroll-Paradoxons vgl. auch Engel 1998, Smith 1994, 178 und Stroud 1979

Ausrichtung, vgl. Searle 1979 und 1983, jeweils Kap. 1) allein handlungswirksam werden kann. Smith fasst das Problem des Externalismus wie folgt zusammen: „it is a contingent and rationally optional matter whether an agent who believes that it is right to act in a certain way is motivated to act accordingly“ (Smith 1994, 71). Dieses Problem umgeht der Emotivismus, der stets auch ein motivationaler Internalismus ist: ein Urteil wie „Alten Damen über die Straße helfen, hurra!“ hat den Status einer Pro-Einstellung.

Nun jedoch zu demjenigen Punkt, in dem sich der Emotivismus in krassem Gegensatz zu unserem intuitiven Moralverständnis positioniert. Emotive moralische Überzeugungen haben schließlich keinen Wahrheitswert und können darum auch nicht Gegenstand rationaler Kritik werden.² Diese Konsequenz scheint nicht akzeptabel, da wir als menschliche Akteure auch intuitiv immer schon davon ausgehen, dass moralischen Urteilen in gewisser Weise Wahrheit oder Falschheit zukommen kann, und dass sie intersubjektiv zur Diskussion gestellt, verifiziert und ggf. berichtigt werden können. Diese Möglichkeit kann jedoch der Emotivismus nicht einräumen; klassische emotivistische Positionen in der Linie von Ayer und Stevenson werden nicht mehr vertreten.³ Die Problemlage jedoch, die sich prägnant im Smith'schen „moral problem“ (in der später zu diskutierenden Form eines Systems von drei miteinander unvereinbaren Propositionen) niederschlägt, ist nach wie vor virulent, da sie m.E. eine grundlegende Fragestellung in der beschreibenden Moraltheorie, der Metaethik also, aufgreift. Die reiche Zahl an wissenschaftlich äußerst hochwertigen Studien, die seit den 90er Jahren des 20. Jahrhunderts zum Problemkreis des „moral problem“ und seinen vielfältigen Manifestationsformen („Internalismus“ vs. „Externalismus“; „Emotivismus“ vs. „Kognitivismus“; „Neo-Humeanismus“ vs. „Neo-Kantianismus“ usw.) sowie zur Metaethik im Allgemeinen erschienen sind⁴, ebenso wie die anhaltende Aktualität der entsprechenden klassischen Referenz-Studien⁵ belegen das dringende, ungebrochene Interesse an der wissenschaftlichen Arbeit an dieser Fragestellung.

² Sie werfen außerdem große Probleme im Bereich der metaethischen Semantik auf, vgl. Geach 1960 und 1965 sowie die Darstellung bei Blackburn 1984, 189ff.

³ Dies gilt auch für die klassische „präskriptivistische“ Position von Hare 1952, die wegen ihrer Nähe zum Emotivismus hier nicht eigens behandelt wurde.

⁴ Vgl. bspw. Audi 1993 u.f.; Bittner 2001; Blackburn 1998; Brady 2000; Cullity & Gaut 1997; Dancy 1993 u.f.; Darwall 1998; Gibbard 1990; Gosepath 1999; Halbig 2002 und 2003; Helm 2001a; Herman 1993; Korsgaard 1996 und 1996a; Raz 1990 u.f.; Slote 2001; Smith 1994 u.f.; Steinfath 2001 u.f.; Velleman 1999; Wallace 1999; Wiland 2002.

⁵ Vgl. bspw. Anscombe 1957 und 1958; Davidson 1963 u.f.; Frankfurt 1971; Kenny 1963; Williams 1973 u.f.

1.2. Emotionstheorien

Es gibt nun eine interessante historische Parallele zwischen dem Projekt der emotiven Moralthorie einerseits und Theorien auf dem Gebiet der Philosophie der Emotionen andererseits. Fast die gesamte neuzeitliche Philosophie, über Descartes, Hume und Kant, vertritt die These, dass Emotionen Affekte oder „passiones“ seien, also nicht-rationale Bewusstseinszustände, die außerhalb unserer Kontrolle liegen und denen wir als Erleidende, also passiv ausgeliefert sind. So schreibt Hume:

„Our authority over our sentiments and passions is much weaker than that over our ideas; and even the latter authority is circumscribed within very narrow boundaries.”
(Hume Enquiry 1748, VII 1, Selby-Bigge S. 68).

In unserer Alltagssprache sagen wir: „Wut steigt in mir auf“, „Traurigkeit überfällt mich“ etc. Dieser Diskurs über Emotionen scheint bis auf die Anfänge der Dichtung in der westlichen Welt zurückzugehen, und er ist in sehr eindringlicher Form in einem Fragment des griechischen Dichters Archilochos aus dem 7. Jhd. v.Chr. greifbar. In einem tetrametrischen Fragment (West I 105) weist das lyrische Ich (Archilochos selbst) den Glaukos darauf hin („Γλαῦχ’, ὄρα“), wie das Meer von Wellen aufgewühlt wird, und Wolken sich am Felsen von Gyrai zusammenballen: es ist ein Zeichen für Sturm („σῆμα χειμῶνος“). Die Natur also ist vorhersehbar, die Angst jedoch überfällt uns unvorhergesehen („κικάνει δ’ἔξ ἀελπίτης φόβος“). Emotionen zeichnen sich dadurch aus, dass sie affektiv und nicht-kognitiv sind, die Opposition von Vernunft und Gefühl ist für uns ein Gemeinplatz, und eben darauf fußen auch die klassischen emotivistischen Theorien. Diesen gemäß stellt eine moralische Überzeugung eine emotive Einstellung dar, die gerade nicht kognitiv und wahrheitsfunktional ist, und sich nicht auf einen Zustand der Welt bezieht. Der Emotivismus speist sich aus der typisch neuzeitlichen These von den Emotionen als Affekten. Das Projekt des Emotivismus wird nun aufgegeben kurze Zeit bevor zum ersten Mal wieder kognitivistische Theorien der Emotionen Anerkennung erlangen. Psychologen (um 1960) und Philosophen (um 1970) gelingt es zu zeigen, dass Emotionen tatsächlich eine kognitive Komponente einschließen, dass sie Kriterien von Rationalität und Wohlbegründetheit folgen und dass sie intentional sind. Ein Anschluss an den Emotivismus scheint unter Einbeziehung der neuen Forschungsergebnisse zur Theorie der Emotionen wieder möglich, darum sollen im Folgenden die Grundzüge der älteren affektiv-physiologischen und der heute vertretenen kognitivistischen Emotionstheorien referiert werden. Ich weise jedoch schon an dieser Stelle darauf hin, dass sich die Zielsetzung meiner Studie von derjenigen der klassischen emotivistischen Projekte insofern unterscheidet, als ich nicht auf eine linguistische oder logisch-semantische Untersuchung moralischer Aussagen abhebe. Ich möchte vielmehr den mentalen Status moralischer Überzeugungen

erhellen und die sich daraus ergebenden Implikationen für die metaethische Handlungstheorie und Moralphychologie klären. Das von mir vorgeschlagene Modell bezieht sich auf moralische Überzeugungen, nicht auf moralische Aussagen.

Die bis weit in das 20. Jahrhundert verbindliche Theorie der Emotionen wurde von William James und Carl Lange unabhängig voneinander im Jahr 1884 bzw. 1885 formuliert (James 1884, Lange 1910). Diese James-Lange-Theorie definiert Emotion als das Erleben einer peripher-physiologischen Veränderung. Ich registriere eine solche peripher-physiologische Veränderung, wie z.B. gesteigerte Transpiration, Herz- und Atemfrequenz, und das Erleben dieser Veränderung *ist* meine Emotion (z.B. der Furcht). Schon 1927 wurden von Walter Cannon wichtige Argumente und Ergebnisse gegen diese Theorie vorgebracht (Cannon 1927), viele weitere folgten. Emotionales Erleben scheint möglich zu sein, selbst wenn die peripher-physiologische Aktivierung fehlt, wie dies z.B. bei Probanden mit Rückenmarksverletzungen der Fall ist (Dana 1921). Cannon bringt zum ersten Mal ein weiteres wichtiges Argument: die peripher-physiologischen Veränderungen, deren Wahrnehmung nach James schon die entsprechende Emotion ist, sind in Wirklichkeit viel zu diffus und undeutlich, um die ganze Bandbreite unseres emotionalen Erlebens erklären zu können. Tatsächlich zeigt schlichte Introspektion, dass die körperlichen Veränderungen bei einem Zustand von äußerst unterschiedlichen Emotionen wie Furcht und intensiv erlebter Freude nicht zu unterscheiden sind.

Die vermeintliche Widerlegung der James-Lange Theorie durch Dana und Cannon wurde später erneut kritisch relativiert. Neuere Untersuchungen an Patienten mit Rückenmarksverletzungen auf unterschiedlichen Niveaus der Wirbelsäule (Hohmann 1966) haben gezeigt, dass emotionales Erleben mit ansteigender Höhe der Läsion an Intensität verliert, bis schließlich nur noch von einem „Als-ob-Erlebnis“ berichtet wird („Es ist, als ob ich Furcht empfinde.“). Die physiologische Komponente, der in der James-Lange-Theorie der zentrale (kausale) Platz zukommt, wird also auch in neueren Emotionstheorien berücksichtigt werden müssen.

Emotionen sind also sicher (1.) keine rein *physiologischen* Phänomene (die behavioristische Emotionstheorie sei an dieser Stelle ausgespart), in der Art wie dies etwa Gefühle des Hungers, der Müdigkeit oder des physischen Schmerzes sind, und sie lassen sich nicht im Sinne von James und Lange auf die Wahrnehmung einer physiologischen Veränderung reduzieren. Reduktionistisch physiologische Emotionstheorien haben sich als nicht plausibel erwiesen und sind empirisch widerlegt worden. Emotionen lassen sich andererseits auch (2.) nicht primär über den emotionalen *Ausdrucks* klassifizieren (d.h.

Mimik, Gestik oder sog. expressive Emotionshandlungen, wie das Zerschlagen einer Vase aus Wut über Hans). Diese Methode kann nur sehr grundlegende Differenzen erfassen wie z.B. „Angst“ vs. „Freude“. Schwieriger ist schon die Differenzierung von „Angst“ vs. „Überraschung“, völlig unmöglich dagegen diejenige von „Scham“ vs. „Verlegenheit“. Gleiches gilt (3.) für eine Klassifikation des affektiven (phänomenalen) *Erlebens*: „Neid“ und „Eifersucht“ sind allein am affektiven Erleben nicht distinktiv zu unterscheiden. Emotionen sind also konzeptuell nicht ohne gewisse kognitive Inhalte zu verstehen.

Im Jahr 1962 formulierten dann Stanley Schachter und Jerome Singer die erste kognitivistische Emotionstheorie in der psychologischen Forschung des 20. Jahrhunderts, ausgehend von einer empirischen Studie, die in der Emotionsforschung noch heute Referenzqualitäten hat (Schachter/ Singer 1962). Gemäß dieser Theorie löst die Wahrnehmung eines Reizes (Bsp.: ein Mann mit Pistole) eine physiologische Aktivierung („arousal“) aus. Diese wird in einem nächsten Schritt kategorisiert („gelabelt“), u.z. entweder als ein physiologisches Symptom, falls keine emotional relevante Situation vorliegt, oder als eine näher bestimmte Emotion, entsprechend einer kognitiven Bewertung der gegebenen Situation (Elemente wie: „es ist Nacht, der Mann lauert in einer dunklen Straßenecke, es handelt sich um einen Unbekannten von zwielichtiger äußerer Erscheinung“ etc. würden als Furcht kategorisiert). Die physiologische Aktivierung geht zwar weiterhin dem emotionalen Erleben voraus, es hängt hier jedoch von unserer kognitiven Bewertung ab, ob und wie diese Aktivierung als Emotion kategorisiert wird.

Aus philosophischer Sicht muss nun die Frage gestellt werden, welcher Stellenwert der von Schachter/ Singer als primär postulierten „Wahrnehmung eines Reizes“ zukommt. Handelt es sich hier schon um eine komplexe Kognition im Sinne eines Urteils oder sogar eines Schlusses? Mit anderen Worten: löst die Wahrnehmung einer Schlange unmittelbar Furcht aus, nach dem Muster einer stimulus-response-Reaktion? Oder bedarf es eines Schlusses wie „Dies ist eine Schlange. (und: Schlangen sind gefährlich.)“? Dieses letztere Beispiel ist tatsächlich schwer zu entscheiden, da es offenbar bestimmte, sog. prinzipiell phobische Objekte gibt, die wir ohne besonderes Wissen über die Welt fürchten. Nach Seligman (1971) haben wir eine gewisse Bereitschaft („preparedness“), mit Furcht auf diejenigen Objekte zu reagieren, die sich in unserer Evolutionsgeschichte als gefährlich herausgestellt haben (vgl. auch Öhman 1987). Diese Objekte bilden demnach die Klasse der potenziell phobischen Objekte. Die gleiche Frage stellt sich auch in Fällen, die sich nicht plausibel auf eine evolutionsgeschichtlich bedingte Bereitschaft zurückführen lassen: Das Gefühl der Erhebung oder Wertschätzung beim Anblick einer schönen Landschaft oder eines

Gemäldes, das uns spontan gefällt, ist womöglich nicht auf ein zugrundeliegendes Urteil zurückzuführen. Vielmehr könnte dieses primäre Gefühl später die Grundlage für ästhetische Urteile über verwandte Objekte darstellen. In diesem Sinne postuliert Steinfath (2002) die Möglichkeit der Genese emotionaler Zustände auf der Grundlage von einfachen Sinneswahrnehmungen. Diese Frage wird in einem späteren Zusammenhang zu klären sein.

Ich halte fest, dass in der zeitgenössischen Psychologie eine syndromische Definition der Emotionen konsensfähig zu sein scheint (Goller 1992, 16ff.). „Emotion“ wird verstanden als ein Zustand, der sich durch drei Komponenten auszeichnet: (1.) das gefühlsmäßige Erleben („Ich-Qualität“ der Emotionen), (2.) bestimmte peripher-physiologische Veränderungen und (3.) ein bestimmtes korrelierendes emotionales Verhalten (Mimik, Gesten, Fluchtverhalten bei Furcht etc.). In der philosophischen Literatur werden viele Spielarten der kognitivistischen Theoriegruppe vertreten, die ich im Folgenden aufschlüsseln werde (vgl. hierzu Nash 1989 und Rapp 2002, 559-570). Die Gemeinsamkeit aller kognitivistischen Emotionstheorien ist zunächst eine Gruppe von Gegnern, nämlich die verschiedenen Körper- oder Gefühlstheorien der Emotionen. Weiterhin besteht die Gemeinsamkeit darin anzunehmen, dass Emotionen einen intentionalen Gehalt haben, wie es z.B. Kenny deutlich ausdrückt: „Emotional attitudes, like other mental attitudes, have formal objects.“ (Kenny 1963, 191). Dies kann dazu führen, dass unterschieden wird zwischen „Emotionen“ einerseits, die den Regeln eines formalen Objekts unterliegen (d.h. ich kann mich z.B. nicht vor einem Unglück *fürchten*, von dem ich weiß, dass es schon geschehen ist), und „Gefühlen“ andererseits, die eben nicht dieser Bedingung unterliegen und damit unter den traditionellen Begriff der „Affekte“ fallen dürften.

Einige Autoren gehen noch einen Schritt weiter, indem sie Emotion lediglich für eine Art von Urteilen halten, d.h. eine „rein kognitivistische“ oder „Urteilstheorie“ der Emotionen vertreten. Solomon etwa hält Emotionen für „relatively intense evaluative judgments“ (Solomon 1976, 187, vgl. auch Solomon 1988). Dieser rein kognitivistische Theorietypus, der auch in der zeitgenössischen Psychologie vorzufinden ist (Beck 1979, Ellis 1982), hat schon eine antike philosophische Tradition (zu der Frage, ob Aristoteles zu der Gruppe der Urteilstheoretiker zu rechnen ist, vgl. Rapp 2002, 554f.): die Stoiker vertraten eine solche „Urteilstheorie“ der Emotionen (vgl. Lloyd 1978). So soll Chrysipp „Trauer“ definiert haben als eine „frische Meinung, dass ein Übel anwesend sei [δόξαν πρόσφατον κακοῦ παρουσίας]“ (SVF 3, 463). Diese Position wird auch noch in der jüngeren Stoa vertreten: wann immer wir durch ein Gefühl affiziert werden, sollen wir nach Epiktet „niemals einem anderen die Schuld daran geben [αἰτιώμεθα] als uns selbst, das heißt: unseren eigenen Auffassungen [τὰ ἑαυτῶν δόγματα]“ (Epict., enchir. 5, Übers. Hülser in LS 65U; vgl. Epict.,

diatr. I 12, 20f.). In gleicher Weise äußert sich Cicero: „Est igitur causa omnis in opinione, nec vero aegritudinis solum, sed etiam reliquarum omnium perturbationum [...]” (Cicero, Tusc. III 11, 24). Es scheint jedoch kein strenges Art-Gattungs-Verhältnis zwischen Emotionen und Urteilen vorzuliegen. Das Element der Affektivität ist sowohl empirisch als auch konzeptionell notwendig, um emotionale Zustände von Glaubenszuständen („Schnee ist weiß.“, „1+1=2“) zu unterscheiden. Gefühls- oder Körpertheorien haben zwar keine Kriterien zur Erklärung der faktisch bestehenden Wohlbegründetheits-Bedingungen von Emotionen, Urteilstheorien dagegen geraten in einen Externalismus in Bezug auf die Erklärung von physiologischen und appetitiven Aspekten von Emotionen. Die Theorie, der ich für meine Studie den Vorzug gebe, ist eine „Komponententheorie“, die Emotionen als ein Phänomen mit (1.) physiologischen, (2.) kognitiven, (3.) appetitiven und (4.) evaluativen Anteilen versteht. Meiner Ansicht nach wird ein solcher Theorietypus u.a. vertreten von Lyons (1980), Stocker (1996) und Steinfath (2000 und 2002). Insbesondere sollen in diesem Zusammenhang an entsprechender Stelle auch die innovativen Ergebnisse Antonio Damasio (1994) zur Neuropsychologie der Emotionen diskutiert werden. Die Unterscheidung von (2) und (4) soll einem Argument begegnen, das Rapp (2002, 561) gegen Sherman (1997, 55) vorbringt: fasst man nämlich den kognitiven Gehalt im gleichen Zug als bewertend auf, verstößt man gegen ein enges Verständnis von „kognitiv“, dementsprechend kognitive Zustände Auskunft darüber geben, wie die Welt faktisch ist, was nicht notwendig eine Bewertung einschließt.

Es gibt also wichtige Gründe dafür, die strikte Opposition zwischen Urteil und Emotion zu relativieren. Im Rahmen einer Komponententheorie der Emotionen wäre nun zu klären, welche Rolle den einzelnen Komponenten zukommt, sowohl im Hinblick auf die Kontrolle bzw. Beeinflussung von emotionalen Zuständen, als auch deren Genese. Ich diskutiere in diesem und im folgenden Absatz besonders die Rolle der Kognitionen, da diese gerade die entscheidende Neuerung im Emotionsmodell ausmachen und damit den besonderen Vorzug gegenüber dem klassischen Emotivismus darstellen. Ich weise darauf hin, dass ich in diesem und im nächsten Absatz von „Kognition“ immer in der Bedeutung von „evaluativem kognitivem Bewusstseinsinhalt“ spreche, also in der nach Rapp „weiten“ Bedeutung von „Kognition“. Zunächst zeigt sich, dass wir mittels unserer Vernunft, d.h. unserer Urteile und Überzeugungen, Zugriff auf unsere Emotionen haben. Sog. pragmatische Paradoxa haben auch hier Geltung. Die Aussage „Ich *bin wütend* auf Hans, weil er mein Stück Torte gegessen hat, aber ich *glaube* nicht, dass Hans mein Stück Torte gegessen hat.“ ist paradoxal. Es ist nicht nur unsinnig, sie zu äußern, sie ist vielmehr sogar performativ-therapeutisch, weil sie im Moment der Äußerung den emotionalen Zustand und damit auch das vermeintliche Paradoxon

aufhebt. Wie ich bereits ausgeführt habe, unterliegen bestimmte Emotionen den konzeptionellen Bedingungen ihrer formalen Objekte (so empfinde ich unter Standardbedingungen keine Wut, wenn ich alleine und nachts in einem Wald von einem wilden Tier angegriffen werde; Furcht wäre hier die angemessene Reaktion). Im vorliegenden Beispiel sind jedoch die Bedingungen des formalen Objekts erfüllt, es geht hier vielmehr um den existentiellen Status des Objekts. Das Haben einer Emotion hängt von einem intentionalen Objekt p ab, von dem ich zumindest glauben muss, dass es der Fall ist. Wenn ich nun überzeugt bin, dass p nicht der Fall ist, und die Emotion dennoch fortbesteht, dann heißt dies, dass p nicht das tatsächliche Objekt meiner Wut war. Vielleicht bin ich ja eigentlich wütend auf Hans, weil er meine Modelleisenbahn kaputt gemacht hat. Um also von einer wohlbegründeten („wellfounded“, Lyons 1980) Emotion sprechen zu können, müssen wir voraussetzen, dass ich ein gewisses Reflexionsniveau erreichen kann, um das Objekt meiner Emotion zu finden. Der wichtige Punkt ist hier, dass das Haben oder Nichthaben einer bestimmten Überzeugung eine zwingende Auswirkung auf einen bestehenden emotionalen Zustand hat. Eine weitere Bedingung ergibt sich jedoch aus dem, was ich den psychopathologischen Status einer Emotion nennen möchte. In der Psychologie definiert man z.B. Phobien als besonders starke Furchtreaktionen, die *nicht* rational zugänglich sind. Die phobische Furcht vor einer Spinne lässt sich nicht aufheben, indem man den Betroffenen darauf hinweist, dass von der Spinne (nehmen wir an, einer europäischen Hausspinne) keine Bedrohung ausgeht. Hier scheint ein prädominanter affektiv-physiologischer Anteil vorzuliegen.

Die meisten emotionalen Zustände sind also prinzipiell dem Einfluss von veränderten Kognitionen oder entscheidenden Neubewertungen eines Sachverhalts zugänglich. Eine stärkere (wie sich zeigen wird: bedingt kausale) Beziehung zwischen Kognition und Emotion würde sich ergeben, wenn sich nachweisen ließe, dass Kognitionen nicht nur bestehende emotionale Zustände beeinflussen, sondern diese auch zuallererst hervorrufen können. Da physiologische Aktivierung präsent sein muss, um einen emotionalen Zustand von anderen nicht-emotionalen Bewusstseinszuständen abzugrenzen, scheint diese Möglichkeit zunächst fraglich. Setzt man nämlich eine Kognition als Ursache für physiologische Veränderung an, so gerät man scheinbar in die gleichen Probleme des Externalismus wie im Falle der Handlungsmotivation. Allerdings beruht das Externalismus-Problem auf der Möglichkeit der praktischen Inkontinenz oder Willensschwäche, das im Fall der Emotionen nicht besteht. Tatsächlich ist die Verursachung einer Emotion durch eine Kognition etwas Alltägliches: wir

erröten in Verlegenheit wegen eines Freudschen Versprechers, selbst wenn vorher keine signifikante Sympathikus-Aktivierung vorlag.

Es bleibt zu klären, warum bestimmte evaluative Kognitionen offenbar nicht Ursache von emotionalen Zuständen werden können, andere jedoch schon. Lyons (1980, 58) bestimmt den Unterschied von emotionalen und nicht-emotionalen Bewertungen als „starke Präferenz“ im ersten Fall. Wenn ich z.B. einem Hund begegne, kann ich verschiedene evaluative Einstellungen ihm gegenüber vertreten, ohne dass ein emotionaler Zustand resultieren würde.

„But if I tell myself that the dog is liable to take a large piece out of me, and that it is giving every indication that it will, and if I would prefer that it did not, *in a very strong sense of prefer*, that is, if I evaluate the dog's approach as threateningly dangerous, then I have reason to get worked up about the situation, and my beliefs have become the plausible cause for my ensuing physiological upset such as a sudden discharge of my sympathetic nervous system and a speeding up of my metabolic processes.“ (loc. cit., Hervorh. v. mir, S.H.)

Goller (1992, 166) benutzt für das Phänomen einer starken Präferenz den Begriff der „heißen“ Kognition (vgl. auch LeDoux 1996, 43; der Begriff stammt von Abelson 1963). Dabei handelt es sich um Evaluationen mit „Ich-Qualität“, die das persönliche Involviertsein des Akteurs in die Situation ausdrücken. In diesem Sinne stellen Emotionen einen persönlichen Weltbezug her, in dem der Akteur erfährt, dass die Sachverhalte der Welt ihn angehen. Ich weise schon an dieser Stelle darauf hin, dass Emotionen mit Werten in Beziehung stehen, nicht nur insofern sie uns auf einen Wert aufmerksam machen (in Lyons' Beispiel auf den Wert der körperlichen Unversehrtheit), sondern auch insofern dem Subjekt im Lichte des affektiven Erlebens einer emotionalen Evaluation zuallererst ein Objekt als werthaftig erscheint. Das Erklärungsmodell der starken Präferenz wurde von Marks (1982) kritisiert für die Klassifizierung der nicht-kognitiven Komponente „Präferenz“ unter eine kognitive Kategorie (im weiten Sinne), nämlich „Evaluation“. Ursächlich für Emotionen sind nach Marks jedoch „desires“. Ich möchte aber anders argumentieren: dank ihrer appetitiven oder orektischen Komponente sind Emotionen zwar in der Lage, physiologische Aktivierung auszulösen und Handlungsmotive abgeben zu können. Es gibt jedoch kein Streben ohne Objekt, d.h. Streben ist stets auf ein Gut gerichtet. Darum muss das entsprechende Werturteil der Strebensregung vorausgehen. Die von Marks vertretene Position ist nur plausibel im Rahmen eines klassisch-humeanischen Modells, das invariante „desires“ schon voraussetzt.

Es wird Gegenstand des gesamten Kapitels 2 der vorliegenden Arbeit sein, auf der Grundlage einer Diskussion der gängigen emotionstheoretischen Positionen ein eigenes gemischt-kognitivistisches Emotionsmodell zu entwickeln, das insbesondere die wichtige Dichotomie von „Aktualität“ und „Disposition“ abzubilden vermag. Diese letztere und einige

andere wichtige Ergebnisse der Untersuchung werden am Ende des Kapitels (in 2.11.) in griffiger Form als Thesenkatalog zusammengestellt.

1.3. Emotionen, Werte und Wünsche

Der Zusammenhang zwischen „Wünschen“ (im Humeschen Sinne von „desires“) und dem von mir vorgeschlagenen Emotionsbegriff wird detailliert zu klären sein. Die Humeanisch ausgerichtete Handlungstheorie baut darauf auf, dass „Wünsche“ als nicht-kognitive Elemente als Motiv in praktische Deliberation eingehen. Auch emotionale Einstellungen sollen einen „motivierenden Grund“ zum Handeln liefern. Besteht jedoch ein strenges Implikationsverhältnis in dem Sinne, (1.) dass das Haben einer bestimmten emotionalen Einstellung immer auch das Haben eines entsprechenden Wunsches einschließt? Oder ist es möglich, (2.) einen Wunsch zu haben ohne die entsprechende emotionale Einstellung, bzw. (3.) eine emotionale Einstellung ohne einen damit zusammenhängenden Wunsch?

Ich beginne mit Fall (3). Aufgrund ihrer affektiven Komponente implizieren Emotionen auch immer entsprechende Wünsche. „Liebe“ impliziert den Wunsch, in der Gesellschaft der geliebten Person zu sein, oder dass der Person ein von ihr geschätztes Gut widerfahre etc.; Furcht impliziert den Wunsch, dass das gefürchtete Objekt gemieden werde; Wut zielt auf Herabsetzung, Beschädigung oder Zerstörung ihres Objekts. Allerdings sind die realisierten Wünsche nicht immer sofort eindeutig auf die richtige zugrundeliegende Emotion zurückzuverfolgen. Wenn ich aus Liebe meiner unheilbar kranken Großmutter Sterbehilfe leiste, so scheint dieser von mir realisierte Wunsch eigentlich meiner Liebe zur Großmutter zu widersprechen. Trotzdem kann ich sinnvoll erklären, ich hätte aus Liebe gehandelt, weil „Liebe“ auch den Wunsch impliziert, der geliebten Person die von ihr geschätzten Güter zu verschaffen (also den erlösenden Tod, den meine Großmutter wünscht), selbst wenn meine Handlung im Gegensatz zu einem anderen aus der Liebe erwachsenden Wunsch steht, nämlich in Gesellschaft der geliebten Person zu sein. Eine weitere Schwierigkeit ergibt sich im Fall von sog. rückwärtsgerichteten Emotionen wie Trauer, die sich auf bereits geschehene und damit unveränderliche Ereignisse beziehen. Trotzdem impliziert die Emotion der Trauer einen entsprechenden Wunsch, ohne den sie nicht verständlich wäre. Es handelt sich aber um einen Wunsch im *modus irrealis*, wie etwa „Ich wünschte, meine Großmutter lebte noch!“.

Bei Fall (2) liegt der Schlüssel zur Lösung in der evaluativen Komponente von Emotionen. Ich will gegen die Annahme argumentieren, dass es möglich sei, einen Wunsch zu haben ohne auch die entsprechende emotionale Einstellung. Fälle, die diese Annahme

illustrieren, sind Wünsche, die mich als Akteur in möglichst geringer Weise angehen, wie etwa „Ich wünschte, 1 plus 1 wäre gleich 3!“. Interessant ist, dass die Äußerung eines solchen Wunsches eine pragmatische Härte einschließt, weswegen ein Zuhörer wohl geneigt wäre zu fragen: „Warum denn das?“. Offenbar ist die Idee eines Wunsches, der uns in keiner Weise angeht (wie etwa eine „kalte“ Kognition nach Goller 1992, 166), ein begriffliches Paradoxon. Es scheint unplausibel, etwas zu wünschen, das ich nicht selbst, bzw. dessen Konsequenzen ich nicht für wertvoll halte. Auch ein weniger hartes Beispiel wie der Wunsch, „dass es der Menschheit im 35. Jahrhundert gut ergehe“ geht uns zwar als Akteur nicht direkt an, weil er den Bereich unserer Erfahrung und unserer Handlungskompetenzen in mehreren Dimensionen übersteigt. Aber ein solcher Wunsch erscheint uns als Zuhörern nicht paradoxal, denn niemand würde ihn äußern, der nicht „die Menschheit“ aus eigenem emotionalen Erleben irgendwie als *wertvoll* erachten würde. Ich „wünsche“ dasjenige, was ein Wert „ist“ bzw. einen Wert „hat“.

Die obige Argumentation spricht eindeutig für ein strenges Implikationsverhältnis (1) von Emotionen auf Wünsche. Natürlich kann es auch hier Konflikte geben, die sich jedoch im Rahmen des von mir vorgeschlagenen Modells leicht erklären lassen. So kann ich mir z.B. aufgrund situationsbezogener Informationen plötzlich bewusst werden, dass ich einen Wunsch habe, der sich nicht sinnvoll auf die Hintergrundidentität meiner emotionalen Einstellungen zurückbeziehen lässt. Die Inkommensurabilität liegt jedoch auf einer tieferen Ebene, denn ein Wunsch schließt natürlich ein, dass ich eine den Wunsch begründende akute Emotion habe, die noch nicht den Prozess eines Integrationsversuchs in meine Hintergrundidentität durchlaufen hat. In den Kapiteln 3.1. bis 3.9. der vorliegenden Arbeit sollen gängige axiologische Theorien und Konzeptionen des „desire“-Begriffs in einigem Detail diskutiert und dabei ständig an meine Konzeption des Emotions-Begriffs zurückgebunden werden. Es wird sich ein wesentlich differenzierteres Bild von dem Zusammenhang zwischen Emotionen, Werten und „desires“ ergeben, als dies im Rahmen dieser Einleitung zu leisten möglich wäre. In den Kapiteln 3.10. bis 3.12. werden abschließend noch einmal dezidiert diejenigen Problemfälle diskutiert, die bereits umrisshaft in dieser Einleitung angesprochen wurden: emotionale Wertungen sollen im Hinblick auf ihre Intentionalität (Kap. 3.10), transpersonale Geltung (Kap. 3.11.) und Zurechenbarkeit (im Sinne von „Verantwortlichkeit eines Akteurs“, Kap. 3.12.) geprüft werden.

1.4. Moralische Überzeugungen als Emotionale Einstellungen

Es scheint also möglich zu sein, eine vertretene moralische Überzeugung als eine emotionale Pro- oder Contra-Einstellung zu verstehen, wobei alle Kriterien gewahrt bleiben, die wir intuitiv als notwendig für eine moralische Theorie ansehen. (1.) Eine solche emotionale Einstellung generiert für ihren Träger einen „internen“ (Williams 1980) oder „motivierenden“ (Smith 1994, 95f.) Grund, also ein Motiv, die entsprechende Handlung auszuführen. Wie oben erläutert, greift das aus der Logik bekannte und von Smith auf die Moralthorie übertragene Carrol-Paradoxon nicht. (2.) Der Vorzug gegenüber dem klassischen Emotivismus ist wiederum, dass diese emotionalen Einstellungen wahrheitsfähig sind, u.z. nicht nur im Hinblick auf ihre „Erfüllungsbedingungen“ (Scarano 2001, 132), sondern auch auf ein Kriterium der Wohlbegründetheit (einer Emotion auf einer Kognition oder einer Emotion als ein Element eines emotionalen Musters). (3.) Die rektifizierenden Momente für emotionale Einstellungen sind akteursinterne Reflexivität (ich muss in der Lage sein, vorgeschobene oder fingierte Objekte zu entlarven), der moralische Diskurs (ein Gespräch mit anderen Personen, die Anstoß nehmen an meinen Einstellungen, ist in der Lage, meine Überzeugungen und damit auch meine emotionalen Einstellungen zu modifizieren) und die aus systeminternen Wertkonflikten erwachsenden Modifikationen der vertretenen Einstellungen. (4.) Daraus ergibt sich, dass wir verantwortlich gemacht werden können für unsere Einstellungen und die aus ihnen resultierenden Handlungen. Emotionen sollen hier also nicht in erster Linie als Affekte verstanden werden, durch die wir bisweilen „disruptive“ Handlungen begehen (die also nicht in einem strengen Sinne in unserer Verantwortung stehen; z.B. „Handlungen im Affekt“). (5.) Da emotionale Einstellungen auf Kognitionen beruhen, entgeht dieses Modell dem Vorwurf des viziösen Subjektivismus. Es geht gerade nicht um „desires“ im Humeschen Sinne, die wir in uns vorfinden, ohne sie verändern zu können. Die Aneignung konsistenter moralischer Vernunftprinzipien ist Voraussetzung für das Haben einer wohlbegründeten emotionalen Einstellungen (wobei hier keine einseitige Implikation, sondern eine Wechselwirkung anzunehmen ist).

Der Emotionsbegriff hinter dem von mir vorgeschlagenen Konzept der „emotionalen Einstellung“ unterscheidet sich in einem wichtigen Punkt von den anderen angeführten Beispielen. Da er sich auf dem Niveau der Akteursidentität situiert, kann es sich nicht um eine akute emotionale Reaktion handeln, sondern nur um eine emotionale Disposition. Der zugrundeliegende Unterschied ist auch im Alltagsverständnis der Emotionen leicht zu fassen. Eine akute Reaktion von „Trauer₁“, die sich in dem Moment manifestiert, da man vom Tod einer nahestehenden Person unterrichtet wird, unterscheidet sich von der vielleicht jahrelang

andauernden „Trauer₂“, die man nach Überwindung des ersten, akuten Eindrucks empfindet. Erstere ist intensiver und hat im affektiven Erleben ein eindeutiges An- und Abklingen, letztere ist gerade nicht eine auf mehrere Jahre ausgedehnte „Trauer₁“-Reaktion. Vielmehr stellt „Trauer₂“ eine „Gefühlshaltung“ (Ulich 1982) dar, die einerseits eine mehr oder weniger starke Erlebnistönung mit sich führt, andererseits eine Art Disposition darstellt, gelegentlich, besonders aber in spezifischen Situationen (z.B. beim Betrachten einer Fotografie der verstorbenen Person) mit „Trauer (1)“ zu reagieren. Beide Typen sind gleichermaßen intentional und kognitiv fundiert, allerdings verstehen wir oft Typ (1) als eine Instantiierung von Typ (2). Denn wenn wir Hans einen „traurigen (fröhlichen) Menschen“ nennen, meinen wir damit, dass er aus einer bestimmten Charakterkonstante vom Typ (2) heraus oft die akute Emotion „Trauer (Freude)“ vom Typ (1) zeigt. Wir sprechen über emotionale Dispositionen in gleicher Weise wie über „Tugenden“ (im Sinne von gr. „aretai“) als Charaktereigenschaften: ein gerechter Mensch ist jemand, der immer oder zumindest meistens gerechte Handlungen ausführt. Bei emotionalen Einstellungen in der Moral handelt es sich also um den Typen (2) der emotionalen Disposition, die sich auf dem Niveau der Akteursidentität situiert, die jedoch stets notwendig auf potenzielle Instantiierungen des Typs (1) verweist. Auf genau diese Weise lassen sich Wahrheitsfunktionalität und Motivationsinternalismus verbinden.

Das zentrale Konzept des Modells, die „emotionale Einstellung“, ist in dieser Hinsicht inspiriert durch die aristotelische Tugend- und Affektenlehre. Aristoteles definiert „Tugend“ (ἀρετή) als eine „Einstellung“ bzw. „feste Grundhaltung“ des Wählens (EN II 6, 1106b36: „ἔξις προαιρετική“; vgl. EE I 10, 1227b8). Die Tugend bezieht sich auf Handlungen und Leidenschaften (EN II 2, 1104b13f.: „περὶ πράξεως καὶ πάθη“, auch 1106b25), wobei diese Bestimmung an späterer Stelle (1106b16f.) auf die „ethische“ (gegenüber der „dianoetischen“) Tugend beschränkt wird. Nun verhalten wir uns durch die Tugend als „Einstellung“ oder „feste Grundhaltung“ („ἔξις“ oder „habitus“ bei späteren Kommentatoren) den Leidenschaften gegenüber („καθ’ ὅς πρὸς τὰ πάθη“) richtig oder falsch (EN II 4, 1105b19-28). Denn in der aristotelischen Ethik gibt es bei allen Leidenschaften ein Zuviel und Zuwenig, keines von beiden ist richtig. Aber Leidenschaften zu empfinden, „wann man soll [δεῖ] und wobei man es soll und wem gegenüber und wozu und wie, das ist die Mitte und das Beste, und dies kennzeichnet die Tugend“ (EN II 5, 1106b21-23, Übers. Gigon). Ausgehend von diesem „Gebotensein“ richtiger Emotionen, muss man „Tugend“ in Bezug auf die Leidenschaften als (reflexionsfähige) „Disposition“ des Typs (2) zu angemessenen Emotionen verstehen. Auf dieser Erkenntnis baut das von mir vorgeschlagene Modell

emotionaler Einstellungen auf. Denn wenn wir als Menschen ausschließlich zu Emotionen des Typs (1) veranlagt wären, machte ein „Gebotensein“ von emotionalen Reaktionen keinen Sinn. Ebenso wenig könnten emotionale Einstellungen dann konsistente, wahrheitsfunktionale und wohlbegründete Handlungsmotive liefern.

Versteht man nun emotionale Einstellungen als Dispositionen (Typ 2) bzw. Charakterkonstanten im Rahmen einer Theorie der Akteursidentität, erscheint es unplausibel, sie auf eine handlungskatalysierende Funktion („gutheißen“-/ „ablehnen“-Einstellung) zu beschränken. Gerade in diesem Punkt unterscheidet sich mein Konzept der emotionalen Einstellung von dem der Pro- oder Contra-Einstellung (nach Davidson 1963 und 1970; vgl. Scarano 2001). Natürlich haben emotionale Einstellungen auch eine handlungstheoretische Funktion: sie liefern uns interne Gründe, d.h. Motive zum Handeln, jedoch mit einer wesentlich reicheren Semantik als dies bei Pro-Einstellungen der Fall sein kann.

Das Modell emotionaler Einstellungen versteht Emotionen nun nicht nur als wertindikativ (vgl. Sherman 1997), sondern v.a. als *wertkonstitutiv* (vgl. Steinfath 2001a und 2002) und vermag daher auch, moralisches Handeln oder Nicht-Handeln kontextsensitiv zu erklären.⁶ Steinfath (2002, 108) macht auf den Unterschied zwischen zwei Klassen von Wertprädikaten aufmerksam. Die erste Klasse der „dichten“ Prädikate (wie „bewundernswert“, „mutig“, „gerecht“) legt schon durch ein semantisches Verhältnis eine Beziehung zu bewertenden Emotionen nahe. Das Modell ist im Bereich der Metaethik situiert und bezieht eine Stellung zum „moral problem“ im Anschluss an den klassischen Emotivismus, jedoch unter der veränderten Perspektive eines handlungstheoretischen, nicht eines logisch-semantischen Ansatzes. Als Vorschlag einer Lösung des Probleme der vermeintlichen Gegensätzlichkeit von Motivationsinternalismus und Wahrheitsfähigkeit moralischer Überzeugungen unternehme ich den Versuch, diese als „emotionale Einstellungen“ zu verstehen. Dabei handelt es sich um kognitiv begründete (und damit reflexiv zugängliche) emotionale Dispositionen (d.h. Typ (2) der oben aufgestellten Typologie) auf der Ebene der praktischen Akteursidentität, die bestimmte Wünsche implizieren und interne Handlungsgründe darstellen können. Weiterhin sind angemessene Emotionen konstitutiv für das Zusprechen axiologischer Eigenschaften, auf denen wiederum moralische Normen fußen. Das Modell erlaubt einen moralpsychologischen Lösungsansatz zur Smith'schen Formulierung des „moral problem“ und gestattet mithin einen neuen Blick auf die dringende Problemlage von Phänomenen wie

⁶ Zum Zusammenhang von Werten und Emotionen vgl. allgemein Scheler 1980, De Sousa 1987 und Johnston 2001 für eine wertrealistische Position, kritisiert von Macki 1977, 38ff.; weiterhin Smith 1992, Mulligan 1998, Lewis 1989, D'Arms/ Jacobson 2000, Wiggins 1998 und Wright 1988.

Willensschwäche und anderen moralischen Dilemmata (wie z.B. Bernard Williams' „Gauguin-Problem“ in Williams 1976a). Eine Diskussion des „moral problem“, seinen Manifestationsformen und metaethischen Hintergründen wird in den Kapiteln 4.1. bis 4.8. der vorliegenden Arbeit unternommen; in Kapitel 4.9. werde ich daraufhin auf der Grundlage aller bis dahin erarbeiteten Ergebnisse meinen alternativen (moralpsychologischen) Lösungsansatz vorschlagen.

2. Was ist „Emotion“?

Kritischer Abriss einer Theoriegeschichte und Entwicklung eines Thesenkatalogs

2.1. Einleitung

Nachdem Vorgehensweise und Ziel der Arbeit dargestellt wurde, soll im vorliegenden Kapitel zunächst eine Theorie der Emotionen erarbeitet werden. Dabei ist der Blick stets auf die Ziele der Gesamtargumentation gerichtet, nämlich die Interpretation von moralischen Überzeugungen als „emotionale Einstellungen“. In diesem Sinne ist also einmal der Handlungs- bzw. Motivationsbezug von Emotionen zu klären, dies jedoch über den Nachweis des Verhältnisses einer ontologischen Begründung von Werten auf Emotionen. Weiterhin wird die prinzipielle Rationalität von Emotionen, bzw. deren kognitive Anteile, herauszustellen sein, da im Laufe der Argumentation eine Theorie moralischen Handelns entworfen werden soll, die einen viziösen Subjektivismus vermeidet. Kognitive Steuerbarkeit von Emotionen liefert hier also einen Rektifikationsmechanismus für deviante Emotionen, andererseits ist in der Forschungsdebatte mehrfach der Vorschlag gemacht worden, die Adäquanz oder Wohlbegründetheit von Emotionen über das Kriterium des kognitiven, intentionalen Gehalts von Emotionen zu definieren. Es wird also nötig sein, ein Modell der individuellen Aktualgenese von Emotionen zu entwickeln, wobei zur Klärung an geeigneter Stelle auch ein Blick auf die onto- und phylogenetische Entwicklung der menschlichen Emotionalität geworfen werden soll. Um die Brücke zu einer Konzeption der „emotionalen Einstellung“ zu schlagen, soll eine Typologie emotionaler Zustände geliefert werden, vor deren Hintergrund die erwähnte Konzeption zu situieren ist. In diesem Rahmen wird auch der Problemfall der Stimmungen zu diskutieren sein. Die Auswahl der Brennpunkte der Behandlung des Themas „Emotion“ in diesem Kapitel wird sich vollständig wohl erst in Retrospektive erhellen, da sie auf noch zu leistende Argumentationsschritte in der weiteren Folge dieser Arbeit hinführt. Anhand eines exemplarischen Durchgangs durch historisch markante Positionen in der Geschichte der Emotionsforschung sollen die angepeilten Themenkomplexe herausgegriffen werden.

2.2. Die Tradition des Kognitivismus

Schon früh in der Geschichte der Theoriebildung zum Phänomen „Emotion“ bildete sich eine Auffassung heraus, die nicht nur in krassem Gegensatz zum Alltagsverständnis stand, sondern auch in der Forschung bis in die heutige Zeit vielfach diskutiert werden sollte. Die Stoiker vertraten eine Theorieform, die später als eine sog. „Urteilstheorie“ (Rapp 2002) der Emotionen bekannt werden sollte.⁷ So soll Chrysipp „Trauer“ definiert haben als eine „frische Meinung, dass ein Übel anwesend sei [δόξαν πρόσφατον κακοῦ παρουσίας]“ (SVF 3, 463). Diese Position wurde auch noch in der jüngeren Stoa vertreten: wann immer wir durch ein Gefühl affiziert werden, sollen wir nach Epiktet „niemals einem anderen die Schuld daran geben [αἰτιώμεθα] als uns selbst, das heißt: unseren eigenen Auffassungen [τὰ ἐαυτῶν δόγματα]“ (Epict., enchir. 5, Übers. Hülser in LS 65U; vgl. Epict., diatr. I 12, 20f.). In gleicher Weise äußert sich Cicero: „Est igitur causa omnis in opinione, nec vero aegritudinis solum, sed etiam reliquarum omnium perturbationum [...]“ (Cicero, Tusc. III 11, 24). Mit diesen Zitaten sind in eindrucksvoller Weise schon zwei Kernthesen des Kognitivismus, spezieller eines „reinen“ oder „strengen“ Kognitivismus, illustriert: Emotionen werden *als* Meinungen, Gedanken oder Urteile aufgefasst, d.h. man könnte sie – im Sinne eines Art-Gattungs-Verhältnisses – als eine *Art* des Urteils verstehen (nach Chrysipp z.B.: ein *rasches* oder *frisches* Urteil). Das Haben von bestimmten Überzeugungen wäre damit also als Kausalfaktor in der Aktualgenese von Emotionen identifiziert.⁸ Und weiterhin wird eine Position bezogen in der Frage nach der Verantwortung für Emotionen, verstanden im Sinne der Zurechenbarkeit von Emotionen zu Subjekten als „Akteuren“ ihres emotionalen Erlebens. In den ausgewählten Passagen verwenden sowohl Cicero als auch Epiktet jeweils einen Begriff der „Ursache“ („αἰτία“, „causa“), um die Verbindung zwischen unseren Überzeugungen und unseren Emotionen zu erfassen. Da Verantwortung für vertretene Überzeugungen ein recht unkontroverser Begriff ist, können wir folgern, dass wir damit auch

⁷ Zur stoischen Affektenlehre vgl. allgemein: Pohlenz, Die Stoa, 4. Aufl. 1970, Bd. I, 141-153; Bd. II, 77-83, Lloyd (1978) und Bormann (1981).

⁸ Der Unterschied zwischen dieser Position und kognitivistischen Theorien im weiteren Sinne besteht gerade darin, dass letztere Kognitionen zwar als eine „Komponente“ oder einen „Aspekt“ von Emotionen ansehen, ihnen jedoch nicht zwangsläufig die Kausalwirkung in der Emotionsgenese zuschreiben. Auch die Identifizierung von Emotionen und Kognitionen findet sich dort i.d.R. nicht. Eine solche Form des Kognitivismus stellt eigentlich schon eine fruchtbare Amalgamierung der Thesen des reinen Kognitivismus mit anderen, noch zu behandelnden Emotionstheorien dar.

für die von uns erlebten Emotionen „verantwortlich“ sind. Bei Epiktet wird diese Schlussfolgerung ja auch explizit gezogen.

Gemäß der stoischen Philosophie handeln Sinnenwesen aufgrund von Impulsen, die durch die Zustimmung zu Vorstellungen entstehen. Diese sog. „Zustimmung“ („synkatathesis“, „adsensus“ bei Cicero, *Academica* 2, 145) ist eine Fähigkeit sowohl der Tiere als auch der Menschen; bei Menschen jedoch beruht sie auf der freien Entscheidung des „logos“, der Vernunft (SVF I, 61), anhand derer die Stoiker schließlich die menschliche Natur definierten. Ein *Affekt* („*pathê*“) entsteht, wenn der Handlungsimpuls nicht mehr unter der Kontrolle der Vernunft steht. Dadurch wird der Impuls unvernünftig und damit – nach Auffassung der Stoiker – unnatürlich (SVF III, 377f.; SVF I, 205). Vgl. dazu Cicero *Tusc.* 4, 11: „Est igitur Zenonis haec definitio ut perturbatio sit, quod pathos ille dicit, aversa a recta contra naturam animi commotio.“ Affekte, als Impulse, die dem Logos nicht gehorchen und sich seiner Kontrolle entziehen, sind „Krankheiten des Logos“ (SVF III, 444) bzw. „Geistesverwirrungen“ („*perturbationes animi*“, Cicero *Tusc.* 4).⁹ Es gibt 4 Hauptarten von Affekten: „*hêdonê*“, „*lypê*“, „*epithymia*“ und „*phobos*“ (weiterhin wird eine Unzahl von Unterarten unterschieden, vgl. dazu: SVF III, 397, 401, 409, 414). Freude und Trauer richten sich auf ein vermeintliches Gut oder Übel in der Gegenwart, Verlangen/ Mut und Furcht auf Gut oder Übel in der Zukunft (SVF III, 444, 447). An dieser Stelle können wir eine dritte Kernthese des Kognitivismus herausheben, nämlich die Auffassung, Emotionen seien zu definieren als typische Reaktionen auf eine je charakteristische Klasse von *intentionalen Objekten*. Damit werden jedoch, wie sich an späterer Stelle zeigen wird, emotionale

⁹ Gemäß der stoischen Affektenlehre müssen diese vier Affekte mitsamt ihrer verschiedenen Unterarten ausgerottet werden, weil sie das Denken verwirren. Sie sind nicht durch eine Zustimmung (*synkatathesis*) des Logos getragen. Allerdings predigen die Stoiker keine Gefühllosigkeit. Sie lassen für ihren idealtypischen Menschen – den „Weisen“ („*sophos*“) – die sog. „*eupatheiai*“, also „gute Affekte“ zu (vgl. SVF III, 431-442). Diese Klassifizierung der Emotionen in a-rationale Affekte („*pathê*“) und rationale Emotionen („*eupatheiai*“) muss aus der Sicht heutiger Emotionstheorien befremdlich erscheinen. Zunächst müsste man die von den Stoikern vertretene Lehre der Seelenteile bzw. Seelenvermögen anerkennen, um differenzieren zu können zwischen Zuständen, denen der Logos zustimmt bzw. nicht zustimmt. Weiterhin werden hier Affektenlehre und Güterlehre vermischt, sodass die Klassifizierung nicht mehr trägt, sobald die stoische Güterlehre aufgegeben wird. Denn „*chara*“ und „*hêdonê*“ z.B. unterscheiden sich vermutlich nicht in ihrer Qualität, den durch sie verursachten physiologischen Veränderungen im Subjekt, und – wenn wir von der stoischen Psychologie abstrahieren – wohl auch nicht hinsichtlich ihrer Genese: beide Emotionen werden durch Meinungen bzw. Urteile verursacht. Der Unterschied ist, dass sich „*hêdonê*“ auf „gleichgültige“ Güter („*adiaphora*“) bezieht und auf einer falschen Meinung beruht, während „*chara*“ sich auf das wahre Gut bezieht und auf einer wahren Meinung beruht.

Phänomene aus dem Beschreibungsbereich der Theorie ausgeschlossen, die zwar faktisch auftreten, jedoch nicht den logischen Kriterien des ihnen entsprechenden intentionalen Objekts genügen. „Angst“ vor Ereignissen in der Vergangenheit wäre im Rahmen einer solchen Theorie nicht zu erfassen, obwohl eine solche Emotion auftreten kann und im englischen Sprachraum sogar mit einem Begriff versehen ist („haunting fear“).

Weiterhin gingen die Stoiker davon aus, dass sich das Seelenpneuma ausdehne oder kontrahiere, je nachdem ob die Seele durch ein vermeintliches Gut oder Übel affiziert wird (SVF I, 209). Furcht ist demnach eine Kontraktion des Seelenpneumas, u.z. im wörtlichen Sinne: es verursacht Blässe, Zittern, Herzklopfen usw. Dies passt zur stoischen materialistischen Konzeption der Seele. Allerdings argumentieren die Stoiker, dass z.B. im Falle von Furcht, die Meinung, ein Übel stehe bevor, falsch sei. Gegenstand von Furcht ist nämlich gar nicht das Übel, sondern es sind die sog. „adiaphora“, die gleichgültigen Dinge, wie üble Nachrede, Schmerz, Tod, finanzieller Ruin. „Epithymia“ wiederum wird durch eine Ausdehnung des Seelenpneumas erklärt, die verursacht wird durch die falsche Meinung, etwas Zukünftiges, das in Wirklichkeit gleichgültig ist, sei für das Subjekt von großer Wichtigkeit und müsse darum erstrebt werden. Allerdings ist Trauer auch durch eine Kontraktion des Seelenpneumas gekennzeichnet, verursacht durch die falsche Meinung, etwas Schlechtes (das in Wirklichkeit gleichgültig ist) sei anwesend, und Freude schließlich kennzeichnet sich wieder durch eine Ausdehnung. Wir finden hier also eine über Kreuz laufende Parallele zwischen Furcht und Trauer, da beide sich durch eine Kontraktion des Seelenpneumas auszeichnen. Interessant ist an dieser materialistischen Konzeption zwar, dass die Stoiker ohne Probleme die physiologischen Aspekte von emotionalen Reaktionen erklären können (wie ich es am Beispiel der Furcht bereits angedeutet habe). Gerade dies bereitet zeitgenössischen (rein) kognitivistischen Theorien die größten Probleme. Aber es wird bei dieser Parallele deutlich, dass diesem Aggregatzustand der Seele (als physiologischer Komponente) keine zentrale Rolle bei der Emotionsgenese zukommt, da die jeweils näher bestimmte „Meinung“ des Subjekts als Kausalfaktor für eine emotionale Reaktion angenommen wird. Und da sich mindestens zwei Zustände durch den gleichen Aggregatzustand der Seele auszeichnen, kann dieser auch nicht zur trennscharfen Distinguierung von Emotionen dienen.

In weiten Teilen der neuzeitlichen Philosophie, über Descartes, Hume und Kant, wird die These vertreten, dass Emotionen Affekte oder „passiones“ seien, also nicht-rationale Bewusstseinszustände, die außerhalb unserer Kontrolle liegen und denen wir als Erleidende,

also passiv ausgeliefert sind. So schreibt Hume: „Our authority over our sentiments and passions is much weaker than that over our ideas; and even the latter authority is circumscribed within very narrow boundaries.“ (Hume Enquiry 1748, Abs. VII 1, Selby-Bigge S. 68). In unserer Alltagssprache sagen wir: „Wut steigt in mir auf“, „Traurigkeit überfällt mich“ etc. Mit den oben diskutierten Thesen bezog die stoische Emotionstheorie vermutlich auch schon in der Antike eine krass kontraintuitive Position. Der griechische Dichter Archilochos (7. Jhd. v.Chr.) weist in einem tetrametrischen Fragment (West I 105) als lyrisches Ich seinen Kameraden Glaukos darauf hin („Γλαῦχ’, ὄρα“), wie das Meer von Wellen aufgewühlt wird, und Wolken sich am Felsen von Gyrai zusammenballen: es ist ein Zeichen für Sturm („σῆμα χειμῶνος“). Die Natur also ist vorhersehbar, zumindest für den, der die Zeichen kompetent zu deuten weiß. Die Angst jedoch überfällt uns nach Archilochos unvorhergesehen („κίχάνει δ’ἔξ ἀελπίτης φόβος“). In dieser eindringlichen Illustration aus der antiken Dichtung finden wir eine *common-sense* Auffassung über Emotionen, die wohl für jedes Zeitalter universal gültig ist. Emotionen werden als „passiones“ verstanden, denen wir als Erleidende ausgeliefert sind, wir sind also nicht die „Akteure“ unseres emotionalen Lebens wie in der stoischen Theorie. Eng damit verbunden ist die Unvorhersagbarkeit, und damit wird das Phänomen „Emotion“ dem wissenschaftlichen Theoriediskurs insgesamt entzogen. Denn was sich nicht vorhersagen, d.h. reliabel mit Kausalfaktoren oder zumindest mit korrelierenden Variablen verbinden lässt, darüber lässt sich keine wissenschaftliche Aussage im Hinblick auf Genese, Verlauf und Kontext des Auftretens machen. Damit ist zwar – auf dem Hintergrund dieses Beispiels aus der Dichtung – noch kein streng logisches Argument, aber doch ein starkes Indiz verbunden, dass wir für unsere Emotionen auch nicht „verantwortlich“ sein können.¹⁰

In der philosophischen Literatur werden viele Spielarten der kognitivistischen Theoriegruppe vertreten, die ich im Folgenden aufschlüsseln werde (vgl. hierzu Nash 1989 und Rapp 2002, 559-570). Die Gemeinsamkeit aller kognitivistischen Emotionstheorien ist zunächst eine Gruppe von Gegnern, nämlich die verschiedenen Körper- oder Gefühlstheorien der Emotionen. Weiterhin besteht die Gemeinsamkeit darin anzunehmen, dass Emotionen einen intentionalen Gehalt haben, wie es z.B. Kenny deutlich ausdrückt: „Emotional attitudes,

¹⁰ Vor dem Hintergrund der hier illustrativ vorgestellten Argumente könnten wir Emotionen nicht willkürlich hervorrufen und ihr Eintreten nicht willkürlich verhindern (da wir ihre Entstehung nicht vorhersagen können). Es bleibt natürlich noch die Frage nach der Verantwortung durch die Möglichkeit zur Kontrolle (Unterdrückung, Sublimierung, etc.) vorhandener emotionaler Zustände („occurrent emotional states“).

like other mental attitudes, have formal objects.“ (Kenny 1963, 191). Dies kann dazu führen, dass unterschieden wird zwischen „Emotionen“ einerseits, die den Regeln eines formalen Objekts unterliegen (d.h. ich kann mich z.B. nicht vor einem Unglück *fürchten*, von dem ich weiß, dass es schon geschehen ist), und „Gefühlen“ und „Stimmungen“ andererseits, die eben nicht dieser Bedingung unterliegen und damit unter den traditionellen Begriff der „Affekte“ fallen dürften. Weiterhin lässt sich eine große Diskrepanz in der Verwendung des Kognitions-Begriffs diagnostizieren (vgl. Solomon 2003; eine ähnliche Situation wird sich in der Folge auch in der psychologischen Debatte zeigen). Diese Vieldeutigkeit im wissenschaftsdiskursiven Gebrauch des Terminus „Kognition“ erschwert die genaue Beantwortung der Frage, was unter einer „kognitivistischen“ Emotionstheorie zu verstehen sei. Drei grundlegende Positionen lassen sich unterscheiden:

- (1.) Emotionen sind „Kognitionen“ im Sinne von Glaubenszuständen (vgl. bspw.: Murphy 1988 und 2003 und Walton 1993).
- (2.) Emotionen sind „Kognitionen“ i.S.v. Gedanken (vgl. bspw. Neu 1978); diese Theorie geht nominell zurück auf Descartes und Spinoza.
- (3.) Emotionen sind „Kognitionen“ i.S.v. (Wert-)Urteilen („[evaluative] judgments“, vgl. bspw. Solomon 1973 u.f.; Calhoun 1984 und Roberts 1996); diese Theorie geht zurück auf die Stoiker

Auch in der heutigen Zeit finden wir noch Vertreter einer quasi-stoischen Emotionstheorie, d.h. einer „rein kognitivistischen“ oder „Urteilstheorie“ der Emotionen, die Emotion lediglich für eine Art von Urteilen halten.¹¹ Solomon etwa hält Emotionen für „relatively intense evaluative judgments“ (Solomon 1976, 187, vgl. auch Solomon 1973). Dieser rein kognitivistische Theorietypus wurde gelegentlich auch in der zeitgenössischen Psychologie vertreten (Beck 1979, Ellis 1982, Neu 1978), allerdings eher von Seiten der psychotherapeutischen als der empirischen psychologischen Forschung. Eine eindeutige Identifikation von Emotionen und Überzeugungen (bzw. Urteilen/ Meinungen/ Gedanken) scheint jedoch nicht plausibel. Zwar spricht für die kognitivistische These, dass beispielsweise sog. pragmatische Paradoxa auch im Bereich der Emotionen Geltung haben (vgl. Calhoun & Solomon 1984, 22). Die Aussage „Ich *bin wütend* auf Hans, weil er mein Stück Torte

¹¹ In der anglophonen Debatte spricht man mit Bezug auf „Urteilstheorien“ auch von „judgmentalism“. Der mittlerweile weithin gebräuchliche Begriff wurde von Gibbard 1990, 128-132 geprägt. Ebenfalls im Umlauf ist die Bezeichnung „propositional attitude school“, die sich in ähnlicher Weise auf einen strengen Kognitivismus bezieht.

gegessen hat, aber ich *glaube* nicht, dass Hans mein Stück Torte gegessen hat.“ ist paradoxal. Es ist nicht nur unsinnig, sie zu äußern, sie ist vielmehr sogar performativ-therapeutisch, weil sie im Moment der Äußerung den emotionalen Zustand und damit auch das vermeintliche Paradoxon aufhebt.¹² Bei vielen Emotionen (sicherlich jedoch nicht bei allen, z.B. phobische Angst) können wir von einer prinzipiellen Kognitionsabhängigkeit bzw. einer kognitiven Beeinflussungsmöglichkeit ausgehen. Es scheint jedoch kein strenges Art-Gattungs-Verhältnis zwischen Emotionen und Urteilen vorzuliegen. Zwar versuchen Theoretiker auf diesem Gebiet, die Menge der in Frage kommenden Urteile oder Überzeugungen näher zu qualifizieren, indem sie von *Werturteilen*, oder noch enger von „relativ *intensiven*“ Werturteilen (vgl. o.) sprechen. Doch auch dies kann nicht überzeugen. Einerseits wird das gesamte pragmatische Gefüge solcher Überzeugungen und Aussagen, von Sprecher über Sprechsituation bis zum Adressaten, zu berücksichtigen sein. Das Urteil „Susanne handelt unmoralisch, sie hintergeht ihren Ehemann“ wird wohl nur dann zu einer starken Emotion führen, wenn gerade *ich* mit Susanne verheiratet bin (mich also in der Rolle des aussagenden Subjekts befinde). Abgesehen von der Vernachlässigung der pragmatischen Redesituation jedoch kann selbst die Einschränkung der emotionsrelevanten Urteile auf „intensive Werturteile“ nicht überzeugen. Warum reagieren wir i.d.R. nicht emotional auf schwer besorgniserregende Meldungen in der Zeitung, ärgern uns aber u.U. darüber, beim Einkauf wieder die Milch vergessen zu haben?

¹² Für eine facettenreichere Deutung dieses Phänomens vgl. DeSousa 1979, 139ff.: Stellen wir uns einen Fall vor, in dem Hans eine Emotion auf der Grundlage einer Überzeugung darbietet, daraufhin aber durch schlüssige Beweise zum Aufgeben seiner Überzeugung bewogen wird. DeSousa unterscheidet drei mögliche Konsequenzen: (1.) „In some cases, the discovery that an emotion was premised on a false belief, instead of simply canceling the emotion, will transmute it into another.“ So kann sich Hans' Empörung in Bedauern oder sogar Scham verwandeln. DeSousa betont jedoch, dass sich dabei nicht der „Fokus der Aufmerksamkeit“ verändert. Diesen Effekt bezeichnet er als „inertia of attention“. Durch dasselbe Aufmerksamkeitsmuster werden die neuen Tatsachen nun als salient markiert, und dementsprechend findet eine Verlagerung des „pattern of concern“ statt. (2.) „In other cases, the emotion's habit of looking for certain sorts of facts and facilitating certain sorts of inference will simply look for more“. Hans kann zwar überzeugt werden, dass seine Empörung gegenüber Hermann nicht gerechtfertigt war, trotzdem aber denken, dass Hermann wohl irgendetwas anderes „auf dem Kerbholz“ haben wird. DeSousa sagt, dass an diesem Beispiel deutlich wird, dass Emotionen eine Disposition darstellen, bestimmte Fragen zu stellen. (3.) Emotionen können sich verändern (bei gleichem Fokus), ohne dass Überzeugungen geändert oder aufgegeben werden. DeSousa liefert ein Beispiel, das auf die Relevanz von interpersonalen Verhältnissen abhebt: ein Professor, der sich von einer vermeintlich fremden Frau eingeschüchtert fühlt, erlebt das Abklingen dieses Gefühls, als er erfährt, dass die Frau tatsächlich einst seine Studentin gewesen war.

Ein weiteres wichtiges Argument ist, dass das Element der *Affektivität* sowohl empirisch als auch konzeptionell notwendig sein dürfte, um emotionale Zustände von Glaubenszuständen (wie z.B. „Schnee ist weiß.“, „1+1=2“) zu unterscheiden. Auch hier ergeben sich Unterschiede der Theorieformen, die oft nicht explizit gemacht werden. Eine kognitivistische Theorie, deren „Kognitivismus“ lediglich in der Annahme eines intentionalen Gehalts von Emotionen besteht, und die damit eine Art Komponententheorie der Emotionen darstellt, wird kein Problem mit der theoretischen Implementierung einer affektiven Komponente haben (so jüngst v.a. Gessner 2004). Für Urteilstheorien stellt dies jedoch eine enorme Schwierigkeit dar. Robert C. Solomon, der prominenteste zeitgenössische Vertreter einer solchen Theorieform hat sich in einem rezenten Aufsatz dieser Herausforderung gestellt (Solomon 2003). Er scheint zunächst erhebliche Zugeständnisse an die konkurrierenden Theorien zu machen, wenn er schreibt: „Belief isn't the right sort of psychological entity to *constitute* emotion. Beliefs are necessarily dispositions, but an emotion is, at least in part, an *experience*. A belief as such is not ever experienced.“ (ebd., 4). Weiterhin räumt er ein, dass Emotionen dynamisch seien, während Glaubenszustände stabil sind (zumindest bis zu ihrer Widerrufung). Darum könnten Glaubenszustände auch keine Konstituenten bzw. Komponenten von Emotionen sein, sondern nur deren „kognitive Vorbedingungen“ („cognitive preconditions“). Dennoch hält er später – im Rahmen der oben wiedergegebenen Überblicksdarstellung – an seiner Position fest, Emotionen mit Werturteilen zu identifizieren.

Solomon räumt ein, dass er in seinen früheren Arbeiten die Rolle des Körpers und dessen Affektivität vernachlässigt hat. Er ist nun überzeugt, dass auch „kinästhetische Gefühle“ notwendig zu Emotionen gehören, u.z. nicht nur als Funktion oder *Effekt* einer Emotion (ich verziehe mein Gesicht, *weil* ich mich ekle), sondern als integraler Teil der *Erfahrung* einer Emotion. Offenbar bezieht er sich damit implizit auch auf die Forschung zur „Facial Feedback“ Hypothese von Ekman (1977 u.a.) und Izard (1990; vgl. auch Buck 1980, Laird 1984, Strack/ Martin/ Stepper 1988, Zajonc/ Murphy/ Inglehart 1989, Cacioppo 1992). Demnach könne dann ein großer Teil dessen, das als Affektivität bezeichnet wird, eher dem Körper zugeschrieben werden, u.z. als „judgments of the body“ (Solomon 2003, 14), die sich in, zu solchen kinästhetischen Gefühlen führenden, körperlichen „Mikro-Praktiken“ ausdrücken und kognitiv in einem vor-sprachlichen (jedoch nicht vor-bewussten) Sinne sind.

Diese Körperurteile will Solomon dennoch als vollwertige Urteile auffassen. Erstens müssen seines Erachtens Urteile nicht immer propositional artikuliert sein, und zweitens besitzen auch Menschen einen großen Anteil an „Knowing how“-Wissen gegenüber einem kleinen Anteil an „Knowing that“ (also Handlungswissen gegenüber Faktenwissen). Die

Voreingenommenheit mit dem Faktenwissen in der Philosophie beruht auf Gründen, über die nur spekuliert werden kann. Solomon denkt, dass „knowing how“ missachtet wird, weil es nicht charakteristisch ist für die menschliche Rationalität, sondern auch bei Tieren vorkommt. Und weiterhin kann „knowing-how“-Wissen nicht auf eine bestimmte Menge von „knowing-that“-Propositionen zurückgeführt werden. Das ist eine weitere Schwierigkeit für die philosophische Untersuchung von Handlungswissen. Es wäre jedoch irreführend, wenn man auf dieser Grundlage behaupten wollte, dass Faktenwissen bzw. propositionales Wissen in irgendeiner Weise erstrangig oder unabhängig von Handlungswissen sei. Ganz im Gegenteil scheint nach Solomon gelegentlich die Art, wie wir durch Handlungswissen, Gewohnheiten und Praktiken auf die Welt reagieren, bzw. der Welt Bedeutung und Sinn verleihen, eine viel kompetentere Reaktion auf die Welt zu sein, als die Art, wie wir die Welt faktisch beschreiben würden bzw. über sie denken.

Nach Solomon kann also eine kognitivistische Emotionstheorie sehr wohl auch die Aspekte der Affektivität und Körperlichkeit erklären, wenn man „Kognition“ als „Werturteil“ versteht, das nicht nur propositionales Wissen, sondern auch Handlungswissen (wie Körperpraktiken und –reaktionen) mit einbezieht. Solomon scheint ein persönliches Amalgam aus seiner Urteilstheorie einerseits und den konkurrierenden psycho-physiologischen Theorien andererseits (die hier an späterer Stelle behandelt werden sollen) herstellen zu wollen. Dabei bleibt jedoch offen, wie Emotionen gerade durch „Handlungswissen“ erklärt werden sollen. Außerdem fällt eine recht typische Strategie ins Auge, den Begriff des Urteils zu erweitern, bis er die notwendige explanatorische Kraft erlangt hat, und wenn dazu auch idiosynkratische Elemente miteinander verbunden werden müssen.¹³ Solange jedenfalls

¹³ Ähnlich – jedoch in weitaus plausiblerer Weise – verfährt Martha Nussbaum in ihrer eindrucksvollen Studie „Upheavals of Thought“. Nussbaum bezeichnet ihre Theorie selbst als „neo-stoisch“, Emotionen sind ihrer Ansicht nach Urteile, u.z. eine Unterklasse von Werturteilen (1996, 30). Die von ihr verfolgte Strategie besteht in der theoretischen Neudeutung von zwei für den strengen Kognitivismus zentralen Begriffen: Nussbaum hält zwar (1.) an der „Intentionalität“ von Emotionen fest, definiert diesen Begriff aber recht weit im Sinne eines „way of seeing“ (27) oder eines „seeing-as“ (129). Sie ist weiterhin der Auffassung, dass (2.) nicht alle Emotionen einen linguistisch strukturierten und sprachlich artikulierbaren Inhalt haben (vgl. 79, 127, 263f.). Damit wird von ihr der Urteilsbegriff auch auf nicht-propositionale Phänomene ausgedehnt. In beiden Fällen (1 und 2) besteht die Strategie also darin, durch definatorische Erweiterung von Begriffen für den Kognitivismus problematische Fälle in die Theorie adäquat aufnehmen zu können. Anders als Solomon jedoch besteht der Impetus von Nussbaums Deutung nicht in dem Desiderat, körperliche Prozesse der emotionalen Dynamik mit einer kognitivistischen Theorie zu vereinbaren. Nach ihrer Auffassung sind körperliche Veränderungen nicht notwendiger Bestandteil von Emotionen. Das Paradigma der körperlichen Veränderungen beschreibt nach Nussbaum die Lust-Unlust-Dichotomie, die Emotionen zu eigen ist. Allerdings, so Nussbaum, hängt diese nicht

Solomons Projekt auf der Stufe der bisher vorgelegten Skizze verbleibt, kann nicht von einer überzeugenden Meisterung dieser Herausforderung gesprochen werden.

An dieser Stelle könnte man spekulieren, ob Affektivität überhaupt bei einem körperlosen Wesen auftreten könnte, ob es sich also um eine rein geistige Leistung handelt. Auch Solomon scheint in den oben referierten Ausführungen diese Annahme ohne weitere Begründung vorauszusetzen. Zunächst wirkt diese Pointe banal, da Leben, wie wir es kennen, ohne Körper überhaupt nicht möglich ist. In diesem Sinne sind auch alle geistigen Leistungen (Urteilen, Schlussfolgern etc.) ontologisch abhängig von einem körperlichen Substrat. Die spekulative Annahme ist jedoch differenzierter zu verstehen: bestimmte rein intellektuelle Denkprozesse, wie z.B. das Lösen einer Mathematik-Aufgabe, sind wahrscheinlich nur von ihrer neuronalen Grundlage abhängig, d.h. von dem Vorhandensein bestimmter neuraler Schaltkreise und von der Fähigkeit unseres Gehirns, bestimmte neurale Prozesse ablaufen zu lassen. Affekt, so könnte man jedoch spekulieren, bedarf darüber hinaus noch der Mitwirkung der Körperperipherie, ganz ähnlich wie etwa ein bestimmter Spiegel der Nährstoffkonzentration im Blut autonom-nervöse Prozesse auslöst, die wir als *Hungergefühl* erleben. Affekt könnte also als eine Interaktion zwischen Gehirn und Körperperipherie verstanden werden. Neben dem Beispiel des Hungergefühls gibt es noch viele weitere nicht-emotionale Gefühle: Durst, Schmerz, Müdigkeit, Mattigkeit etc.¹⁴ Einerseits kann an dieser Stelle die Frage nach der „mental causation“ gestellt werden: wie kann – einer kognitivistischen Emotionstheorie folgend – ein mentaler Zustand (ein Werturteil) ein solches emotionales Körpergefühl auslösen? Diese metaphysische Frage soll jedoch hier ausgeblendet werden. Im Rahmen der stoischen Philosophie wäre eine Lösung problemlos möglich (s.o.), im Rahmen der zeitgenössischen kognitivistischen Emotionstheorie wird diese Frage jedoch überhaupt nicht thematisiert. Andererseits können wir weniger pointiert fragen, ob eine kognitivistische Emotionstheorie adäquat – im Sinne ihrer explanatorischen Kraft und ihres theoretischen Beschreibungsbereichs – Auskunft über Affektivität bei Emotionen geben kann, besonders wenn wir die – an dieser Stelle noch spekulative – Annahme der affekttheoretischen Relevanz der Körperperipherie ernst nehmen wollen.

mit einem physiologisch messbaren Zustand zusammen: wer Freude empfindet, zeigt kein bestimmtes Muster physiologischer Aktivierung, sondern tritt in einen bestimmten *Modus* des Handelns und Denkens ein (63).

¹⁴ An späterer Stelle wird in differenzierterer Weise zu klären sein, ob es sich tatsächlich um strikt nicht-emotionale Gefühle handelt, oder ob nicht emotionaler Affekt zum Erleben dieser Körpergefühle beiträgt.

2.3. *Psycho-physische Emotionstheorien*

Die bis weit in das 20. Jahrhundert verbindliche Theorie der Emotionen wurde von William James und Carl Lange unabhängig voneinander im Jahr 1884 bzw. 1885 formuliert (James 1884, Lange 1885). Diese James-Lange-Theorie definiert Emotion als das Erleben einer peripher-physiologischen Veränderung. Ich registriere eine solche peripher-physiologische Veränderung, wie z.B. gesteigerte Transpiration, Herz- und Atemfrequenz, und das Erleben dieser Veränderung *ist* meine Emotion (z.B. der Furcht). Im Folgenden sollen einige Ausführungen von William James näher verfolgt werden.¹⁵ Dass die Wahl auf James, an Statt auf Lange gefallen ist, hat zwar willkürliche Gründe, diese jedoch nehmen keinen Einfluss Art oder Ergebnis der daraus zu entwickelnden psycho-physischen Emotionstheorie nach James-Lange. In der Tat ist die Ähnlichkeit zwischen beiden Theoretikern bei näherem Hinsehen so verblüffend stark,¹⁶ dass sowohl der eine als auch der andere problemlos als Statthalter für die mit gemeinsamen Namen bezeichnete Theorie dienen könnte. James bietet sich letztlich aus praktischen Gründen an, da mit ihm ein Autor vorgestellt wird, der eine breite Rezeption in der Geschichte der Philosophie und Psychologie erfahren hat.

Die James'sche Emotionstheorie, wie sie sich uns heute darstellt, lässt sich in zwei Fassungen aufteilen (vgl. Kochinka 2004, 214; Meyer/ Schützwohl/ Reisenzein 2001, 92): die erste, die gleichzeitig die vorrangig rezipierte Fassung darstellt, geht zurück auf den berühmten Aufsatz in Mind „What is an emotion?“ (James 1884), und findet sich in ähnlicher Form auch in den Principles of Psychology (James 1950) oder anderen Sammelausgaben (z.B. James 1984). Die zweite Fassung findet sich im Aufsatz „The physical basis of emotion“ (James 1894), wo der Autor schon auf einige Einwände reagiert und seine Theorie überarbeitet darstellt. Diese überarbeitete Fassung wurde jedoch im Vergleich zur ersten kaum rezipiert.

¹⁵ Gute Überblicksdarstellungen zur Emotionstheorie von William James (und damit zur James-Lange-Theorie) bieten Kochinka 2004, 213-260, Meyer/ Schützwohl/ Reisenzein 2001, 90ff. und Barbalet 1999.

¹⁶ Beide Theorien sind erstaunlich deckungsgleich, weshalb im Allgemeinen von der James-Lange-Theorie gesprochen wird (vgl. Kochinka 2004). Ein (unwesentlicher) Unterschied besteht darin, dass Lange (1885) die Kategorie der peripher-physiologischen Veränderungen als Veränderungen in drei „Muskelsystemen“ fasste: (1.) Veränderungen des Muskeltonus, was Lange als „latente Innervation“ bezeichnete. (2.) Veränderungen in den Muskeln der Eingeweide („viszera“) und (3.) Kontraktion oder Dilatation der Gefäßmuskulatur. In letzterem erblickt er den primären Faktor in der Emotionsgenese: „die vasomotorischen Störungen, die Veränderungen der Weite der Blutgefäße und damit der Blutmenge in den einzelnen Organen, [sind] die eigentlichen, primären Elemente der Affekte [], während die übrigen Erscheinungen [...] nur sekundäre Effekte sind, die ihre Ursache in Variationen der Gefäßinnervation haben“ (Lange 1885, 43).

Die These von James, die in pointierter Weise unser Alltagsverständnis vom Zusammenhang zwischen emotionales Erleben und physiologischer Veränderung umdreht, bezieht dieser selbst zunächst auf primäre („gröbere“: „coarser emotions“) Emotionen wie „grief, fear, rage, love, in which every one recognizes a strong organic reverberation“ (James 1950 II, 449):

„Our natural way of thinking about these coarser emotions is that the mental perception of some fact excites the mental affection called emotion, and that this latter state of mind gives rise to the bodily expression. My theory, on the contrary, is that *the bodily changes follow directly the perception of the exciting fact, and that our feeling of the same changes as they occur IS the emotion* [...] and that the more rational statement is that we feel sorry because we cry, angry because we strike, afraid because we tremble, and not that we cry, strike, or tremble, because we are sorry, angry, or fearful, as the case may be.“ (James 1950 II, 449)

James versteht hier implizit emotionale Reaktionen als Analogon zu Instinktreaktionen, was er an früherer Stelle in den Principles auch erwähnt: „*Instinctive reactions and emotional expressions [] shade imperceptibly into each other. Every object that excites an instinct excites an emotion as well*“ (James 1950 II, 442). Instinktreaktionen wiederum sind nach James angeborene, spontan ablaufende, adaptive Reaktionsmuster auf bestimmte Objekte, die in der phylogenetischen Geschichte ein bestimmtes Reiz-Reaktions-Muster in der Gattung Mensch niedergelegt haben. „Every living creature is in fact a sort of lock, whose wards and springs presuppose special forms of key, – which keys however are not born attached to the locks, but are sure to be found in the world near by as life goes on“ (James 1884, 191). Instinkte – und wir können daher ergänzen: auch Emotionen – sind also biologisch adaptive Mechanismen, die sich auf bestimmte relevante Objekte der Umwelt beziehen. „Now among these nervous anticipations are of course to be reckoned the emotions, so far as these may be called forth directly by the perception of certain facts. In advance of all experience of elephants no child can but be frightened if he suddenly find one trumpeting and charging upon him. No woman can see a handsome little naked baby without delight, no man in the wilderness see a human form in the distance without excitement and curiosity“ (ebd.). Die Relevanz kann hier als ein Fördern oder Hemmen des Selektionsvorteils gesehen werden. Dabei betont James, dass diese Objekte nicht mehr unbedingt in der heutigen Lebenswelt anzutreffen sein mögen. Stellen wir uns vor, dass Hans nie zuvor eine Schlange gesehen hat, dann jedoch, bei einer Reise nach Australien, mit panischer Angst auf eine solche reagiert. Den „Schlüssel“ (die Schlange) hatte Hans zwar noch nie angetroffen, aber als dies in unserem Beispiel der Fall war, passte er auf ein entsprechend codiertes „Schloss“ (dem Reiz-Reaktions-Schema der Emotion „Furcht“). Man kann daher einräumen, dass sich die

Adaptivität von Instinkten und Emotionen evtl. auf eine Lebenswelt bezieht, die wir als heute lebende Menschen nicht mehr vorfinden.

Nach James Vorstellung löst also die primitive Sinneswahrnehmung eines relevanten Reizes eine emotionale Reaktion aus, ganz so wie die Wahrnehmung eines schnell heranfliegenden Objekts instinktiv eine körperliche Schutz- oder Abwehrreaktion auslöst. Dabei ist mit „emotionaler Reaktion“ zunächst einmal die physiologische Aktivierung gemeint, die James als die grundlegende Phase in der Emotionsgenese annimmt: nämlich typischerweise Reaktionen, die mit Funktionen des vegetativen Nervensystems korrelieren und z.B. Durchblutung, Herz- und Atemfrequenz, Transpiration, Peristaltik des gastrointestinalen Trakts, Muskeltonus, endokrine Prozesse usw. betreffen. (vgl. James 1884, 191f.; 1950 II, 447 u. 482). Hinzuzufügen ist, dass James weiterhin postuliert, dass eine solche peripher-physiologische Veränderung aktual auch bewusst empfunden wird, damit ein emotionales Erleben (also ein Gefühl) entsteht (James 1884, 192; 1950 II, 450f.).

James' Theorie erfasst in prägnanter Weise die intuitiv von uns unterstellte Tatsache, dass Emotionen mit solchen peripher-physiologischen Veränderungen korrelieren. Kochinka konstatiert, dass Beispiele dafür

„in reicher Auswahl in der Alltagssprache abgebildet sind: Da kann einem etwa (aus diesem oder jenem Gefühl) die Luft weg- oder das Herz stehen bleiben, oder jemand schnappt nach Luft, vielleicht weil sein Herz rast oder bis zum Halse schlägt. Manchmal wird einem schwindlig, er wird schwach, vielleicht errötet er oder wird blass, mitunter wird ihm dabei heiß und kalt, womöglich, wenn er zu derberen Formulierungen neigt, könnte er gar kotzen. Zum Glück zeigt sich der Magen nicht immer so feindselig, auch wenn man, ehe man sich über die Schmetterlinge im Bauch freuen darf, nicht selten mit dem Kloß im Hals fertigwerden muss. Manchmal bleibt einem auch einfach die Spucke weg oder man bricht in Schweiß aus, wenn nicht in Tränen“ (Kochinka 2004, 218).

Die zentrale Frage ist natürlich, worin genau diese Korrelation besteht: unser Alltagsverständnis will, dass die Wahrnehmung solcher Reaktionen *Begleiterscheinungen* unserer Emotionen sind, James jedoch postuliert, dass sie gerade unsere Emotionen *sind*.

Erinnern wir uns an das eindringliche Gedankenexperiment, mit dem James versuchte zu illustrieren, dass Emotionen eben nicht nur durch körperliche Prozesse begleitet werden, sondern dass allein diese Prozesse eine Emotion überhaupt zu einer Emotion machen, dass sie also ein notwendiges Kriterium dafür liefern, etwas als Emotion zu klassifizieren:

„If we fancy some strong emotion and then try to abstract from our consciousness of it all the feelings of its bodily symptoms, we find we have nothing left behind, no ‘mind-stuff’ out of which the emotion can be constituted, and that a cold and neutral state of intellectual perception is all that remains. [...] What kind of an emotion of fear would be left if the feeling neither of quickened heart-beats nor of shallow breathing, neither of trembling lips nor of weakened limbs, neither of goose-flesh nor of visceral

stirrings, were present, it is quite impossible for me to think. Can one fancy the state of rage and picture no ebullition in the chest, no flushing of the face, no dilatation of the nostrils, no clenching of the teeth, no impulse to vigorous action, but in their stead limp muscles, calm breathing, and a placid face?“ (James 1884, 193)

Zur Methodologie dieses Gedankenexperiments bemerkt Kochinka:

„Das Unvorstellbare muss nicht inexistent sein, es kann ja auch schlicht unvorstellbar sein. – Was bleibt bspw. von einem Auto, wenn wir uns jede mögliche Farbe, die es haben kann, wegdenken? Ein Auto, das dem Auge nicht gelb oder grün, nicht blau oder rot, nicht eisenfarben (wie unlackiert) noch Perleffekt-Sonstwas erscheint, ein Auto, das überhaupt keine Farbe hat, wenn wir es in die Sonne stellen, können wir uns nicht vorstellen. Aber das heißt doch noch lange nicht, dass ein Auto *nichts anderes ist* als sein Farbe“ (Kochinka 2004, 221, Hervorh. v. mir., S.H.)

James' Argument, dass bei Abstraktion von jeglicher Empfindung körperlicher Aktivierung nur ein kalter Zustand intellektuellen Begreifens übrigbleibe, der nicht Emotion heißen könne, ist sicher richtig (zumindest bezogen auf starke emotionale Reaktionen, wie James selbst einschränkend bemerkt). Was Kochinka hier meint, ist jedoch, dass dieses Argument in keiner Weise die *Identität* von Körperempfindung und Emotion beweist. Auf logisch-metaphysische Weise illustriert besagt der Einwand: stellen wir uns die Art „Mensch“ und die artbildende Eigenschaft (*differentia specifica*) „Vernunftbegabung“ vor. Wir können uns kein Partikulare der Art „Mensch“ ohne prinzipielle Vernunftbegabung vorstellen; ein solches Individuum würde nicht in diese Art gehören.¹⁷ Dennoch bedeutet das nicht, dass der Mensch Vernunftbegabung „ist“, d.h. dass „Mensch“ und „Vernunftbegabung“ identisch sind, jedenfalls nicht im intensionalen Sinne.

James schlägt nun zwei Verfahren vor, wie sich seine Hypothese überprüfen lassen soll: (1.) müsste bei Simulation bestimmter körperlicher Veränderungen – auch bei Abwesenheit eines emotional relevanten Reizes – ein entsprechendes emotionales Gefühl erzeugt werden können, da nach James ja dieses emotionale Erleben nichts anderes ist als die Wahrnehmung bestimmter musterhafter Körperveränderung, und eben nicht eine Wahrnehmung von, Reaktion auf, oder gar Stellungnahme zum emotionalen Reiz selbst ist. Dieser bleibt unbewusst und spontan nur als Auslöser einer instinkthaften Körperveränderung verstanden. Und (2.) müssten Personen, deren Propriozeption in spezifischer Weise ausgeschaltet oder anästhesiert ist, d.h. Personen, die emotionstypische peripher-physiologische Veränderungen nicht wahrnehmen können, auch kein emotionales Erleben haben, also keine entsprechenden Gefühle.

¹⁷ Dieses Argument ist nur im Sinne einer logisch-metaphysischen Illustration gemeint und enthält keinerlei ethische oder anthropologische Implikationen.

Zu (1) konstatiert James, dass durch die Erfahrung innerhalb der Grenzen, innerhalb deren eine Verifikation möglich ist, diese Schlussfolgerung eher bestätigt als widerlegt werde. James nennt in diesem Zuge das Beispiel von Schauspielern, die bei authentischer Nachahmung bestimmter Gefühlsausdrücke angeben, auch die entsprechenden Gefühle zu erleben (James 1950 II, 464). Allerdings betont er die Grenzen der Verifizierbarkeit zu Recht, da die meisten vegetativen Prozesse bekanntlich nicht, oder nur in sehr eingeschränktem Maße, unter unseren bewussten Verfügbarkeit stehen.

Zu (2) bemerkt James: „A positive proof of the theory would, on the other hand, be given if we could find a subject absolutely anaesthetic inside and out, but not paralytic, so that emotion-inspiring objects might evoke the usual bodily expressions from him, but who, on being consulted, should say that no subjective emotional affection was felt. Such a man would be like one who, because he eats, appears to bystanders to be hungry, but who afterwards confesses that he had no appetite at all. Cases like this are extremely hard to find“ (1950 II, 455). Oder in James 1884, 203: „A case of complete internal and external corporeal anaesthesia, without motor alteration or alteration of intelligence except emotional apathy, would afford, if not a crucial test, at least a strong presumption, in favour of the truth of the view we have set forth“. Die grundsätzliche Schwierigkeit in der empirischen Nachprüfbarkeit, die James zu Recht sieht, ist, dass ein Proband mit einem Syndrom völliger „interner Deafferenz“ (vgl. Kochinka 2004, 224), d.h. einer Unterbrechung der nervösen Leitung von Informationen aus dem vegetativen Nervensystem und den Skelettmuskeln zum Gehirn, herangezogen werden müsste, der aber – wie James sagt – nicht paralytisch sein darf. Denn einerseits reagieren wir nach James instinkthaft mit spezifischen körperlichen Veränderungen auf emotionsrelevante Reize, und diese Reaktivität dürfte natürlich nicht von vorneherein ausgeschaltet sein, wollte man die Theorie empirisch nachprüfen. Der Proband müsste also noch in der Lage sein, bestimmte externe Reize zu verarbeiten: optisch, auditiv etc., aber auch barisch-haptisch über die Haut. Es erscheint unwahrscheinlich, ein solches intern deafferenziertes Muster vorzufinden bei gleichzeitiger Intaktheit des Sinneswahrnehmungsapparats einschließlich der Haut- und Muskelpropriozeption. Andererseits, und noch wichtiger, beschränkt James schließlich seinen Begriff von Emotion nicht nur auf die Wahrnehmung von vegetativ-physiologischen Körperveränderungen, sondern schließt auch muskuläre Reaktionen, ja ganze Handlungen als solche instinktiven Reaktionen auf emotionsrelevante Reize mit ein. Im Wortlaut heißt es schließlich: „we feel sorry because we cry, angry because we strike, afraid because we tremble“ (vgl. oben). Wahrscheinlich wäre solch differenzierte Reaktionen überhaupt nicht möglich bei einem

Probanden, bei dem wir das Vorliegen einer solchen völligen internen Deafferenz im Gedankenexperiment unterstellen würden. James selbst berichtet (in 1884, 204f.; 1950 II, 455) von drei Fällen, in denen angeblich Probanden mit Symptomen untersucht wurden, die der von ihm postulierten internen Deafferenz zumindest nahe kamen. James selbst konstatiert jedoch, dass die Ergebnisse entweder nicht gesichert, oder aber widersprüchlich seien. Später (1894, 527ff.) führt er dann drei weitere solcher Fälle an, die angeblich seine Hypothese belegen. Dabei ist jedoch zu beachten, dass nur einer der Probanden entsprechende Läsionen der Nervenbahnen aufwies, während den beiden anderen lediglich durch einen gew. Dr. Sollier eine interne Empfindungslosigkeit hypnotisch suggeriert wurde.

Schon 1927 wurden von Walter Cannon wichtige Argumente und Ergebnisse gegen diese Theorie vorgebracht (Cannon 1927). Auch danach wurde immer wieder kritisch auf die James-Lange-Theorie Stellung genommen, teils um sie als Gegner aufzubauen, teils um sie als Inspiration und Vorläufer eigener Überlegungen darzustellen. Ein schon zu Lebzeiten James' viel diskutiertes Beispiel betrifft die prinzipielle Offenheit von emotionsrelevanten Stimuli für ambivalente Interpretationen. James wurde schon im Rahmen persönlicher Kritik auf seine Theorie darauf hingewiesen, dass sein Modell nicht erklären könne, warum wir z.B. in der Wildnis bei der Begegnung mit einem Bären intensive Angst, begleitet mit starken peripher-viszerale Empfindungen verspüren. Andererseits können wir im Zoo relativ gleichmütig einem Bären im Käfig gegenüberstehen. Der Einwand trifft natürlich einen zentralen Schwachpunkt in James' Theorie, der ja bekanntlich emotionale Reaktionen in Analogie zu Instinktreaktionen verstand, wodurch zunächst einmal die aktive Interpretation eines Reizes durch das Subjekt ausgeschlossen wäre.

Das Beispiel mit dem Bären wurde schon von einem Zeitgenossen James' benutzt, um eben diesen Einwand zu formulieren, dass eine „bloße Wahrnehmung“ durchaus unterschiedliche emotionale Reaktionen bzw. das gänzliche Ausbleiben einer solchen zur Folge haben kann: „A chained or caged bear may excite only feelings of curiosity, and a well armed hunter might experience only pleasurable feelings at meeting one loose in the woods. It is not, then, the perception of the bear that excites the movements of fear“ (Worcester 1883, 287). Damit ist also mit kritischem Bezug auf James nahegelegt, dass anstelle einer bloßen

Wahrnehmung sinnvoller eine in irgendeiner Form *interpretierende* Wahrnehmung an die Stelle des Emotionsauslösers zu setzen wäre.¹⁸

Im Rahmen der eingangs erwähnten „überarbeiteten“ Fassung seiner Theorie, die sich im Aufsatz „The physical basis of emotion“ findet (James 1894), begegnet James diesem Einwand wie folgt. Dabei müssen wir jedoch zwei Fälle unterscheiden, die sich ansatzweise schon in der oben zitierten Passage von Worcester finden: (1.) zeigt das Beispiel vom Bären im Zoo, dass wir offenbar gelegentlich auf den gleichen Reiz stark emotional reagieren, gelegentlich überhaupt nicht. Und (2.) zeigt das Beispiel von Worcester, dass der eine große Furcht vor einem Bären empfinden mag, der andere (der „gut bewaffnete Jäger“) jedoch Freude angesichts des Antreffens eines Bären. Wir reagieren also mit qualitativ unterschiedlichen Reaktionen auf den gleichen Reiz.

James räumt nun zu (1) ein, zwar tatsächlich zu postulieren, dass emotionale Reaktionen analog zum Schema von Instinktreaktionen gewissermaßen „reflexartig“ durch ein elementares Objekt ausgelöst würden (z.B. durch den Bären). Dabei sei jedoch nicht zu übersehen, dass diese elementaren Objekte auch stets Teil eines organisierten, ganzheitlichen Zusammenhangs seien, einer „Situation“: „But whatever be our reaction on the situation, in the last resort it is an instinctive reaction on that one of its elements which strikes us for the time being as the most vitally important“ (ebd., 518). Die Frage, ob eine solche Konzeption, die doch recht viel Aktivität und Interpretationsarbeit seitens des Subjekts verlangt, noch mit dem Instinktbegriff zu vereinbaren sei, der ja explizit der James'schen Emotionstheorie zu Grunde liegt, wird auch von Kochinka aufgeworfen (2004, 227). Offenbar scheint James diese in seinem Argument inhärente Aktivität vielmehr in den äußeren Reiz hineinverlagern zu wollen, indem er davon spricht, dass es die Elemente sind, die uns in irgendeiner Form „angehen“ („which strikes us“).

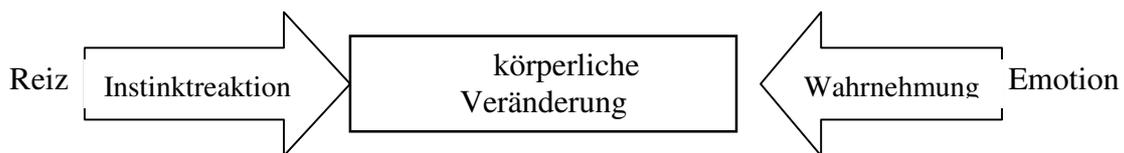
Diese Argumentationsrichtung wird vor allem in James' Stellungnahme zu (2) virulent: „The same bear may truly enough excite us to either fight or flight, according as he suggests an overpowering ‚idea‘ of his killing us, or one of our killing him“ (James 1894, 518). Gerade hier scheint diese Argumentation problematisch zu werden. Es mag noch vorstellbar sein, dass in der Gemengelage einer ganzheitlichen Situation ein bestimmtes Element (und kein anderes) salient wird, ohne dass es dabei unserer bewussten und aktiven Anstrengung bedarf. Dies ließe sich z.B. dadurch erklären, dass wir – wie Seligman (1971) sagt – eine gew.

¹⁸ In diesem Sinne auch Lange 1885, 69: „Fange ich an zu zittern, wenn ich mit einer Pistole bedroht werde, so ist es offenbar nicht der Sinneseindruck, der die Furcht hervorruft; denn die geladene Pistole sieht nicht anders aus, als die leere, die ich gar nicht beachtet hätte.“

angeborene „Bereitschaft“ („preparedness“) haben, auf bestimmte Elemente zu reagieren (wenngleich dieses Argument schon allein beim Bären im Käfig scheitern würde). Wie jedoch soll nun im Sinne von James' Stellungnahme zu (2) eine unbewusste Instinktreaktion eine solch differenzierte Bewertung oder Interpretation der Situation vornehmen können, der entsprechend wir den Bären dann entweder als Gefahr oder als willkommene Jagdtrophäe interpretieren?

Dieser Punkt ist m.E., wenn auch kritisch zu vermerken, keine Widerlegung der James'schen Theorie. James selbst scheint offenbar Anstrengungen unternommen zu haben, den Instinktbegriff zu erweitern und das Konzept der „Situation“ in sein Modell aufzunehmen, um dieser Kritik gerecht zu werden. Zwar wird dabei der Instinktbegriff über den Rahmen des Umgangssprachlichen und (Natur-)Wissenschaftlichen hinaus erweitert, eine solche Strategie ist jedoch nicht von vorneherein als illegitim zu betrachten. James scheitert letztlich an den fehlenden begrifflichen und konzeptionellen Mitteln, die nötig wären, um sein Modell hinreichend zu explizieren.

Ein weiteres Problem scheint die Diversität, oder gar Idiosynkrasie der Elemente zu sein, die James als instinktive Reaktionen auf einen emotionsrelevanten Reiz ansieht. Erinnern wir uns noch einmal daran, dass die „Emotion“ nach James eben keine Reaktion auf den *Reiz selbst* ist, sondern eine Wahrnehmung unseres *Körperzustands*, der nun selbst die Folge einer spontanen, unbewussten, instinktiven Reaktion auf einen solchen Reiz darstellt. Man könnte folgendes Schema vorschlagen:



James hatte explizit formuliert: „we feel sorry because we cry, angry because we strike, afraid because we tremble“, und dabei neurologisch und physiologisch ganz unterschiedliche Phänomene zusammengebracht. Das Zuschlagen ist sicherlich als eine Handlung zu werten, das Weinen können wir evtl. noch unterdrücken, es liegt also zumindest zu einem bestimmten Grad im Bereich unserer Verfügbarkeit, das Zittern jedoch lässt sich nicht kontrollieren, ebenso wenig wie – hier nicht explizit angesprochene – viszerale Veränderungen, die den Gastro-Intestinaltrakt oder Herz-Lungen-Funktionen betreffen. James stellt also als Korrelat der von ihm als primär postulierten instinktiven Reaktion ein Amalgam aus vegetativ-

nervösen Prozessen, halbwillkürlichen Emotionsexpressionen (Weinen, Zittern, Lachen¹⁹) und genuinen Handlungen (zuschlagen, weglaufen) zusammen. Auch dies hatte schon Worcester kritisch angemerkt:

„If I see a shower coming up, and run for a shelter, the emotion is evidently of the same kind, though perhaps less in degree, as in the case of the man who runs from the bear. According to Professor James, I am afraid of getting wet because I run. But supposing that, instead of running, I step into a shop and buy an umbrella. The emotion is still the same. I am afraid of getting wet. Consequently, as far as I can see, the fear, in this case, consists in buying the umbrella.“ (Worcester 1893, 291)

Gemäß James' Theorie würde also ein und dieselbe Emotion (Furcht) in jeweils verschiedenen Körperzuständen bzw. Handlungen bestehen (laufen bzw. Schirm kaufen). Weiterhin bleibt hier das Problem bestehen, wie ein so differenziertes und reflektiertes Phänomen wie eine Handlung Folge eines instinktiven Reaktionsschemas sein soll. Auch hierzu nimmt James in seinem jüngeren Aufsatz Stellung und bezieht sich auf seine Formulierung „we are frightened because we run“ : „Yet let the word ‚run‘ but stand for what it was meant to stand for, namely, for many other movements in us, *of which invisible visceral ones seem by far the most essential* [...]. Whatever the fear may be in such a case it is not constituted by the voluntary act“ (James 1894, 519, Hervorh. v. mir, S.H.)

James scheint in der Tat zu übersehen, dass die menschliche Psyche jenseits von elementarer Angst, spontanem Zorn und überwältigender Freude auch höhere Emotionen kennt, die durchaus kognitiv modelliert sein können. Der Großteil der Emotionen eines erwachsenen Menschen in einer westlichen Industriegesellschaft wird vermutlich von letzterem Typus sein, wenngleich auch solche Emotionen – wie Damasio (1994, 137 fig. 7-2) sagt – stets „Huckepack reiten“ („piggy-backed“) auf dem Mechanismus der elementaren Primäremotionen. Allerdings lässt sich eines doch auf die Kritik an James wiederum einwenden. Wenn wir mit einem (zugegeben zunächst recht schwachen Argument) kontern wollten, könnten wir darauf hinweisen, dass wir vis-à-vis dem Bären im Käfig zwar keine voll

¹⁹ Hier sei an Befunde von Ekman & Friesen 1982, Ekman/ Davidson/ Friesen 1990 und Damasio 1994 erinnert, denen gemäß sich willkürliches Lächeln von genuin-emotionalem Lächeln unterscheidet. Nach Damasio (1994, 139-143) sind jeweils unterschiedliche motorische Areale und unterschiedliche Nervenbahnen für die Darbietung eines Lächelns zuständig (vgl. auch Trepel 1999), wodurch sich einige kuriose wie theoretisch hochinteressante Befunde ergeben: Bei bestimmten Formen der Gesichtsparalyse (Goller 1992, 134: Pseudobulbärparalyse) wie auch bei einseitiger Paralyse durch Schlaganfall (Damasio, loc. cit.) können Vpn. ihre Gesichtsmuskeln nicht mehr willkürlich in eine bestimmte Konfiguration bringen. Sie können jedoch alle spontanen Emotionsausdrücke bilden. Umgekehrt können Parkinson-Kranke zwar willentlich ihre Gesichtsmuskeln bewegen, allerdings keinen emotionalen Ausdruck mehr zeigen, wenngleich sie noch emotionales Erleben haben (Ploog 1987).

ausgebildete Angst empfinden, aber dennoch sicherlich ein seltsames Gefühl in der Magenröhre haben. Wir fühlen uns, *als ob* wir ängstlich seien, sind es aber nicht, weil wir *wissen*, dass die Situation ungefährlich ist. Dieses Gegenargument tut mehr als die Behauptung der James-Kritiker lediglich zu relativieren. In einer detaillierteren Betrachtungsweise des Phänomens „Emotion“ unterscheiden Psychologen heute vielfach eine erste Phase der *Emotionselicitatation* von einer späteren *Coping*-Phase. Während James ein Modell der Emotionselicitatation vorschlug, bezieht sich die referierte Kritik eher auf die Relevanz der Coping-Phase, in der die eigenen emotionalen Reaktionen eines Subjekts durch dieses bewertet und vermittelt werden. Mit entsprechender Erfahrung können damit evtl. sogar inchoative emotionale Tendenzen sogleich unterdrückt oder gemildert werden, so wie wenn wir gleichmütig dem Bären im Käfig gegenüberstehen. Wir können spekulieren, dass kleine Kinder, auch wenn sie über das gleiche Wissen verfügen, dass die Situation gefahrlos ist, dies nicht völlig souverän vollbringen.

Ein weiterer, schon von Cannon (1927) vorgebrachter Kritikpunkt bezieht sich auf die unspezifische Diversität der menschlichen viszeralen Reaktionen. Es scheint in der Tat problematisch zu sein, bestimmte Muster viszeraler körperlicher Aktivität zu bestimmen, die dann wiederum spezifische emotionale Zustände auslösen würden. Gemäß der James-Lange-Theorie werden aber nun emotionale Zustände als Reaktionen auf Körperprozesse *unter Absehung* vom auslösenden Reiz definiert.²⁰ Wenn Hans also Furcht empfindet, dann müsste dies gemäß James-Lange allein dadurch zu erklären sein, dass Hans ein spezifisches Muster der körperlichen Furcht-Erregung wahrnimmt, das sich trennscharf vom Muster von bspw. Zorn unterscheidet. Bei dieser Wahrnehmung, die letztlich über die Qualität von Hans' Emotion entscheidet, darf Hans gerade nicht auf eine kognitive Bewertung des auslösenden Reizes zurückgreifen. Wir können diesen Einwand als die These vom un-eindeutigen Verhältnis zwischen Emotion und (Mustern von) körperlicher Veränderung bezeichnen.

Im Detail besagt sie folgendes: (a) Wenn wir einen Emotionstyp für sich betrachten (z.B. Angst), dann scheint sich kein eindeutiges Muster von körperlicher Veränderung finden zu lassen, das allen Instanziierungen dieses einen Typs gemeinsam wäre. Kenny formuliert diese Betrachtung wie folgt:

²⁰ Weil James körperlichen Prozessen den Primat gegenüber emotionalem Erleben eingeräumt hat, wurde seine Theorie auch dafür kritisiert, dass sie Emotion nur als ein Epiphänomen verstehen könne (Gordon 1987, 88), metaphorisch gesprochen als eine „Schaumkrone“: nach Keith Oatley bezeichnet der James'sche Emotionsbegriff lediglich „froth on top of the real business of behavior“ (1992, 133).

„*What is the feature of experience, recognizable by introspection without reference to content, which is common to fear of mice and fear of waking the baby, fear of overpopulation and fear of being over dressed, fear of muddling one's tenses and fear of Hell?*“ (Kenny 1963, 14)

Auch Lyons (1980, Kap. 7) kommt zu diesem Ergebnis. Mit dieser Teilthese zusammen hängt (b) die Beobachtung, dass phänomenal erlebte (gefühlte) Körperveränderungen nicht die Qualität einer Emotion bestimmen oder zur Distinguierung von Emotionstypen herangezogen werden können: es gibt offenbar kein eindeutiges Muster von Körperveränderungen im Verhältnis zu je einem spezifischen Emotionstyp steht.²¹ Ich denke, die Thesen (a, b) bedürfen einer kurzen Illustration. Nehmen wir als Beispiel den Emotionstyp „Angst“. Und nehmen wir weiterhin ein Muster (d.h. eine Menge) von Körperveränderungen *K* an, mit den Elementen: {erhöhter Herzschlag; weiche Knie; kalter Schweiß}. Teilthese (a) besagt nun, dass nicht bei allen Instanziierungen von „Angst“ die Menge *K* vollständig realisiert sein muss: es können Elemente wegfallen, evtl. sogar andere hinzukommen. Teilthese (b) besagt wiederum, dass *K* nicht charakteristisch für „Angst“ ist, sondern auch bei anderen Emotionstypen auftreten kann, wir könnten uns z.B. „Liebe“ vorstellen, vielleicht auch „Empörung“ usw. Die These (a, b) dissoziiert also charakteristische Körperveränderungen vollständig von bestimmten Emotionstypen („types“) sowie deren Instanziierungen („tokens“), wir können auch sagen: zwischen beiden besteht ein mehr-mehrdeutiges Verhältnis.

In den gängigen Spielarten der Komponententheorien der Emotionen wird meistens stillschweigend ein Verhältnis von (*gefühlter*) *Körperveränderung* einerseits und *Emotion* andererseits angenommen, das Marks (1982) explizit macht (vgl. in gleicher Weise: Lyons 1976, 610ff.), und das die von James postulierte Sichtweise dieses Verhältnisses nun wieder in Richtung der common-sense Auffassung zurückdreht: da im Rahmen dieser Auffassung die (gefühlte) Körperveränderung nicht die Grundlage der Emotion ist, bleiben als Alternativen noch, erstere entweder als Assoziation oder als Kausaleffekt von letzterer anzusehen. Da jedoch Emotionen in der Regel charakteristisch mit (gefühlten) Körperveränderungen verbunden sind (wenn auch nicht mit ein-ein-deutiger inhaltlicher Korrespondenz), wird meistens die Hypothese von Körperveränderungen als Kausaleffekt von emotionalen

²¹ Als historischer Einwand gegen die James-Lange-Theorie von Cannon 1927; das Argument wurde seitdem häufig repliziert, z.B. von Solomon 1973, 158: „What is the difference in feeling between a common pair of emotions – for example embarrassment and shame? People are rarely, if ever, confused about which emotion they ‚have‘, but when asked to differentiate between them [the feelings, S.H.] they find themselves speechless“; weiterhin Bedford 1962, 111; vgl. auch: Koch 1987, 69f.

Zuständen vertreten. So schreibt Marks (loc. cit., 229): „we say ‚His fright *caused* him to shiver‘, ‚He blanched *from* fear‘, ‚His fear *made* him tense up‘. [...] The strong suggestion here is that the locus of the emotion is (causally) prior to the various manifestations of the emotion“.

Cannon (1927) (vgl. Meyer/ Schützwohl/ Reisenzein 2001) liefert im Ganzen fünf Argumente gegen die James-Lange-Theorie²²: (1.) Cannons eigene Beobachtungen an sympathiko-ektomisierten Katzen²³ zeigten, dass emotionales Verhalten unverändert dargeboten wird, obwohl kein sympathisches „feedback“ aus den Viszera mehr vorhanden ist. Dazu ist anzumerken, dass der Schluss vom unveränderten Verhalten auf unverändertes Erleben nicht gültig ist: Katzen können letzteres nicht mitteilen und bleiben daher als Probanden defizitär.²⁴ Weiterhin spricht Cannon nur von einer Ektomie des sympathischen, nicht des parasympathischen Anteils. Auch hier wäre eine bessere Versuchsanordnung wünschenswert, wenngleich einzuräumen ist, dass die typischen Korrelate des Sympathikus – im Allgemeinen nämlich viszerale *Erregungszustände* – eher dem zu untersuchenden Phänomen entsprechen als die des Parasympathikus – nämlich Dämpfung und Energiekonservierung. (2.) Die gleichen Muster viszeraler Körperveränderung treten bei verschiedenen emotionalen Reaktionen wie auch bei nicht-emotionalen Zuständen

²² Vgl. die in vielen Punkten ähnliche Kritik von Wilhelm Wundt, fast zwei Jahrzehnte vor Cannon: „Wenn freilich C. Lange und W. James in [den] begleitenden Erscheinungen die ausschließliche Ursache der Affekte selbst erblicken wollten [...], so ist diese paradoxe Annahme aus drei Gründen unhaltbar. Erstens treten die entscheidenden äußeren Symptome der Affekte erst in einem Moment hervor, wo die psychische Natur des Affekts schon deutlich differenziert ist: dieser selbst geht also denjenigen Innervationswirkungen voran, die hier als seine Ursachen in Anspruch genommen werden. Zweitens ist es absolut unmöglich, die Mannigfaltigkeit der psychischen Affektzustände dem verhältnismäßig einfachen Schema der Innervationsänderungen einzuordnen: die psychischen Vorgänge selbst sind um vieles reicher als ihre spezifisch verschiedenen Ausdrucksformen. Endlich stehen drittens die physischen Begleiterscheinungen der Affekte in durchaus keiner konstanten Beziehung zu der *psychologischen Qualität* derselben. Dies gilt namentlich von den Puls- und Atmungswirkungen, aber auch z.B. von den pantomimischen Ausdrucksbewegungen. Affekte, die einen sehr verschiedenen, ja entgegengesetzten Gefühlsinhalt haben, können unter Umständen in bezug auf diese physischen Erscheinungen zu der nämlichen Klasse gehören.“ (Wundt 1909, 211).

²³ Es handelt sich hier also um Tiere, bei denen die afferenten Bahnen des sympathischen Anteils des autonomen Nervensystems ausgeschaltet wurden, mithin um „vollständig deafferente“ Versuchswesen, wie wir sie im Rahmen der James-Lange-Theorie diskutiert haben.

²⁴ Auch Barbalet 1999, 253 merkt an, dass die von Seitens Walter Cannons vorgetragene Kritik an der James'schen Theorie hinkt auf Grund der Tatsache, dass James emotionales *Erleben* untersucht, während Cannon in seiner Kritik auf emotionales *Verhalten* abhebt.

(körperliche Tätigkeit, Fieber etc.) auf. (3.) Viszerale Reaktionen liefern nicht das Niveau von Differenziertheit, das zur Erklärung unseres vielfältigen und facettenreichen emotionalen Erlebens notwendig wäre.²⁵ (4.) Die Latenzzeit von viszeralen Reaktionen ist größer als die von emotionalen Reaktionen.²⁶ (5.) Die künstliche Induktion peripher-physiologischer Erregung führt nicht zu emotionalen Reaktionen. Dies hatte Marañon bereits 1924 untersucht.

Können wir abschließend festhalten, dass die James-Lange-Theorie in ihrem explanatorischen Umfang ein angemessenes Bild der menschlichen Emotionalität zeichnet, oder müssen wir – wie Damasio – folgendes Fazit ziehen?

„[James’] account works well for the first emotions one experiences in life, but it does not do justice to what Othello goes through in his mind before he develops jealousy and anger, or to what Hamlet broods about before exciting his body into what he will perceive as disgust, or to the twisted reasons why Lady Macbeth should experience ecstasy as she leads her husband into a murderous rampage.“ (Damasio 1994, 130)

Was Damasio der James’schen Theorie hier vorwirft, ist, dass sie die subtileren, feineren oder höherentwickelten Phänomene der menschlichen Emotionalität nicht zu fassen vermag. In der Tat ist ja schon in James’ Beispielen immer nur die Rede von primären Emotionen wie Trauer, Furcht, Zorn, und dies auch immer in Bezug auf recht einfache (weil gegenständliche externe) Objekte. Doch auch James räumt ein, dass wir über solch feineren Gefühle verfügen: „ästhetische“, „moralische“ und „intellektuelle“ Gefühle, wie er sagt: „We have, then, or some of us seem to have, genuinely *cerebral* forms of pleasure and displeasure, apparently

²⁵ Dieses Argument, das häufig repliziert wurde und im Allgemeinen als „das“ schlagende Argument gegen Aktivierungs-Theorien der Emotionen angesehen wird, muss nach heutigem Kenntnisstand relativiert werden. Nach Ekman/ Levenson/ Friesen (1983) und Levenson (1992) können innere Organe durch das ANS in der Tat recht spezifisch angesteuert werden, sodass sich einigermaßen typische viszerale Aktivierungsmuster für jeweils bestimmte Emotionen ergeben. Dennoch wird man das Argument von Cannon, wenn auch nicht kategorisch, so doch graduell weiter gelten lassen. Besonders hoch differenzierte, komplexe, kognitiv gehaltvolle Emotionen (wie z.B. „Scham“ und „Schuld“) entziehen sich einer solchen Klassifikationsmethode, wenngleich diese evtl. bei primären Emotionen recht reliabel funktioniert.

²⁶ Dieses Argument, traditionell eher nebensächlich behandelt, scheint heute gegenüber dem oben diskutierten Argument (3) eine höhere Relevanz zu gewinnen. Der *nervöse* Weg der Rückmeldung einer Körperveränderung (also von der Aktivierung, der Efferenz des Impulses in die Viszera, und der afferenten Rückmeldung ins Gehirn) nimmt ca. 1-2sek in Anspruch (LeDoux 1996, 314). Das liegt nicht an der vergleichsweise langen Strecke in die Körperperipherie, sondern an der langen Reaktionszeit der für die inneren Organe typischen „glatten Muskulatur“. Bei einer solchen Latenzzeit ist jedoch ohne Weiteres anzunehmen, dass eine nervöse Rückmeldung in die erlebte Gefühlsdynamik eines emotionalen Zustands mit einght. Der *hormonelle* Weg einer Rückmeldung kann bis zu mehrere Stunden dauern (ebd.).

not agreeing in their mode of production with the ‚coarser‘ emotions we have been analyzing“ (James 1950 II, 468). Der Wahl des Terminus „zerebrale“ Gefühle scheint also auf eine Variation im Weg der Emotionsgenese zu deuten, dies jedoch nur augenscheinlich, wie der mit „apparently“ ansetzende, letzte Halbsatz nahe legt. Denn nach nüchterner Untersuchung („sober scrutiny“) solcher Fälle vermeintlich „zerebraler“ Emotionen macht man eine erstaunliche Beobachtung:

„Bei allen sentimental und impressionablen Leuten treten körperliche Wirkungen [zu diesen ‚feineren‘ Emotionen, S.H.] hinzu: die Stimme verändert sich und die Augen werden feucht, wenn moralische Wahrheit empfunden wird usw. Wo man irgendwie stärker hingerissen wird, der Grund mag noch so intellektuell sein, da treten diese sekundären Prozesse deutlich zutage. Wenn wir nicht tatsächlich lachen über die Nettigkeit einer Beweisführung oder einen witzigen Einfall; wenn wir nicht zittern gegenüber der Gerechtigkeit, oder ein Prickeln empfinden gegenüber einem Akt der Großmut, dann kann der psychische Zustand, in dem wir uns befinden, überhaupt kaum ein emotionaler genannt werden. Er ist dann in Wirklichkeit nichts als eine intellektuelle Wahrnehmung davon, wie gewisse Dinge zu nennen sind – nett, richtig, witzig, großmütig usw. Ein derart urteilender Geisteszustand müsste eher der Klasse der erkennenden als jener der emotionalen Akte beigezählt werden.“ (James 1920, 385f.)

Sein Lösungsvorschlag in dieser, aus der deutschen Ausgabe zitierten Passage, ist, dass vermeintlich „intellektuelle“ Emotionen entweder (1.) den gleichen Mechanismus durchlaufen, wie er ihn für die primären Emotionen postuliert,²⁷ oder dass sie (2.) – wenn wir eben nicht „lachen [...], nicht zittern [...], oder ein Prickeln empfinden“ etc. – nichts anderes sind als „erkennende Akte“ (und damit keine eigentlichen Emotionen), die zwar durch bestimmte phänomenale Gefühle begleitet sein können, nicht jedoch auf tatsächlichen Körperveränderungen (lachen, zittern, prickeln) beruhen. James hat sich – und das belegen exemplarisch die angeführten Textstellen – theoretisch mit der Integration der komplexeren Emotionen in sein Modell der Emotionsgenese bemüht, und diese Anstrengung wird häufig übersehen in kritischen Kommentaren zu James. Der Vorwurf von Damasio trifft dennoch zu. James scheint ebenfalls für die Kategorie der „sekundären“ (d.h. feineren) Emotionen seinen Elizitationsmechanismus für nicht adäquat gehalten zu haben, weshalb er die meisten dieser Phänomene schlicht aus dem Bereich der Extension des Emotionsbegriffs ausschließt. Da er jedoch offenbar weiterhin zulässt, dass solche „erkennenden Akte“ von Gefühlen begleitet werden, stellt dies ein gewisses Makel seiner Theorie im Ganzen dar: solche intellektuellen

²⁷ Wengleich hier darauf hinzuweisen ist, dass die Formulierung „treten hinzu“ und „sekundäre Prozesse“ etwas unglücklich ist, da James schließlich den Primat der physiologischen Aktivierung vertritt. Evtl. beziehen sich diese Formulierungen aber auch auf die oben unter (2) angeführten „un-eigentlichen“ Emotionen, die in Wirklichkeit „erkennende Akte“ sind, die lediglich von körperlichen Symptomen begleitet werden.

Gefühle überschneiden sich wohl nicht nur in unserer Alltagserfahrung, sondern auch in der Sache wesentlich mit genuinen Emotionen. Nachdem jedoch James diese intellektuellen Gefühle aus dem Gegenstandsbereich seiner Emotionstheorie ausgeschlossen hat, stehen nun deren Begriffe und Modelle nicht mehr zur Erklärung jener Gefühle zur Verfügung, obwohl sie dazu sicherlich einiges beitragen könnten. Wenngleich wir also nicht sagen können, dass James' Theorie an dieser Stelle streng genommen „scheitert“, so wäre es doch wünschenswert, eine engere theoretische Verbindung zwischen der Analyse primärer und sekundärer Emotionen herzustellen, um einen angemessenen Blick auf das Gesamtphänomen der Emotionalität zu gewinnen. Diese Stoßrichtung hat sicherlich auch die zitierte Kritik von Seiten Damasio, und es ist wohl einer der Vorzüge seiner Theorie, eben jene Verbindung zu verwirklichen.

Die vermeintliche Widerlegung der James-Lange Theorie durch Dana und Cannon wurde später erneut kritisch relativiert. Dana (1921) hatte geglaubt, empirisch nachgewiesen zu haben, dass emotionales Erleben auch möglich ist, selbst wenn die peripher-physiologische Aktivierung fehlt, wie dies z.B. bei Probanden mit Rückenmarksverletzungen der Fall ist. Hierbei ergibt sich jedoch das oben diskutierte Problem, dass bei solchen paralytischen Patienten nicht nur die afferente Verbindung von der Körperperipherie zum Gehirn unterbrochen ist, sondern dass sie durch die völlige Ausschaltung der Nervenbahnen des Rückenmarks auch keine physiologischen Veränderungen oder Körperpraktiken darbieten können, die ja nach James zunächst auftreten müssen, damit ein emotionaler Zustand qua Wahrnehmung ausgelöst werden kann. Neuere Untersuchungen an Patienten mit Rückenmarksverletzungen auf unterschiedlichen Niveaus der Wirbelsäule (Hohmann 1966) haben gezeigt, dass emotionales Erleben mit ansteigender Höhe der Läsion an Intensität verliert, bis schließlich nur noch von einem „Als-ob-Erlebnis“ berichtet wird („Es ist, als ob ich Furcht empfinde.“).²⁸ Die physiologische Komponente, der in der James-Lange-Theorie

²⁸ In neuerer Zeit wurde versucht, die Studie von Hohmann zu replizieren (Bermond et al. 1991). Nach diesen Ergebnissen zeigen die Rückenmarksläsion-Patienten keine Beeinträchtigung des emotionalen Erlebens. Dabei wurde jedoch nicht per Selbstaussage in aktuellen emotionalen Zuständen getestet, sondern die Vpn. wurden gebeten, ihr jetziges emotionales Erleben mit dem von früher zu vergleichen. Natürlich sind Untersuchungen durch Selbstaussage (wie auch Hohmann) offen für zahlreiche Fehlerquellen. Andererseits müssen erstens weitere Wege des Körperfeedbacks berücksichtigt werden, z.B. über Hormone und den nervus vagus (zu letzterem vgl. den entsprechenden Abschnitt unten), zweitens die Rückmeldungen z.B. aus der Gesichtsmuskulatur, die nicht über das Rückenmark innerviert ist, und drittens die Möglichkeit von sog. as-if-feelings (Damasio 1994, vgl. u.). Besonders in Hinsicht auf den letzten Punkt wird wieder die

der zentrale (kausale) Platz zukommt, wird also auch in neueren Emotionstheorien berücksichtigt werden müssen. Schon bei Hohmann zeigt sich jedoch die Verschiebung des Paradigmas, die dann leitendes Programm der Schachter-Singer-Studie werden sollte: möglicherweise erklären physiologische Veränderungen gar nicht *kategoriale* oder qualitative Unterschiede von emotionalen Zuständen (Angst vs. Zorn), sondern deren *dimensionale* Ausprägung, d.h. die Intensität des emotionalen Erlebens. Zu fragen wäre dann freilich auch, ob bei völliger Abwesenheit physiologischer Aktivierung dennoch ein emotionaler Zustand auftreten kann. Evtl. handelt es sich im Fall der physiologischen Aktivierung um eine skalare Größe, die jedoch bei völliger Abwesenheit (bzw. im homöostatischen Zustand des Organismus) dennoch eine binäre Dichotomie abbildet und emotionales Erleben damit letztlich doch „notwendig“, d.h. als notwendige Voraussetzung, bedingt.

Halten wir fest, dass Emotionen also sicher (1.) keine rein *physiologischen* Phänomene sind, in der Art wie dies etwa Gefühle des Hungers, der Müdigkeit oder des physischen Schmerzes sind, und sie lassen sich nicht im Sinne von James und Lange auf die Wahrnehmung einer physiologischen Veränderung reduzieren. Emotionen lassen sich andererseits auch (2.) nicht primär über den emotionalen *Ausdrucks* klassifizieren (d.h. Mimik, Gestik oder sog. expressive Emotionshandlungen, wie das Zerschlagen einer Vase aus Wut über Hans). Gleiches gilt (3) für eine Klassifikation des affektiven (phänomenalen) *Erlebens*: „Neid“ und „Eifersucht“ sind allein am affektiven Erleben vermutlich nicht distinktiv zu unterscheiden. Emotionen sind also konzeptuell nicht ohne gewisse kognitive Inhalte zu verstehen.

2.4. Moderne kognitivistische Emotionstheorien in der empirischen Psychologie

Im Jahr 1962 formulierten Stanley Schachter und Jerome Singer die erste kognitivistische Emotionstheorie²⁹ in der psychologischen Forschung des 20. Jahrhunderts, ausgehend von

Versuchsordnung problematisiert: da Rückenmarksverletzungen häufig die Folge von Unfällen sind, werden die entspr. Probanden schon vor dem Unfall einen funktionierenden Mechanismus des emotionalen Körperfeedbacks aufgebaut haben, und demnach (gemäß Damasio) vermutlich auch schon die Variante der as-if-Schleife, die unter Umgehung des Körpers emotionales Gefühl gewissermaßen „zerebral simuliert“.

²⁹ In der emotionstheoretischen Forschung findet natürlich auch eine breite Debatte um die Begriffe „Kognition“ und „Emotion“ statt (die später referiert werden soll), weshalb es auch strittig ist, welche Theorie (nach welchen Kriterien) als „kognitivistisch“ zu gelten habe. So zählt z.B. Gessner (2004, 46) die Theorieskizze von Schachter & Singer zu den sog. „Theorien zur kognitiven Attribution“ (und mithin nicht zu den kognitivistischen Theorien

einer empirischen Studie, die in der Emotionsforschung noch heute Referenzqualitäten hat (Schachter/ Singer 1962).³⁰ Da sich auch heute noch wertvolle grundlegende Überlegungen aus dem Schachter-Singer-Experiment ableiten lassen, soll dieses im Folgenden kurz dargestellt werden. Im Anschluss daran werden gängige Interpretationen referiert, sowie ein eigener Standpunkt des Verfassers der vorliegenden Arbeit, der bis dato offenbar noch nicht aus den Experimenten abgeleitet wurde.

Schachter und Singer führten die von Gregorio Marañon (1924) angestellten Vorüberlegungen weiter und versuchten vor allem, eine Erklärung für die unbefriedigenden Ergebnisse seines Experiments zu finden. Ein Ansatzpunkt war dabei die Überlegung, dass die Versuchspersonen im Marañon-Experiment darüber informiert waren, dass sie eine Injektion erhielten, und auch von welcher Art die injizierte Substanz war (Schachter & Singer 1962, 381). Der Impetus der Schachter-Singer Theorie war die Überlegung, die als unzureichend empfundene James-Lange Theorie um einen notwendigen Erklärungsfaktor zu erweitern, ohne jedoch die körperliche Grundlage der Emotionalität, ausgedrückt durch peripher-physiologische Aktivierung (d.h. erhöhte Sympathikus-Erregung), aufzugeben. Schachter und Singer wollten die Emotionsgenese als einen Zwei-Faktoren-Prozess, konstituiert durch peripher-physiologische Aktivierung und Kognitionen, verstehen. Dabei bestimmen Kognitionen die jeweilige *Qualität* des emotionalen Erlebens, während die physiologische Aktivierung dessen *Intensität* bestimmt.³¹ Beide Faktoren stellen nach

im eigentlichen Sinne). An späterer Stelle soll eine Arbeitsdefinition geliefert werden, mit der für die vorliegende Untersuchung eine Unterscheidung von kognitivistischen und nicht-kognitivistischen Theorien getroffen werden kann.

³⁰ Die Schachter-Singer Experimente basieren auf knapp 40 Jahre vorher durch Marañon (1924) angestellten Vorüberlegungen und eigenen empirischen Versuchen. In der Linie der James-Lange Theorie hatte Marañon 210 Probanden Adrenalin injiziert, um Symptome peripher-physiologischer Erregung (Beschleunigung des Herzschlags, Transpiration, Beschleunigung der Atemfrequenz, veränderte Peristaltik des Gastro-Intestinaltrakts etc.) hervorzurufen, die nach James-Lange die Grundlage unseres emotionalen Erlebens darstellen. Sein Experiment – in vielen Hinsichten eine krude Vorform des berühmten Schachter-Singer Experiments – lieferte jedoch keine Bestätigung der James-Lange Theorie: 71% der Versuchspersonen berichteten lediglich körperliche Symptome, 29% berichteten von Zuständen, die sich anfühlten, „als ob“ die Vpn eine Emotion erlebten.

³¹ Später wurde v.a. die Hypothese kritisiert, dass die physiologische Erregung *allein* die Intensität des emotionalen Erlebens bestimme. Marshall & Zimbardo (1979) replizierten die Euphorie-Bedingung des Schachter-Singer Experiments unter Einsatz höherer Adrenalin-Dosen und stellten bei den Vpn trotz euphorischer Situationsbedingung negative Emotionen fest. Zu den gleichen Ergebnissen kommt auch eine Studie von Maslach (1979). Hohe Grade physiologischer Aktivierung scheinen auch die Qualität des Emotionserlebens zu verändern, u.z. in Richtung des negativen Spektrums zu verschieben.

Schachter und Singer notwendige Bedingungen der Emotionsgenese dar, jedoch für sich je einzeln keine hinreichende Bedingung. In ihrem berühmten Experiment versuchten sie also, emotionale Zustände unter Laborbedingungen künstlich durch Manipulation der beiden Faktoren zu induzieren.

(1.) Der Faktor der physiologischen Erregung wurde auf pharmakologischem Wege manipuliert: Schachter und Singer verabreichten ihren 184 Vpn eine subkutane Injektion. Diese enthielt entweder Epinephrin (ein Adrenalin-Derivat zur Stimulierung des sympathischen Anteils des vegetativen Nervensystems) oder ein Placebo (Kochsalzlösung). Alle Vpn erhielten zunächst die Information, dass die Wirkung des Vitaminpräparats „Suproxin“ auf das Sehvermögen getestet werden solle. Dies ist eine in der empirischen Psychologie gängige Strategie zur Verschleierung des eigentlichen experimentellen Ziels. Es sollen dadurch sog. „demand characteristics“ verhindert werden, nämlich das (bewusste oder unbewusste) Verhalten von informierten Probanden, gezielt dem experimentellen Ziel konforme Ergebnisse zu liefern. Weiterhin wurden die Probanden in Gruppen unterteilt, die nun zusätzliche, unterschiedliche Informationen erhielten: (a) „Epinephrine Informed (Epi Inf)“: Diese Gruppe wurde mit korrekten Informationen über die „Nebenwirkung“ der Injektion versorgt, wobei auch hier die Illusion gewahrt wurde, dass es sich um ein Vitaminpräparat handele, und damit lediglich um Nebenwirkungen. Demnach könnten die Vpn die (für adrenalin-induzierte Sympathikus-Aktivierung typischen) Symptome von beschleunigter Herz- und Atemfrequenz, einem leichten Handtremor, Wärmegefühlen, v.a. im Gesicht, Transpiration, etc. erwarten. (b) „Epinephrine Ignorant (Epi Ign)“: diese Gruppe erhielt die (falsche) Information, die Injektion habe keine Nebenwirkungen. (c) „Epinephrine Misinformed (Epi Mis)“: diese Gruppe wurde über die auftretenden Nebenwirkungen falsch informiert. Den Vpn wurde gesagt, sie hätten mit Effekten wie tauben Füßen, Jucken, leichten Kopfschmerzen etc. zu rechnen. (d) Die Placebo-Gruppe, deren Injektion ja tatsächlich keine Nebenwirkungen hervorrufen sollte, erhielten die gleiche Information wie die Epi Ign Gruppe (nämlich dass tatsächlich keine Nebenwirkungen auftreten). Zumindest für Gruppe (a) Epi Inf konnte auf der Basis der Marañon-Experimente schon die Prognose angestellt werden, dass die Vpn ihre erlebte physiologische Aktivierung korrekt der Injektion zuschreiben könnten und demnach kein emotionales Erleben auftreten sollte.

(2.) Der Faktor „Kognition“ wurde nun durch Manipulation situativer Bedingungen verändert. Die Vpn sollten in eine nach Schachter und Singer vermutlicher Weise emotionsrelevante Situation versetzt werden. Während sie auf den angekündigten Sehtest (im Rahmen der vermeintlichen „Suproxin“-Studie) warteten, wurde ein sog. „stooge“, also ein

als Proband getarnter Vertrauter der Experimentatoren, zur Manipulation diese emotionsrelevanten Situation eingesetzt. (a) Euphorie: der stooge beginnt im Warteraum mit allerlei Albernheiten (Werfen mit Papierkugeln und Papierflugzeugen, Spielen mit Hula-Hoop-Reifen etc.) und versucht, die Vpn dabei miteinzubeziehen. (b) Ärger: der stooge diskutiert mit den Vpn die Fragen eines auszufüllenden Fragebogens und empört sich dabei eindrucksvoll darüber, dass diese lächerlich oder beleidigend seien. Der zugrundeliegende Fragebogen wurde entsprechend präpariert, damit eine solche, als beleidigend dargestellte Interpretation auch tatsächlich als gerechtfertigt erscheinen konnte.

Aus den Gruppen (1a-d) und (2a,b) stellten Schachter und Singer sieben experimentelle Bedingungen zusammen (s. Tabelle unten), auf die die insges. 184 Vpn (männlich College-Studenten) zu Gruppen von 23-27 Vpn pro Einzelbedingung aufgeteilt waren.

(2a) Euphorie	(2b) Ärger
(1a) Epi Inf	(1a) Epi Inf
(1b) Epi Ign	(1b) Epi Ign
(1c) Epi Mis	n.a.
(1d) Placebo (= Epi Ign)	(1d) Placebo (= Epi Ign)

Die Reaktionen der Vpn auf die Manipulation der beiden Variablen „Aktivierung“ und „Kognition“ wurden einerseits durch instrumentelle Messungen festgestellt (Messung der Pulsfrequenz als Indikator der physiologischen Erregung), teilweise durch Verhaltensbeobachtungen durch einen Einwegspiegel (Einschätzung der emotionalen Reaktionen durch die Experimentatoren), sowie durch Selbsteinschätzung in einem Fragebogen (Angaben zur momentanen Stimmung wie auch zu körperlichen Symptomen).

Die Ergebnisse des Schachter-Singer Experiments sind beeindruckend, wenn auch problematisch in ihrer Interpretation und durchaus kontrovers aufgenommen. Die Messungen zum Faktor „Aktivierung“ ergaben zunächst völlig erwartungskonform, dass die physiologische Aktivierung bei den Epi-Gruppen signifikant höher war als in den Placebo-Gruppen. Die diesbezügliche Wirkung von Epinephrin ist schlicht pharmakologisch sehr reliabel und zeigte die erwartete Wirkung. Die Ergebnisse zum Faktor „Kognition“ jedoch liefern nur nach intensiver Filterung interessante Aufschlüsse.

(I.) In der Ärger-Bedingung (2b) zeigten die Vpn der Epi-Ign-Gruppe statistisch signifikant höhere Werte emotionaler Reaktionen als die Vpn der Epi-In- oder Placebo-Gruppe (beachte dass in 2b keine Epi-Mis-Gruppe vorlag); dies jedoch *nur* nach den Ergebnissen der Verhaltensbeobachtung durch den Einwegspiegel, und *nicht* nach den Ergebnissen der Selbsteinschätzung im Fragebogen! Wie ist es nun zu erklären, dass die

Experimentatoren in der Verhaltensbeobachtung emotionale Ärger-Reaktionen erkennen konnten, die Vpn dies jedoch offenbar nicht im Fragebogen über sich selbst berichten? Fasst man die beiden Messergebnisse zusammen, wird der Effekt natürlich herausgemittelt, weswegen Schachter und Singer sicherlich an einer Erklärung dieses Phänomens gelegen war. Diese bietet sich nun m.E. tatsächlich recht zufriedenstellend an, indem Schachter und Singer auf die Tatsache verweisen, dass es sich bei den Vpn um Collegestudenten handele, die an dem Experiment teilnahmen, um zusätzliche Bonus-Punkte für ihren Psychologie-Kurs zu erhalten. Wie bereits referiert wurde, sollte sich der situative Ärger-Faktor auf die von den Experimentatoren ausgeteilten Fragebögen richten. Wir können daher durchaus dem plausiblen Hinweis von Schachter und Singer folgen, dass die Vpn ihren Ärger auf die Experimentatoren nicht im Fragebogen zur Selbsteinschätzung zum Ausdruck brachten, um eben jene Bonus-Punkte nicht zu gefährden, während sie sich in Anwesenheit des „stooges“ unbeobachtet wähnten und „gehen ließen“. Die Erklärung ist m.E. plausibel, wenngleich hierin natürlich ein Fehler in der Versuchsanordnung zu sehen wäre.

(II.) Ein vergleichbares Ergebnis für die Euphorie-Bedingung (2a) konnte jedoch zunächst nicht gefunden werden. Eine genaue Durchsicht der Fragebögen zur Selbsteinschätzung ergab jedoch, dass einzelne Vpn der Epi-Ign- und Epi-Mis-Gruppen in den Angaben zu körperlichen Symptomen diese explizit auf die Injektion zurückführten. Damit würde sich für jene Vpn – die Schachter und Singer „self-informed subjects“ (394) nennen – die gleiche Ausgangslage ergeben wie für die Epi-Inf-Gruppe, für die ja in Anlehnung an Marañon keine emotionale Reaktion zu erwarten war. Die Hypothese von Schachter und Singer basiert entscheidend darauf, dass die erlebte Aktivierung nicht als künstlich induziert, sondern als in Zusammenhang mit dem emotionsrelevanten Reiz stehend wahrgenommen wird. Schließt man nun diese einzelnen „self-informed subjects“ aus, so ergeben sich nach Schachter und Singer tatsächlich signifikante Unterschiede in der emotionalen Reaktion zwischen den – gefilterten – Epi-Ign- & Epi-Mis-Gruppen einerseits und zwischen den Epi-Inf- & Placebo-Gruppen andererseits.

Unter Einbeziehung der beiden oben erwähnten Filter (I und II), interpretieren die Experimentatoren ihre Ergebnisse sehr zuversichtlich als Bestätigung ihrer Hypothese:

„Given a state of sympathetic activation, for which no immediately appropriate explanation is available, human subjects can be readily manipulated into states of euphoria, anger, and amusement.“ (Schachter & Singer 1962, 396)

Auf der Grundlage ihrer experimentellen Ergebnisse formulieren Schachter und Singer einen Theorievorschlag zur Erklärung des Prozesses der Emotionsgenese. Während unter den

experimentellen Labor-Bedingungen die Probanden mit einer zunächst unerklärlichen physiologischen Aktivierung konfrontiert werden, für die sie dann eine emotionsrelevante Erklärung zu finden versuchen (bzw. die Symptome als rein körperlich verstehen, sofern eine solche Erklärung fehlt), so machen Schachter und Singer in Analogie dazu einen Vorschlag zur Erklärung der „alltäglichen“ Emotionsgenese. Gemäß dieser Theorie löst die Wahrnehmung eines Reizes eine physiologische Aktivierung („arousal“) aus. Ich gehe z.B. eine dunkle Straße entlang und plötzlich taucht ein Mann mit Pistole auf (vgl. 380). Der Reiz löst einen Zustand physiologischer Aktivierung aus, der in einem nächsten Schritt kategorisiert („gelabelt“) wird, u.z. entweder als ein physiologisches Symptom, falls keine emotional relevante Situation vorliegt, oder als eine näher bestimmte Emotion, entsprechend einer kognitiven Bewertung der gegebenen Situation (Elemente wie: „es ist Nacht, der Mann lauert in einer dunklen Straßenecke, es handelt sich um einen Unbekannten von zwielichtiger äußerer Erscheinung“ etc. würden als Furcht kategorisiert). Es ist wichtig festzuhalten, dass hier kognitive Elemente im engeren Sinne, nämlich Wissen, Urteile und Schlussfolgerungen, in die Emotionsgenese hineinspielen. Schachter und Singer gehen allerdings davon aus, dass die primären Prozesse der Bewertung eines Reizes (und damit die Auslösung der physiologischen Aktivierung, die in einem zweiten Schritt gelabelt wird) automatisch und unbewusst ablaufen. Die physiologische Aktivierung, die auf spontanen Reiz-Reaktions-Schemata beruht, geht weiterhin dem emotionalen Erleben voraus, es hängt hier jedoch von unserer kognitiven Bewertung ab, ob und wie diese Aktivierung als Emotion kategorisiert wird. Die Schachter-Singer Theorie beschreibt den Zusammenhang zwischen Kognition und Emotion zwar schon wesentlich genauer, allerdings finden sich auch hier Prozesse der primären Reizbewertungen, bei denen die Frage offen bleibt, inwiefern (und nach welchen Kriterien) sie als „kognitiv“ oder „nicht-kognitiv“ zu verstehen sind.

Kehren wir zum Schluss noch einmal zur Interpretation der Ergebnisse zurück. Wie bereits gesagt, betrachteten die Epi-Mis- oder Epi-Ign-Gruppen ihren Zustand lediglich als körperliche Erregung, so lange keine adäquate emotionsrelevante Kognition zur Verfügung stand. Als sie jedoch geeigneten Stimuli in der Euphorie- oder Ärger-Bedingung ausgesetzt waren, wurden sie leicht durch diese Reize in den entsprechenden emotionalen Zustand versetzt. Lyons (1976, 611) unterstreicht, dass offenbar die Illusion wichtig ist, dass die physiologische Erregung von der emotionsrelevanten Situation herrührt. Denn die Epi-Inf-Gruppe konnte – wie zu erwarten und durch Marañon vorgeführt – nicht emotional beeinflusst werden und schrieb ihren Zustand einfach der Epinephrin-Injektion zu.

Ich denke, dass das Experiment eine weitere Interpretation zulässt, die meines Wissens bis dato übersehen wurde. Es scheint in der Tat eine empirische Entscheidung zu erlauben in der schon diskutierten Frage, ob eine peripher-physiologische Körperveränderung ein Kausaleffekt oder aber eine Vorbedingung bzw. Auslöser von Emotion sei. Die Epi-Inf-Probanden ließen sich nicht in einen emotionalen Zustand hineinversetzen, da sie ihre Erregung auf Grund ihrer Vorab-Informationen der Injektion zuschreiben konnten. Dies sollte jedoch nicht auftreten, sofern die These richtig wäre, dass Emotionen (kausal) physiologische Erregung hervorrufen. In diesem Fall hätten die Epi-Inf-Probanden dennoch emotional reagieren müssen, da sich die von der Emotion hervorgerufene Erregung mit der experimentell induzierten hätte überlagern müssen. Nun kann man einwenden, dass evtl. die emotionsrelevanten Reize nicht stark genug gewesen seien (mit anderen Worten: die Euphorie-Stimuli waren nicht wirklich lustig, und die Ärger-Stimuli nicht wirklich ein Anlass für Ärger). Den Probanden hätte mit der Injektion einfach eine bessere Erklärung zur Verfügung gestanden als ein emotionsrelevanter Reiz von geringer Intensität. Dieser Einwand kann jedoch nicht gelten, da hierfür die Epi-Misinf- und Epi-Ign-Gruppen sozusagen die Kontrollgruppe bilden. Denn wären die vermeintlich emotionsrelevanten Reize tatsächlich nicht relevant gewesen (bzw. von geringer Intensität oder Relevanz), dann würde man im Zuge des vorgestellten Einwands auch zugeben müssen, dass die Probanden dieser zwei Gruppen nur darum emotionale reagiert haben, weil eben physiologische Erregung vorhanden war. Und dies spräche ebenfalls eindeutig für die Hypothese, dass physiologische Erregung in irgendeiner Form einer Emotion vorgängig sei und gegen die Hypothese, dass diese Erregung ein Kausaleffekt von Emotion sei. Unter diesem Blickwinkel betrachtet scheinen die Schachter-Singer Experimente noch eine – an dieser Stelle evtl. noch spekulative – Perspektive auf einen möglichen Mechanismus der Emotionsgenese zu eröffnen. Diese Interpretation spricht dafür, dass es ein spontan und vorbewusst ablaufendes Bewertungssystem geben müsste, das Reize vorbewertet und peripher-physiologische Prozesse auslösen kann. Ein zweites Bewertungsprogramm lässt uns körperliche Erregung – also Gefühle – bewusst *als* eine bestimmte Emotion erleben, sofern wir über geeignete Informationen verfügen, die unseren Zustand erklären können, und durch die wir ihn kognitiv als eine Emotion eines bestimmten Typs etikettieren können (die Schachter-Singer-Theorie ist im englischen Sprachraum auch als „Labeling-Theory“ bekannt).

Wie bereits angedeutet, muss aus philosophischer Sicht nun die Frage gestellt werden, welcher Stellenwert der von Schachter und Singer als primär postulierten „Wahrnehmung

eines Reizes“ zukommt. Handelt es sich hier schon um eine komplexe Kognition im Sinne eines Urteils oder sogar eines Schlusses? Mit anderen Worten: löst die Wahrnehmung einer Schlange unmittelbar eine Aktivierung des sympathischen Anteils des vegetativen Nervensystems aus, nach dem Muster einer stimulus-response-Reaktion, oder ist schon diese primäre Bewertung in einem stark kognitiven Sinne als ein „Wahrnehmen-und-Bewerten als“ („appraisal“) oder sogar als ein Urteil oder Schluss zu verstehen? In letzterem Fall würde dann auch in diese primäre Bewertung eines Reizes unser Weltwissen einbezogen werden, weil ohne dieses keine kognitiv gestützte Evaluation denkbar wäre. Dagegen spricht das Beispiel sog. „prinzipiell phobischer“ Objekte, die wir ohne besonderes Wissen über die Welt fürchten. Nach Seligman (1971) haben wir eine gewisse Bereitschaft („preparedness“), mit Furcht auf diejenigen Objekte zu reagieren, die sich in unserer Evolutionsgeschichte als gefährlich herausgestellt haben (vgl. auch Öhman 1987). Diese Objekte bilden demnach die Klasse der potenziell phobischen Objekte.³² Die gleiche Frage stellt sich auch in Fällen, die sich nicht plausibel auf eine evolutionsgeschichtlich bedingte Bereitschaft zurückführen lassen: Das Gefühl der Erhebung oder Wertschätzung beim Anblick einer schönen Landschaft oder eines Gemäldes, das uns spontan gefällt, ist womöglich nicht auf ein zugrundeliegendes Urteil zurückzuführen. Vielmehr könnte dieses primäre Gefühl später die Grundlage für ästhetische Urteile über verwandte Objekte darstellen. In diesem Sinne postuliert Steinfath (2002) die Möglichkeit der Genese emotionaler Zustände auf der Grundlage von einfachen Sinneswahrnehmungen. Es scheint erforderlich zu sein, zur Entscheidung dieser fundamentalen Frage auf eine Klärung des Kognitions-Begriffs hinzuwirken. Dazu soll zunächst die berühmte „Emotions-Kognitions“-Debatte in der Psychologie in ihrem Abriss nachgezeichnet werden. Im Anschluss daran soll versucht werden, mittels einer Arbeitsdefinition die weitere Untersuchung auf eine Grundlage zu stellen, die zumindest im

³² Watson & Rayner 1920 stellten einen Versuch zur Furchtkonditionierung an: der kleine Albert (Versuchsperson, 11 Monate) bekam beim Spielen mit einer Ratte ein unangenehmes Geräusch zu hören und „lernt“ dadurch, mit Furcht gegenüber der Ratte zu reagieren. Anzumerken ist hier, dass Watson & Rayner nur den emotionalen Ausdruck untersuchten (Weinen, Vermeidungsreaktion etc.), und nicht das Erleben. Die konditionierte Furchtreaktion lässt sich relativ leicht wieder löschen, indem die Vpn. ohne Darbietung des unkonditionierten Reizes (Geräusch) mehrfach der Ratte ausgesetzt wird. Seligman 1971 diskutiert die Tatsache, dass der Versuch nicht erfolgreich repliziert werden konnte. Dabei ist auffällig, dass die Replikationsversuche neutrale Reize als konditionierte Reize nutzten, während Watson ein Tier benutzte. Das sieht Seligman als einen Beleg für seine Bereitschafts-Theorie. Dennoch zeigt sich bei Furchtkonditionierung auf ein potenziell phobisches Objekt, dass die Löschung sehr schwierig, wenn nicht unmöglich ist.

Rahmen dieser Studie eine Unterscheidung zwischen kognitiven und nicht-kognitiven mentalen Vorgängen zu treffen erlaubt.

2.5. Die „emotion-cognition“-Debatte in der Psychologie: Richard Lazarus und Robert Zajonc

Robert Zajoncs Artikel „Feeling and thinking: Preferences need no inferences“ (Zajonc 1980) war der Auslöser der sog. Emotions-Kognitions-Debatte in der Psychologie. Nachdem das von Schachter und Singer entwickelte kognitivistische Paradigma jahrzehntelang die Lehrbücher zur Emotionspsychologie dominiert hatte (vgl. Goller 1992, 47), stellte Zajonc in seinem programmatischen Artikel den Kognitivismus wieder in Frage. Richard Lazarus (1982 und 1984) als herausragender Vertreter einer kognitivistischen Emotionstheorie widersprach heftig Zajoncs Ansatz, wobei durch die Beiträge der beiden Kontrahenten eine Debatte ausgelöst wurde, in der es schließlich mehr und mehr um die Frage nach der Bestimmung des Kognitions-Begriffs überhaupt ging. Eine solche Bestimmung stellt auch heute noch in vielen Studien zu Emotionen ein Desiderat dar. Zajonc postulierte in seinem Artikel die weitgehende Unabhängigkeit der beiden Systeme „Emotion“ und „Kognition“. Kognitive Bewertungen seien nicht notwendig für die Emotionsgenese und nicht als primäre Instanz im jeweiligen Prozess der Emotionselizitation anzunehmen. *Emotionen* seien unausweichlich, unwillkürlich, und treten ohne psychische Anstrengung auf. D.h. wir können den Emotionsausdruck kontrollieren, nicht jedoch unser Emotionserleben; wir reagieren mit unseren Emotionen spontan; und wir müssen unser Bewusstsein nicht aktiv auf ein Objekt richten, um eine Emotion bezüglich dieses Objekts zu haben, wir müssen nicht deliberieren und schlussfolgern. *Kognitionen* hingegen, mit der Ausnahme von einfachen Sinneswahrnehmungen, müssen willkürlich gemacht werden und bedürfen einer bewussten intentionalen Ausrichtung auf ein Objekt, mithin also einer bewussten psychischen Aktivität durch das Subjekt.

Nach Zajonc muss ein Reiz nur minimal wahrgenommen werden, d.h. schematisch oder oberflächlich, um eine emotionale Reaktion auszulösen. Diese Instanz – eine Sinneswahrnehmung also – könne jedoch kaum als „Kognition“ im eigentlichen Sinne bezeichnet werden, da der Aspekt der *Information*, der Kognitionen auszeichne, nachrangig sei. Interessant ist, dass Zajonc im von ihm postulierten „Primat des Affektiven“ gerade einen adaptiven, und nicht einen disruptiven Mechanismus unserer Emotionalität sieht. Er geht davon aus, dass ein Organismus primär über sein emotionales System auf einen Reiz reagiert, und

sieht darin einen evolutionären Unterschied zwischen Tieren und Pflanzen, der zugleich einen Selektionsvorteil bildet. Dieser Mechanismus ist adaptiv, weil er Reaktionen auf Umweltreize bereitstellt, die schneller ablaufen als bewusste, kognitive Prozesse und unabhängig von ihnen sind. Um bei der Begegnung mit einer Schlange überleben zu können, muss die bedrohte Maus einen Fluchtimpuls erzeugen können, noch bevor kognitive Prozesse bezüglich der vorliegenden Bedrohung abgeschlossen werden können. Vielmehr muss sie sogar schon dann einen Fluchtimpuls verspüren, wenn nur ein Rascheln im Unterholz wahrgenommen wird, also noch bevor irgendeine Kognition über die Bedrohung zur Verfügung steht. Dieses Argument bezieht Zajonc vor allem auf die Existenz des limbischen Systems, das bei vielen Tieren einen beträchtlichen Anteil an der gesamten Hirnmasse ausmacht. Auch in der phylogenetischen Geschichte des Menschen stellte dieses System den zentralen adaptiven Mechanismus dar, noch bevor Sprache oder höhere Kognitionsprozesse entwickelt waren. Es gibt also einen unbestreitbaren phylogenetischen Primat des Affektiven. Ich denke zwar, dass das Argument prinzipiell auf Menschen übertragbar ist, da unser höher entwickeltes Hirn doch auf alten Strukturen wie dem limbischen System aufbaut. Allerdings darf dabei nicht der faktisch ebenfalls vorhandene Anteil des Kortex an emotionalen Reaktionen des Menschen übersehen werden. Zajonc vertritt jedoch weiterhin die These von der ontogenetischen Vorrangstellung der Emotionen gegenüber Kognitionen, wobei er sich auf Befunde von Izard (1984) beruft. Säuglinge könnten schließlich weinen und lächeln, lange bevor Sprache oder höhere Kognitionen möglich sind. Dennoch würde ich dieses Argument einschränken, z.B. durch die Beobachtung, dass Säuglinge sicher keine Scham empfinden können. Es scheint höhere Emotionen zu geben, die notwendig von bestimmten kognitiven Leistungen abhängen, und die sich durch ihre charakteristische Kopplung mit bestimmten Urteilen oder Gedanken auch unterschiedlich „anfühlen“. Auch wenn Scham, Schuld und Empörung vielleicht jeweils bestimmten anderen Emotionen sehr nahe stehen bzw. deren affektive Aspekte in sich aufnehmen, fühlen sie sich dennoch anders an als ihre jeweils affektiv verwandten Emotionen. Offenbar können Kognitionen auch einen Einfluss auf die Ausdifferenzierung von Gefühlsqualitäten nehmen.

Einige Beobachtungen aus dem Alltagsleben unterstützen den von Zajonc postulierten Primat des Affektiven: So lassen sich „emotionale Urteile“ z.B. nicht leicht aufgeben und scheinen kraft ihrer „affektiven Potenz“ kühl räsionierte Einwände auszustechen. Selbst bei schlagenden Argumenten gegen unser bisheriges Urteil ändern wir nicht leicht unseren emotionalen Eindruck von einer Person. Wenn wir Hans aus eigener Erfahrung für einen liebenswerten Menschen halten und erfahren, dass er angeblich seine Frau schlägt, wären wir

dennoch geneigt, bei unserem bisherigen emotionalen Eindruck zu verbleiben. Wir würden wohl argumentieren, dass eine solch vorbildliche Person wie Hans zu so etwas gar nicht fähig wäre etc., wobei es keine Rolle spielt, dass unser emotionaler „Eindruck“ ein wesentlich schwächeres Argument ist als der faktisch berichtete Fall einer Misshandlung. Im Gegensatz dazu hätten wir keinerlei Neigung, den Satz des Pythagoras zu verteidigen, wenn uns ein Mathematiker ein überzeugendes Argument über seine Falschheit vorführt. Eine ähnliche Beobachtung führt Zajonc an: Stellen wir uns vor, dass wir einer unbekannt Person vorgestellt werden. In der Regel wissen wir innerhalb von Sekundenbruchteilen, ob wir die Person mögen oder nicht, u.z. vermittelt durch unsere Gefühle dieser Person gegenüber. Wir vertreten meist eine große Sicherheit in unserer gefühlsmäßigen Bewertung, der jedoch ein erstaunlicher Mangel an Verbalisierungsfähigkeit von Gründen oder bewussten (kognitiv gestützten) Erwägungen gegenübersteht. Wenn wir gefragt werden, warum wir Hans mögen bzw. nicht mögen, können wir meist nur antworten, dass wir Hans „sympathisch“, „nett“, „angenehm“, oder – im Fall von potentiellen Sexualpartnern – „attraktiv“ finden (bzw. gegenteilig bewerten). Wir finden meist nur *weite* axiologische Prädikate.

Jenseits solcher Beobachtung, die natürlich nur als Illustration dienen können, sieht Zajonc seine Hypothese auch durch experimentelle Ergebnisse gestützt. Bei seinen emotionspsychologischen Experimenten machte sich Zajonc einen zuvor von ihm entdeckten Effekt zu Nutze, nämlich den Effekt von der „bloßen Darbietung“ („mere exposure“, vgl. Zajonc 1968). Dabei konnte gezeigt werden, dass Vpn. Präferenzen setzten, ohne zu wissen, worum es sich bei den zu Grunde liegenden Objekten handelte. Allein die Vertrautheit mit bestimmten Reizen („exposure“) reichte dazu aus. Wie wir in einer Studie von Kunst-Wilson & Zajonc (1980) erfahren, wurden in einem Experiment 24 Vpn geometrische Figuren (unregelmäßige Vielecke) für die Dauer von 1 Millisekunde dargeboten. Dabei handelt es sich um einen sog. „subliminalen Wahrnehmungstest“, d.h. die Darbietungszeit des Stimulus ist so kurz gewählt, dass keine bewusste Wahrnehmung möglich ist. In einem ersten Durchlauf absolvierten die Vpn daraufhin einen Wiedererkennungstest unter normalen Bedingungen. Den Vpn wurden dabei sowohl vorher dargebotene Figuren als auch völlig neue Stimuli zur Unterscheidung durch Wiedererkennen vorgelegt. Dieser erste Test lieferte Ergebnisse auf Zufallsniveau, die Vpn konnten also nicht mit signifikant über Zufall liegender Wahrscheinlichkeit die bereits subliminal präsentierten Figuren von den neuen unterscheiden, das subliminal dargebotene Material wurde also tatsächlich nicht bewusst von den Vpn wahrgenommen. In einem weiteren Durchlauf wurden die Vpn nun aufgefordert, auf der Grundlage des gerade vorgelegten Materials zu bestimmen, welche der geometrischen

Figuren ihnen „besser gefallen“. Die Bewertungen als „besser gefallend“ fielen bei 17 von 24 Vpn (also bei 71% der Probanden) zusammen mit der Menge der alten (subliminal dargebotenen) Stimuli. Die Probanden konnten also sozusagen unbewusst wahrgenommene Reize über ihr emotionales Gedächtnis „wiedererkennen“. Mit anderen Worten: das System der affektiven Etikettierung eines Reizes als „gut/ schlecht gefallend“ liefert einen Wiedererkennungs-Algorithmus, der spontan und unbewusst abläuft. Auch die Experimentatoren bewerten das Ergebnis so, dass Vpn anhand einer emotionalen Bewertung bekannte Stimuli „erinnern“ können, selbst wenn keine bewusste Wiedererkennung möglich ist.³³ Im Sinne der Argumentation von Zajonc weisen auch Birbaumer & Schmidt (1989, 594) darauf hin, dass Stimmungen durch subliminale Wahrnehmung beeinflusst und induziert werden können. Wenn Vpn. emotionsrelevante Reize mit nur 1-5 Millisekunden Dauer dargeboten werden, führen diese zur Veränderung der Stimmung und sogar zum Auftreten starker Emotionen (bei der subliminalen Darbietung von bedrohlichen Gesichtern), ohne dass das dargebotene Material bewusst wahrgenommen wurde oder wiedererkannt werden konnte. Weiterhin unterstreichen sie, dass emotional getönte Geschmacks- und Geruchsaversionen

³³ Bei Kunst-Wilson & Zajonc scheint jedoch eine verdeckte Prämisse vorzuliegen, auf deren Grundlage eine solche Interpretation überhaupt Sinn macht, nämlich dass die subliminal wahrgenommenen Stimuli auf Grund ihrer „Vertrautheit“ im Wahrnehmungstest als „besser gefallend“, d.h. emotional angenehm getönt, markiert werden. In dieser Hinsicht ist die Versuchsanordnung des Experiments von Marshall & Halligan (1988) wesentlich ausgereifter. Damit wäre schließlich zunächst nur die Fähigkeit zu subliminaler Wahrnehmung bewiesen, und nicht der Weg einer unbewussten Emotionselizitation. LeDoux (1996, 58f.) konstatiert, dass für Zajonc offenbar „Präferenzen“ „einfache emotionale Reaktionen“ sind, wobei es sich eben um die von mir erwähnte verdeckte Prämisse handelt. In jüngerer Zeit führten allerdings Murphy & Zajonc (1993) eine weitere Studie durch, die ebenfalls als Beleg für Zajoncs These herangezogen wurde, und die ohne die verdeckte Prämisse auskommt: wieder wird den Vpn ein subliminaler Reiz (ein lächelndes oder ein finsternes Gesicht) dargeboten, auf das dann mit einiger Verzögerung des Zielreiz-Muster folgt, das wiederum mehrere Sekunden, also zur bewussten Wahrnehmung, dargeboten wird. Wiederum sollen die Vpn bewerten, wie ihnen die Zielreize gefallen haben. Dabei zeigt sich nach Murphy & Zajonc, dass diese Bewertung signifikant mit der subliminalen emotionalen Aktivierung zusammenhing. In einem Blindversuch zur Verifikation, in dem sowohl der Aktivierungs- als auch der Zielreiz zur bewussten Wahrnehmung dargeboten wurde, stellte sich der Effekt in der Tat nicht ein. Wir könnten also folgern, dass Zajonc nicht nur nachgewiesen hat, wie subliminale „mere exposure“ – und nicht bewusste Wahrnehmung – unsere Präferenzen beeinflusst, sondern auch, dass dieser Effekt ebenfalls in der Emotionselizitation zum Tragen kommt, was ein Argument gegen den Weg einer bewussten oder deliberativen Emotionsgenese darstellt. Anzumerken bleibt, dass die Zajoncsche Interpretation von Präferenzen als Emotionen nur einen sehr eingeschränkten Emotionsbegriff abgibt.

von Vpn. sogar im anästhesierten Zustand gelernt werden können. Zu diesen Belegen³⁴ können wir noch die Beobachtung hinzufügen, dass die Induktion von Stimmungen und sogar starken akuten Emotionen durch Drogen, Hormone und direkte elektrische oder haptische Stimulation von Hirnarealen in der Hirnforschung völlig unkontrovers belegt ist.

Die Schlussfolgerung, die Zajonc aus seiner Hypothese vom präkognitiven Status von Emotionen zieht, lautet, dass Emotionen nicht nur phylogenetisch das primäre Reiz-Reaktions-System darstellen, sondern auch beim Menschen und auch in ontogenetisch spätere Lebensphasen unsere Reaktionen auf die Umwelt bestimmen. Unbewusste, nicht kognitiv reflektierte emotionale Einstellungen bestimmen nach Zajonc viel häufiger unser Handeln als bewusste rationale Überlegung:

„Der evolutionäre Ursprung emotionaler Reaktionen, der auf ihren Überlebenswert hinweist, ihr Freisein von Aufmerksamkeitskontrolle, ihre Schnelligkeit, die Wichtigkeit emotionaler Unterscheidung für das Individuum, extreme Handlungen, die ein Affekt hervorrufen kann, all dies deutet darauf hin, dass es mit den Emotionen etwas Besonderes auf sich hat. Menschen heiraten nicht oder lassen sich scheiden, morden nicht oder begehen nicht Selbstmord, noch opfern sie ihr Leben der Freiheit lediglich aufgrund einer detaillierten kognitiven Analyse des Für und Wider ihrer Handlungen.“ (Zajonc 1980, 172; Übers. v. Goller 1992)

Als Antwort auf die Herausforderung des Kognitivismus durch Zajonc betont Richard Lazarus (1982 und 1984), dass jeder Emotion eine kognitive Bewertung von Sinneseindrücken voraus gehe. Nach Lazarus sind Emotionen „complex, organized states [...] consisting of *cognitive appraisals*, action impulses, and patterned somatic reactions“ (Lazarus et al. 1980, 198; Hervorh. v. mir, S.H.). Bewertungen und Einschätzungen sind jedoch nicht nur notwendig für das Auftreten einer Emotion, sondern bilden auch einen integralen Bestandteil des emotionalen Zustands selbst. D.h. bestimmte kognitive Zuschreibungen

³⁴ Zajonc führt neben der zitierten experimentellen Studie (Kunst-Wilson & Zajonc 1980) auch einen neuroanatomischen Befund als Beleg für seine These an. Die Existenz einer direkten neuralen Verbindung zwischen der Retina (Netzhaut, also dem Rezeptor von visuellen Sinneswahrnehmungen) und dem Hypothalamus (als wichtigstes vegetatives Hirnzentrum), die sog. „retinohypothalamische Bahn“ (vgl. Nauta & Haymaker in Haymaker/ Anderson/ Nauta 1969 und Moore 1973), sieht Zajonc als Beleg dafür, dass emotionale Reaktionen aufgrund rein sensorischen Inputs, ohne den Einfluss von Denk- und Gedächtnisleistungen des Kortex, direkt vermittelt werden können. Nach heutiger Auffassung dient diese Nervenbahn jedoch hauptsächlich dazu, Informationen über Tageslichteinstrahlung an einen Hypothalamus-Kern (Ncl. suprachiasmaticus) zu vermitteln, auf deren Grundlage dieser den sog. zirkadianen Rhythmus (Wachen-Schlafen-Rhythmus und die davon abhängigen Veränderungen im vegetativen Nervensystem) steuert (vgl. Trepel 1999, 171). Eine Verschaltung zu emotionsrelevanten Hirnzentren liegt nicht vor.

stehen nicht nur am Anfang als Auslöser einer Emotion, sie werden zu einer kritischen Dynamik im emotionalen Zustand selbst. Hier wird zum ersten Mal ein feineres Niveau der theoretische Analyse des Phänomens „Emotion“ erreicht, indem Lazarus Emotionen als dynamische Phänomene begreift und an das Stadium der akuten Emotionselizitation noch die Betrachtung der sog. „coping“-Phase anschließt. Jede Emotionsqualität entsteht durch ein ihr eigenes Bewertungsmuster. Eine Emotion ergibt sich aus der wertenden Wahrnehmung einer tatsächlichen, vorgestellten oder vorweggenommenen Person-Umwelt-Beziehung (die Lazarus „transaction“ nennt). „Jede Emotion hat ihr eigenes kognitives Thema“ (Lazarus et al. 1980, 200): negativ getönte Emotionen entstehen im allgemeinen, wenn eine Transaktion (reziproke Interaktion von Person und Umwelt) als schädlich, bedrohlich oder auf irgendeine Weise als mangelhaft bewertet wird, und wenn die eigenen Coping-Optionen als unangemessen oder ineffizient angesehen werden. Positiv getönte Emotionen entstehen, wenn eine Transaktion als angenehm oder günstig bewertet wird oder wenn die eigenen Mittel bereits zur Kontrolle oder Bewältigung von Schaden erfolgreich eingesetzt worden waren. Weiterhin bilden Verpflichtungen („engagements“) und Überzeugungen über sich und die Welt wichtige Vorbedingungen für emotionale Reaktionen. Überzeugungen und motivationale Konstrukte beeinflussen den Bewertungsprozess. Andererseits können auch motivationale Konstrukte nicht ohne Berücksichtigung kognitiver (und emotionaler) Variablen definiert werden.

Lazarus unterscheidet zwischen drei Formen der Bewertung: (1.) primäre, (2.) sekundäre Bewertung und (3.) Neubewertung (vgl. Lazarus 1982 und 1984: „primary appraisal“, „secondary appraisal“, „reappraisal“).

(1.) Bei einer primären Bewertung wird eine Individuum-Umwelt-Transaktion in Hinsicht auf ihre Relevanz für das Subjekt bewertet, u.z. entweder als „irrelevant“, „angenehm-positiv“ oder „stressbezogen“. Eine irrelevant bewertete Transaktion hat keine Relevanz für das Individuum, eine „angenehm-positiv“ repräsentiert einen für das Individuum wünschenswerten Zustand. Eine stressbezogene Bewertung untergliedert sich weiter in „Schaden-Verlust“-Bewertung (bereits erlittene Verletzungen), „Bedrohung“ (Antizipation von Verletzung) und „Herausforderung“ (tatsächliches, aktuelles schädliches Ereignis).

(2.) Die sekundäre Bewertung ist die Einschätzung der eigenen Bewältigungsmöglichkeiten angesichts einer irgendwie relevant bewerteten Transaktion. Sie spielt natürlich eine große Rolle für die emotionale Reaktion, da die Bewertung einer Situation in besonderem Maße von der Bewertung der eigenen Coping-Optionen abhängt. Dadurch übt die sekundäre Bewertung Einfluss auf die primäre aus.

(3.) Neubewertungen ergeben sich aus Veränderungen in den Kontextbedingungen. Die Umwelt reagiert auf die Aktionen des Individuums, und diese neuen Daten werden erneut Gegenstand eines Bewertungsprozesses, der nach Lazarus kontinuierlich abläuft und nicht notwendig bewusst sein muss. Die Neubewertung kann jedoch auch als Folge einer intrapsychischen Stressbewältigungsstrategie auftreten. In diesem Fall reagiert das Individuum nicht auf Veränderungen der Umwelt, sondern es versucht, eine z.B. als stressbezogen bewertete Transaktion zu bewältigen durch Verleugnung, Abwendung der Aufmerksamkeit oder Rationalisierung. Diese Form der Neubewertung nennt Lazarus „defensive Neubewertung“.

Nach Lazarus kann es keine Emotionstheorie geben ohne Berücksichtigung der Coping-Prozesse, da diese zwischen der Transaktion und der emotionalen Reaktion vermitteln.

Lazarus vs. Zajonc: Die Emotions-Kognitions-Debatte. Die Debatte dreht sich um die Frage, ob Kognitionen Emotionen beeinflussen oder hervorrufen können (Lazarus), oder ob Emotionen unabhängig von Kognitionen sind (Zajonc). Nach Lazarus geht jeder emotionalen Reaktion eine kognitive Bewertung voraus, sei sie auch unbewusst und sehr unspezifisch. Zajonc vernachlässigt die Möglichkeit spontaner, automatischer und unbewusster kognitiver Bewertungen, durch die alle von Zajonc angeführten Phänomene auch im Rahmen einer kognitivistischen Theorie erklärt werden können. Zajonc wiederum hält den von Lazarus verwendeten Kognitions-Begriff für zu weit: wenn dieser Begriff selbst primitive Sinnesreizungen einschließt, kann er selbstverständlich alle emotionalen Phänomene erklären, da in jedem Fall zunächst ein Reiz wahrgenommen werden muss (vgl. auch Spies & Hesse 1986, 76). Damit sei jedoch keine wissenschaftliche Erkenntnis gewonnen, es werde lediglich ein Streit um Definitionen geführt. Zajonc (1984a) plädiert in diesem Sinne für eine empirische Lösung der Debatte, wohingegen Lazarus (1984) anmerkt, dass der gegenwärtige Stand der empirischen Forschung dies nicht zulasse. Ein schlagender empirischer Beweis, ob Emotionen tatsächlich durch irgendeine Form von Bewertungen oder Kognitionen akut ausgelöst werden, erschien zur Zeit der Debatte schwierig zu erbringen. Der Streit dreht sich im Wesentlichen um den Begriff der „Kognition“. Zajonc benutzt „Kognition“ als Bezeichnung für Informationsverarbeitung und deren Ergebnisse, die Transformation des sensorischen Inputs in eine subjektiv verfügbare Form (wenngleich er das Merkmal „bewusst“ nicht für eine notwendige Bedingung von Kognitionen hält):

„I will understand by ‘cognition’ those internal processes that are involved in the acquisition, transformation, and storage of information. These processes always derive from a transformation of some sensory input – immediate or temporally distal –

according to a specifiable code. The code allows the input to be transformed into 'representations'. We sometimes mean by 'cognitions' the products of the above processes, and there tends to be a confusion between 'representations' and 'cognitions' because they are often synonymous." (Zajonc 1984b, 241)

Und an anderer Stelle:

"'Pure' sensory input, untransformed according to a more or less fixed code, is not cognition. It is just 'pure' sensation. Cognition need not be deliberate, rational, or conscious, but it must involve some minimum 'mental work'. This 'mental work' may consist of operations on sensory input that transform that input into a form that may become subjectively available, or it may consist of the activation of items from memory." (Zajonc 1984a, 118)

Lazarus hingegen verwendet einen weiten Kognitions-Begriff, der auch Wahrnehmungen einschließt. Nach Lazarus sind auch Tiere fähig, Kognitionen nach seinem Verständnis des Begriffs zu haben. Kognitive Bewertungen müssen nicht notwendig willentlich, rational oder bewusst sein. Jede Form von Informationsverarbeitung ist Kognition (vgl. Dörner & Stäudel 1990, 298).³⁵ Der von Spies & Hesse (1986) erhobene Vorwurf trifft wohl zu (s.o.).

„We do not have to have complete information to react emotionally to meaning. We can react to incomplete information, which in fact we do in most ordinary transactions. The meaning derived from incomplete information can, of course, be vague; we need to allow for this type of meaning as well as for clearly articulated and thoroughly processed meaning. [...] Cognition cannot be equated with rationality. The cognitive appraisals that shape our emotional reactions can distort reality as well as reflect it realistically. [...] Cognitive appraisal does not necessarily imply *awareness* of the factors in any encounter on which it rests." (Lazarus 1982, 1021f.)

Zu den beiden Verwendungsweisen des Kognitionsbegriffs ist von mehreren Seiten Stellung genommen worden. Dörner & Stäudel (1990, 294) betonen, dass der von Zajonc benutzte Begriff der „Informationsverarbeitung“ nicht hinreichend zur Klärung des Kognitionsbegriffs ist, er ist viel allgemeiner als dieser. Sogar der Lauf eines Automotors könne als Informationsverarbeitungsprozess beschrieben werden, wenn man von den materiellen und energetischen Aspekten des Prozesses abstrahiert. Allerdings ist dies sicher kein kognitiver Prozess. Pekrun (1988) schlägt Begriffsdefinitionen zur klaren Unterscheidung von „Kognition“ und „Wahrnehmung“ vor: (1.) „Kognition“ soll eine Sachverhaltsrepräsentation bezeichnen, die nicht direkt auf sensorischem Input basiert, sondern vom zentralen Nervensystem intern produziert wird (ebd., 107). (1.1.) „Kognition“ im Sinne eines Ergebnisses ist demnach eine interne Sachverhaltsrepräsentation, die nicht

³⁵ Zum Kognitionsbegriff in der Psychologie vgl. Kochinka 2004, 88-110; Goller 1992, 174ff.; weiterhin: Dörner & Stäudel 1990; Pekrun 1988; Mandler 1979 und Neisser 1974.

unmittelbar auf sensorischen Inputs beruht, (1.2.) „Kognition“ im Sinne des Prozesses ist die Produktion von 1.1. auf der Basis von Wahrnehmungen, vorangegangenen Kognitionen oder Gedächtnisinhalten. (2.) „Wahrnehmung“ bezeichnet eine zeitlich und physikalisch direkte Repräsentation sensorisch aufgenommener Information (ebd., 104). Pekrun wendet sich also andersherum gegen einen im Sinne Lazarus „weit“ angelegten Kognitionsbegriff, der auch Wahrnehmungen einschließen würde. Nach Pekrun sind weiterhin auch Vorstellungen Kognitionen, da sie im Gegensatz zu Wahrnehmungen zeitlich und physikalisch nicht direkt repräsentiert, sondern intern produziert werden.

Im Vergleich mit der philosophischen Terminologie scheinen hier zwei sich entsprechende Gegensatzpaare vorzuliegen: Erkenntnis (bzw. Wahrnehmung) vs. Urteil in der Philosophie; Wahrnehmung vs. Kognition in der Psychologie. Diese Paare bilden jedoch *nicht* die gleiche Opposition ab. Wenn wir im Sinne eines weiten Begriffs nach Lazarus eine Vorstellung als Kognition verstehen, dann können Kognition (*psychologisch*) und Urteil (*philosophisch*) nicht dem gleichen Paradigma angehören: eine Vorstellung muss kein Urteil sein. Das wichtige Unterscheidungsmerkmal der von Pekrun vorgeschlagenen Opposition ist „direkte Repräsentation“ vs. „intern produzierte Repräsentation“. Denn das zu Grunde liegende theoretische Interesse zielt auf die Erklärung der kognitiven Leistung eines Organismus in der Emotionsgenese, und eine im Subjekt hervorgerufene Vorstellung, die also nicht aus einem direkten Input der Sinnesorgane stammt, muss eindeutig als eine kognitive Leistung gewertet werden. (1.) Wir können also die psychologische Sprache benutzen für eine erste Stufe der Kognitivismus-Debatte: „Wahrnehmung vs. Kognition“ ist das Begriffspaar, das wir bei der Argumentation für oder gegen eine schwache kognitivistische These benutzen können. (2.) „Imagination vs. Urteil“ würde ich als Begriffspaar vorschlagen, um für oder gegen eine starke kognitivistische These, die sog. Urteilstheorie, zu argumentieren. Denn auf dieser zweiten Stufe gilt zu klären: wenn das Subjekt gewisse kognitive Leistungen in der Emotionsgenese erbringt, welcher Art muss dann die entsprechende Kognition sein? Kann die Imagination eines Objekts oder eines Sachverhalts eine Emotion auslösen (die Vorstellung eines Wolfes), oder bedarf es eines bewertenden Urteils („Wölfe sind gefährlich.“)? Die auf dieser Stufe gestellten Frage scheint wesentlich relevanter zu sein für eine Argumentation für oder gegen den Kognitivismus. Das, was nach Pekrun im Sinne eines psychologischen Kognitionsbegriffs unterschieden wird, nämlich direkte Wahrnehmungsrepräsentation und Imagination, stellt sich unter Anwendung eines genaueren philosophischen Instrumentariums als identisch heraus. Beides steht auf der gleichen Stufe der propositionalen Strukturierung. Die eigentlich relevante Frage, nämlich ob im Prozess der Emotionselizitation propositional

höher strukturierte mentale Phänomene kausal beteiligt sind, lässt sich nur durch das unter (2) angeführte Begriffspaar erfassen.

Alexander Kochinka hat in einem rezenten meta-theoretischen Überblick über verschiedene Kognitionsbegriffe zwischen drei verschiedenen Begriffen in der psychologischen Forschung unterschieden (Kochinka 2004, 88-94):

(1.) Der „enge“ Kognitionsbegriff entstammt der Denkpsychologie, wobei hier zu unterstellen ist, dass Denken und Kognition deckungsgleiche oder zumindest stark verwandte Begriffe sind. In diesem Kontext wird Denken ausschließlich als „Problemlösen“ verstanden (Schaub 1997). In diese Tradition sind weiterhin noch Piagets Arbeiten einzureihen: die bei ihm postulierten „kognitiven Fähigkeiten“, die ein Individuum nach und nach erwirbt, sind alle auf das Problemlösen bezogen.

„Ein enger Begriff der Kognition (mit langer Geschichte) fokussiert also auf das Lösen von Problemen und sieht damit von Fähigkeiten und Phänomenen ab, die wir landläufig ebenfalls mit dem Denken in Verbindung bringen würden.“ (Kochinka 2004, 91)

(2.) Dem „mittleren“ Kognitionsbegriff gemäß werden Kognitionen nicht nur ausschließlich als Akte des Problemlösens verstanden, sondern haben auch ihre Funktion bei der Wahrnehmung, beim Gedächtnis, beim Lernen, bei der Aufmerksamkeit und beim Verstehen und Produzieren von Sprache.

„Der Begriff Kognition wird als Sammelbezeichnung für die geistige Aktivität von Menschen verwendet. In der kognitionspsychologischen Forschung bezeichnet Kognition die Gesamtheit der informationsverarbeitenden Prozesse und Strukturen eines intelligenten Systems [...]. Menschliche intelligente Systeme umfassen Prozesse und Strukturen für Wahrnehmung und Aufmerksamkeit, für Gedächtnis, Denken und Problemlösen, für Lernen sowie für Sprachverstehen und Sprachproduktion.“ (Kluwe 2001, 352)

(3.) Und der „weite“ Kognitionsbegriff wird schließlich wie folgt diskutiert:

„'Kognition' wird hier mehr oder minder gleichgesetzt mit dem Inhalt des Bewusstseins überhaupt. Auf meist nicht eigens bestimmte Weise wird alles zur Kognition, was einem irgendwie ‚durch den Kopf zieht‘ – neben den im mittleren Kognitionsbegriff hinzugekommenen Phänomenen mitunter auch Gefühle und Stimmungen, Motivationen oder der Wille, Träume (Tagträume ebenso wie Alpträume) oder woran auch immer sich noch denken lässt.“ (Kochinka 2004, 92)

Im gleichen Sinne definiert auch Reber (1995, 133) den Kognitionsbegriff („cognition“) als:

„a broad (almost unspecifiably so) term which has been traditionally used to refer to such activities as thinking, conceiving, reasoning, etc. Most psychologists have used it to refer to any class of mental ‚behaviors‘ (using that term very loosely) where the underlying characteristics are of an abstract nature and involve symbolizing, insight, expectancy, complex rule use, imagery, belief, intentionality, problem-solving, and so forth“.

Wir haben also gesehen, wie einerseits die Begrifflichkeiten aus dem semantischen Feld „Kognition“ querzuliegen scheinen in den philosophischen und psychologischen Diskursen. Weiterhin jedoch hat sich gezeigt, dass auch innerhalb eines Fachdiskurses natürlich Differenzen bestehen. Betrachten wir die zuletzt diskutierte Differenzierung des Kognitionsbegriffs in der Psychologie. Der weite Kognitionsbegriff (ad 3) liefert uns kein informatives Klassifizierungskriterium in der Kognitivismus-Debatte, da er allein auf phänomenologische Salienz abhebt. Dieser Begriff würde zur Klassifizierung von Theorien als kognitivistisch bzw. nicht-(anti-)kognitivistisch nichts beitragen, er würde stattdessen die Unterscheidung zwischen aktuellen emotionalen Zuständen („kognitiv“ im Sinne von phänomenologisch salient) und emotionalen Dispositionen („nicht-kognitiv“ im Sinne von phänomenologisch nicht-salient bzw. dormant) erlauben. Dies ist jedoch nicht der Gegenstand der Kognitivismus-Debatte.

Das Unterscheidungskriterium zwischen dem engen (ad 1) und dem mittleren (ad 2) Kognitionsbegriff in der Psychologie nach Kochinka wiederum scheint für den philosophischen oder auch interdisziplinären Diskurs nicht relevant zu sein: während sich beide Begriffe lediglich darin unterscheiden, dass sich (1) dezidiert nur auf „Problemlösen“ bezieht, so haben beide (1,2) dennoch die Gemeinsamkeit, sich auf die „geistige Aktivität“ von Menschen zu beziehen, und damit auf die mentalen Systeme Gedächtnis, Wahrnehmung, Denken, Urteilen und Schließen, sowie linguistische Systeme. Ich denke, dass die genannten Elemente eine sinnvolle Anschlussmöglichkeit für eine Arbeitsdefinition von „Kognition“ für die Zwecke der vorliegenden Studie bieten. Dabei möchte ich v.a. noch einmal auf den Begriff der „geistigen Aktivität“ hinweisen, den wir weiterhin auch bei Reber (s.o.: „mental ‚behavior‘“) finden.

Um ein klares Bild davon zu gewinnen, worum es eigentlich in der Kognitivismus-Debatte geht, sollten wir als erstes Kriterium für unsere Arbeitsdefinition festhalten, dass kognitive Prozesse eine mentale Aktivität seitens des Subjekts implizieren. Betrachten wir die Parteien des Kognitivismus-Disputs in der Psychologie (bspw. Lazarus und Zajonc) oder in der Philosophie (bspw. Solomon und Steinfath), so wird deutlich, dass das Kriterium Aktivität vs. Passivität als latentes Fundament dem eigentlichen Kriterium kognitivistisch vs. nicht-kognitivistisch unterliegt. Zajonc und Steinfath plädieren dafür – auf je ihre Weise –, dass wir den Mechanismus der Emotionselizitation auf der Basis von einfachen Sinneswahrnehmungen verstehen können, während Lazarus und Solomon dafür Bewertungen bzw. Urteilsprozesse reklamieren. In der Tat scheinen Wahrnehmungsprozesse ohne „mentale Aktivität“ des Subjekts abzulaufen, u.z. insofern diese spontan, mühelos – d.h. ohne Anstrengung –, und

unwillkürlich ablaufen. Wenn Hans' Blick auf die Schlange vor seinen Füßen fällt, und Hans entsprechend disponiert ist, dann wird er mit Furcht reagieren: spontan und unwillkürlich. Im Gegensatz dazu erfordern Prozesse des Bewertens und Urteilens „mentale Aktivität“, insofern diese willkürlich ablaufen und vom Subjekt „angestoßen“ werden müssen. Zwar können sich auch diese Prozesse hochgradig habitualisieren: Ideen können uns „kommen“, scheinbar ohne dass wir uns anstrengen, indem wir schlussfolgerten, überlegten oder urteilten. Dabei ist jedoch nicht zu übersehen, dass auch Ideen niemals „uneingeladen“ kommen. Wenn mentale Aktivitäten habituell ablaufen, dann ist die willkürliche Aktivität des Subjekts in jedem Fall auf einer die Habitualität vorbereitenden Ebene zu verorten.

Das Kriterium des „strengen Kognitivismus“ soll für unsere Zwecke also eine willkürliche (und nicht spontane) mentale Aktivität des Subjekts bedeuten, also Prozesse des aktiven Bewertens, Urteilens und Schlussfolgerns umfassen, nicht jedoch mentale Aktivität im Sinne von einfachen Sinneswahrnehmungen.

2.6. Logisch-konzeptuelle Kritik des Kognitivismus

Nach den bisherigen Ausführungen zu Fragen der Aktualgenese der Emotionen, in denen immer auch die jeweiligen Probleme kognitivistischer und nicht-kognitivistischer Theorien zur Sprache gekommen sind, scheint an dieser Stelle eine Zusammenfassung und Bündelung der dargestellten Kritik des Kognitivismus nötig zu sein. Dabei sollen auch mit Bezug auf diagnostizierte Problemlagen Alternativvorschläge präsentiert werden. Die im Folgenden aufgeführten Musterfälle sollen als ein Argumentkatalog gegen einen „schwachen“ Kognitivismus dienen. Für den weiteren Argumentationsgang der vorliegenden Arbeit wäre nichts gewonnen, Einwände gegen den starken Kognitivismus, also gegen eine Urteilstheorie, zu formulieren. Zur Debatte steht vielmehr die Frage, ob Kognitionen als in der Emotionselization kausal wirksame Faktoren angenommen werden können. Dabei sollen nun die begrifflichen Probleme eines solchen Standpunktes gesammelt und erörtert werden.

Anhand von Aristoteles' Rhetorik, den Ausführungen von Christof Rapp (2002) folgend, lassen sich einige begriffliche Probleme des Kognitivismus treffend illustrieren. Aristoteles liefert insofern ein gut geeignetes Beispiel, als ihm in der Forschung oft eine streng-kognitivistische Urteilstheorie zugeschrieben wird. Wenngleich Aristoteles' Emotionstheorie (in der Rhetorik) klar kognitivistisch zu nennen ist, scheint jedoch die Zuschreibung einer Urteilstheorie unhaltbar, da Aristoteles selbst explizit eklatante Problemfälle einer solchen Theorieform vorführt.

(1.) *Problem der Kontrarietät bzw. Kontradiktion* bestimmter Emotionen. Nach Aristoteles schließen sich bestimmte Emotionen, und damit auch das Haben entsprechender Urteile seitens des Subjekts im Falle einer kognitivistischen Emotionstheorie, gegenseitig aus. So empfiehlt Aristoteles in seiner Rhetorik, einer Schrift, die sich an die Gerichtsredner wendet, dass ein Redner, der Mitleid mit dem Angeklagten ausschließen möchte, vorsorglich die Richter in Empörung versetzen solle:

„Wenn daher die Rede die Richter in eine solche Stimmung versetzt und den Nachweis erbringt, dass die, die den Anspruch auf Mitleid erheben, wie auch die Dinge, worüber sie Mitleid fordern, unwürdig sind, es zu erlangen, bzw. würdig, es nicht zu erlangen, so ist es unmöglich, Mitleid zu empfinden.“ (Rhet. II.9, 1387b14ff.)

Hier handelt es sich noch nicht um einen besonders harten Fall für eine kognitivistische Theorie, denn man kann zu Recht annehmen, dass ein wohlbegründetes Gefühl von Empörung zugleich ein Gefühl von Mitleid ausschließt, ebenso wie das Urteil, dass es dem Angeklagten unverdientermaßen gut ergeht (Empörung), das Urteil, dass es ihm unverdientermaßen schlecht ergeht (Mitleid) ausschließen würde. Hier handelt es sich schließlich um kontradiktorische Urteile. Allerdings behauptet Aristoteles weiter, dass sich auch emotionale Zustände gegenseitig ausschließen können, selbst wenn sie *nicht* auf die gleichen Objekte gerichtet sind:

„Weiterhin ist man sanftmütig gegen die, die man fürchtet oder vor denen man sich schämt; denn *solange man sich in solcher Verfassung befindet, vermag man nicht zu zürnen*; es ist nämlich unmöglich, gleichzeitig zu fürchten und zu zürnen.“ (Rhet. II.3, 1380a31ff., Hervorh. v. mir, S.H.)

Ich möchte dies durch ein anderes Emotionspaar illustrieren: Die Emotionen Liebe und Furcht schließen sich gegenseitig aus, und zwar auch dann, wenn sie sich auf unterschiedliche Objekte beziehen. Es ist einleuchtend, dass sich die beiden Urteile „Person P will mir Gutes“ (Liebe) und „P stellt eine Bedrohung dar“ (Furcht) gegenseitig ausschließen. Aristoteles These ist jedoch stärker: Wenn ich P fürchte, kann ich nicht gleichzeitig akut „Liebe“ für S empfinden, da beide Emotionen sich gegenseitig ausschließen. Implizit habe ich in meiner eigenen Illustration „akut“ ergänzt. Denn diese Kontrarietät von emotionalen Zuständen bezieht sich in der Tat nur auf akute Emotionen, und nicht auf emotionale Dispositionen. Hans kann sich also problemlos vor dem Ansteigen des Ölpreises fürchten (weil er der Betreiber einer Tankstelle ist), und dennoch Susanne „lieben“, in dem Sinne wie man „lieben“ hier für eine Disposition bzw. ein Persönlichkeitsmerkmal benutzt. Dennoch stellt die Beobachtung von Aristoteles eine enorme Herausforderung an den Kognitivismus dar. Selbst im Rahmen eines gemäßigten Kognitivismus, der z.B. Überzeugungen oder Urteile als die

auslösenden Momente in der Emotionselizitation versteht, lässt sich nicht leicht theoretisch erklären, warum die emotionalen Reaktionen, die mit zwei logisch-propositional völlig kompatiblen Überzeugungen verbundenen sind, sich gegenseitig ausschließen. Wenn tatsächlich Kognitionen allein emotionale Zustände auslösen, dann fragt sich, warum eine Bewusstmachung entsprechender Kognitionen nicht zur entsprechenden emotionalen Reaktion führt, wenn keine unvereinbaren Kognitionen vorliegen. Allein die Tatsache, dass sich Hans schon im Zustand der Angst befindet, scheint in unserem Beispiel zu verhindern, dass Hans aktuell Liebe für Susanne empfinden kann. Zurückzuführen ist dieses Phänomen sicherlich auf die bewusstseinssteuernde Funktion von Emotionen. Wilhelm Wundt hatte von einer „phänomenologischen Potenz“ von Affekten gesprochen, kraft derer diese stets in den Vordergrund des Bewusstseins drängen (Wundt 1909, 243f.). Was sich jedoch im phänomenologischen Vokabular lediglich beschreiben lässt, das vermag Ronald DeSousas Theorie von Emotionen als „paradigm scenarios“ sehr überzeugend zu erklären.

Nach DeSousa besteht die oft beobachtete Ähnlichkeit zwischen Emotionen und Überzeugungen darin, dass erstere eine *urteilsbildende Funktion* haben. Insofern hängen sie mit Glaubenszuständen zusammen. Emotionen „stellen diejenigen Fragen, die ein Urteil mit Überzeugungen *beantwortet*“ (vgl. u.). Insofern vergleicht DeSousa Emotionen mit wissenschaftlichen Paradigmen im Sinne von Thomas Kuhn (1962): sie sind etwas, in dessen Sinne wir die Welt sehen. Emotionen hängen also eng mit Urteilen und Überzeugungen zusammen, sie sind jedoch weniger von solchen getragen, als sie selbst ein Paradigma darstellen, das bestimmte Überzeugungen unter sich zu versammeln tendiert bzw. bestimmte Überzeugungen nach sich zieht. Emotionen bedingen, was wir wahrnehmen und welche Urteile wir fällen, indem sie unsere Aufmerksamkeit steuern: sie sind „patterns of salience“. Darum erlaubt also ein aktueller Zustand von Furcht keine emotional wirksame Kognition, die mit Liebe korreliert, weil eben das Aufmerksamkeitsmuster – und damit auch der Modus des Urteilens, Wahrnehmens und Denkens – durch den Furcht-Zustand geprägt ist.

(2.) *Problem der ungerechtfertigten Beilegung von Emotionen.* Aristoteles erwähnt hier folgenden Fall:

„Ebenso setzt die an einem anderen vorher genommen Rache dem noch größeren Zorn gegenüber einem Dritten ein Ende. [...] Denn man wird sanftmütig, wenn man seinen Zorn an einem anderen ausgelassen hat, was dem Ergophilos zugute kam, denn obwohl man ihm heftiger zürnte als dem Kallisthenes, erließ man ihm doch die Strafe, weil man am Tage zuvor den Kallisthenes zum Tode verurteilt hatte.“ (Rhet. II.3, 1380b7ff.)

Die Beilegung des Zorns gegenüber Ergophilos folgt hier nicht aus dem Fassen eines Urteils, dass dieser sich doch nichts habe zu Schulden kommen lassen, oder aber dass seine Tat angemessen gesühnt worden sei. Die Richter vertreten wohl nach wie vor die gleichen Überzeugungen gegenüber der Tat und dem Beklagten. Zudem zürnten sie angeblich dem Ergophilos sogar noch mehr als dem Kallisthenes, der schließlich zum Tode verurteilt wurde. Eine kognitivistische Theorie sähe sich hier vor dem Problem, erklären zu müssen, wie ein emotionaler Zustand beendet werden kann, ohne dass irgendeine Veränderung in den zugrundeliegenden Kognitionen aufgetreten wäre.³⁶ Offenbar liegt hier ein Fall ähnlich dem aus der Psychologie bekannten Erregungstransfer-Paradigma vor: das Abklingen des Zorns in den Richtern (gegenüber Kallisthenes, der seine gerechte Strafe erfahren hatte; wobei jedoch in diesem Erklärungsmodell das Objekt der Emotion keine Rolle mehr spielt), das einem Abklingen der affektiven und physiologischen Erregung entspricht, scheint irgendwie nachzuwirken, sodass am folgenden Tag der Zorn gegenüber Ergophilos plötzlich nicht mehr in der vorherigen Schärfe erlebt wird, u.z. eben aus dem Mangel an Erregung, bzw. der Wahrnehmung dieses Mangels. Das veränderte Gefühl wirkt sich also – umgekehrt, als dies die kognitivistischen Emotionstheorien behaupten – auf die Urteile der Richter gegenüber Ergophilos aus, sodass dieser freigesprochen wird. Auch dieser Fall ist im Rahmen einer kognitivistischen Theorie nicht zu erklären.

(3.) *Problem der notwendigen physiologischen Voraussetzungen von Emotionen.* Aristoteles vertritt die These, dass emotionale Reaktionen auch vom Charakter des Subjekts abhängen. Unter den charakterbildenden Faktoren diskutiert er auch das Alter. Demnach haben die Jungen starke Emotionen, die bei ihnen auch handlungswirksam werden, die Alten dagegen erleben ihre Emotionen schwächer, legen sie schnell bei und handeln nicht mehr aus ihnen heraus. Dies liege u.a. daran, dass die Alten „erkaltet“ seien, während die Jungen „erhitzt“ sind (Rhet. II.13, 1389b29-32). Ein alter Mensch wird also nicht die gleichen Emotionen empfinden wie ein junger, evtl. wird er gar keine Emotion empfinden, wo ein junger dies sehr wohl tut, selbst wenn beide die gleichen Urteile bzw. Überzeugungen über das Objekt der Emotion vertreten. Es gibt demnach keine hinreichende Kausalbeziehung zwischen Kognitionen und Emotionen, wenn nicht die notwendige Bedingung gewisser

³⁶ Hier handelt es sich also um den umgekehrten Fall von „emotionaler Trägheit“, nämlich dem Fortdauern eines emotionalen Zustands angesichts veränderter Kognitionen, die diesen Zustand nicht mehr tragen dürften. Hans kann Hermann weiter zürnen selbst wenn sich herausstellt, dass Hermann überhaupt nicht verantwortlich ist für eine Missetat, die Hans diesem zuschreibt.

physiologischer Faktoren gegeben ist. Wir haben im Laufe der obigen Ausführungen andere Beispiel für diese Fragestellung diskutiert, nämlich die sog. deafferenten Patienten (im Rahmen der James-Lange-Theorie, aber auch nach Dana 1921 und Hohmann 1966 bspw.; weiterhin bestimmte ähnliche gelagerte Fälle wie Damasio „frontal-lobe-patients“, s.u.). Auch dies übersteigt den explanatorischen Rahmen einer kognitivistischen Theorie.

(4.) *Problem der Persönlichkeitsmerkmale als Mediatorvariablen.* Unterschiedliche Persönlichkeitsmerkmale führen zu unterschiedlichen emotionalen Reaktionen. Bei Aristoteles spielt hier der (moralische) Charakter eine große Rolle: die emotionale Reaktion eines tugendhaften unterscheidet sich von der eines schlechten Menschen, denn, wie Aristoteles sagt, „man muss denen gegenüber, die unverdient Unglück erleiden, Mittrauer und Mitleid empfinden, aber ebenso Unwillen empfinden über die, denen es unverdienterweise wohl ergeht“, und weiter: „z.B. wird kein Rechtschaffener über Vaternörder und Mordsüchtige Leid empfinden, wenn sie ihre Strafe erhalten“ (Rhet. II.9, 1386b11ff.). Auf den ersten Blick scheint es hier doch nur darum zu gehen, welche Urteile man gegenüber den entsprechenden Personen vertritt: ob sie Unrecht begangen haben oder nicht, und ob es sich um ein Vergehen von besonderer Schwere, um eine Schandtät handelt, wie im Beispiel des Vaternörders. Diese Sichtweise wäre aber allzu simplizistisch und abstrahiert von dem Subjekt, um das es hier geht. Denn um souverän und korrekt beurteilen zu können, ob es sich um ein Verbrechen oder eine Schandtät handelt oder nicht, das hängt für Aristoteles davon ab, ob das urteilende Subjekt ein tugendhafter Mensch ist oder nicht. Wer durch und durch verdorben ist, wird vielleicht einen Mord zur Erreichung irgendeines Zieles gar nicht für ein Vergehen halten. Emotionale Reaktionen korrelieren also gar nicht direkt mit bestimmten Urteilen, sondern hängen auch noch von den Persönlichkeitsmerkmalen des urteilenden Subjekts ab. Wir können es also als ein Desiderat der gängigen kognitivistischen Emotionstheorien ansehen, auch Elemente wie die Sozialisation, den moralischen Charakter und die Persönlichkeitsmerkmale des Subjekts zu berücksichtigen, um emotionale Reaktionen auf der Grundlage von Kognitionen erklären zu können. Damit scheint sich ein weites Feld aufzutun. Dennoch zeigen Forschungen zum Erwerb emotionaler Dispositionen, wie z.B. DeSouzas Modell von Schlüsselszenarien („paradigm scenarios“), narrative oder psychoanalytische Ansätze bspw. von Wollheim (1999) oder Goldie (2002), oder aber Modelle des neuronalen Lernens von emotionalen Mustern (nach Descartes 1649, Damasio 1994 und LeDoux 1996 bspw.), dass diese hochkomplexe Grundlage unserer Emotionalität mit angemessenen theoretischen Werkzeugen zu bewältigen ist.

(5.) *Das Intentionalitätsproblem.* In der kognitivistischen Emotionsanalyse herrscht breiter Konsenz bezüglich der Auffassung, dass Emotionen in irgendeiner Form auf Objekte „gerichtet“ sind. Dies wird häufig auch als Unterscheidungsmerkmal zwischen Emotionen und (Körper-) Gefühlen angeführt. Das körperliche Gefühl der Müdigkeit bswp. ist auf kein Objekt gerichtet, also nicht intentional, während die Emotion „Angst“ auf ein Objekt gerichtet ist, z.B. bei der „Angst vor der Abschlussprüfung“. Lassen wir zunächst quasi-pathologische Fälle wie ungerichtete Angstzustände, Depressionen o.ä. außer Acht, so scheint diese Annahme auf den ersten Blick sehr überzeugend zu sein.

„Intentionality is a property of actions and mental states. It is the property of being directed at or toward something. Emotions typically have this property. When one is angry or afraid, for example, one is angry at someone or something, afraid of someone or something. This someone, this something is the emotion's intentional object, that at or toward which it is directed. By contrast, bodily sensations of pleasure and pain, the comforting feeling or a warm bath, say, or the aching feeling of sore muscles, are not directed at or toward anyone or anything. They are not intentional states. Hence, a conception of emotion that identifies the phenomenon with feelings like these misrepresents it.“ (Deigh 1994, 826)

Nach Gordon (1987, 23f.) kann man weiterhin unterscheiden zwischen zwei Typen von emotionalen Zuständen, nämlich $E1(S,p)$, der eine Emotion $E1$ bezeichnet, die ein Subjekt gegenüber dem propositionalen Inhalt p erlebt: Hans bedauert, *dass er das Auto seines Nachbarn verkratzt hat*. Beim Typ $E2(S,x)$ handelt es sich um einen vom Subjekt erlebten emotionalen Zustand gegenüber einem Objekt x :³⁷ Hans hat *Angst vor der Abschlussprüfung*. Zu beachten ist hier, dass Gordons Beobachtungen auf einer sprachanalytischen Grundlage zu verstehen sind. In einem weiteren Sinne können wir natürlich auch propositionale Inhalte als „Objekt“ einer Emotion verstehen.³⁸ Was nun die Kategorie des Objekts angeht, so wurden viele Vorschläge zu dessen Kategorisierung gemacht. Einen dezidierten Kognitivismus

³⁷ Vgl. in diesem Sinne auch Birnbacher (1998, 104). Dort wird zunächst die aus der neueren analytischen Philosophie bekannte Unterscheidung zwischen „knowledge by acquaintance“ und „knowledge by description“ aufgegriffen, wobei es sich beim Letzteren um propositionales Wissen, beim Ersteren um nicht-propositionales Wissen handelt. Seiner Ansicht nach gibt es auf der Grundlage dieser Unterscheidung drei Klassen von Emotionen in Bezug auf deren intentionales Objekt: (1.) solche, die nur gegenständlich Objekte haben (z.B. Liebe), (2.) solche, die nur propositionale Objekte haben (z.B. Ärger) und (3.) solche, die Objekte beider Arten haben können (z.B. Angst).

³⁸ Vgl. in diesem Sinne auch Petrus 2000, 131: „Emotionale Eigenschaften sind intentionale Eigenschaften.“ Sie haben einen Modus (z.B. „... ist traurig“) und einen Inhalt („dass es regnet“). Mit Bezug auf Gordon denkt Petrus, dass der Inhalt nicht immer als „dass“-Satz auftreten muss, dass es sich hierbei jedoch nur um (oberflächen-) sprachliche Unterschiede handele.

zeichnet nun die Annahme vom „formalen Objekt“ eines bestimmten Emotionstypus aus. Dabei wird angenommen, dass jedem Typ von Emotion ein entsprechendes formales Objekt zugeordnet werden könne, „Angst“ bspw. richtet sich demnach auf „das Gefährliche“.³⁹ Über eine solche Kategorisierung des formalen Objekts sollen nun nicht nur trennscharfe Katalogisierungen von Emotionstypen vorgenommen werden können, sondern auch Rationalitätskriterien für Emotionen formuliert werden. Erinnern wir uns, dass der starke Kognitivismus (die Urteilstheorie) durch ihre Gleichsetzung von Emotionen und Urteilen die Rationalität von Emotionen schlicht an die Rationalität der entsprechenden Kognitionen zurückbinden kann. Der schwache Kognitivismus vermag dies nun über die Kategorie des formalen Objekts zu leisten, indem von der „Angemessenheit“ („appropriateness“) oder der „Wohlbegründetheit“ („well-foundedness“, Lyons 1980) gesprochen wird.

Das *intentionale* Objekt – so Deigh (1994, 834) ist demnach eine konzeptuelle Beschränkung der Klasse von Objekten, die das *materiale* Objekt einer spezifischen Emotion bilden können (vgl. Kenny 1963, 187ff. und Döring & Peacocke 2002). Setzen wir diesen Gedanken fort, müssen wir sagen, dass die Instanz, die etwas als ein adäquates Objekt einer spezifischen Emotion qualifiziert, der *Glaube* des Subjekts ist, dass jenes Objekt bestimmte Eigenschaften hat. Damit wäre also propositional strukturiertes Denken eine notwendige Vorbedingung für die Aktualgenese von Emotionen.

Deigh formuliert eine Kritik an dieser Auffassung von Kenny, die ich als eine Art „Meta-Theorie-Vorwurf“ verstehe: was diese Theorieschule als Wirkfaktor in der Emotionselizitation annimmt, ist in Wirklichkeit Teil unserer *Interpretation* von Emotionen, jedoch weder konstitutiv noch kausal für eine Emotion als solche. Ein Objekt kann ein intentionales Objekt sein, ohne dass das Subjekt irgendwelche Annahmen oder Glaubenszustände diesem gegenüber vertreten müsste. Ein Säugling bspw. kann mich anschauen, und ich bin dann das Objekt dieses Blicks. Daraus folgt jedoch nicht, dass der Säugling irgendwelche Annahmen über mich vertritt. Zwar muss ich – als Objekt eines Blickes – *sichtbar* sein, das ergibt sich aus der logisch-begrifflichen Analyse der Tatsache, dass ich angeschaut werde. Und dennoch kann mich der Säugling anschauen, ohne dass *dieser selbst* die Annahme vertritt, dass ich sichtbar bin. Hier zeigt sich m.E. die Vermischung von Theorie und Meta-Theorie.

In der Debatte um die Rationalitätskriterien des formalen Objekts wird also m.E. (und offenbar auch gemäß Deighs Standpunkt) nicht unterschieden, ob es dabei um eine

³⁹ In diesem Sinne Solomon 1984, 57: „Emotions are distinguished, at least in part, by what they are ‚about‘, their so-called intentionality.“

nachträglich Bewertung der Rationalität von fremden oder (retrospektiv betrachteten) eigenen emotionalen Zuständen handelt (also einer *hermeneutischen* Perspektive), oder aber ob die Kriterien des formalen Objekts Einfluss auf die *Aktualgenese* von emotionalen Zuständen selbst haben. Mit anderen Worten: muss bspw. ein Säugling ein Objekt kompetent als „gefährlich“ bewerten oder kategorisieren können, damit wir ihm die Fähigkeit zuschreiben, kompetente und angemessene Furchtreaktionen zeigen zu können? Der Impetus dieser Kritik verweist auf eine andere Auffassung der Intentionalität von Emotionen, nicht im Sinne eines rationalistisch verstandenen „formalen Objekts“, sondern im Sinne einer einfachen mentalen Repräsentation. Nach dieser Auffassung sind Emotionen intentional, weil sie qua mentaler Repräsentationen auf ein Objekt, nämlich den auslösenden internen oder externen Reiz gerichtet sind.

„Intentionality, in the sense of being directed at an object, is a property of emotions whether or not their subject possess reason. The emotions of antelope, for instance, though neither rational nor irrational, are not objectless states; [...] Hence, cognitivists must be using the word ‘intentionality’ in some altered sense when they exclude emotions like these from their studies on the grounds that they lack the necessary intentionality. In this altered sense, the word describes a property of mental states that only rational creatures can experience.” (Deigh 1994, 848)

Die Kategorie des formalen Objekts beschreibt demnach keine genetische Perspektive von Emotionen, sondern eine hermeneutische; das formale Objekt wäre demnach durch diesen „schwachen“ Intentionalitätsbegriff nicht widerlegt, sondern nur auf einen anderen Bereich der Gesamtheorie verwiesen.

Das, was ich als eine Vermengung von Theorie und Meta-Theorie bezeichnet habe, nennt Deigh einen „sokratischen Kognitivismus“ (Deigh 1994, 840; den Bezug auf den platonischen Sokrates sieht Deigh in Protagoras 358d-e). Der Fehlschluss liegt dabei in folgender Annahme: Wenn wir mit einer Fähigkeit zur Empfindung von Angst geboren werden, dann offenbar auch mit der Fähigkeit, Gefahr zu detektieren. Als Schlussfolgerung müssen wir demnach *wissen*, was „Gefahr“ bedeutet. Nach Deigh impliziert das Haben eines bestimmten Empfindungsvermögens bzw. einer bestimmten dispositionalen Reaktivität jedoch *nicht* notwendig das Haben entsprechender Begriffe. Schließlich könnten Wildgänse Wetterveränderungen spüren, ohne darum auch den Begriff von „Jahreszeit“ zu haben.

Perler (2000, 83f.) macht den Vorschlag, dass man Emotionen so verstehen könnte, dass sie zwar einen phänomenalen Gehalt haben (also „Qualia“) sind, aber keinen intentionalen. Allerdings können Emotionen sehr wohl mit einem intentionalen Zustand verbunden sein, etwa indem ich wütend bin und gleichzeitig an Hans denke, auf den ich meine Wut richte, oder sie können auch mit einem intentionalen und kognitiv gehaltvollen Zustand verbunden

sein, etwa indem ich wütend bin und gleichzeitig schlussfolgere, dass das wohl an den unverschämten Bemerkungen von Hans liegt.⁴⁰

Im Gegensatz zu Perlers Vorschlag hat sich Peter Goldie jüngst in der Linie einer phänomenologischen Sichtweise positioniert, ohne jedoch den analytischen Anspruch der Intentionalität aufzugeben:

„I will argue that emotions involve two kinds of feeling, and that these two kinds are sometimes run together in our ordinary way of speaking of emotional experience, as for example when we say that I felt very afraid. First, there is what I will call bodily feeling, the feeling from the inside of the condition of one's body as being a certain way or as undergoing certain changes. For example, in fear I feel the hairs go up on the back of my neck. Secondly, there is what I will call feeling towards, the feeling one has towards the object of one's emotion. For example, in fear I feel the dangerousness of the lion. [...] Both, I will argue, are intentional.“ (Goldie 2002a, 235)

Während üblicherweise zwischen einem nicht-intentionalen (phänomenalen) Gefühlsanteil und einem anderen, intentionalen Anteil im Phänomen der Emotion unterschieden wird, fasst Goldie den Begriff der Emotion als vollständig durch Gefühle konstituiert auf. Damit wird ein Anschluss an phänomenologische Gefühlstheorien (vgl. den Überblick bei Coriando 2002) unternommen, der dennoch eine Form von Intentionalität im Sinne eines Gerichtet-Seins („directedness“) bereitstellen kann. Auch hierin kann eine mögliche Lösung des kognitivistischen Intentionalitätsproblems gesehen werden, allerdings erfordert ein solcher Ansatz wesentlich weitreichendere Explikationen, als dies im vorliegenden Kapitel geschehen kann. An späterer Stelle soll noch einmal vertieft die mögliche Intentionalität von Gefühlen, Sinneswahrnehmungen und Körperempfindungen diskutiert werden (vgl. Kap. 3.10.).

(6.) *Das Problem der Inkommensurabilität von Kognitionen und Emotionen.* Es leuchtet unmittelbar ein, dass eine Identifizierung von emotionalen und kognitiven Zuständen (Urteilen oder Überzeugungen) an zwei wichtigen Unterschieden scheitern muss: Emotionen sind dynamisch (Kognitionen nicht), und Emotionen werden „erlebt“ (Kognitionen ebenfalls nicht). Sowohl Kupperman als auch Solomon (letzterer als der prominenteste Vertreter eines starken Kognitivismus) unterstreichen, dass Emotionen einen phänomenalen Erlebnisaspekt haben, während Kognitionen nicht in dem gleichen Sinne als „qualia“ erlebt werden. Doch ich will mich in diesem Argumentkatalog auf den schwachen Kognitivismus beschränken. Auch hier tun sich Probleme auf. Wie ist nämlich die Klasse derjenigen Kognitionen zu spezifizieren, die überhaupt eine emotionsauslösende Wirkung haben können? Warum

⁴⁰ Perler merkt jedoch einschränkend an, dass es nicht „die“ Emotionen gibt (2000, 85); vgl. auch Amélie Rorty's Bemerkung (1980a, 1 und 1988, 104): „Emotions do not form a natural class.“

reagiert niemand emotional auf die aktuelle Repräsentation der Kognition „Schnee ist weiß“, während ein verheirateter Mann auf die Überzeugung „Meine Frau hintergeht mich“ wohl mit großer Wahrscheinlichkeit emotional reagieren wird. Eine bereits vorgestellte Strategie, bspw. von Solomon, Lyons oder Nussbaum, besteht darin, (1.) die fragliche Klasse auf die Werturteile einzuschränken, und dies noch weiter auf die relativ *intensiven* Werturteile.⁴¹ Eine weitere Möglichkeit besteht darin, (2.) Mengen von Kognitionen zu benennen (so in einem hochkomplexen ontologischen Theorierahmen bei Gessner 2004). Hierbei bleibt jedoch fraglich, ob eine solche Katalogisierung prinzipiell erschöpfend geleistet werden kann. Ein weiterer Vorschlag (der vermeintlich mit (1) identisch ist) wurde in jüngerer Zeit von Goller und LeDoux vertreten, (3.) Emotionen als Reaktionen auf sog. „heiße Kognitionen“ zu verstehen (Goller 1992, 166; LeDoux 1996, 43; der Begriff stammt von Abelson 1963). Nur oberflächlich wird man „heiße Kognitionen“ mit „intensiven Werturteilen“ gleichsetzen wollen. Intensive Werturteile lassen sich auch ad hoc fällen. Wie ich den Begriff der heißen Kognition verstehe, muss es sich dabei um eine Überzeugung, Einstellung oder ein Urteil handeln, das erstens *habitualisiert* vorliegt, also aus einem gewissen (spirituellen, nicht mathematischen) Lernprozess des Subjekts zu erklären ist, und das zweitens eine *Adressatenqualität* hat, also in Inhalt (in Bezug auf die involvierten Personen oder Objekte) oder Struktur (als Imperativ, Sollen-Satz o.ä.) auf das Subjekt gerichtet ist. Wir können also in diesem Punkt (6) eindeutig eine Parallele zu Punkt (4) erkennen: beide heben auf die „Vorgeschichte“ des Subjekts eines emotionalen Zustands ab, bei beiden geht es um die Erklärung von Prozessen des Lernens oder der Habitualisierung. Ich denke, dass diese beiden Punkte die vielversprechendsten Voraussetzungen bieten, bei angemessener Überarbeitung die Probleme des Kognitivismus zu überwinden.

Betrachten wir nun v.a. noch einmal die unter (5) und (6) dargestellten Problemfälle des Kognitivismus, die sich darauf bezogen zu bestimmen, wie Emotionen genetisch mit Kognitionen zusammenhängen, und aufzuzeigen, worin genau die unleugbare Intentionalität

⁴¹ Diese z.Zt. sehr beliebte Strategie wird von Vertretern kognitivistischer Theorien oft als Ausweg aus den Problemen eines strengen Kognitivismus im Sinne einer Urteilstheorie angesehen. Die Umdeutung einer Urteils- in eine Werturteils-Theorie kann jedoch nicht als eine Lösung gelten. Denn, wie Clarke (1986, 664) scharfsinnig und bisher offenbar unrezipiert angemerkt hat, verpflichtet die obige Annahme kraft der Supervenienz von Werten auf natürliche Eigenschaften im gleichen Zug dazu, nicht-wertende Urteile als notwendige Voraussetzung für Emotionen anzusehen: „To judge that apples are good requires that one judge them to have certain properties in virtue of which they are good. Thus, if emotions are shown not to require any non-evaluative judgments, they cannot require evaluative judgments either“.

vieler emotionaler Zustände besteht. Dabei wird schnell deutlich, dass es evtl. Sinn machen würde, statt der „Kognitionsabhängigkeit“ von Emotionen oder deren „Kognitivismus“ lieber die „Rationalität“ von Emotionen zu untersuchen. Begrifflich macht diese Umformulierungen keine nennenswerten Fortschritte, ist doch der Vernunftbegriff noch wesentlich schwieriger zu fassen als der Begriff der „Kognition“. Und dennoch erfasst ein solcher Paradigmenwechsel eine Stoßrichtung der Forschung, die immer schon impliziert zu sein scheint: spricht man vom „Kognitivismus“ in der Emotionstheorie, dann geht es meistens darum nachzuweisen, dass Emotionen adaptiv statt disruptiv sind, dass wir bedingt verantwortlich für unsere Emotionen sind, und dass diese die typischen Vollzüge unserer Vernunft (v.a. der praktischen Vernunft), also zielgerichtetes Überlegen, Fassen von Absichten, Wählen von Mitteln und Festlegen von Zielen, unterstützen oder sogar tragen. Darin – als einer ersten Annäherung – könnte ein Aspekt der „Rationalität“ von Emotionen bestehen. Und einige Autoren haben versucht, dieses Phänomen zu erklären, ohne sich auf die Debatte um den Kognitivismus in der Aktualgenese der Emotionen zu vertiefen.

Für Ronald DeSousa (1979, 132f.) besteht die „Rationalität“ von Emotionen – ähnlich wie bei anderen rationalen Zuständen – in ihrem „Erfolg“ („success“), d.h. die Rationalität von Emotionen ist deren Angemessenheit („appropriateness“):

„There is some temptation to think that rationality properly applies only to *transitions*. These, however, can only be assessed in the light of something like the probability of adequacy which they confer to the resultant state. In the case of inferences, ‚adequacy‘ is simply truth. We do not rate a transition – an argument or inference – rational if and only if it leads to true belief, but we do require at least that it be of a sort that generally tends to increase the likelihood of truth. It is therefore useful to stipulate the following extension of the notion of rationality: call ‚intrinsic rationality,‘ or more briefly *success*, the property of the end state which rational transitions make more probable. In the case of actions, ‚success‘ is generally called just that: success. In belief it is called *truth*. In wants, success is actual desirability or goodness. I propose to extend this to emotions, and to call the success of an emotion ‚appropriateness‘.”

DeSousa versucht also eine allgemeingültige Deutung des Rationalitätsbegriffes im Rahmen seines Modells von „success“, wobei dieser Begriff nur bei Handlungen wörtlich zu nehmen sei. In anderen Fällen, in denen wir von Rationalität sprechen, müsse „success“ jeweils anders gedeutet werden, so spricht er im Falle von Überzeugungen von „Wahrheit“, bei Emotionen von „Angemessenheit“. Nach DeSousa ist eine Emotion angemessen, wenn das sie auslösende Objekt eine solche Emotion „rechtfertigt“ (ebd., 133: „if and only if the evoking object or situation *warrants* the emotion“) Für die Angemessenheit von Emotionen bezieht sich DeSousa nun nicht auf subjektive Zustände im Inneren des Trägers eines emotionalen Zustands (angemessen angesichts eines bestimmten Wunsches o.ä.), sondern auf die objektiv verstandene Situation der Emotionselizitation. Dabei stellt er eine Analogie zur

Sinneserfahrung an. Auch dort gibt es die Möglichkeit der Halluzination oder Illusion, und auch dort müssen wir Adäquanz über sog. „sekundäre Eigenschaften“ (wie „rot“, „blau“ etc.; anders als die Eigenschaft der Masse z.B.) bestimmen. Doch auch diese – wenngleich sie auch von unseren Sinnesrezeptoren abhängen – unterliegen objektiven Standards, wenn wir z.B. „rot“ als die Wahrnehmung von Blut eines normalen Beobachters unter Standardbedingungen definieren. Nach DeSousa ergänzen Emotionen andere mentale Phänomene wie Wünschen und Glauben darin, das Fassen von Zielen und das Bilden von Überzeugungen zu ermöglichen. Emotionen sind: „*determinate patterns of salience among objects of attention, lines of inquiry, and inferential strategies*“ (ebd., 137). Sie operieren auf einem „Meta-Level“ in Bezug auf Überzeugungen (ebd., 140), DeSousa sagt auch:

„On my view, emotions ask the questions which judgment answers with beliefs: but as every committee chairman knows, questions have a good deal to do with the determination of answers: the rest can be done with innocuous facts.“ (ebd., 138)

In dieser sehr kurz gefassten Hinführung zu DeSousas Modell von „paradigm scenarios“ haben wir uns immer mehr von einem atomaren Verständnis des intentionalen Objekts als „Überzeugung“, „Urteil“ oder „Gedanken“ entfernt. Wenn wir von Emotionen als „Urteilen“ sprechen, so DeSousa, dann müssen wir dies viel mehr in dem Sinne verstehen, wie wir wissenschaftliche Paradigmen (nach Thomas Kuhn 1962) als eine Art von „Urteilen“ auffassen könnten: „they are what we see the world ,in terms of’.“ Emotionen bestimmen also, wie in der oben zitierten einschlägigen Definition, welche Aspekte salient gemacht werden an einem Objekt unserer Aufmerksamkeit, an einer Richtung des Nachforschens oder Nachdenkens bzw. an einem Prozess schlussfolgernden Überlegens. Sie sind in diesem Sinne also keine Urteile, sondern befinden sich auf einer anderen Stufe als diese.⁴²

DeSousas Begriff der „paradigm scenarios“ (ebd., 142f.) beschreibt einerseits den Erwerb von Emotionsbegriffen und emotionalen Dispositionen und erklärt andererseits die Aktualgenese eines emotionalen Zustands aus einer Disposition heraus. Nach DeSousa erwerben wir zunächst ein Emotionsvokabular, das wir eben mit bestimmen

⁴² „Although emotions are manipulators of reasoning, the experience of emotion tends to be an ‚intuitive’ one: it is not easy to formulate reasons for one’s shifts of attention. [...]An explanation sometimes offered for this is that emotions are somatic phenomena; that they consist largely in autonomic reactions, of which the phenomenological component consists in dim awareness. A preferable explanation can be given in terms of the model I have sketched: paying attention to certain things is a source of reasons, but comes before them. Similarly, scientific paradigms, in Kuhn’s sense, are better at stimulating research in certain directions than at finding compelling and fair reasons for their own adoption. They are too ‚deep’ for that, too, unlike specific, easily formulated beliefs.” (DeSousa 1979, 139)

Schlüsselszenarien unseres Alltagslebens in der frühen Kindheit assoziieren. Solche Schlüsselszenarien⁴³ werden später gefestigt und unterfüttert durch Geschichten und Märchen, die Kindern erzählt werden, und in einem noch späteren Stadium verfeinert und erweitert durch unseren Umgang mit Kunst und Literatur.⁴⁴ In Schlüsselszenarien erlernen wir die Zuordnung einer standardisierten („normalen“) Reaktion auf ein bestimmtes „Objekt“ (wobei DeSousa dieses eben nicht im Sinne eines formalen Objekts verstehen will, sondern in einem weiten Sinne als „Situation“, „Ziel“ [target] oder „Anlass“ einer Emotion).⁴⁵ Die Varietät kognitiver Komplexität von Emotionen lässt sich durch dieses Modell erklären durch den Bezug auf je spezifische Schlüsselszenarien: „If the paradigm scenario cannot be apprehended without complex linguistic skills, for example, we shall not expect to find in whoever lacks those skills an emotion specifically tuned to that scenario“ (ebd., 143).

Ebenso wie DeSousa hat vor einiger Zeit auch Annette Baier (1990) die Ansicht vertreten, dass einer Theorie des intentionalen Objekts von Emotionen nicht Genüge getan ist, indem man einfach eine Klasse von formalen Objekten benennt, worunter dann z.B. ein Einzelding x fallen und die Emotion erklären würde. Die verstehende Rekonstruktion eines solchen intentionalen Objekts kann auch nach Baier nur geleistet werden, indem man auf eine individuelle biografische *Geschichte* abhebt, die dann assoziativ das tatsächliche „Tiefenobjekt“ einer Emotion herauskehrt. Intentionale Objekte von Emotionen unterliegen also einem narrativen Konstruktivismus.

Auch Jesse Prinz hat in einigen rezenten Publikationen versucht, den Erwerb von Emotionsbegriffen zu erklären. Dabei stellt er auch eine eigene kritische Auseinandersetzung mit dem Kognitivismus an und schlägt ebenfalls eher eine Wende zum Begriff der

⁴³ Nach DeSousa beruht jede Emotion auf einem solchen Schlüsselszenario, in dem das Subjekt der Emotion eine „Rolle spielt“ (so z.B. „Mut“: ich werde angegriffen und muss mich wehren; „Furcht“: mir droht Unheil, ich muss entkommen; usw.). Durch das Erleben solcher Szenarien einerseits und die Zuordnung unserer Gefühle sowie die Leitung der Gefühle durch unsere Bezugspersonen andererseits bilden wir aus solchen typischen Szenarien sozialer Interaktion emotionale Dispositionen. Adam Morton (2002) hat m.E. richtig darauf hingewiesen, dass dieses Modell auf einer *narrativen* Grundlage beruht.

⁴⁴ In jüngster Zeit hat u.a. auch Martha Nussbaum (1996) die Ansicht vertreten, dass Emotionen maßgeblich durch Kunst und Literatur beeinflusst werden.

⁴⁵ Daraus lasse sich – so DeSousa – auch die Ansicht erklären, dass Emotionen motivational wirksam seien: schließlich verbinden sie Standardreaktionen mit den ihnen zugeordneten Objekten. In Wirklichkeit verhält es sich – zumindest aus der Perspektive des Erwerbs emotionaler Dispositionen – andersherum: die Zuordnung (also das motivierte Handeln) findet nach logischer Notwendigkeit *vor* der Bildung einer Disposition statt.

„Rationalität“ vor, wobei dieser bei Prinz durch den Begriff der „Bedeutung“ zu fassen versucht wird, anders als die oben dargestellten Gleichsetzungen DeSousas als „success“, „Wahrheit“ oder „Angemessenheit“. Prinz (2003, 78) versucht in diesem Sinne, die Fragestellung, an der kognitivistische Theorien meistens eher implizit als explizit arbeiten, klar herauszustellen. Für Prinz stellt diese der Emotionsforschung inhärente Fragestellung das sog. „Emotions-Problem“ dar.⁴⁶ Das Problem lässt sich prägnant im Sinne eines Dilemmas formulieren:

- (1.) Emotionen sind bedeutungsvoll, vernunftsensitiv und intentional. Also müssen sie kognitiv sein.
- (2.) Einige Emotionen entstehen unabhängig von der Beteiligung des Neokortex, in dem höhere kognitive Funktionen neuronal repräsentiert sind. Also können nicht alle Emotionen kognitiv sein.
- (3.) Die Emotionen unter 2. sind auch bedeutungsvoll. Also kann etwas Bedeutung haben, ohne kognitiv zu sein.

Prinz zieht daraus das Fazit, dass nonkognitive Theorien „zu wenig erklären“ (nämlich: wie kann etwas Bedeutung haben, ohne kognitiv zu sein?), und dass kognitive Theorien demgegenüber zu viel erklären (nämlich indem sie alle Emotionen als „kognitiv“ verstehen, um deren Bedeutungshaftigkeit erklären zu können).

Prinz wendet Begriffstheorien zur Lösung dieses Problems an, speziell die sog. „informational theories“ von Dretske (1981) und Fodor (1990). Der Ausgangspunkt dieser Überlegungen ist folgender: betrachten wir das Erlernen von alltagssprachlichen Begriffen wie z.B. „Tisch“. Wie lässt sich erklären, auf welche Weise wir kompetent einen „token“ zu einem „type“ zuordnen? Wenn wir nämlich versuchen, „Tisch“ intensional zu definieren, stellen wir schnell fest, dass wir in einen Regress verfallen: ein Begriff definiert sich durch andere Begriffe, und das bis ins Unendliche fort. Dieses Problem versteht Prinz analog zum Problem der Kognitivisten, hinreichende Kriterien für die Entstehung einer Emotion anzugeben. Dabei wird jedoch eine Menge von Überzeugungen oder Urteilen vermutlich nie vollständig sein, um einen Emotionstypus mit allen seinen Instanzierungen genetisch erklären zu können. Prinz geht nun im Sinne Dretskes davon aus, dass mentale Zustände auf Dinge

⁴⁶ Im englischen Original heißt es „the emotion problem“. Der Begriff verweist evtl. seinem Gestus und seinen theoretischen Klärungsansprüchen nach auf den 1994 von Michael Smith geprägten Begriff „the moral problem“, mit dem eine (wenn nicht „die“) zentrale Fragestellung beschreibender Moraltheorie angesprochen wurde, nämlich das theoretische Dilemma zwischen Motivationsinternalismus und Wahrheitsfunktionalität moralischer Überzeugungen.

verweisen, die ursächlich dafür waren, dass dieser Zustand geprägt („tokened“) wurde. In der Terminologie von Dretske schreibt Prinz: „a mental state refers to what it is *set up to be set off by*“. Ein Feuermelder z.B. ist eingerichtet („set up“), um Feuer zu entdecken, aber manchmal wird er ausgelöst („set off“) durch eine Zigarette. Das Attribut „set up“ ist also nötig, um die Menge der möglichen Referenten von „set off“ einzuschränken auf die wesentlichen.

„Primitive mental representations represent that which they have the function of detecting. [...] An internal state *detects* some thing if it is reliably activated by that thing. [...] An internal state has the *function* of detecting some thing if it is created to detect that thing. Internal states can be created in two ways: by evolution or learning.“ (Prinz 2002, 141)

Nach der Jamesianischen Theorie sind Emotionen nun interne Zustände, die körperliche Veränderungen registrieren. Solche Zustände werden also ausgelöst („set off“) durch körperliche Veränderungsmuster („patterned bodily changes“), und sie sind auch eingerichtet („set up“), dies zu tun, weil solche Veränderungsmuster normalerweise mit Angelegenheiten zusammenfallen („co-occur“), die uns besonders angehen. Wie James also behauptet hat, sind Emotionen körperlich („embodied“), *und* wie der Kognitivismus behauptet, repräsentieren sie zentrale Verweise auf Angelegenheiten, die uns angehen („they represent core relational themes of concern to us“). Wenn also „appraisal“, also eine Wertschätzung oder Ablehnung (ein Werturteil mit entsprechender Einstellung des Subjekts), richtig definiert ist als ein „mentaler Zustand, der eine Organismus-Umwelt Beziehung repräsentiert, die sich auf das Wohlergehen des Organismus bezieht“, dann können Prinz's „embodied states“ auch als appraisals gelten.

Damit argumentiert Prinz gegen den Kognitivismus, insofern er behauptet, dass komplexe Inhalte nicht notwendigerweise komplexe Repräsentationen brauchen. Das Konzept der „embodied appraisals“ im semantischen Modell von Prinz kann Emotionen und deren kognitiven Gehalt repräsentieren, ohne sie zu beschreiben. Angst z.B. ist ein Gefahrendetektor, der durch eine Menge von Elementen ausgelöst wird, die die Eigenschaft teilen, gefährlich zu sein. Kognitive Urteile über diese Elemente (x ist gefährlich) kommen evtl. später hinzu, allerdings ist dies ein sekundärer Fall. „After concepts are acquired, fear can be triggered by cognitive appraisals [...] But these cases are derivative.“ In einem anderen Aufsatz fasst Prinz sein Argument in folgender prägnanten Weise zusammen:

„Prevailing theories of emotion present us with a false dichotomy. Cognitive theorists would have us believe that emotions always involve judgments, and non-cognitive theorists would have us believe that emotions carry no information at all or, at best, carry information about changes in our bodily states. If emotions are mere feelings or modes of processing, they are, strictly speaking, meaningless. They do not represent anything. If emotions are states that register changes in the body, they may be said to represent such things as heart rate or respiratory patterns rather than core relational

themes. We are presented with a choice. Either emotions involve judgments representing core relational themes, or they do not involve judgments and fail to represent core relational themes. If they fail to represent core relational themes, we cannot explain the roles they play in our lives, and, if they involve judgments, we cannot explain their neural circuitry. The solution is simple. Emotions can represent core relational themes without explicitly encoding such themes in the form of judgments. A mental state can represent a demeaning offense without being built up from the concepts *demeaning* and *offense*." (Prinz 2002, 140)

Prinz's Theorie kann dem Nonkognitivismus folgend die Körperlichkeit von Emotionen erklären, dem Kognitivismus folgend erklären, wie Emotionen „core relational themes“ darstellen, d.h. sich auf ein formales Objekt beziehen. Prinz nimmt dabei an, dass – nach James und Damasio – Emotionen reliabel durch körperliche Veränderungen ausgelöst werden, dass sie diese also detektieren. Diese Annahme bildet eine Art minimaler Basis für die Argumentation. Denn wenn Emotionen – im Sinne James' – lediglich eine wahrnehmende Reaktion auf Körperzustände sind, dann kann man weiter annehmen, dass sie auch nichts anderes „repräsentieren“ als diese Körperzustände, also keinen kognitiven Inhalt haben. Wenn Prinz also erfolgreich nachweisen kann, dass auch solche Jamesianischen Emotionen Bedeutung haben, dann dürfte das Argument a fortiori auch für kognitiv komplexere Emotionen gelten. Andererseits, so Prinz, werden Emotionen auch durch die von den Kognitivisten postulierten „core relational themes“ (Smith & Lazarus 1993; Lazarus 1991b) ausgelöst, d.h. sie detektieren ebenso reliabel Instanzen von „Gefahr“ und „Beleidigung“. Nach Prinz können wir hier eine Antwort finden in Anlehnung an den evolutionären „Zweck“, zu dem Emotionen geschaffen wurden. Der Mechanismus der Angst wurde demnach wohl in der Geschichte der Evolution weitergegeben, weil er zuverlässig „Gefahr“ detektieren kann. Demnach repräsentieren Emotionen also „core relational themes“. Dennoch erfordert diese Schlussfolgerung nicht, dass Emotionen – entsprechend einem strengen philosophischen Kognitivismus – urteilsförmig seien: „All that matters is that they are set up and set off in a certain way“ (Prinz 2002, 142). Dadurch können wir der falschen Dichotomie entgehen. Prinz' Vorschlag lässt sich auch in Damasio's Theorie (also in ihrer Grundlage eine rezente Jamesianische Theorie) integrieren: nehmen wir an, Emotionen detektieren Körperzustände und –veränderungen. Nehmen wir weiter an, dass solche Zustände und Veränderungen auftreten, wenn wir uns in einer Situation befinden, die einem „core relational theme“ entspricht: Gefahr, Verlust, Repugnanz etc. Auf diese Weise repräsentieren Emotionen – als Detektoren von Körperprozessen – solche Bewertungsthemen. Emotionen sind also keine Urteile, weil sie diese Bewertungsthemen nicht beschreiben, oder Begriffe mit einbeziehen, die sich auf diese Themen richten. Emotionen werden aber – evtl. nach James

und Damasio über den Umweg des Detektierens von Körperprozessen – zuverlässig durch solche Bewertungsthemen ausgelöst („set off“).

„On this view emotions are not cognitive states, but they carry the kind of information that cognitive states carry.“ (ebd.)

Wie können kognitiv hoch komplexe Emotionen mit dem Modell von Prinz erklärt werden, also Schuld, Scham, Empörung etc.? Prinz beruft sich auf Oatley & Johnson-Laird (1997): komplexe Emotionen sind „Basis-Emotionen + Gedanken“. Diese Gedanken sind keine konstituierenden Teile der Emotion, sondern mit den Worten von Prinz sog. „calibrating causes“. Emotionen können Kalibrierungsdateien („calibration files“) haben, die bestimmte Urteile, Gedanken, Wahrnehmungsreize enthalten, die als Auslöser einer Emotion fungieren können.

Erste Ansätze zu einer Analogie zwischen dem Erwerben emotionaler Dispositionen und dem Erlernen von Begriffen finden sich weiterhin auch schon bei Heller (1980, 167ff), wo dies als „Koordinierungsprozess“ beschrieben wird.⁴⁷ Nach Heller kann diese Koordinierung auf zwei Wegen erfolgen:

(1.) Das Subjekt besitzt einen emotionalen Begriff, ohne zuvor die koordinierte Emotion empfunden zu haben. Wenn nun das Gefühl der Emotion auftritt, kann es auf den bereits vorhandenen Emotionsbegriff bezogen, d.h. mit diesem „koordiniert“ werden. U.z. entweder indem das Gefühl eruptiv in mir ausbricht, und ich dieses stark affektive Phänomen mühevoll mit einem bekannte Begriff verbinden muss. Oder indem ich versuche, die einem Begriff korrelierende Emotion willkürlich herbeizuführen, sie kennen lernen will. Oder indem ich mich einer bestimmten Emotion entsprechend verhalte, und andere den ausgedrückten Zustand als die Emotion x deuten. Für eine solche „Emotion-zu-Begriff“ Zuordnung bringt Heller ein anschauliches Beispiel:

„Praktisch jedes Kind weiß, dass es Liebe gibt, ehe es jemals verliebt gewesen ist; die Existenz dieser emotionellen Disposition und ihre, je nach Zeitalter veränderte Bedeutung wird quasi vor die Augen der Kinder ‘geschoben’. Es erübrigt sich zu sagen, wie entscheidend im Leben eines jeden Menschen diese bestimmte ‘Koordinierung’ ist. ‘Ob ich wirklich verliebt bin?’ ‘Ob das, was ich fühle, Liebe ist?’ [etc.]“ (ebd.)

⁴⁷ Ebd.: „Das Erlernen des Gefühls ist immer zugleich der ‘Koordinierungs’-Prozess des emotionellen Begriffs und des Gefühls. Den Begriff: ‘koordinat’ haben wir aus der Geometrie entlehnt; Emotion und emotioneller Begriff können nämlich nie deckungsgleich (da jede Emotion auch individuell idiosynkratisch ist), nur koordinat sein.“

Dazu ist noch anzumerken, dass die erste Zuordnung nicht endgültig ist, und dass weder die gefühlte Emotion noch deren Begriff statisch sind: z.B. kann auch die Bedeutung des Begriffs selbst durch eine Reihe erlebter emotionaler Zustände modifiziert werden, obwohl die Zuordnung hier in die umgekehrte Richtung verläuft.

(2.) Der zweite Weg der Koordinierung ist verläuft folgender Maßen (vgl. ebd., 170f.): Wir fühlen eine Emotion, ohne zu wissen welche, und suchen nach dem passenden Begriff. Wir führen also (in Umkehrung des ersten Weges) eine „Begriff-zu-Emotion“ Zuordnung durch. Heller schlägt folgende Illustration vor:

„Wenn wir dem Kinde eine schöne Geschichte erzählen und es in Tränen ausbricht und sagt: ‘ich bin aber sehr traurig’, dann kann der Erwachsene ihm ruhig antworten: ‘du bist jetzt nicht wirklich traurig, eher bist du gerührt’. Es ist möglich, dass auch das Kind das nächste Mal sagt: ‘ich bin gerührt’; in diesem Fall differenziert es zwei Gefühle voneinander, und passt sie verschiedenen Begriffen an.“ (ebd.)

2.7. Neurologische Grundlagen von Emotionen

Macht es Sinn, angesichts der breiten und komplizierten Debatten um den Zusammenhang von „brain states“ und „mental states“, nach den neurologischen Grundlagen von Emotionen zu fragen? Ich denke, diese Absicht ist in mehrerer Hinsicht gerechtfertigt. Wir würden wohl sagen, dass im Falle des Urteilens oder des rein intellektuellen Problemlösens (z.B. einer komplizierten Mathematik-Aufgabe) unsere mentalen Vorgänge und Zustände phänomenal als unabhängig von unserem Körper (ausschließlich Gehirn) erscheinen, und dass weiterhin eine Identifikation von bestimmten involvierten Hirnarealen zur Zeit noch recht aussichtslos erscheint, sofern eine solche Aussage informativ sein soll (und nicht etwa den „gesamten Großhirnkortex“ aufweist). Anders verhält es sich bei Emotionen. Hier scheint unser Körper notwendig mitbeteiligt zu sein, entweder durch emotionale Gefühle oder durch die gewissen Emotionen typischen phänomenalen Körperbildrepräsentationen. Damit ist ein erster Weg zur Neuropsychologie der Emotionen gebahnt, da man z.B. nach der zentralnervösen Verschaltung fragen kann, die bestimmte Informationen aus der Körperperipherie transportieren und im Gehirn verschalten. Im Zusammenhang damit konnten auf empirischem Wege schon relativ sicher emotionsrelevante Hirnareale oder Nervenkernelnzentren und – schaltkreise identifiziert werden. Natürlich sind begrifflich-analytische (also im strengen Sinne philosophische) Ergebnisse nicht direkt mit empirischen Ergebnissen aus der naturwissenschaftlichen Hirnforschung oder der Psychologie zu verrechnen. Ich denke in der Tat, dass es sich hierbei um zwei verschiedene Erklärungsebenen handelt, die jedoch komplementär zueinander zu einer umfassenden und ganzheitlichen Theorie der Emotionen

beitragen können. Elizabeth Anscombe (1957) ist sicher rechtzugeben, dass keine empirische Forschung im Stande wäre, eine begriffliche Verwirrung aufzuklären. Andersherum können empirische Ergebnisse Leitlinien für eine begriffliche Theoriebildung darstellen. Und kein begrifflich-analytisches Ergebnis (sofern es sich nicht um eine normative Aussage handelt) sollte im krassen Widerspruch zu einem allgemein akzeptierten Ergebnis empirischer Forschung stehen.

2.8. Rezente neuro-psychologische Emotionstheorien: Antonio Damasio

Der Vertreter der z.Zt. ausgefeiltesten neuro-psychologischen Emotionstheorie ist zweifellos Antonio Damasio (1994 u.f.). Wie Daniel Dennett in einer in der Times erschienen Rezension urteilte, handelt es sich bei Damasio's Studie um einen der „originellsten Beiträge“ zu den Fragen von „Geist“ und „Bewusstsein“ in den (bis dato) vergangenen Jahren (Dennett 1995). Entscheidend beigetragen zur Entwicklung einer solch umfassenden Theorie der Emotionen hat wohl einerseits die breite Verfügbarkeit bildgebender Verfahren in der medizinischen Forschung, v.a. die Anwendung verschiedener Formen der Computertomographie (MRT [Kernspin-Tomographie], fMRT [funktionelle MRT], PET [Positronen-Emissions-Tomographie] etc.). Zwar fand auch in früheren Jahrzehnten der Hirnforschung Spekulation über Emotionszentren im Gehirn statt,⁴⁸ eine empirische Unterstützung der Annahmen bzw. eine Anleitung der Hypothesenbildung konnte jedoch nur über relativ krude Untersuchungsformen wie die Hirnstrommessung (Elektro-Enzephalogramm) erfolgen. Andererseits konnte Damasio, in seiner Funktion als Direktor des Departments für Neurologie an der Universität Iowa, nach eigenen Aussagen (Damasio 1994, 300) auf über 1800 Patientendaten zurückgreifen. Wenige Theorien in der empirischen psychologischen Forschung zu Emotionen können sich auf eine solch umfangreiche Menge von Probandendaten stützen.

⁴⁸ Descartes stellt in dieser Hinsicht ein sehr frühes Beispiel dar, indem er in den *Passiones* die Epiphyse (auch Glandula pinealis, Zirbeldrüse, wegen ihrer Zapfenform) als das „Emotionszentrum“ des Gehirns identifiziert. Der Grund für diese Wahl kann nur noch Gegenstand von Spekulation sein, interessant ist jedoch, dass die Epiphyse die *einzig*e Struktur im Gehirn ist, die nicht paarig vorliegt (LeDoux 1996, 97). Das Modell ist in seiner Struktur bahnbrechend und kann als Grundlage aller modernen Theorien der Hirn-Körper-Interaktion in der Emotionsgenese angesehen werden. Inhaltlich irrte sich Descartes jedoch: heute ist bekannt, dass die Epiphyse fast ausschließlich für die hormonelle Steuerung des zirkadianen Rhythmus zuständig ist (Trepel 1999, 176).

Was können wir von einer neuro-psychologischen Theorie als fruchtbaren Beitrag für eine philosophische Untersuchung von Emotionen erwarten? Zunächst wird ein solcher Ansatz den Vorzug haben, dass er bestimmte Herangehensweisen zu verbinden vermag, die in der Tradition getrennt vorlagen. Das Phänomen der peripher-physiologischen Aktivierung hat sich als notwendiger, aber streitbarer Aspekt im Konstrukt „Emotion“ herausgestellt. Jenseits der Hypothesen von James und Lange, wenngleich daran anschließend, gelingt es Damasio zu zeigen, wie im Gehirn unter dem Erleben eines emotionalen Zustands Körperprozesse und – zustände verarbeitet werden. Die heutzutage recht gute Kenntnis der Verschaltung der Nervenbahnen im Gehirn und der Funktion bestimmter Hirnzentren und Nervenkerne erlaubt Hypothesen darüber, welche neuralen und mentalen Kapazitäten beim Erleben von Emotionen aktiviert werden und beteiligt sind. Natürlich lassen sich durch empirische Ergebnisse keine begrifflichen Fragestellungen z.B. zur Intentionalität von Emotionen klären. Andererseits ist der Neurologie relativ gut bekannt, welche Prozesse in welchen Hirnzentren zum Bewusstsein gelangen, und welche nicht. Stellen wir uns z.B. vor, eine solche empirische Untersuchung würde zeigen, dass beim Prozess der Aktualgenese von (bestimmten Typen von) Emotionen allein unbewusst operierende autonome oder vegetative Hirnzentren beteiligt wären, dann wäre damit sehr wohl ein Indiz gegen die Hypothese einer bestimmten Spielart des Kognitivismus verbunden. Die Behauptung, Emotionen entstünden aufgrund von bewussten, deliberativen Urteilsprozessen wäre in diesem Licht zumindest stark in Zweifel gezogen. Abgesehen von solch „harten“ argumentativen Fakten kann eine Beschäftigung mit der Neuropsychologie der Emotionen uns ganz allgemein über das Zusammenspiel von mentalen, neuralen und körperlichen Prozessen aufklären und viele partikuläre Streitpunkte lösen oder zumindest plausible Hypothesen anbieten.

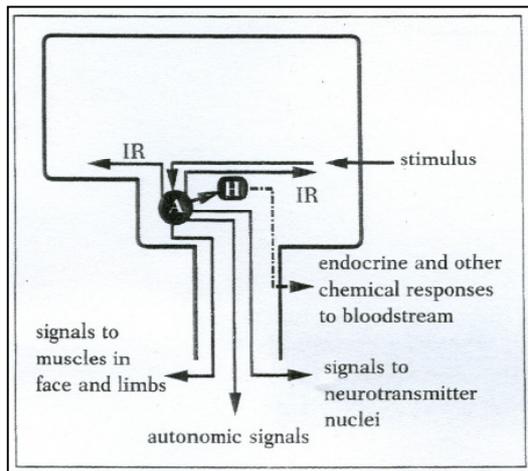
Angestoßen wurden Damasios Überlegungen zum Mechanismus der Emotionselizitation durch Beobachtungen an Patienten mit einem bestimmten Muster von Hirnläsionen. In der Geschichte der medizinischen Dokumentation tauchen immer wieder Fälle von Patienten auf, die nach erlittenen Hirnverletzungen durch soziopathisches Verhalten auffallen, während sie in allen gängigen Intelligenz- und Persönlichkeitstests Normalwerte erreichen. Hier ist der berühmte Fall des Phineas Gage zu nennen, den Damasio rekonstruktiv untersucht: ein Eisenbahnarbeiter aus dem 19. Jhd., dessen Schädel bei einem Unfall von einer Metallstange durchbohrt wurde. Der Patient konnte noch viele Jahre weiterleben mit einer Läsion am präfrontalen Lappen des Großhirnkortex. Die Fallstudie des Patienten „Elliot“, den Damasio untersuchte und diskutiert, ist aufschlussreich: Elliot zeigte in etwa die gleiche Präfrontallappen Läsion wie der medizinhistorische Fall des Phineas Gage, der von Damasio und

seiner Frau untersucht wurde. Elliot zeigte normale Ergebnisse bei Tests zu Intelligenz, Persönlichkeit, Urteilen und Entscheidungsfindung, hatte jedoch eine gestörte Emotionalität. Das wichtigste Merkmal seiner Störung war jedoch eine „dissociation between real-life failure and laboratory normalcy“. Wie passen diese Ergebnisse zusammen? Damasio spekuliert: „the cold-bloodedness of Elliot’s reasoning prevented him from assigning different values to different options in the decision-making process“ (51). Ähnliche Fallstudien stellen der sog. Brickner’s „patient A“ (1932), der „Herb-Penfield-patient“ (1940) und der „Ackerly-Benton-patient“ (1948) dar (vgl. Damasio 1994, 54-57). Dem soziopathischen und irrationalen Verhalten dieser Patienten stand man perplex gegenüber, da sie typischerweise normal intelligent blieben, auch nach der erlittenen Verletzung. Damasio spekuliert jedoch, dass bei diesem Verletzungsmuster emotionsrelevante Hirnareale ausgeschaltet wurden, und konnte diesen Befund an eigenen Patienten bestätigen. Dieses Syndrom, das sich durch soziopathisches Verhalten zusammen mit einer ausgeprägten Emotionsarmut und Apathie auszeichnet, tritt auf bei bilateralen Läsionen der präfrontalen Großhirnkortizes und/ oder bei Läsionen des somatosensorischen Zentrums der rechten Hemisphäre, verbunden mit einer Läsion des Corpus amygdaloideum (Amygdala oder „Mandelkern“) (61f.). Die Rolle des somatosensorischen Kortex im emotionalen Erleben konnte bestätigt werden durch Untersuchung von sog. anosognostischen Patienten: das Syndrom der Anosognosie, bei dem Patienten kein Bewusstsein mehr von ihrem eigenen Gesundheitszustand haben und sich z.B. trotz einer durch Schlaganfall bedingten halbseitigen Lähmung nicht als gelähmt anerkennen, zeichnet sich gerade durch die Ausschaltung somatosensorischer Kapazitäten des Großhirns aus. Interessant ist nun, dass Anosognostiker die gleichen Defizite im sozialen Verhalten und emotionalen Erleben zeigen wie die sog. (bereits vorgestellten) „prefrontal-lobe-lesion“ Patienten (65-68).

Damasio (70f.) identifiziert drei emotionsrelevante Bereiche im Hirn, mit je eigenem spezifischen Befund bei Läsionen, nämlich: (1.) (ventromediale) präfrontale Großhirnkortizes, (2.) die Amygdala, und (3.) rechte somatosensorische Kortizes.

Bei der Aktualgenese von Emotionen unterscheidet Damasio zwischen „early (primary) emotions“ (die er auch „Jamesian emotions“ nennt, womit sein theoretischer Rückbezug auf die James-Lange Theorie unterstrichen wäre) einerseits, und „adult (secondary) emotions“ andererseits. Diese Unterscheidung ist unkontrovers und ein Gemeinplatz in der modernen Psychologie. Auch die Beispiele, die Damasio (2001, 102) für die jeweils gemeinte Klasse gibt, können als Standard gelten: unter Primäremotionen fallen z.B. „Freude“, „Trauer“,

„Angst“, „Zorn“, „Überraschung“ und „Ekel“; Beispiele für Sekundäremotionen sind: „Scham“, „Eifersucht“, „Schuld“, „Stolz“. Einerseits fällt auf, dass diese Unterscheidung nach dem Vorliegen einer klaren Annäherungs-Vermeidungs-Tendenz differenziert: im Fall der Primäremotionen ist meist (zumindest für Freude, Angst, Zorn, Überraschung und Ekel) ein starker Aspekt von Attraktion oder Repugnanz enthalten, der sich nicht so leicht bei den genannten Sekundäremotionen feststellen lässt. Andererseits wird deutlich nach der Kategorie des intentionalen Inhalts der betreffenden Klasse von Emotionen differenziert: wengleich die genannten Primäremotionen evtl. auch jeweils ein linguistisch oder propositional höher strukturiertes Objekt zulassen, sind zumindest für alle genannten Beispiele nicht-propositionalen Objekte denkbar: z.B. ekele ich mich *vor* dem Spinat auf meinem Teller, freue mich *über* Hans' Geschenk, habe Angst *vor* dem Bären etc. In allen Fällen scheint nur die mentale Repräsentation eines materiellen oder imaginierten Objekts involviert zu sein. Jenseits von modellhaftem Lernen oder Lernen von typisierten, „paradigmatischen Szenarios“ (nach DeSousa) scheint keine weitere Anforderung an kognitive Vermögen des Subjekts gestellt zu werden, wie etwa propositionales Wissen, Überlegen oder Schlussfolgern. Bei den genannten Beispielen von Sekundäremotionen verhält es sich anders. Schon sprachanalytisch betrachtet gibt es wohl keine Schuld *über* oder *vor* etwas. Ich kann jedoch Schuld empfinden, *weil* ich es versäumt habe, einen schwerkranken Freund zu besuchen. Variationsmöglichkeiten auf der Ebene der Oberflächengrammatik liefern hier kein schlagendes Gegenargument: Zwar lässt z.B. Scham eine Formulierung zu wie „Ich schäme mich *über* meinen faux-pas“. Das Objekt „faux-pas“ ist jedoch nicht kategorial vergleichbar mit Objekten von Primäremotionen, wie z.B. „Bär“, „Geschenk“, „Spinat“ etc. Hier handelt es sich um die sprachliche Verallgemeinerung eines Sachverhalts, der linguistisch und propositional komplex strukturiert ist und tatsächlich auf höheren kognitiven Leistungen des Subjekts aufbaut. Nur wer versteht, dass es gesellschaftliche Normen gibt, und dass bestimmte Konsequenzen bei deren Übertretung drohen, kann auf der Grundlage solchen Wissens eine Emotion wie „Scham“ empfinden.



Damasio 1994, 132, fig. 7-1

Primäre Emotionen sind hauptsächlich durch die Amygdala bzw. das limbische System zu erklären. Entsprechend des hier wiedergegebenen Schaubilds⁴⁹ beschreibt Damasio die Aktualgenese von Primäremotionen wie folgt: Die sensorische Wahrnehmung eines Stimulus aus der Außenwelt oder der Körperperipherie (d.h. Körper ausschließlich Gehirn) wird im limbischen System,

v.a. in der Amygdala verarbeitet. Damasio scheint hier implizit meine Kritik an einer intentionalen

Objektbeschreibung in der Aktualgenese von Emotionen zu teilen, da er explizit darauf hinweist, dass die Amygdala Stimuli auf das Vorliegen bestimmter „features“ hin prüft:

„[...] we are wired to respond with an emotion, in preorganized fashion, when certain features of stimuli in the world or in our bodies are perceived, alone or in combination. Examples of such features include size (as in large animals); large span (as in flying eagles); type of motion (as in reptiles); certain sounds (such as growling); certain configurations of body state (as in the pain felt during a heart attack).“ (131)

Nach Damasio verfügen die Nervenkerne der Amygdala über eine „dispositionale Repräsentation“ solcher features, die eine Mustererkennung im sensorischen Input erlaubt. Aus philosophisch-terminologischer Sicht mag der Begriff der „Repräsentation“ hier problematisch sein. Wir könnten im gleichen Sinne auch formulieren, dass die Nervenkerne und Synapsen der Amygdala in einer solchen Weise fest verschaltet („hard-wired“) sind, dass die Erkennung bestimmter Muster zu festgelegten („preorganized“) Aktivierungsmustern anderer Hirnregionen führt. Die Amygdala setzt daraufhin verschiedene Prozesse in Gang, zuvörderst die Aktivierung des sympathischen Anteils des vegetativen Nervensystems sowie die Aktivierung des Hypothalamus, der für eine peripher-physiologische Aktivierung über die Blutbahn (durch endokrine, also hormonelle und chemische Prozesse) zuständig ist:⁵⁰ wie

⁴⁹ Die Ausrichtung des Schaubilds ist: frontal rechts, okzipital links. Die starke Linie umreißt also schematisch den Hirnstamm und den Großhirnkortex einschließlich limbischem System. Okzipital liegt unterhalb des Umrisses (nicht dargestellt) das Cerebellum (Kleinhirn).

⁵⁰ Das Schaubild von Damasio ist, was die Verschaltung der genannten Prozesse angeht, nicht vollständig zutreffend. Zwar erklärt Damasio (132, fig. 7-1) selbst, dass das Diagramm eine vereinfachte Darstellung der Prozesse bietet. Meiner Meinung nach sind jedoch einige Richtigstellungen nötig, die sich auf mehr als nur vereinfachte Darstellungsweisen beziehen. Sowohl seinem Text als auch dem Diagramm nach gibt Damasio zu verstehen, dass die Amygdala direkt eine Aktivierung des vegetativen (autonomen) NS initiiert, im Schaubild als „autonomic signals“ bezeichnet. Das Corpus amygdaloideum ist jedoch nicht direkt mit vegetativen Zentren im

Damasio sagt, löst die Amygdala die „Aufführung“ („enactment“) eines Körperzustandes aus, eine Formulierung, die analog zu seiner Metapher vom Körper als dem „Theater“ der Emotionen zu verstehen ist. Primäremotionen erfordern weder eine „Erkenntnis“ ihres Objekts, noch propositional strukturiertes Wissen darüber:

„Note that in order to cause a body response, one does not even need to ‚recognize‘ the bear, or snake, or eagle, as such, or to know what, precisely, is causing pain. All that is required is that early sensory cortices detect and categorize the key feature or features of a given entity (e.g. animal, object) [...]. A baby chick in a nest does not know what eagles are, but promptly responds with alarm and by hiding its head when wide-winged objects fly overhead at a certain speed.“ (131f.)

Eine Emotion wie Furcht wäre durch diesen Prozess als ein adaptiver Mechanismus ausgewiesen, indem einerseits durch die Initiierung peripher-physiologischer Aktivierung der Organismus auf bestimmte Reaktionen in einer Situation höchster Dringlichkeit versetzt wird, und andererseits durch interne neurale Reaktionen (im Schaubild „internal responses“= „IR“) bestimmte Reflexe ausgelöst werden und eine allgemeine Veränderung des Bewusstseinsmodus erreicht wird, durch die z.B. Vigilanz steigt, Aufmerksamkeit fokussiert wird auf die durch die Emotion als salient markierten Reize (vgl. Christiansen & Loftus 1991), und i.A. schnellere Entscheidungs- und Deliberationsalgorithmen präferiert werden (vgl. Keinan 1987). An dieser Stelle ist hervorzuheben, dass der beschriebene Prozess noch kein spezifisches phänomenales Erleben der Vorgänge *als* Emotion einschließt. Die Vorgänge selbst verlaufen unbewusst und unwillkürlich, wenngleich sie den Bewusstseinsmodus und bestimmte kognitive Haltungen des Subjekts verändern. Bis zu dieser Stelle kann der Prozess jedoch noch nicht erklären, wie ein „Gefühl“ entsteht.

Damasio folgt in dieser Hinsicht der James-Lange Theorie und versteht „**emotionale Gefühle**“ („feeling an emotion“) als die Wahrnehmung bzw. das Erleben der zuvor spontan initiierten Körperveränderungen. Diese Veränderungen werden als „feedback“ über

Hirnstamm (der Impulse über die medulla oblongata, bzw. das System der formatio reticularis, an das Rückenmark weitergibt) verschaltet. Die Amygdala weist afferente und efferente Bahnen zu anderen Zentren des limbischen Systems auf (nicht zum Hirnstamm), weiterhin Efferenzen zum Hypothalamus (Trepel 1999, 173, 196 und 202). Der Hypothalamus, als „oberstes Integrationsorgan vegetativer Funktionen“ (ibid., 169) ist efferent mit vegetativ-nervösen Kernen im Hirnstamm und Rückenmark verschaltet. Man kann also folgern, dass nicht nur der endokrine Weg, wie Damasio postuliert, sondern auch der nervöse Weg dieser Prozesse über den Hypothalamus vermittelt wird. Dieser Weg würde jedoch nichts an einem Prinzip von Damasio's Theorie ändern: Prozesse des vegetativen Nervensystems und seines Steuerungsorgans (Hypothalamus) sind grundsätzlich „unbewusst“ und „unwillkürlich“ (ibid., 1 und 261ff.).

viszerosensible (aus den inneren Organen) und somatosensorische Bahnen (aus der Skelettmuskulatur und dem Gesicht) an entsprechende Hirnzentren vermittelt und zum Bewusstsein gebracht.⁵¹ In diesem Lichte ist auch seine Metapher vom Körper als dem „Theater“ der Emotionen zu verstehen: die peripher-physiologischen Veränderungen werden dort „aufgeführt“ und in einem zweiten Schritt von sensiblen und sensorischen Hirnzentren registriert, was wiederum zum emotionalen Gefühl führt:

„As body changes take place, you get to know about their existence and you can monitor their continuous evolution. You perceive changes in your body state and follow their unfolding over seconds and minutes. 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 feeling. If an emotion is a collection of changes in body state connected to particular mental images that have activated a specific brain system, the essence of feeling an emotion is the experience of such changes in juxtaposition to the mental images that initiated the cycle.“ (Damasio 1994, 145)⁵²

Das Verhältnis zwischen Emotionen und Gefühlen ist nicht ein-ein-deutig: alle Emotionen erzeugen Gefühle (Damasio: „if you are awake and alert“), nicht alle Gefühle jedoch lassen sich auf Emotionen zurückführen. Schmerz, Hunger, Müdigkeit und sexuelle Erregung sind allesamt Gefühle, sie entsprechen jedoch nicht jeweils bestimmten Emotionen.

⁵¹ Damasio lässt offen, über welchen Weg die viszerosensiblen und somatosensorischen Afferenzen an Hirnzentren vermittelt werden. Somatosensorische Afferenzen aus der Skelettmuskulatur werden in der weißen Substanz des Rückenmarks transportiert, viszerosensible Afferenzen im Seitenhorn der grauen Substanz des Rückenmarks. Weiterhin führt der Nervus vagus, der als einziger Hirnnerv bis weit in die Viszera hinein innerviert (von „vagus“, lat.= „weit umherschweifend“), viszerosensible Afferenzen des Sympathikus und Parasympathikus. Die Informationen werden in verschiedenen Kernen des Hirnstamms (genauer: medulla oblongata) verschaltet (vgl. Trepel 1999, 108 Abb. 5.3), z.B. werden die viszerosensiblen Afferenzen des N. vagus über die Nuclei tractus solitarii zum Thalamus vermittelt. Da „nahezu alle sensiblen und sensorischen Bahnen“ in den Thalamus projizieren (mit Ausnahme der olfaktorischen Bahnen z.B.) (ibid. 162), und da dieser für die Integration und Selektion der eintreffenden Informationen zuständig ist, liegt die Vermutung nahe, dass das bewusste gefühlsmäßige Erleben einer Emotion über den Thalamus – als sprichwörtliches „Tor zum Bewusstsein“ – vermittelt wird.

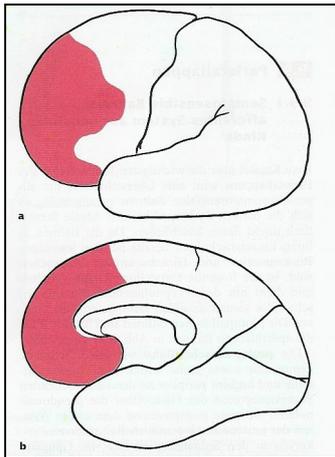
⁵² Auch in einer rezenten Publikation hält Damasio an dieser Definition fest: Damasio 2001, 103: “If an emotions is a collection of body state and brain state changes connected to particular mental images whose evaluation activated a specific brain system, *feeling an emotion is the imaging of those changes in juxtaposition to the mental images which initiated the cycle.*” Vgl. weiterhin ebd.: “What I call a *feeling* is (a) a mental image of our body state while thoughts about specific contents are processed, and (b) the mental imaging of the cognitive mode with which such contents are processed (e.g. the rate at which images are generated, how intensely they are attended, how they participate in inferential processes and choice of actions, and so on).”

Damasio nennt die Klasse der Primäremotionen auch „Jamesian emotions“ (vgl. o.). Wenn wir uns die in Kap. 2.3. referierte Kritik Damasio an der James-Lange Theorie in Erinnerung rufen, die v.a. darauf abstellte, dass James' Emotionsbegriff der Komplexität und Diversität der menschlichen Emotionalität nicht gerecht wird, dann verwundert es nicht, dass Damasio im Rahmen seiner Theorie einen komplexeren Elizitationsmechanismus für Sekundäremotionen postuliert. Während er davon ausgeht, dass Emotionen bei Menschen wie auch bei Tieren auftreten,⁵³ emotionale Gefühle jedoch wahrscheinlich nur beim Menschen, so unterliegen Sekundäremotionen einer weiteren Einschränkung: sie entwickeln sich erst im Laufe der Ontogenese eines Menschen, kommen also beim Neugeborenen gar nicht vor, während sie beim erwachsenen Menschen sogar den Großteil aller emotionalen Erfahrung ausmachen. Sekundäremotionen entstehen – so Damasio – „once we begin experiencing feelings and forming *systematic connections between categories of objects and situations, on the one hand, and primary emotions, on the other*“ (134). Hier ist zunächst festzuhalten, dass (1.) Sekundäremotionen die Fähigkeit zu emotionalen Gefühlen voraussetzen und damit die Fähigkeit des Subjekts, emotionale Zustände *bewusst* zu erleben und Reflexionen über vermutete Objekte, Kontexte oder Szenarios anzustellen. Nur ein Subjekt, das Emotionen bewusst erleben kann und über Objekte und Szenarios zu generalisieren vermag, ist – so können wir schlussfolgern – auch in der Lage, die von Damasio postulierten „systematischen Verbindungen“ zwischen Objektkategorien und Primäremotionen herzustellen. (2.) Weiterhin versteht Damasio Sekundäremotionen nicht als fest verschaltete Mechanismen von Emotionszentren im Gehirn, sondern als Ausdruck von Lernprozessen. Darum können Sekundäremotionen auch nicht mehr allein durch die relativ begrenzten kognitiven Kapazitäten des limbischen Systems erklärt werden. (3.) Dennoch ist festzuhalten, dass Sekundäremotionen sich nicht fundamental von Primäremotionen unterscheiden: indem sie schließlich systematische Verbindungen zwischen Objektkategorien bzw. Situationen und Primäremotionen darstellen, gehen letztere gleichsam in das Phänomen der Sekundäremotionen mit ein. Damasio betont mehrmals, dass sich Sekundäremotionen über den Mechanismus der Primäremotionen ausdrücken, bzw. auf diesem „Huckpack reiten“ (137 fig. 7-2: „piggy-backed“).

⁵³ Vgl. das Interview „Auch Schnecken haben Emotionen“. Der Neurobiologe Antonio Damasio über seine Theorie der Gefühle, das Zusammenspiel von Hirn und Körper und die Nachteile eines Lebens ohne Angst, in: Der Spiegel (01.12.2003).

Wie beschreibt Damasio nun den Mechanismus der Aktualgenese von **Sekundäremotionen**? Zunächst leitet er seine Ausführungen mit einem Beispiel ein: er fordert den Leser auf, sich vorzustellen, einen alten Freund zu treffen, den man lange nicht gesehen hat, oder vom unerwarteten Tod eines engen Kollegen zu erfahren. Damasio vermutet, dass der Leser bei einem solchen Selbstexperiment eine Emotion erleben wird, die wahrscheinlich auch mit milden Formen der physiologischen Erregung einhergeht. Bei Sekundäremotionen beginnt das Reaktionsschema (1.) mit der Aktualisierung mentaler Bilder, also mit einem kognitiven Prozess: „The process begins with the conscious, deliberate considerations you entertain about a person or situation. These considerations are expressed as mental images organized in a thought process“ (136); es handle sich um eine „kognitive Evaluation“ des Objekts. Diese mentalen Bilder können non-verbal (Situationen, Schlüsselszenarien, Objekte, Personen) oder verbal (Überzeugungen, Erwägungen, Überlegungen, Urteile, Schlussfolgerungen etc.) sein. In einem nächsten Schritt (2.), der wiederum unbewusst und spontan abläuft, werden Hirnareale im präfrontalen Großhirnkortex (im Schaubild „VM“= ventromedialer präfrontaler Großhirnkortex) aktiviert, sofern dort Muster von gelernten Assoziationen der aktualisierten Bilder mit erlebten Primäremotionen abgelegt sind. VM führt – in der gleichen Weise wie die Amygdala im Mechanismus der Primäremotionen – einen Mustervergleich durch. Dieses neurale Substrat *erworbener* Dispositionen (im Gegensatz zu *angeborenen* im Falle von Primäremotionen) verkörpert also gelerntes Wissen eines Subjekts darüber, wie bestimmte Szenarien typischerweise mit bestimmten erlebten (d.h. gefühlten) Emotionen verbunden waren. Einerseits ist die Fähigkeit zu und die tatsächliche vorgängige Realisierung von erlebten Primäremotionen die notwendige Bedingung dafür, dass ein Individuum Sekundäremotionen erleben kann. Ein Individuum, das zu keiner elementaren Emotionalität fähig ist, kann schließlich auch nicht die emotionale Tönung von solchen paradigmatischen Szenarien erlernen. Andererseits ist die Ausprägung solcher Dispositionen zu Sekundäremotionen irreduzibel individuell: sie hängt ab von der konkreten Lebenserfahrung eines Individuums, auch wenn sie auf einen Mechanismus von angeborenen dispositionalen Reaktionen aufbaut. Zwar werden die Dispositionen zu Sekundäremotionen intersubjektiv sehr ähnlich sein, da zumindest die Mitglieder derselben Gesellschaft prinzipiell ähnliche Lebenserfahrungen machen. So können wir vermuten, dass wir fast alle mit Freude reagieren, wenn wir uns das Wiedersehen mit einem alten Freund vorstellen. Dennoch sind durch die Berücksichtigung individueller Erfahrungen und durch die Verbindung mit einem dynamischen Denkprozess (das Aktualisieren von Bildern) unendlich viele feine Nuancen möglich: so kann z.B. bei Hans

regelmäßig eine subtile Traurigkeit zu dieser Freude gemischt sein, weil evtl. in seiner Erfahrung auf Begegnungen immer eine spürbare Trennung folgte etc. Schließlich (3.) wird auf der Grundlage des Mustererkennungsprozesses in VM eine Aktivierung entsprechender Reaktionsmuster im limbischen Systems (v.a. Amygdala) vermittelt und der oben beschriebene Mechanismus der Primäremotionen ausgelöst. Das Erleben des Feedbacks über die entstandenen Körperveränderungen löst (4.) wiederum ein emotionales Gefühl aus.



Diese Funktion des präfrontalen Kortex (vgl. dessen Lage im Schaubild links nach Trepel 1999; oben Ansicht temporal (a), unten medial auf Höhe des corpus callosum (b)) wird auch durch neuroanatomische Erkenntnisse gestützt: er empfängt Afferenzen aus fast allen Großhirnkortexarealen, dem limbischen System (Amygdala), dem Thalamus und der formatio reticularis; Efferenzen fließen in zahlreiche kortikale Areale und den Thalamus. Nach Trepel (1999) hat eine bilaterale Läsion des präfrontalen Kortex „schwerste Persönlichkeitsveränderungen“, „Herabminderung intellektueller Fähigkeiten und der psychischen und motorischen Motivation“ und Veränderung der ethischen Grundhaltung (scham- und taktloses, enthemmtes Verhalten bei völliger Gleichgültigkeit des Patienten) zur Folge. Damit ist im Grunde das gleiche Syndrom beschrieben, das Damasio bei den von ihm untersuchten „soziopathischen“ Patienten festgestellt hatte. Patienten wie „Elliot“ oder der historische Phineas Gage zeigen eine solche Läsion. Die von Damasio beschriebene und untersuchte Defizienz im Ausführen von individuellen Handlungsplänen und die allgemeine emotionale Rohheit und Unfähigkeit zu geregelten sozialen Kontakten (vgl. Damasio 1994, Kap. 9) sind wohl der auf die Läsion zurückgehenden Unfähigkeit des Erlebens von Sekundäremotionen geschuldet. Das Aktualisieren und Erwägen von Bildern, die mit typischerweise emotionsrelevanten Szenarien korrelieren, ist bei diesen Patienten nicht mehr mit emotionaler Tönung versehen. Einerseits fallen dadurch sozial erwartete Emotionen aus, andererseits stellen sich z.B. beim deliberativen Erwägen von Handlungsoptionen alle mental repräsentierten Informationen als eintönig und flach dar. Solche Patienten können jedoch immer noch Primäremotionen erleben. Wenngleich ein solcher Patient also im sozialen Umgang affektiv verroht oder gar apathisch wirkt, würde er dennoch Furcht empfinden, wenn er nachts allein im Wald ein plötzliches Geräusch aus dem Unterholz vernehmen würde. Umgekehrt ist die Apathie von Patienten mit Läsionen des limbischen Systems (v.a. der Amygdala) viel grundlegender: sie können weder Primär- noch Sekundäremotionen haben.

Damasio nimmt auch die Existenz von „**as-if feelings**“ an (vgl. Damasio 1994., 155f.), bei denen die Körperschleife umgangen wird. Hierin liegt nun eine theoretische Möglichkeit zu erklären, wie sog. „intellektuelle Emotionen“ entstehen, eine Erklärung, die im Rahmen der James-Lange-Theorie z.B. für schwere Probleme gesorgt hatte. Zur Frage steht, wie „feinere“ Emotionen entstehen, die einerseits ohne signifikante physiologische Aktivierung auftreten, teilweise auch nur als gefühlsmäßige Tönung anderer Bewusstseinsinhalte (wie Gedanken, Überlegungen, Vorstellungen, Überzeugungen oder Einstellungen), und die andererseits allein durch die Tätigkeit des Geistes, ohne ein akutes Reiz-Reaktions-Schema hervorgerufen werden. Aus neurologischer Sicht entwickelt Damasio eine Erklärung mittels der Annahme einer „Als-ob-Schleife“ der Emotionsgenese, die die oben erläuterte „Körperschleife“ umgeht. Erst nachdem ein Organismus in seiner ontogenetischen Geschichte emotionale Erfahrungen gemacht hat, können sich neuronale Muster im Sinne von synaptischen Verschaltungen ausbilden, die im Stande sind, Gehirnregionen zu aktivieren, die eigentlich durch ein entsprechendes Körperfeedback aktiviert würden. Fast prophetisch mutet daher James' Formulierung von „zerebralen Emotionen“ an, die er als Synonym für seine Ausdrücke „feinere“ oder „intellektuelle“ Emotionen benutzte. Nach Damasios Ansicht verhält es sich in der Tat so, dass „Als-ob-Gefühle“ eine emotionale Erfahrung im Gehirn („zerebral“) simulieren, ohne dass die Körperperipherie zuvor in einen Aktivierungszustand versetzt wurde.

Nun könnte man einwenden, dass wir an diesem Punkt dort angekommen sind, wo wir vor dem Exkurs zur Neuropsychologie der Emotionen begonnen hatten. Ich denke jedoch, nicht ganz. Nicht unbedeutend erscheint mir die Annahme, dass sog. intellektuelle Emotionen allein aus der Perspektive einer Ontogenese der emotionalen Entwicklung verstanden werden können: sie sind insofern ein Produkt der primären und sekundären Emotionsmechanismen, das zwar synchronisch betrachtet einen kategorialen Unterschied markiert (emotionales Erleben ohne physiologische Aktivierung), allerdings diachronisch betrachtet ontologisch von den primitiveren Emotionsmechanismen abhängig ist. Denn nur eine persönliche Geschichte emotionaler Erfahrungen eröffnet in der ontogenetischen Entwicklung die Möglichkeit zur Anlage eines neuronalen Mechanismus zur Generierung intellektueller Emotionen.

Damasio (ebd., 157) stellt zwei Theorien zum Zusammenhang von Körper und Gefühl vor, wobei er schließlich für die erstere optiert. (1.) Gefühle entstehen durch das „Auslesen“ des Körper-Zustands. Davon ausgenommen sind die als-ob Gefühle, die also einen abgeleiteten Fall darstellen. Oder: (2.) Eine Hirnregion setzt physiologische Veränderungen in

Gang und informiert das somatosensorische System darüber, das wiederum ein Gefühl entstehen lässt. Hier wäre Gefühl und Körperveränderung lediglich parallel und nicht sukzessive-kausal. Die als-ob Schleife wäre dann nicht abgeleitet, sondern der Standardfall. Damasio optiert jedoch für die erste Variante. Denn nach (2) wäre jedes Gefühl identisch mit einem prototypischen phänomenologischen Muster, weil schließlich keine Körperrückmeldungen einbezogen werden, die individuelle Vorgänge, dynamische Erlebnisveränderungen usw. in das Gefühlserleben einbringen könnten. Außerdem muss der Körper in jedem Fall im Gefühlserleben repräsentiert sein, weil neben der neuralen Verbindung zwischen Hirn und Körper noch die chemische Route existiert. Sie transportiert Hormone und andere Botenstoffe zwischen Hirn und Körper, und selbst wenn das Gefühlserleben autonom innerhalb des Hirns ausgelöst würde, während parallel die Körperveränderungen abliefen, würde die chemische Route die Veränderungen dennoch zurückmelden, wenn auch mit erheblicher Verzögerung, da sie gegenüber der neuralen Route erheblich höhere Latenzzeiten aufweist. Diese beiden empirischen Befunde sprechen also eindeutig für einen Primat der Körperschleife: Die von einer best. Hirnregion ausgelösten körperlichen Veränderungen werden von einer anderen Hirnregion ausgelesen und führen zu unseren emotionalen Gefühlen. Im Ausnahmefall (im Sinne eines Derivats) der Als-ob-Schleife kann dieser Weg umgangen werden.

Die Annahme einer Als-Ob-Schleife, die den Weg des Körperfeedbacks umgeht und soz. die emotionale Erregung „zerebral simuliert“, wird gestützt durch einen Befund von Stuart Valins (1966). Bei dem Experiment wurden männlichen Probanden Bilder von halb entblößten Frauen gezeigt. Während des Betrachtens der Bilder erhielten sie eine akustische Einspielung ihres angeblich aktuellen Herzschlags. In Wirklichkeit wurde jedoch die Frequenz der angeblichen Herztöne willkürlich manipuliert. Nach der Betrachtungsserie wurde die Vpn aufgefordert, die Bilder nach ihrer Attraktivität zu beurteilen. Dabei wurden diejenigen Bilder für attraktiver gehalten, bei denen Töne einer hohen Herzfrequenz eingespielt worden waren, selbst dann, wenn die eingespielte von der tatsächlichen (ebenfalls gemessenen) Herzfrequenz signifikant dissoziiert war. Das Gehirn bildet also nicht nur Mechanismen aus, Gefühle der Körpererregung unabhängig von der Beteiligung des Körpers zu erzeugen, wir können uns sogar bewusst einen Zustand der Erregung vorgaukeln, der sich auf unser emotionales Erleben auswirkt. Allerdings sprechen weder das Valins-Experiment noch Damasio's Modell der Als-Ob-Schleife gegen eine Beteiligung des Körpers im Prozess der Emotion. In diesem Sinne lautet auch LeDoux's (1996, 318) Bewertung des Valins-Experiments:

„[Dieses Ergebnis] war natürlich nur möglich bei einem Gehirn, das schon viele reale Rückmeldungen erlebt hatte, sodass die Art und Weise, wie eine Rückmeldung sich anfühlt und funktioniert, imaginiert und dazu benutzt werden konnte, die subjektiven Erlebnisse tatsächlich zu beeinflussen. Dies spricht eher für als gegen die Rolle körperlicher Zustände bei emotionalen Gefühlen.“

Um die Ausführungen zu Damasio zu beschließen, können wir zusammenfassend seine Theorie anhand eines fünf Punkte umfassenden Thesenkatalogs charakterisieren, den Damasio in einem rezenten Aufsatz (2001) aufstellt: (1.) Emotionen sind komplexe, musterhafte Syndrome von chemischen und neuralen Reaktionen. Sie haben eine adaptive Funktion und beziehen sich auf die (körperliche) Existenz ihres Subjekts. (2.) Emotionen sind biologisch determinierte Prozesse, die durch Lernen, Sozialisation und Kultureinflüsse verändert und mit variierender Bedeutung versehen werden können. (3.) Emotionsrelevante Hirnareale umfassen eine relativ eingeschränkte Gruppe von subkortikalen Regionen und gehören zu übergeordneten Strukturen, die Körperzustände regulieren und repräsentieren. (4.) „All the devices can be engaged automatically, without conscious deliberation“ (ebd., 103). (5.) Emotionen benutzen die Körperperipherie als ihr „Theater“, verändern aber auch Verarbeitungsmodi und neurale Prozesse im Gehirn. Beides – Körperveränderungen und neurale Vorgänge – bilden das Substrat eines eventuell wahrgenommenen Gefühls.

2.9. Neurologische Grundlagen von Emotion:

Exkurs zu Schachter & Singer 1962, Hohmann 1966 und Damasio 1994

Nachdem nun alle wichtigen Schulen der Emotionstheorien des 20. Jahrhundert vorgestellt und kritisch gewürdigt wurden, können an dieser Stelle einige Fäden aufgenommen werden im Rahmen eines Exkurses zur Frage des Körperfeedbacks. Seit William James und Carl Lange das Konzept der peripher-physiologischen Aktivierung in besonders pointierter Form in die Debatte um Emotionen eingebracht haben, ist immer wieder darauf Stellung genommen worden. Wie sich im Laufe meiner Ausführungen gezeigt hat, lässt sich von vielen Seiten Kritik gegenüber der James-Lange-Theorie formulieren. Doch trotz aller Kritik (die sicherlich auch die Forschung in dieser Richtung vorwärts getrieben hat) hat man bis heute am Konzept der physiologischen Aktivierung festgehalten. Das Anliegen von Schachter & Singer (1962) war es, durch kognitive Etikettierungsprozesse („labeling“) Schwachpunkten in der James-Lange-Theorie gerecht zu werden. Und das rezente Modell von Damasio (1994 u.f.) versucht, die instinkthaften Körperreaktionen von James in eine

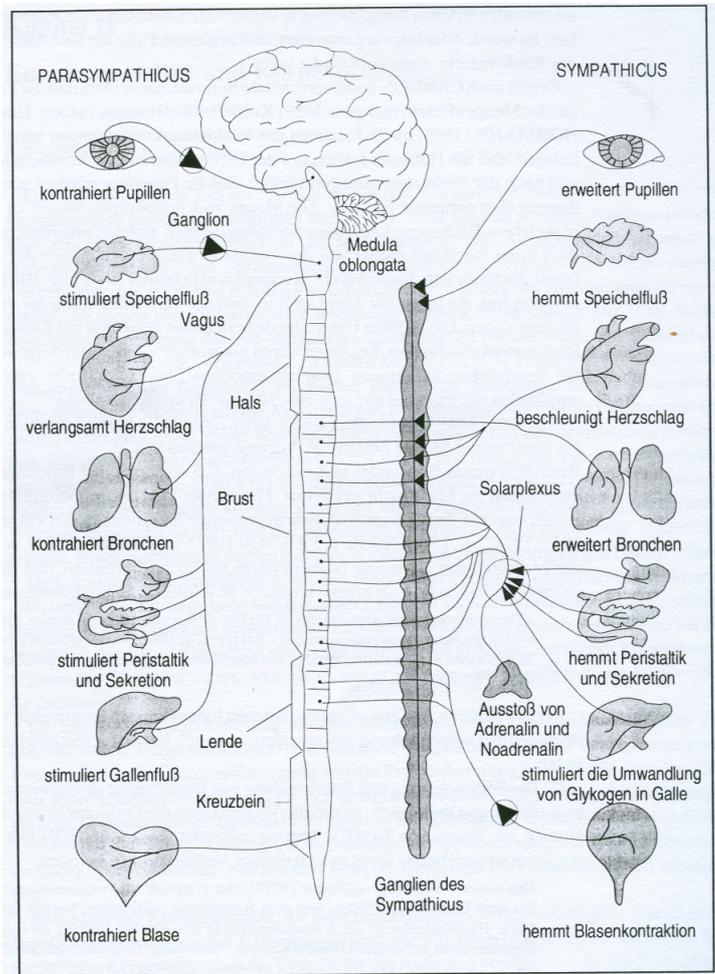
umfassende neuropsychologische Emotionstheorie zu integrieren, die auch der Rationalität der „feineren“ Emotionen gerecht wird.

Wie wir gesehen haben, fasst James den Begriff der Körperreaktionen (das spätere Paradigma der „physiologischen Aktivierung“) sehr weit, sodass vegetativ-nervöse (Blutdruck, Herzfrequenz, Peristaltik des Magen-Darm-Traktes) und halbwillkürliche (Lachen, Weinen) Reaktionen ebenso dazu gehören wie komplexe Handlungen (zuschlagen, weglaufen). Bei Schachter & Singer ist dann nur noch die Rede von peripher-physiologischer Aktivierung, womit eben die Parameter des vegetativen (oder autonomen) Nervensystems gemeint sind. Damasio fasst das Paradigma nun wieder etwas weiter und schließt auch die Skelettmuskulatur mit ein. Damit ist also einerseits der Muskeltonus der gestreiften Muskulatur gemeint (im Gegensatz zur glatten Muskulatur der Eingeweide), aber auch Rückmeldungen von allgemeinen Körperbildern, die Konfigurationen der Extremitäten abbilden (Kauern, „weiche Knie“, Abwehrgestik etc.). Andererseits schließt Damasio explizit die durch Izard und Ekman bahnbrechend und umfangreich erforschte Rückmeldung der Gesichtsmuskulatur mit ein. Angesichts der Komplexität der beteiligten Organe, Extremitäten und Muskelgruppen erscheint die Frage interessant, wie die Ergebnisse von Hohmann (1966) in diesem Rahmen zu interpretieren seien.

Wie bereits diskutiert, hatte Hohmann Patienten mit einer Läsion des Rückenmarks untersucht und dabei festgestellt, dass die Intensität des emotionalen Erlebens mit steigender Höhe der Läsion auf dem Niveau der Wirbelsäule abnimmt. Dies wurde als eine Bestätigung der Schachter-Singer-Theorie angesehen, der gemäß die physiologische Aktivierung („arousal“) die Intensität einer emotionalen Reaktion, die kognitive Etikettierung deren Qualität bestimmt. Nun ist jedoch das Rückenmark (die medulla spinalis) nicht der einzige Weg, über den vegetativ-nervöse Reaktionen vermittelt werden.

Das zentrale Nervensystem (ZNS) des Menschen gliedert sich in ein somatisches Nervensystem, das z.B. die willkürliche Steuerung der Skelettmuskulatur und die bewusste Wahrnehmung der Körperperipherie vermittelt, und ein vegetatives (oder: autonomes) Nervensystem (ANS), das wiederum die unbewusste und unwillkürliche Steuerung vegetativer Prozesse wie Atmung, Herzschlag, Verdauung steuert sowie die unbewusste Wahrnehmung viszeraler Prozesse vermittelt. Das ANS hat einen sympathischen und einen parasympathischen Anteil (vgl. Schaubild unten). Die jeweiligen Korrelate des Sympathikus bzw. Parasympathikus lassen sich dem Schaubild entnehmen, zusammenfassend kann man jedoch ohne Weiteres formulieren, dass der Sympathikus eine energiemobilisierende Funktion hat, während der Parasympathikus eher energiekonservierend wirkt. Eine physiologische

Aktivierung im Sinne von „arousal“ wäre also schon begrifflich eher dem sympathischen Anteil zuzuordnen. Angemerkt sei schließlich, dass sich die Aufteilung des ANS in Sympathikus und Parasympathikus allein auf die efferenten Nervenfasern bezieht, also auf jene, die Nervenimpulse vom Gehirn in die Körperperipherie leiten. In umgehrter Richtung, also afferent, wird diese Unterscheidung sinnvollerweise nicht getroffen (vgl. Trepel 1999).



Schema des ANS
nach Goller 1992, 32

Die Theorien von James, Schachter & Singer und Damasio verlangen nun allesamt zweierlei: die Vermittlung von efferenten Impulsen in die Körperperipherie zur Aktivierung bestimmter Organe und Muskelgruppen, sowie das afferente Feedback dieser Aktivierung. Welche Situation liegt nun bei den Hohmann-Patienten vor? Das Rückenmark, das in diesem Fall auf einem bestimmten Niveau durchtrennt ist, gliedert sich in graue (Hinter-, Seiten- und Vorderhorn) und weiße Substanz (sensible und motorische Bahnen). Im Hinter- und Seitenhorn werden nun propriozeptive Afferenzen, vegetative Afferenzen der Viscera und viszerale Efferenzen des Sympathikus und Parasympathikus vermittelt. Die Bahnen der weißen Substanz vermitteln die sensible (afferente) und motorische (efferente) Versorgung der Skelettmuskulatur und des Schmerz- und Tastempfindens. Tatsächlich sind also afferent

und efferent alle organischen Korrelate einer physiologischen Aktivierung hier repräsentiert. Die mir bekannten Studien⁵⁴ diskutieren in dieser Hinsicht jedoch nicht die Relevanz alternativer Innervierungswege. Zu nennen ist hier v.a. der Nervus vagus. Er gehört zu den 12 Hirnnerven (X. Hirnnerv), nimmt unter diesen aber eine Sonderstellung ein, insofern er als einziger Hirnnerv über den Schädel- und Kopfbereich hinaus bis weit in die Körperperipherie hinein innerviert. Der N. vagus übernimmt durch seine Verschaltung in verschiedenen Nervenkerne des Hirnstamms u.a. die gesamte *parasympathische* allgemein-viszeromotorische Versorgung des Körpers vom Hals über Thorax (Herz, Bronchien) bis ins Abdomen (Gastrointestinaltrakt),⁵⁵ sowie die Vermittlung allgemeiner und spezifischer viszerosensibler Afferenzen aus Lungen, Trachea, Ösophagus (Speiseröhre), Gastrointestinaltrakt, rechtem Herzvorhof und Aortenbogen.⁵⁶ Festzuhalten ist hier, dass der N. vagus zwar viszerosensible Afferenzen transportiert (die ja nicht nach Sympathikus und Parasympathikus unterschieden werden), jedoch efferent nur Anteile des Parasympathikus. Stellen wir uns einmal vor, bei den Hohmann Patienten sei der N. vagus intakt (wovon wohl auszugehen ist), während die Versorgung durch das Rückenmark auf einem bestimmten Niveau unterbrochen ist. Die Herstellung eines physiologischen Aktivierungszustands kann nicht über das Rückenmark vermittelt werden. Der Weg über den N. vagus ist jedoch ebenfalls ausgeschlossen, da dieser efferent nur *parasympathische* Anteile führt, also nur dämpfend, und nicht erregend auf die inneren Organe wirken kann. Andersherum jedoch wäre eine *Rückmeldung* des Körperfeedbacks über den N. vagus ohne Weiteres zu erklären, wenn wir an die oben ausgeführten Details der Innervierung denken. Die Interpretation der Hohmann-Studie, nämlich dass sie zeige, Patienten mit Rückenmarksläsionen könnten kein Körperfeedback mehr *verarbeiten*, ist in diesem Sinne also ungenau. Tatsächlich kann das ANS der betroffenen Patienten keine physiologische (sympathische) Aktivierung mehr *herstellen*. Der Befund der Studie ist jedoch in beiden Fällen derselbe, sodass die hier aufgezeigte Ungenauigkeit nicht die Relevanz der Untersuchung in Frage stellt.⁵⁷

⁵⁴ Mit Ausnahme von LeDoux (1996, 316), der explizit auf den Nervus vagus hinweist, ohne jedoch im Detail zu diskutieren, welche vegetativen Anteile dieser überhaupt vermittelt.

⁵⁵ Verschaltung im Nucleus dorsalis nervi vagi.

⁵⁶ Verschaltung im Nucleus tractus solitarii.

⁵⁷ Interessant wäre also u.U. eine Wiederholung des Schachter-Singer-Experiments, wobei als Probanden dann Hohmann-Patienten (also Patienten mit Rückenmarksläsionen) zu wählen wären. Da ja nicht die Fähigkeit zur Rückmeldung, sondern zur Aktivierung von peripher-physiologischen Prozessen ausgeschaltet ist, sollte diese Probandengruppe völlige normale Ergebnisse im Schachter-Singer-Experiment liefern, da ja gerade die physiologische Aktivierung künstlich durch Epinephrin-Gabe induziert wird.

Wir können darüber hinaus jedoch annehmen, dass bei den Hohmann-Patienten noch weitere Wege des Körperfeedbacks bzw. zur efferenten Aktivierung intakt waren. Zum einen existiert parallel zur nervösen Bahn der Informationsvermittlung noch die hormonelle. In dieser Hinsicht sind jedoch zahlreiche Bedenken vorgetragen worden, unter denen das wichtigste wohl die Tatsache ist, dass hormonelle Prozesse sehr langsam ablaufen, bisweilen mit einer Latenzzeit von mehreren Stunden. Zwar sind hormonelle Einwirkungen auf Affekt- (im Sinne von Stimmungs-) Störungen teilweise recht gut untersucht, man denke z.B. an die Korrelation von Depressionserkrankungen mit Serotonin-Niveaus. Dort werden auch die langen Latenzzeiten dieser Prozesse nicht ins Gewicht fallen, da Stimmungen (und „Verstimmungen“) sich schließlich durch eine gewisse Ebenmäßigkeit in ihrem Verlauf und eine allgemein längere Dauer auszeichnen. Andersherum laufen hormonelle Prozesse evtl. in der Tat zu langsam ab, um in die kritische Dynamik von akuten emotionalen Zuständen eingehen zu können. Wie man auch immer darüber spekulieren will, der Zusammenhang ist jedenfalls bisher nicht ausreichend untersucht.

Neben den oben geschilderten Wegen existieren schließlich noch weitere Nervenbahnen, über die ebenfalls Rückmeldungen über Körperzustände vermittelt werden können. Dabei handelt es sich jedoch – und darum seien sie hier am Schluss behandelt – um Bahnen, die Muskelgruppen innervieren, die z.B. Schachter & Singer nicht als Korrelate ihres Paradigmas der peripher-physiologischen Aktivierung gelten lassen, nämlich die Skelettmuskulatur. Wie oben erwähnt, wird diese jedoch von James und Damasio explizit mit einbezogen. Zunächst muss man, bezogen auf das Beispiel der Hohmann-Patienten, wohl die Rückmeldung von skelett-muskulären Vorgängen in der Körperperipherie ausschließen. Wie oben ausgeführt wurde, werden die bewussten sensiblen und motorischen Prozesse durch die weiße Substanz der medulla spinalis vermittelt, deren Informationsfluss bei den vorliegenden Läsionen ebenfalls unterbrochen sein muss. Dadurch zeichnet sich ja gerade das Krankheitsbild der „Querschnittslähmung“ aus (und nichts anderes liegt bei den Hohmann-Patienten vor), dass weder willkürliche Ansteuerung der Skelettmuskulatur noch bewusstes Empfinden der Körperperipherie unterhalb der Läsion noch möglich ist. Allerdings schließen sowohl James als auch Damasio die Emotionsmimik der Gesichtsmuskulatur mit ein. Bei James ist dies implizit, indem er das Beispiel des Lachens anführt. Bei Damasio wird die Annahme theoretisch (neuroanatomisch) expliziert mit Rückbezug auf die grundlegenden Forschungen von Izard und Ekman. Erinnern wir uns, wie James pointiert formulierte: „wir sind fröhlich, weil wir lachen“. Die Gesichtsmuskulatur wird durch zwei Hirnnerven innerviert, u.z. sensibel über den Nervus trigeminus (V. Hirnnerv) und motorisch über den N. facialis (VII.

Hirnnerv). Letzterer wird im Pons u.a. im Nucleus nervi facialis verschaltet, der motorische Impulse zur Versorgung sowohl der willkürlichen als auch der emotionalen Mimik erhält, jedoch über getrennte efferente Bahnen (Trepel 1999, 110).⁵⁸ Auch allgemein viszeromotorische Impulse der Tränen-, Nasen- und Gaumendrüsen werden durch den N. facialis versorgt und im Nucleus salivatoris superior verschaltet. Wir können also die sensible und motorische Versorgung verschiedener Prozesse wie Mimik, Weinen, Trockenheit des Mundes etc. über die Hirnnerven V und VII erklären, u.z. unabhängig von den Bahnen des Rückenmarks. Auch bei Hohmann-Patienten wäre also sowohl eine (efferente) Ansteuerung solcher Prozesse, als auch eine (afferente) Rückmeldung ohne Weiteres möglich. Gemäß der Theorie vom „facial feedback“ (wie auch nach James im Übrigen) haben solche Rückmeldungen einen Einfluss auf Qualität und Intensität des erlebten emotionalen Zustands. Damit sind wohl auch die widersprüchlichen Ergebnisse der Studie von Bermond et al. (1991) zu erklären. Beim Versuch, die Hohmann-Studie zu replizieren, konnten keine der dort erzielten Befunde festgestellt werden. Rückenmarksläsionen schienen keinen Einfluss auf das emotionale Erleben zu haben. Wenn wir dies nun den alternativen Wegen des Körperfeedbacks z.B. über die Nervi trigemini und facialis zuschreiben, kann man natürlich zu Recht fragen, warum sich dieser Effekt nicht auch bei Hohmann eingestellt hatte. Darum unterstellen die Autoren der jüngeren Studie diesem auch eine fehlerhafte Ausführung. Zu berücksichtigen sind in dieser Hinsicht natürlich die Unzahl von Variationen, die sich bei qualitativen Studien ergeben können, besonders bei einer Untersuchung durch Selbstauskunft der Probanden, an einem empirisch so schwer greifbaren (weil irreduzibel subjektivem) Phänomen wie dem emotionalen Erleben. Eine weitere Variable können schließlich beide Studien nicht ausschließen, nämlich den von Damasio postulierten Mechanismus der Als-Ob-Schleife. Über diesen Mechanismus kann das Gehirn, nachdem sich durch viele emotionale Erfahrungen die neuralen Muster bestimmter körperlicher Aktivierung eingepägt haben, eine solche Aktivierung auch „zerebral simulieren“. Der Mechanismus, wenn er denn existiert, ließe sich bei Probanden mit Unfallverletzungen nicht ausschließen, und bei Probanden mit angeborenen Rückenmarksschädigungen wäre intrasubjektiv keine Möglichkeit zum Vergleich gegeben, da nie ein anderer als der status quo (und damit der vermeintlich defizitäre Zustand) des emotionalen Erlebens bestand.

⁵⁸ Dadurch sind die vielfach beobachteten Effekte des selektiven Ausfalls entweder der einen oder der anderen zu erklären.

2.10. Ausblick: Emotionen und Akteure

Warum sind Emotionen relevant für Prozesse des Urteilens und Entscheidens, und damit für eine Theorie des Handelns überhaupt? Wir haben im Verlauf unserer Untersuchung zur Emotionstheorie verschiedene Standpunkte kennengelernt, ob und wie Emotionen mit Handlungstendenzen oder Motivation zusammenhängen. Eine prominente These zu dieser Fragestellung entwickelt Damasio (1994, Kap. 8) auf der Grundlage seiner bereits vorgestellten neuropsychologischen Emotionstheorie. Nach Damasio (1994, 168f.) finden Prozesse des Urteilens und Entscheidens entweder (1) im persönlichen bzw. sozialen Raum oder (2) im abstrakten Raum statt. Während die unter (1) bezeichneten Szenarien keiner Erläuterung bedürfen, muss hinzugefügt werden, dass Damasio mit (2) etwa das Lösen von Mathematik- und Logik-Aufgaben oder auch die Arbeit an wissenschaftlichen Fragestellungen ohne direkte Handlungsimplicationen für das Subjekt selbst innerhalb seines sozialen Bezugsrahmens versteht. Die von Damasio untersuchten Vpn mit pathologischen Veränderungen in den emotional relevanten Hirnregionen zeigen im Falle von (2) eine normale Performance, bei (1) jedoch scheitern sie. Wir kennen wohl auch aus unserer (freilich klischeehaft überzeichneten) Alltagserfahrung nicht-pathologische Persönlichkeiten mit vergleichbaren Symptomen. Der „zerstreute Professor“ etwa, der zwar ein brillanter Denker ist, aber durch die Planung von sozialen Verpflichtungen gründlich überfordert wird. Auf der Grundlage der empirischen Befunde mit seinen VM-Patienten formuliert Damasio die sog. „Somatic Marker Hypothesis“ (vgl. auch Bechara et al. 1997 und Damasio 1996):

„In short, somatic 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. This is the essence of the somatic marker hypothesis. [...] On occasion somatic markers may operate covertly (without coming to consciousness) and may utilize an ‘as-if-loop’.“ (Damasio 1994, 174)

Nach Damasio werden im aktuellen Entscheidungsprozess die aufgerufenen und erwogenen mentalen Bilder der verschiedenen Optionen durch bestimmte, früher erlebte und mit diesen Bildern assoziierte emotionale Zustände „markiert“. Dieser Mechanismus versucht gewissermaßen, unser Gefühl der „Intuition“, oder das Erlebnis bestimmter „Bauchgefühle“ hinsichtlich einer Entscheidungsoption oder eines bevorstehenden Ereignisses als einen unbewussten somatischen Markierungsprozess zu verstehen. Ein intuitiv erlebtes Gefühl der Vorsicht kann dann bspw. dazu führen, dass Handlungstendenzen abgeschwächt werden. Wichtig ist jedoch, wie Damasio (ebd.) anmerkt: „Somatic markers do not deliberate for us“.

Die assoziierten Gefühle markieren lediglich bestimmte Optionen mit einer affektiven Attraktions- bzw. Repugnanz-Qualität. Damasio testet seine Somatic Marker Hypothesis empirisch mit Hilfe des von ihm selbst entwickelten „Iowa Gambling Task“.⁵⁹

Somatische Marker sind nach Damasio prinzipiell adaptive Mechanismen. Die Folgen des Fehlens eines entsprechenden Mechanismus – bei VM-Patienten durch Hirnläsionen begründet – illustriert Damasio anhand einer Anekdote: so berichtet er (Damasio 1994, 192f.) von einem Patienten mit ventromedialer präfrontaler Läsion, der zu einer Sitzung ins Labor kam und dafür an einem Wintertag durch sehr ungünstige Wetterbedingungen mit dem Auto anreisen musste. Auf die Frage Damasios, wie die Fahrt gewesen wäre, erwiderte der Patient, dass es kein Problem sei, solange man die entsprechenden Vorschriften zum Fahren auf glatter Straße beachte. Er hatte unterwegs eine Fahrerin beobachtet, die genau vor ihm über einen vereisten Straßenabschnitt fuhr, ausscherte und anstatt gegenzulenken auf die Bremse trat. Diese instinktive Reaktion auf Grund eines emotionalen Impulses müssen wir für gewöhnlich willkürlich unterdrücken, um rational richtig reagieren zu können (also ohne in Panik zu geraten gegenzulenken). Damasios Patient war in der Lage, ohne jegliche emotionale Beeinflussung mit klarem Kopf über die vereiste Fläche zu fahren und die entsprechenden Manöver auszuführen, obwohl er genau vor sich erlebt hatte, wie jemand aus Angst disruptiv reagiert hatte. Hier war das Fehlen eines entsprechenden somatischen Markers eindeutig vorteilhaft. Damasio berichtet jedoch weiter, dass er dem selben Patienten am Ende jener Sitzung zwei Termine für die nächste Sitzung vorschlug. Daraufhin prüfte der Patient in seinem Kalender, ob einer der beiden ausgeschlossen werden könnte, und als dies nicht der Fall war, soll er eine halbe Stunde damit zugebracht haben, Pro und Contra beider Termine aufzustellen: Koordinierung mit anderen Terminen in der Gegend, Rahmenbedingungen wie Wetter und andere Aktivitäten in der entsprechenden Zeitspanne etc. Als Damasio selbst ungeduldig wurde und seinem Patienten nach der besagten halben Stunde mitteilte, dass er einfach zu dem zweiten der beiden Termine kommen solle, antwortete dieser ohne zu zögern: „That’s fine.“

⁵⁹ Vgl. Bechara et al. 2005, 160 fig. 1: „The participants are given four decks of cards, a loan of \$2000 facsimile US bills, and asked to play so as to win the most money. Turning each card carries an immediate reward (\$100 in decks A and B and \$50 in decks C and D). Unpredictably, however, the turning of some cards also carries a penalty (which is large in decks A and B and small in decks C and D). Playing mostly from decks A and B leads to an overall loss. Playing from decks C and D leads to an overall gain. The players cannot predict when a penalty will occur, nor calculate with precision the net gain or loss from each deck. They also do not know how many cards must be turned before the end of the game (the game in fact ends after 100 card selections).“

Wir finden also am Beispiel von Damasio Untersuchung eine psychologische Perspektive auf den Zusammenhang zwischen Emotion und Motivation, die jedoch für unsere weitere Argumentation fruchtbar sein kann, indem sie uns einerseits hypothesenbildend anleiten und andererseits ein implementierungsfähiges Modell auch für philosophische Handlungstheorien liefern könnte. Dies trifft jedoch nicht auf alle psychologischen Befunde zu dieser Fragestellung zu. So zeigen bspw. mehrere Studien (Isen/ Daubmann/ Nowicki 1987; Otto & Schmitz 1993), dass positive Stimmung kreatives Problemlösungsverhalten fördert. Die hier festgestellten statistischen Korrelationen jedoch können uns lediglich eine Richtung weisen. Sie sind m.E. jedoch nicht implementationsfähig in eine philosophische oder interdisziplinäre Studie.

Aus philosophischer Perspektive lassen sich ebenfalls Standpunkte zum Zusammenhang von Emotion und Motivation festhalten. Dies geschieht in der Regel über den „desire“-Begriff, etwa indem entweder angenommen wird, (1.) dass desires (im Sinne von „Präferenzen“, vgl. Lyons 1980) Elizitationsbedingungen von Emotionen sind, oder (2.) dass desires zu den Komponenten eines emotionalen Zustands gehören (Rey 1980, Marks 1982, Searle 1983, 29-36, Gordon 1987, Davis 1987 und Green 1992), oder (3.) dass „desire“ eine Manifestation von Emotionen ist (Wollheim 1999 und 2003). Eine typische „belief-desire-Analyse“ von Emotionen finden wir bspw. bei Robert Gordon (1987), dem gemäß der emotionale Zustand „A ärgert sich, dass p“ zu analysieren ist als: (1.) A glaubt dass p und (2.) A wünscht („desires“), dass nicht-p. Erinnern wir uns daran, dass Gordon zwischen zwei Typen von emotionalen Zuständen unterscheidet, nämlich $E1(S,p)$, der eine Emotion $E1$ bezeichnet, die ein Subjekt gegenüber dem propositionalen Inhalt p erlebt, und $E2(S,x)$, bei dem es sich um einen von S erlebten emotionalen Zustand gegenüber einem Objekt x handelt. Das daraus sich ergebende Problem ist nun, dass die belief-desire-Analyse vollständig plausibel nur auf den Typ $E1(S,p)$ zutrifft. Denn wie sollte eine Emotion des Typs (S,x) , bspw. „Hans verachtet Hermann“, analysiert werden in „Hans glaubt, dass p, und: Hans wünscht, dass nicht-p“? Emotionen involvieren mit Sicherheit ein Konstrukt in der Art von „desires“, es ist jedoch bisher noch nicht klar, in welcher Weise. Eine eindeutige sprachanalytische Lösung wie Gordon sie vorzuschlagen scheint, dürfte sich als problematisch erweisen.

Auch DeSousa (1979) vertritt die These der motivationalen Potenz von Emotionen, sein Modell der „paradigm scenarios“ jedoch dreht den Erklärungsansatz interessanterweise um. DeSousa betrachtet den ontogenetischen Prozess des Erwerbs von Emotionsbegriffen und beschreibt diesen – in aller Kürze – wie folgt:

„We are made familiar with the vocabulary of emotion by association with *paradigm scenarios*, drawn first from our daily life as small children, later reinforced by the stories and fairy tales to which we are exposed, and, later still, supplemented and refined by literature and art.“ (De Sousa 1979, 142)

Paradigmatische Szenarien (oder auf deutsch auch: Schlüsselszenarien) haben zwei Aspekte: (1.) eine Situation, die die charakteristischen Objekte einer Emotion zur Verfügung stellt (wobei man evtl. besser von „Zielen“ oder „Anlässen“ einer Emotion sprechen sollte) und (2.) eine geordnete Menge von charakteristischen Reaktionen auf dieses Situation. Die motivationale Potenz von Emotionen leitet sich gewöhnlich aus der Tatsache her, dass diese mit einer Menge von typischen Reaktionen korrelieren. DeSousa schlägt nun eine umgekehrte Lesart vor: wir reagieren in einer bestimmten Situation auf eine bestimmte Weise, und erwerben dadurch einen Emotionsbegriff. Nach dieser Zuordnung nehmen wir dann an, dass es die Emotion ist, die solche Verhaltensweisen *verursacht*. In Wirklichkeit jedoch ist die Emotion (als Begriff oder „Verständnis“ der Schlüsselsituation) dem korrelierenden Verhalten nachgeordnet.⁶⁰

De Sousa versteht Emotionen als „determinate patterns of salience among objects of attention, lines of inquiry, and inferential strategies“ (De Sousa 1979, 137). Demnach haben Emotionen Ähnlichkeiten sowohl mit Glaubenszuständen (*beliefs*), also auch mit Wünschen (*desires*). Sie haben einerseits eine *urteilsbildende Funktion*: insofern hängen sie mit Glaubenszuständen zusammen. Emotionen „*stellen* diejenigen Fragen, die ein Urteil mit Überzeugungen *beantwortet*“ („emotions ask the questions which judgment answers with beliefs“, ebd.). Insofern vergleicht De Sousa Emotionen mit wissenschaftlichen Paradigmen im Sinne von Thomas Kuhn (1962): sie sind etwas, in dessen Sinne wir die Welt sehen. Emotionen hängen also eng mit Urteilen und Überzeugungen zusammen, sie sind jedoch weniger von solchen getragen, als sie selbst ein Paradigma darstellen, das bestimmte Überzeugungen unter sich zu versammeln tendiert bzw. bestimmte Überzeugungen nach sich

⁶⁰ Das Konzept von Schlüsselszenarien bietet eine interessante neue Perspektive auf die Kognitivismus-Frage in der Emotionstheorie. Anhand dieses Konzepts kann überzeugend erklärt werden, warum bestimmte Emotionen – Schuld bspw. – offenbar kognitiv „gehaltvoller“ sind als andere – z.B. Furcht. Bestimmte Schlüsselszenarien lassen sich nicht *verstehen*, ohne dass das Subjekt über bestimmte kognitive Fähigkeiten verfügt. Ein Schuldgefühl kann nicht aufkommen, wenn ich nicht verstehe, dass es moralische Regeln gibt, die ich verletzen kann. Das heißt jedoch nach DeSousa nicht, dass die Emotion „Schuld“ bestimmte Kognitionen als Komponenten oder Aspekte enthält, sondern nur, dass ich die entsprechende Schlüsselsituation erkennen und verstehen muss. Wenn ich das nicht kann, werde ich keine Emotion erleben, die auf genau diesen Typus von Situation abgestimmt ist.

zieht.⁶¹ Emotionen haben eine *motivationale Funktion*. Sie sind zwar selbst keine Wünsche (desires), aber indem sie Aufmerksamkeitsfokussierung und Schlussfolgerungsstrategien eines Akteurs bestimmen, dirigieren sie unsere Motivation in den Fällen, wo ein gegebener Wunsch uns mehrere Möglichkeiten zu seiner Erfüllung lässt. DeSousa vertritt in dieser Hinsicht ein wesentlich differenzierteres Modell des Zusammenhangs von Emotionen und desires. Wir wollen diesem Zusammenhang im folgenden Kapitel weiter nachgehen. DeSousas Emotionsbegriff ist sicherlich nicht die Ausbuchstabierung des notorischen „*besire*“ (nach der von Altham 1986 geprägten Wortschöpfung aus „*belief*“ + „*desire*“), mit dem sich auf einen Schlag das von Smith (1994) aufgestellte „*moral problem*“ lösen ließe. Wir können vielmehr sagen: gäbe es „*besires*“ in der Welt, dann wäre per se auch das *moral problem* nicht existent. Ich denke dennoch, dass ein Lösungsansatz zum *moral problem* sich auf der Grundlage unserer gewonnenen Erkenntnisse zur Theorie der Emotionen wird erarbeiten lassen. Beschließen wir das vorliegende Kapitel daher mit einer Beobachtung von DeSousa und mit einem Resümee der bisherigen Ergebnisse im Rahmen eines kurzen Theoriekatalogs. De Sousa unternimmt ein Gedankenexperiment, wie wir als menschliche Akteure wohl wären, wenn wir keine Emotionen hätten:

„Let us try to imagine what it would be like to be without emotions. It is not easy to make the supposition clear in view of the close connection between emotion and

⁶¹ Eine eindrucksvolle Illustration von DeSousas These, dass Emotionen „*patterns of salience*“ seien, liefert ein Experiment von Keltner, Ellsworth und Edwards (1993, Experimente 1 und 3). Die Vpn erhalten die Anweisung „*Stellen Sie sich vor, Sie werden in einem Verkehrsunfall verletzt!*“. Untersucht wurden zwei Gruppen von Vpn, bei Gruppe A wurde vorher eine traurige Stimmung induziert, bei Gruppe B eine wütende/ ärgerliche. Als danach die Vpn befragt wurden, welche Ursache sie für wahrscheinlicher halten, aufgrund derer der Unfall sich ereignet habe (unter zwei Optionen: 1. vereiste Straße, 2. betrunkenen Fahrer), korreliert die Auswahl in der Zuordnung A-1 und B-2: die Vpn mit trauriger Grundstimmung gaben an, dass sie einen Unfall aufgrund einer vereisten Straße für wahrscheinlicher hielten, die wütenden Vpn nannten als Ursache einen betrunkenen Fahrer. Wir sehen an dieser Stelle ab von den offensichtlichen Punkten, in denen sich das Experiment kritischen Nachfragen stellen muss. Ergibt sich evtl. ein „*bias*“ auf der Grundlage der stimuli, die zur Induktion der jeweiligen Stimmung herangezogen wurden? Wenn man weiterhin mit Birnbacher (1998) annimmt, dass sich Emotionen bezüglich ihrer Transitivität in 3 Klassen einteilen lassen (transitiv, intransitiv und gemischt), dann kann sich eine statistische Korrelation von B-2 allein aus dieser begriffsanalytischen Erkenntnis erklären lassen, weil Trauer eine gemischte Transitivität hat, Wut jedoch streng transitiv ist (sich also immer auf ein Objekt, spezieller: eine Person, bezieht). In diesem Sinne haben auch Christiansen & Loftus (1991; vgl. Schmidt-Atzert 1996, 180f.) empirisch gezeigt, dass Emotionen die Aufmerksamkeit auf ein emotional salientes Element fokussieren. Den Vpn wurde ein Videofilm eines Verkehrsunfalls dargeboten. Eine signifikant höhere Zahl der Vpn konnten auf spätere Befragung hin zwar die Farbe des Mantels des Unfallopfers angeben, nicht jedoch die Farbe des Unfallwagens.

motivation: it must not be confused with the supposition that we might be without wants. For I am still speaking of agents. I see two forms for this speculation to take. A being without emotion would either be some kind of Kantian monster with a computer brain and a pure rational Will, or else a Cartesian animal-machine, an ant, perhaps, in which every 'want' is preprogrammed and every 'belief' simply a releasing cue for a specific response. My hunch here is that it is because we are neither one nor the other, 'neither beast nor angel', that we have emotions as well as beliefs and desires." (De Sousa 1979, 135)

Schauen wir uns dieses Gedankenexperiment an mit aller Vorsicht, die geboten ist bei derlei Unternehmungen. Ich denke, dass sich in diesem Fall besonders deutlich zeigt, wie wenig nützlich eine solche Überlegung ist, aus mehreren Gründen. De Sousa will eine Person imaginieren, die ohne Emotionen, jedoch nicht ohne „wants“ ist. Interessant ist seine Wortwahl, da wir „wants“, als „Bedürfnisse“, im Sinne von atomaren, evtl. biologisch determinierten Antrieben interpretieren müssten, im Gegensatz zum weiteren Begriff des „desire“, der auch in De Sousas diskursivem Kontext vorkommt, und dessen er sich ebenfalls hätte bedienen können. Die These, dass wir ohne Emotionen auch Bedürfnisse hätten, die sich z.B. im Hunger oder Sexualtrieb ausdrücken, kann ohne weiteres akzeptiert werden. Spannender wäre die Frage, ob wir ohne Emotionen auch Wünsche im Sinne von „desires“ hätten: nicht nur „Hunger“ *tout court* eben, sondern Appetit auf ein gegrilltes Steak bspw. Bleibt nicht der Sexualtrieb überhaupt als ein körperliches Bedürfnis unterdeterminiert ohne die Hinzuziehung von sexuellem Affekt (Lust, Begehren etc.)? Auch Wünsche nach beruflichem Erfolg bspw. oder die aktuell empfundene Lust aufs Fußballspielen lassen sich wohl nicht als „desires“ verstehen, ohne ihren Bezug auf die sie unterliegende Emotionalität. De Sousa wehrt jedoch vorgreifend ab, indem er hinzufügt: „For I am still speaking of agents...“. Doch wären wir tatsächlich noch „Akteure“, wenn wir zwar noch über körperliche Bedürfnisse und – wie auch immer zu verstehende – „desires“ verfügen würden, jedoch bar jeder Emotionen wären? Ich denke, dass genau diese Frage den Kern der Debatte um die motivationale Potenz von Emotionen bildet. Sie wird uns im folgenden Kapitel beschäftigen.

2.11. „Emotion“. Ein Thesenkatalog

Bevor wir uns den Ausführungen von Kapitel 3 zuwenden, sei im Folgenden in aller Kürze ein Resümee unserer bisherigen Ergebnisse dieses Kapitels in Form eines Thesenkatalogs zum Phänomen „Emotion“ angestellt. Die einzelnen Thesen liefern nun keine neue Argumentation mehr, sondern sollen auf prägnante Weise einige Punkte zusammenfassen, die sich als besonders wichtig und relevant für unseren weiteren Argumentationsgang darstellen.

T1: Die in der philosophischen Forschung weit verbreitete Rede von Emotionen als „Zuständen“, „Phänomenen“, „Syndromen“ oder „mentalen Qualia“ ist ungenau. Emotionen sind Prozesse. Aus dieser methodischen Ungenauigkeit speist sich die Überzeugungs- oder Widerlegungskraft vieler kognitivistischer oder anti-kognitivistischer Argumente, deren genaue Einordnung im Rahmen eines statischen Emotionsmodells nicht möglich ist. Zumindest müsste in Anlehnung an die psychologische Forschung zwischen zwei Phasen von Emotion unterschieden werden, nämlich (1.) der Emotionselizitation und (2.) dem Coping-Prozess (nach Lazarus 1982; 1984 u.a. sowie Lazarus et al. 1980). Das Argument des *pragmatischen Paradoxons* (etwa: „Ich bin wütend auf meine Schwester, glaube aber nicht, dass sie sich irgendetwas hat zu Schulden kommen lassen“, vgl. Calhoun & Solomon 1984) bspw., das oft für einen Kognitivismus beliebiger Form bemüht wird, gilt jedoch nur für den Coping-Prozess, da es sich auf Beeinflussung und Steuerung (evtl. Unterdrückung oder Dissipation) von bestehenden emotionalen Zuständen bezieht. Es bezieht sich hingegen *nicht* primär auf den Prozess der Emotionselizitation selbst. Erinnern wir uns an das Beispiel vom Bären im Käfig, das wir im Rahmen der James-Lange-Theorie diskutiert haben. Die Tatsache, dass wir eine emotionale Erregung selbst in der besagten Situation erleben (also konfrontiert mit einem Bären, der keine „Gefahr“ darstellt), wurde als Standard-Argument gegen James und Lange eingesetzt. Die Unterscheidung von mindestens zwei Phasen im Prozess der „Emotion“ ermöglicht hingegen eine präzisere Analyse: wir könnten bspw. einräumen, dass Emotionen – oder auch eine bestimmte Klasse von Emotionen (die Primäremotionen) – spontane, unbewusste Reaktionen außerhalb unserer willentlichen Kontrolle darstellen. Selbst wenn wir also einen Kognitivismus auf der Ebene der Emotionselizitation ausschließen würden (bzw. eingeschränkte Geltung einräumen würden, nämlich lediglich für Sekundär- und Als-ob-Emotionen), so ist damit noch Nichts über das Phänomen „Emotion“ im Ganzen gesagt: das Beispiel mit dem Bären im Käfig zeigt, dass einer kognitivistischen Interpretation der Coping-Phase große Plausibilität eignet. Das Ziel der hier vorgeschlagenen Herangehensweise soll einerseits eine adäquatere Analyse des Phänomens „Emotion“ sein, andererseits ein reflektierter Gebrauch der Qualifikation „kognitivistisch“. Viele vermeintliche Problemfälle für den Kognitivismus (aber auch für diverse anti-kognitivistische Modelle) stellen sich im Lichte einer detaillierteren Analyse nicht als Probleme für einen reflektierten Kognitivismus dar.

T2: Emotionen entstehen in der sozialen Interaktion oder in Derivatformen davon.⁶²

Der Begriff der sozialen Interaktion an sich bedarf m.E. keiner weiteren Erläuterung. Emotionen entstehen in den Situationen der sozialen Welt, in denen wir in Interaktion mit anderen Akteuren Ziele verfolgen, Gespräche oder Verhandlungen führen, Aufgaben erfüllen, Konfrontationen austragen, Befehle geben, Bitten und Fragen stellen, Spielen, Lieben, Hassen etc. Unter Derivatformen der sozialen Interaktion verstehe ich z.B. *imaginierte* oder *antizipierte* Formen der Interaktion (Angst vor einer Konfrontation), oder aber soziale Interaktion auf einer *Schwundstufe* (Leiden, im Sinne von Trauer oder Depression, unter Einsamkeit). Dadurch wird zumindest eine offene Frage des strengen Kognitivismus, nämlich nach welchem Kriterium ein Werturteil (beliebiger Intensität) zu einer Emotion führt, von der Ebene des logisch-linguistischen Kalküls auf die pragmatische Ebene verlegt, die m.E. implizit stets mitenthalten ist. Nur wenn es in einer (realen oder imaginierten) Situation um *mich als Akteur unter anderen Akteuren (realen, imaginierten oder ermangelten)* in meinem sozialen, körperlichen Dasein geht, können wir relativ verlässlich die Entstehungsbedingungen von Emotionen bestimmen.⁶³ Mit anderen Worten: die Frage, ob

⁶² Lazarus benutzt das Modell der „Transaktion“ (im Sinne einer reziproken Interaktion von Person und Umwelt) als Paradigma der Emotionselizitation (z.B. in Lazarus et al. 1980, 200). Mit Ausnahme von vernachlässigbaren begrifflichen Nuancen entspricht dies meinem Begriff der „Interaktion“. Gestützt auf empirische Ergebnisse geht auch Schmidt-Atzert (1996, 32) davon aus, dass die Interaktion mit anderen Menschen die „wichtigste Quelle des emotionalen Erlebens“ ist.

⁶³ Als Konsequenz dieser These könnte auch eine vorsichtige In-Frage-Stellung des Kantischen Modells von der „Achtung vor dem Sittengesetz“ postuliert werden. Kants Gefühl der Achtung ist nicht auf Personen bezogen, sondern auf das moralische Gesetz. Emotionen haben ihren Platz jedoch zunächst in der Person-Umwelt-Beziehung. Gäbe es also ein moralisch relevantes, selbstgewirktes Gefühl der Achtung vor dem Sittengesetz, dann würde die Fähigkeit des Subjekts, ein solches Gefühl in sich autonom hervorzubringen, eine Art Als-ob-Emotion darstellen (also einen abgeleiteten Fall der Emotionsgenese), da sie sich auf ein abstraktes Objekt bezieht, nicht auf die reale Umwelt oder andere Personen. Es stellt sich zunächst die Frage, ob ein solcher Prozess vorstellbar und möglich ist. Aus einer anderen Perspektive (dem von Scheler konstatierten Dualismus zwischen dem „Primat des Affekts“ und dem „Primat des Kognitiven“ in der Geschichte der Emotionsphilosophie folgend; in diesem Sinne ist auch seine Klassifikation der Gefühle von den „sinnlichen“ bis zu den „geistigen“ zu verstehen (Scheler 1954, 341ff.)) kommt auch Coriando zu dem Befund, dass das Kantische Gefühl der Achtung „Apriorität sowie – damit einhergehend – Vernunftbezogenheit aufzuweisen vermag, ja ein durch die Vernunft ‘selbstgewirktes’ Gefühl ist –also letztlich *kein* Gefühl“ (Coriando 2002, 22). Was jenseits sozialer Interaktion übrig bleibt, ist der privateste Raum des solitären Rasonierens, Denkens, Imaginierens und Schlussfolgerns. Und hier reagieren wir eben nur dann emotional, wenn der Inhalt dieser Aktivitäten im Sinne Lazarus’ eine Bewertungsdimension aufweist, die sich auf uns als Teil der sozialen Welt bezieht, auf die Realisierung unserer Ziele, auf unser Leben in der Gemeinschaft und unsere von der Gesellschaft

eine gegebene Kognition ([Wert-]Überzeugung, Gedanke, Urteil, Maxime etc.) wie bspw. „Frau Müller hintergeht ihren Ehemann“ emotionsrelevant ist oder nicht, hängt nicht (oder zumin. nicht nur) von der Frage ab, wie diese Proposition evtl. zu klassifizieren sei, sondern vielmehr davon: ob das aussagende (oder mental repräsentierende) Subjekt selbst Herr Müller ist. Und wenn ja, ob Herr Müller viel an seiner Ehefrau liegt oder nicht etc. Und wenn nein, ob das Subjekt der Aussage evtl. mit Herrn Müller befreundet ist etc. Logisch-konzeptuelle Emotionsanalysen müssen stets um eine pragmatische Dimension, m.E. nach auch um eine narrativ-hermeneutische Dimension (Wollheim 1999; Morton 2002) ergänzt werden.

T3: Gehirn, periphere Muskel- und Organsysteme sowie ein zentrales Nervensystem (mit autonomen Anteilen) sind ontologische Voraussetzungen für Emotionalität.⁶⁴ Diese These schließt zwei weitere Implikationen mit ein. (1.) benennt sie ein anthropologisches „puzzle“, nämlich die spekulative Beobachtung, dass offenbar zur Ausführung streng kognitiver, d.h. logisch-propositionaler Operationen wie logisches Urteilen, deduktives Schließen oder mathematische Operationen allein das Gehirn mit dem Apparat von Sinnesorganen die ontologische Grundlage des Prozesses bildet. Anders gesagt: wir können uns relativ problemlos ein körperloses Gehirn mit Sinnesapparat vorstellen, das in der Lage wäre, solche kognitiven Operationen auszuführen. Das gilt jedoch nicht für Emotionen. Nach unseren bisherigen Untersuchungen ist der Körper das „Theater“ unserer Emotionen, die aus ontogenetischer Perspektive nur durch eine gewisse Interaktion zwischen neuronalen und physiologischen Prozessen entstehen können. LeDoux formuliert diese Beobachtung – in der Sprache der Kognitionswissenschaften – wie folgt:

„Wenn die biologische Maschine der Emotion – nicht aber der Kognition – notwendig den Körper einschließt, dann muss die Maschine, auf der die Emotion laufen kann, eine andere sein als jene, auf der die Kognition laufen kann. Selbst wenn man das funktionalistische Argument (dass die Art der Hardware gleichgültig sei) für den kognitiven Aspekt des Geistes gelten lassen könnte (und es ist nicht klar, dass man das kann), so scheint es für die emotionalen Aspekte des Geistes jedenfalls nicht zu gelten

zu bestätigenden Persönlichkeitsmerkmale. In bestimmten anderen Fällen kann dieser private Bereich überhaupt nicht Ort des Emotionalen sein, etwa bei der Lösung einer Rechenaufgabe (es sei denn, dies wäre Teil meines Projekts, der nächste Gewinner der Fields-Medaille zu werden).

⁶⁴ Diese These ist – wie der Leser leicht mitvollziehen wird – auf meinem auf Damasio aufbauenden Theorierahmen gegründet. Ein vergleichbares Ergebnis lässt sich interessanterweise auch in einem aristotelischen Theorierahmen erreichen, besonders unter Berücksichtigung von *De Anima*. So schlussfolgert Adamos (2001, 228): „It is plain [...] that emotions, being affections (*pathê*) of the soul, can never exist without the body, or be only in the soul. For if they existed either only in the soul or only in the body, it would mean that the form or the matter of a living thing could exist by itself, separately“.

(da es bei Emotionen sehr wohl auf die Art der Hardware ankommt).“ (LeDoux 1996, 46)

(2.) erlaubt diese These eine spekulative Überlegung zu der Frage, welche Lebensformen über Emotionen verfügen würden, und welche nicht. Dem gemäß würden folgende empirisch subsistierenden oder spekulativen Lebensformen *nicht* über Emotionalität verfügen: Pflanzen (kein Element erfüllt), körperlose Wesen (kein Element erfüllt), Einzeller (kein Gehirn oder höheres Nervensystem), höhere Maschinen (kein Körper mit peripherem Nervensystem). Emotionalität in welcher Form auch immer wäre hingegen wohl allen höheren Tieren zuzusprechen, zweifelhaft sind evtl. Gattungen wie Insekten, Spinnen etc. Emotionales Erleben, d.h. *Gefühle*, können wir jedoch u.U. nur Menschen zusprechen, da diese das Vermögen der Bewusstmachung von emotionalen Reaktionen voraussetzen. Die emotionalen Reaktionen, die wir an Tieren beobachten, sind lediglich instinktive oder gelernt-konditionierte Verhaltensweise, die durch das Aktivierungsprogramm von Emotionen ausgelöst werden. Dieser Aspekt der These soll nicht verfolgt werden, da er für die Ziele der vorliegenden Arbeit irrelevant ist.

T4: Die Definition bestimmter geschlossener Mengen von (Basis-)Emotionen für bestimmte Gattungen (z.B. Menschen), zusammen mit der inhaltlich erschöpfenden Definition der enthaltenen Emotionstypen (z.B. bei Darwin 1872 und Izard 1971,⁶⁵ Tomkins 1984, 167f.⁶⁶, Plutchik 1984, 200f.⁶⁷ etc., rezenter auch bei Gessner 2004), kann als aussichtsloses Projekt angesehen werden. Dies wird vor allem durch Analogie

⁶⁵ Carrol Izard übernimmt die bei Darwin grundlegende Unterscheidung zwischen 10 Basis-Emotionen. Die 10 Grundemotionen beschreiben Variationen auf einer qualitativ distinkten Intensitätsskala, die durch zwei Endpunkte begrifflich fixiert wird: 1. Interesse – Erregung, 2. Vergnügen – Freude, 3. Überraschung – Schreck, 4. Kummer – Schmerz, 5. Ärger – Wut, 6. Ekel – Abscheu, 7. Geringschätzung – Verachtung, 8. Furcht – Entsetzen, 9. Scham/Schüchternheit – Erniedrigung, 10. Schuldgefühl – Reue. Hier wird also dezidiert nicht mit uneindeutigen Emotionsbegriffen operiert, sondern es werden Emotionstypen postuliert, die sich begrifflich – je nach unterschiedlicher Intensität – unterschiedlich fassen lassen.

⁶⁶ Tomkins unterscheidet neun angeborene Emotionen, drei positive und sechs negative: 1. Interesse oder Erregung (interest or excitement), 2. Vergnügen oder Freude (enjoyment or joy), 3. Überraschung oder Schreck (surprise or startle), 4. Kummer oder Qual (distress or anguish), 5. Furcht oder Entsetzen (fear or terror), 6. Scham oder Demütigung (shame or humiliation), 7. Verachtung (contempt), 8. Ekel (disgust), 9. Ärger oder Wut (anger or rage).

⁶⁷ Plutchik geht von 8 Primäremotionen aus, die sich – wie Farben – mischen lassen, u.z. zu 28 dyadischen und 56 triadischen Kombinationen, womit sich 84 Emotionen auf einer bestimmten Intensitätsstufe ergeben. Unterscheidet man mit Plutchik weiter nach vier Intensitätsstufen, kommt man auf 336 verschiedene Emotionen.

augenfällig. Der Versuch, den Raum der möglichen menschlichen Emotionen definatorisch zu erfassen, gleicht dem Versuch, einen Katalog aller jemals möglichen Urteile, Wünsche und Gedanken zu erstellen. Die Vielfalt von Emotionsbegriffen, die teilweise sogar interkulturell unterschiedlich bzw. inkongruent sind, erklärt sich aus der Tatsache, dass viele sog. „sekundäre“ Emotionen ein komplexes Phänomen aus affektiven (primär-emotionalen) und kognitiven Anteilen darstellt. Auch wenn das Argument noch nicht gegen die Typisierung der Emotionen spricht, so lässt es sich in einem weiteren Schritt darauf ausweiten: es ist schwer vorhersagbar, welche Typen von Kognitionen (bestimmte Repräsentationen von „Selbst vs. Umwelt“, bestimmte Urteile und Gedanken) eine emotionale Relevanz erlangen, sodass sie z.B. über assoziatives Lernen mit bestimmten Gefühlen verbunden werden und einen neuen Emotionsbegriff begründen. Auf einer sehr frühen phylogenetischen Entwicklungsstufe können Scham und Schuld als Beispiele dienen für Emotionsbegriffe, die sich sicher nicht durch eigene emotionstypische (d.h. affektive) Qualitäten ausdifferenziert haben, sondern durch die Verbindung von geteilten kognitiven Kategorien mit affektiven Erlebnisqualitäten entstanden sind und soziolinguistisch einheitlich bezeichnet wurden. In jüngerer Zeit könnte man an das Gefühl des „ennui“ denken, das sich als Ausdruck der als dekadent wahrgenommenen Lebensweise der zivilisierten Gesellschaften entwickelt hat. Auf dem Gebiet der interkulturellen Forschung sind in der Emotionspsychologie Beispiele bekannt wie der in Japan gängige Begriff „amae“. Dieser von Morsbach & Tyler (1986) aus der Sicht westlicher Emotionsforscher für die interkulturelle Forschung untersuchte Begriff bezeichnet das „angenehme Gefühl der Abhängigkeit“ von einer anderen Person (Schmidt-Atzert 1996, 88) oder auch die Praxis, „die Liebe eines anderen auszunutzen oder sich an der Freundlichkeit eines anderen gütlich zu tun“ (LeDoux 1996, 125). Ein solches Gefühl ist – zumindest mit eigenem Begriff – in der westlichen Welt nicht bekannt.

T5: In der Emotionstheorie können wir sinnvoll zwischen einer *genetischen* Beschreibung der *Emotionselizitation* und einer *hermeneutischen* Beschreibung der *Emotionsinterpretation*. Damit verbunden ist die These, dass die Kategorie des „formalen Objekts“ (v.a. Kenny 1963) kein genetisches, sondern ein hermeneutisches Paradigma bildet. Kognitivistische bzw. Appraisal-Theorien der Emotionen müssen sich einer kritischen Frage stellen: Ist die für die Emotionsgenese reklamierte Bewertung (Evaluation, Werturteil, appraisal etc.) tatsächlich die Ursache eines emotionalen Zustands (oder zumindest ein Aspekt dieses Zustands), oder handelt es sich eher um eine Interpretation des emotionalen Zustands aus der Außenperspektive bzw. aus einer retrospektiven

Rationalisierung einer eigenen Emotion? D.h. löst mein Werturteil meinen emotionalen Zustand aus (bzw. gehört als Aspekt notwendig zu ihm)? Oder ist ein entsprechendes Werturteil eine Interpretation eines fremden emotionalen Zustands aus der 3.-Person-Perspektive, bzw. eines eigenen emotionalen Zustands in Retrospektive? Wir haben im Verlauf des vorliegenden Kapitels gesehen, dass das kognitivistische Paradigma im Rahmen einer differenzierten Relativierung zu sehen ist, die jedoch wiederum nicht den Kognitivismus per se in Frage stellt. Die Dependenz von emotionalen Zuständen auf Kognitionen als ihren Elizitationsbedingungen stellt sich bspw. anders dar bei primären als bei sekundären Emotionen und Als-ob-Emotionen. Auch die Unterscheidung zwischen aktuellen emotionalen Zuständen und emotionalen Dispositionen spielt hier eine Rolle. Auf dieser Grundlage erscheint es nicht als berechtigt, das „formale Objekt“ als ein Paradigma der „Auslösebedingungen“ von emotionalen Zuständen zu verstehen. Die Kategorie des formalen Objekts beschreibt hermeneutisch die standardisierte Klasse von materialen Objekten, auf die wir mit je bestimmten Emotionstypen reagieren. Sie kann jedoch keine reliablen Aussagen über den genetischen Faktor in der Emotionselizitation machen. Dies ist vielmehr wahrscheinlich nur auf der Grundlage einer ästhetischen, nicht einer ontologischen Klassifikation von Phänomenen möglich. Einfache Primäremotionen könnten als ein unbewusstes, spontanes Verarbeiten von bestimmten ästhetischen Merkmalen der Außenwelt bzw. von Einzelobjekten verstanden werden. Ein tentativer Katalog solcher Eigenschaft könnte u.a. umfassen: Schlangen und Insekten, plötzliche Geräusche, wütende Gesichter, plötzlicher Haltverlust, bedrohlich aufgehängte Objekte, völlige Dunkelheit (Prinz 2003), weiterhin Ausgesetzttheit (Raum, große Höhe, Menschenmenge), schnelle Bewegungen und die atmosphärische Nähe anderer Lebewesen. Emotionale Reaktionen korrelieren nicht mit best. konstanten Sachverhalten, sondern mit deren Eigenschaften, oder in einer Sprache der Ontologie ausgedrückt: sie korrelieren nicht mit der Substanz einer Entität, sondern mit deren Akzidenzien.

T6: Emotionen korrelieren mit Handlungstendenzen (Bechara et al. 2005; Adolphs et al. 1996; Busemeyer/ Townsend/ Stout 2002; Schwartz 2002; Lazarus 1991a und 1991b; Frijda 1986; Tomkins 1984), können sich in „desires“ manifestieren (Wollheim 2003 und 1999) bzw. enthalten eine motivational wirksame Komponente (Stocker 2002; Steinfath 2001; Goleman 1996; deSousa 1987; Greenspan 1988 und 1995; Wolf 1984). Die These, dass Emotionen motivationale Implikationen haben, lässt sich auf vielfältige Weise formulieren. Richard Wollheim hat argumentiert, dass Emotionen sich in dem mentalen

Zustand „desire“ manifestieren können. In eben diesem Sinne haben wir bereits in der Einleitung zu der vorliegenden Arbeit den Zusammenhang zwischen Emotionen und Wünschen angerissen. Dabei konnten wir als vorläufigen Konsens festhalten, dass (fast) alle Emotionen eine motivationale Tendenz implizieren, und sei es auch in einem *modus irrealis* (wenn wir bspw. in der Trauer um einen Verwandten „wünschen“, dass diese Person nicht tot sei). Dieselbe These – gefasst als Korrelation zwischen „Emotion“ und „Motivation“ – kann als *common sense* in der empirischen Psychologie und auch in den Neurowissenschaften gelten. Ein Beispiel für einen Nachweis dafür bilden Damasio's Iowa-Gambling Experimente (vgl. Bechara et al. 2005). In anderer Richtung argumentiert (wie wir v.a. noch in Kapitel 3 sehen werden) Silvan Tomkins, der „Emotion“ kraft der affektiven Qualität als das grundlegende motivationale System im Menschen ansieht, das amplifizierend und dämpfend auf die vermeintlich grundlegenden Systeme der Triebe und Körperbedürfnisse einwirkt. Im weiteren Verlauf der vorliegenden Arbeit soll näher untersucht werden, worin genau aus philosophischer und interdisziplinärer Sichtweise die motivationalen Implikationen von „Emotion“ bestehen könnten.

T7: Wir sprechen von „Emotionen“ sowohl in Bezug auf akut erlebte Zustände („occurrent emotional states“), als auch auf emotionale Dispositionen. Die entsprechende Typologie, die dieser These unterliegt, hatten wir schon im Rahmen der Einleitung diskutiert. Dabei war am Beispiel der Emotion „Trauer“ deutlich geworden, dass wir zwischen akuter, phänomenologisch als salient erlebter Trauer („Trauer₁“) im Sinne eines „occurrent emotional states“ einerseits und zwischen Traurigkeit als emotionaler Disposition („Trauer₂“), wofür zumeist das Adjektiv gebräuchlich ist, und die wir als Teil oder Merkmal des Charakters einer Person auffassen, unterscheiden. Zwischen den beiden Instanzen von „Trauer“ besteht eine Akt-Potenz-Beziehung. Wir können diese Beziehung also in gewisser Hinsicht als eine kausale Beziehung verstehen, nämlich insofern Trauer₁ als eine Instanzierung von Trauer₂ zu verstehen ist. Bestimmte Kognitionen, Schlüsselszenarien oder andere „kalibrierende“ Kontextbedingungen erklären die Auslösung einer Typ(1)-Reaktion auf der Grundlage einer Typ(2)-Disposition. Analog verhält es sich mit Charaktermerkmalen, spezieller mit habituellen Charaktermerkmalen, die wir unter den Begriff „Tugenden“ (im Sinne von gr. „aretai“) fassen: auch in diesem Falle können wir das Darbieten einer Typ(1)-Handlung, etwa das Eintreten für ein durch Gewalt bedrohtes, unschuldiges Opfer, als die Instanzierung eines Typ(2)-Habitus, also einer Tugend, nämlich in unserem Beispiel der „Tapferkeit“, verstehen. Auch diese Differenzierung erlaubt eine adäquatere Analyse von „Emotion“ im Hinblick auf

bestimmte Begrifflichkeiten. So wird v.a. die Frage nach der *Verantwortung* für Emotionen meist völlig analog zur Frage nach dem *kognitiven Gehalt* emotionaler Zustände beantwortet, wohingegen ich jedoch denke, dass beide Fragestellungen zwar miteinander in Verbindung stehen, in Wirklichkeit jedoch quer zueinander liegen. Die Frage nach der Verantwortung für Emotionen wird von zentraler Relevanz für unser Konzept der „emotionalen Einstellung“ sein. Zwar ist es sicher richtig, dass diese Frage verhältnismäßig leicht zu beantworten sein dürfte, wenn sich nachweisen ließe, dass alle Emotionen auf Urteilen als ihren Elizitationsbedingungen beruhen. Die Frage nach der Verantwortung für Emotionen wäre demnach gleichzusetzen mit der Frage nach der Verantwortung für von einem Akteur vertretene Urteile. Unsere Ausführungen im vorliegenden Kapitel haben es jedoch als zweifelhaft erscheinen lassen, dass tatsächlich *alle* Emotionen auf Urteilen als ihren Elizitationsbedingungen beruhen. Nicht nur kann also die Differenzierung von Aktualität und Disposition die adäquate Grundlage für eine Entscheidung der Verantwortlichkeits-Fragestellung liefern. Darüber hinaus kann sie uns für die Zwecke der vorliegenden Untersuchung auch in die Lage versetzen, eine a fortiori-Argumentation zu führen, indem wir nachweisen, dass sich Verantwortung für alle Emotionen (Primär- wie Sekundär-) in jedem Fall auf der dispositionalen Ebene (Typ(2)) nachweisen lässt. Primäre Emotionen entstehen zwar im Rahmen eines nicht-willkürlich, spontan ablaufenden Mechanismus, können jedoch zu einem gewissen Grad willkürlich kontrolliert oder unterdrückt werden. Sekundäre Emotionen entstehen im Gedankenprozess bzw. im Bewusstseinsstrom, sodass ihre Entstehungsbedingungen auch in einem kognitivistischen Vokabular relativ reliabel zu benennen sein werden. Emotionale Dispositionen wiederum unterliegen sogar einer willkürlichen Kontrolle, die sich jedoch nicht auf das Hervorrufen akuter Zustände, sondern auf die Beeinflussung der dispositionalen Grundlage bezieht.

3. Emotionen, Werte und „desires“

3.1. Einleitung

“[...] emotion seems to be a consequence of valuing certain things, rather than part of the process of coming to do so. The latter can be, and often is a dispassionate process.”
Wright (1988, 11)

Im vorliegenden Kapitel werden wir der Frage nachgehen, worin „Werte“ bestehen. Dabei soll einerseits unser Begriff des „Wertes“ analysiert und verschiedene Ansätze zu seiner Erklärung vorgestellt werden; andererseits wird es um das Phänomen des „Wertes“ gehen, genauer: um eine Genealogie des Wertens. Hierbei soll es weiterhin gezielt um den Bezug zu Emotionen gehen. Das oben zitierte Motto dieses Kapitels, in dem eine Position von Crispin Wright zum Ausdruck kommt, stellt schon zwei Alternativen vor, wie wir den Zusammenhang von Werten und Emotionen erklären können: nach Wright basieren Emotionen auf unseren Werten. Die Tatsache, dass wir in bestimmten Situationen auf bestimmte Objekte emotional reagieren (u.z. zumeist konstant), ließe sich demnach dadurch erklären, dass solche emotionalen Dispositionen in von uns unterhaltenen Wertüberzeugungen bestehen. Man könnte jedoch auch annehmen, dass Werte hingegen auf Emotionen beruhen, dass also Emotionen (in welcher Form auch immer) Wertüberzeugungen entstehen lassen, also konstitutiv für das Phänomen des „Wertes“ sind. Diese These soll im Folgenden vertreten werden, wobei jedoch auch die Einschränkungen und Bedingungen für das Zutreffen dieser These diskutiert werden sollen.

Fragen wir zunächst: „was sind Werte“?, oder anders formuliert: „was hat einen Wert?“, wobei ich die ontologischen Implikationen dieser Fragestellungen zunächst ausblenden möchte. Vollziehen wir also mögliche Antworten nach, die sich bei Kupperman (1999) und Steinfath (2002) finden lassen. Nach Kupperman bezieht sich der Begriff „Wert“ (bzw. „value“) auf: „what is worth having or being, taken purely for its own sake, or what is such that (taken by itself, apart from anything it causes) it is preferable that it exist rather than not exist“. Negative Werte wären damit: „what it is advisable, if one discounts all causal relations, not to have or be, or what is such that it is better that (taken by itself) it not exist“ (1999, 7f.). Wir finden hierin schon den Bezug von Werten auf „Präferenzen“ („it is preferable that it

exist“)⁶⁸, womit eine wichtige begriffliche Verbindung angesprochen ist, die später in unserer Argumentation wichtig werden wird. An dieser Stelle sei lediglich darauf aufmerksam gemacht.

Nach Steinfath (2002) sind Werte dasjenige, worüber wir mit Hilfe von Wertprädikaten sprechen. Es gibt zwei Arten von Wertprädikaten: (1.) allgemeine, wie bspw. „gut“ und „schlecht“, und (2.) spezifische, wie bspw. „langweilig“, „bewundernswert“, „fürchterlich“, etc. Interessant ist nun, dass – würden wir die Liste weiter fortsetzen – fast alle spezifischen Wertprädikate sich auf Emotionsbegriffe zurückbeziehen. Können wir also annehmen, dass wir unsere Wertprädikate der Existenz der jeweils entsprechenden Emotionsbegriffe verdanken? Und falls ja, wie würde sich eine solche Verbindung darstellen? Gemäß der von Steinfath vertretenen Position sind unsere Emotionen konstitutional für unsere Werte. Eine ähnliche Argumentationslinie soll auch hier vertreten werden. Dennoch wird weiter zu fragen sein: was ist gemeint mit „unseren Emotionen“? Aufbauend auf unserer Analyse in den vorangegangenen Kapiteln können wir schon vorab zumindest differenzierend fragen: sind es *aktuelle persönliche Emotionen*, die Werte konstituieren, oder *persönliche emotionale Dispositionen*, oder menschliche Emotionalität im Allgemeinen, womit nichts anderes als *Affektivität* gemeint sein kann. Und weiterhin: bedeutet diese konstitutionale Beziehung, dass alle Werte daher ausschließlich phänomenal zu verstehen sind und in einem Erlebnis bestehen? Oder können wir annehmen, dass nicht nur ein Erlebnis von Wert sein kann, sondern dass auch das involvierte Objekt (bzw. die Objekte) durch ein primordiales Erlebnis axiologisch markiert werden?

In rezenten Studien zur Emotionstheorie ist der Bezug zur Axiologie oft implizit, nämlich bei den sogenannten „streng kognitivistischen“ oder „judgmentalistischen“ Emotionstheorien. Schon seit der Stoa wird schließlich die Position vertreten, Emotionen beruhen auf unterhaltenen Wertüberzeugungen (worin wir die oben zitierte Position von Wright wiederfinden). Andererseits wird im Rahmen solcher Theorien oft behauptet, dass eine Emotion zu erleben bedeute, ein implizites Werturteil abzugeben. Dies kann entweder in der Weise verstanden werden, dass eine Emotion ein unterhaltenes Werturteil explizit macht. Wir könnten diese Feststellung jedoch auch anders lesen und annehmen, dass eine Emotion zu erleben bedeutet: sich zu sich selbst und/ oder zu anderen in einer Weise zu verhalten, die der

⁶⁸ Natürlich lässt sich Kuppermans Formulierung für negative Werte mühelos umformen von „what it is advisable“ in „what it is preferable“. Die Variation lässt sich wohl damit erklären, dass Kupperman in seiner Studie nicht auf die Verbindung von Werten und Präferenzen (d.h. „desires“) abhebt und daher problemlos sprachlich variieren kann.

Äußerung eines Werturteils entspricht. Diese Werturteil wäre insofern implizit, als es nicht propositional geäußert und daher im Subjekt vermutlich auch nicht als sprachliches Wissen abgelegt ist. Insofern könnte eine Emotion also uns selbst oder anderen gegenüber ein innovatives Werturteil bzw. eine wertende Perspektive auf einen Teil der Welt begründen. Für diese Lesart hat es in jüngster Zeit reiche Belege seitens der neurowissenschaftlichen Forschung gegeben. Bevor wir jedoch in die Diskussion des Zusammenhangs zwischen Werten und Emotionen eintreten, wollen wir uns zunächst kursorisch einiger bekannter Positionen in der Bestimmung des Begriffs „Wert“ versichern.

Ich möchte im Folgenden einer sinnvollen terminologischen Unterscheidung von Gaus folgen:

“Value language is grammatically complex, having a verb form (where someone *values* something), an adjectival form (where something is said to be *valuable*), and an abstract noun form (where something is said to be *a value*). Let us call these, respectively, *valuing*, *valuableness*, and *a value* (or *values*).” (Gaus 1990, 2)

Wir können also nach Gaus unterscheiden zwischen „werten“ als Adjektivform, das gewissermaßen eine subjektive Seite des Wertbegriffs darstellt. Hier kann die Frage auftreten, welche psychologischen Prozesse im Werten involviert sind, was der mentale Zustand des Wertens sei, oder allgemeiner gesagt: wie Werte überhaupt entstehen.⁶⁹ Dieser Aspekt wirft Fragen der Wertepistemologie und –psychologie auf und kann zu einer bestimmten wertontologischen Position führen, wenn wir annehmen, dass gerade dieser Prozess des „Wertens“ für die Entstehung von „Wert“ (in der Nominalform) verantwortlich ist. Die eher objektive Seite des Wertbegriffs, die Konzepte von „Wert“ und „Werthaftigkeit“, können im Rahmen einer wertrealistischen Position die privilegierte Position einer wertontologischen Fragestellung einnehmen. Unabhängig vom Wertrealismus jedoch verbinden wir mit diesem Aspekt den *normativen* Aspekt von *geteilten* Werten, d.h. die Frage nach Wertesystemen und gesellschaftlichen Werten bspw.

Es wird deutlich geworden sein, dass meine Fragestellung besonders auf den Aspekt des „Wertens“ im Sinne der Adjektivform abhebt. Denn in der Tat mag der Begriff der Emotion nicht das geeignete Instrument zur Analyse des objektiven Aspekts des Wertbegriffs sein:

⁶⁹ Die Frage nach der Entstehung von Werten via dieses „subjektiven“ Aspekts des Wertbegriffs stellt sich freilich nur, wenn wir „Wert“ auf das Leben, Denken und Tun von wertenden Lebewesen zurückbeziehen, und wäre daher hinfällig, wenn wir einen strikten Realismus vertreten, dem gemäß Werte einfach Entitäten in der Welt darstellen, die wir – durch dialektische Prozesse, Überlegen oder ein spezielles Sensorium – einfach „entdecken“ müssten.

gesellschaftliche Debatten oder Vernunftkritik sind evtl. besser geeignet, um Fragen von gesellschaftlichen oder gemeinschaftlichen Werten und deren Normativität zu klären.

Dennoch werde ich zu argumentieren versuchen, dass „Wert“ in die Welt kommt vermittelt von Lebewesen, die *zu werten fähig* sind. „Werten“ im Sinne der Adjektivform halte ich also für ein Tun oder eine Praxis von Lebewesen, die in der Fähigkeit besteht, emotionale Zustände zu erleben. Damit soll jedoch nicht zugleich behauptet werden, dass Werte jeweils aktual nur dem Erleben selbst zukommen. Die Formulierung soll gerade verdeutlichen, dass schon die Fähigkeit zur Emotionalität konstitutiv für die Genealogie von Werten ist. Weiterhin können wir sagen, dass damit (in Bezug auf frühere Ausführungen zur Aktualgenese der Emotionen) ein Rückbezug zu bestimmten Tatsachen in der Welt angestellt wird: diese Fähigkeit, emotionale Zustände zu erleben, besteht in einer konstanten Disposition gegenüber bestimmten Zuständen in der Welt. Wir haben gesehen, dass Primäremotionen auf reaktiven Dispositionen beruhen. Und wie wir ebenfalls gesehen haben, ist dieser primäre Mechanismus der Emotionsgenese primär im ontogenetischen Sinne und wird daher im Laufe einer normalen Entwicklung verfeinert und ausgebaut.

Dieser Entwicklungsstufe, die sich durch die Vorrangigkeit von Primäremotionen auszeichnet, wird jedoch auch ein Primat darin zukommen, wie wir den Begriff des Wertes erlernen. Mit den Worten von Helm (2001) können wir sagen: wir erlernen, was es bedeutet, dass bestimmte Dinge oder Angelegenheiten „Bedeutung“ („import“) für uns haben. Nach DeSousa (1979) stellen solche Erlebnisse „Salienzmuster“ („patterns of salience“) für unser Denken, Urteilen, Überlegen und Handeln zur Verfügung. Darüber hinaus überträgt dieses Entwicklungsstadium durch solch axiologisch relevante emotionale Erlebnisse die spezifische Qualität der *Affektivität* auf alle späteren Instanzen des Wertens. Wir könnten also – an dieser Stelle noch eher plakativ – gegen Wright einwenden, dass (im Erwachsenenalter) Emotionen nur deswegen eine Konsequenz von Werturteilen sein können, weil wir (in einem frühen Lebensalter) Werte überhaupt erst durch emotionale Reaktionen erfahren haben, und sich die für Emotionen spezifische Qualität der Affektivität mit der Praxis des Wertens assoziiert hat.

3.2. Der Begriff des „Wertes“

Wir haben weiter oben Kuppermans Ausgangsdefinition des Wertbegriffs gegeben. Dabei ist augenfällig, dass der Begriff des „Wertes“ mit dem Begriff des „Guten“ in enger Verbindung steht, ja dass wir beide als Synonyme verwenden können. Schließlich lautet die von Kupperman benutzte Variation für negative Werte: „what is such that it is *better* that

(taken by itself) it not exist“ (Hervorh. v. mir, S.H.). Beide Begriffe erlauben eine weitere Unterscheidung: nämlich im Falle von „Gut“ diejenige zwischen einem Gut für sich und einem Gut als Mittel, im Falle von „Wert“ die analoge Unterscheidung zwischen intrinsischem und extrinsischem Wert. Letzteres (also extrinsischer Wert oder zweckrationale Güte) wird gelegentlich auch als technische Güte verstanden (worin auch einer der vielen Aspekte des antiken aretê-Begriffs besteht): ein „gutes“ Messer ist ein Messer, das seinen Zweck gut erfüllt, also Brot zu schneiden bspw. Solche technische Güte korrespondiert natürlich nicht direkt mit instrinsischer Güte, und hierin finden wir eine leichte Diskrepanz zwischen dem Begriff des „Guten“ und dem Begriff des „Wertes“ in unserer alltagssprachlichen Verwendung. Zur Illustration dieses Punktes möchte ich ein Beispiel von Becker (1973, 34ff.) anführen, das Gaus (1990, 120f.) ausführlich diskutiert. Wir könnten bspw. sagen dass eine Folterbank technische (also extrinsische) Güte haben kann, insofern sie auf eine Weise entworfen und gebaut ist, durch die sie einen Zweck (Menschen zu foltern) hervorragend zu erfüllen ermöglicht. In diesem Sinne verhält es sich wie mit dem Brotmesser, das eine Funktion hervorragend erfüllt. Natürlich kommt darum der Folterbank nicht zugleich intrinsische Güte zu, da wir die Praxis des Folterns nicht für moralisch erlaubt oder gar empfehlenswert halten. Wir können also eine Folterbank zwar „gut“ nennen, insofern sie einen bestimmten Zweck erfüllen mag, den wir jedoch nicht unbedingt als für sich gut ansehen mögen. Andererseits scheinen wir mit Blick auf solche extrinsische Güte nicht leicht von einer „wertvollen“ Folterbank zu sprechen. Mehr als noch bei den Begriffen „gut“ oder „schlecht“ scheint in der Werte-Begrifflichkeit stets auch eine Stellungnahme enthalten zu sein, die über eine rein zweckrationale Bewertung hinausgeht. Die Werte-Begrifflichkeit scheint also eher schon von selbst die stärkere Behauptung von intrinsischem Wert einzuschließen. Hierbei scheint es sich jedoch um Bedeutungsnuancen zu handeln. Zu erwähnen ist noch die von Kant prominent gemachte Unterscheidung zwischen einem intrinsischen Gut *in einem bestimmten Kontext*, und einem intrinsischen Gut *in Absehung vom Kontext*. Nach Kant stellt nur der „gute Wille“ ein solches intrinsisches Gut dar, das ohne weitere Qualifizierung stetig gut ist, in Absehung vom Kontext. Andere intrinsische Güter, sogenannte „Glücksgüter“ wie Wohlstand, Gesundheit oder körperliche Vorzüge, sind intrinsische Güter, jedoch nur im Rahmen eines bestimmten Kontextes, nämlich der Bedingung, dass sie mit dem guten Willen verbunden sind.

3.3. „Valuing: desiring or believing?“

Betrachten wir nun cursorisch einige bekannte Ansätze in der Werttheorie und stellen wir die obige, programmatische Frage, die Michael Smith (1992) in prominenter Weise als Titel eines seiner Artikel benutzt hat. Wir beziehen uns damit – und im Folgenden – auf den Begriff des „Wertens“ in seiner Adjektivform, wie ich ihn ausgehend von Gaus erläutert habe. Im Rahmen einer Kartographie des menschlichen Geistes, die aus der analytischen Philosophie stammt (vgl. Wollheim 2003), können wir nun fragen: wenn wir etwas werten bzw. wertschätzen⁷⁰, bedeutet dies, dass wir uns mit einem *Wunsch* („desire“) auf das entsprechende Objekt beziehen? Oder unterhalten wir vielmehr einen bestimmten *Glaubenszustand* hinsichtlich des entsprechenden Objekts?

3.3.1. Werte und „desires“

Dementsprechend wäre eine Position in der Bestimmung der Weise, wie Werte begründet werden, anzunehmen, dass dasjenige für eine Person (intrinsisch) gut ist, was diese Person wünscht (im Sinne von „what this person desires“). Dieser Ansatz, der prominent von J.S. Mill vertreten wurde (rezent auch Brandt 1996), erlaubt es uns, einen intuitiv im Wertbegriff angelegten Aspekt zu erfassen: Werte stehen in Beziehung zum Handeln bzw. zu Handlungstendenzen. Wenn Hans bekennt, dass er einen hohen Wert mit Erdbeereis verbindet, dann können wir annehmen, dass Hans – *ceteris paribus* – in geeigneten Situationen versuchen wird, Erdbeereis zu erlangen. Intuitiv scheint also einiges für den Rückbezug von Werten auf „desires“ zu sprechen. Dabei wird jedoch leicht übersehen, dass einige Schwierigkeit darin besteht, überhaupt den Begriff des „desire“ zu fassen. Was ist ein „desire“?

(1.) Ein „desire“ kann als ein körperliches Bedürfnis, ein „Schmachten“ („craving“) oder ein Trieb aufgefasst werden. Das Phänomen des „desire“ wird damit auf körperliche Zustände wie bspw. Durst, Hunger, Müdigkeit oder sexueller Lust bezogen. Es handelt sich also gewissermaßen um einen „reflektiven“ desire-Ansatz insofern „desire“ als das Erleben eines Bedürfniszustands verstanden wird. Diese Körperzustände, zusammen mit ihrer mentalen Repräsentation, machen das entsprechende Objekt wünschenswert („desirable“). Wenn wir

⁷⁰ Im Folgenden werde ich auch „wertschätzen“ im Sinne von „werten“ oder „bewerten“ verwenden, in Abweichung von der Alltagssprache auch in Bezug auf negative Werte. Ich denke, dass dadurch noch genauer der Bezug auf eine Tätigkeit des Subjekts (eben im Sinne einer Adjektivform) hergestellt werden kann.

also durstig sind, dann wünschen wir oder verlangen wir nach („we desire“) trinkbarer Flüssigkeit, um unseren Durst zu stillen, zumindest nachdem wir in einer sehr frühen Entwicklungsstufe die Verbindung zwischen dem Trinken und dem Stillen des Durstes gelernt haben. Interessant ist jedoch, dass bei allen möglichen Körperzuständen, die vermeintlich desires begründen, Mediatorfaktoren ins Spiel kommen. So ist die Menge an Flüssigkeit, die wir konsumieren, nicht nur durch unseren „körperlichen Durst“, d.h. durch den Grad unserer Dehydrierung, bestimmt, sondern auch durch die sensorischen Qualitäten der konsumierten Flüssigkeit: d.h. dadurch, ob das Getränk wohlschmeckend ist oder nicht (Pfaffman 1960). Ebenso verhält es sich mit Hunger: je hungrier wir sind, umso mehr steigt die Bereitschaft, Nahrung zu konsumieren, die wir für sensorisch und geschmacklich nicht angenehmen halten. Noch weiter führt die Tatsache, dass wir gelegentlich auch Appetit auf (ein „craving for“) Getränke und Nahrungsmittel haben, obwohl wir eben nicht hungrig oder durstig sind. Wir könnten bspw. Appetit auf einen Gin and Tonic haben, weil wir den Geschmack dieses Getränks wertzuschätzen gelernt haben, ebenso lässt sich der Appetit auf ein Erdbeereis erklären. Das Bedürfnis (im Sinne von „desire“) nach Essen und Trinken hat also sicherlich eine körperliche Basis, die jedoch nach Abzug kultureller Einflussnahme noch einmal einer Modellierung durch andere Mediatoren unterworfen ist. Auch im Rahmen eines instrumentalistischen desire-Ansatzes, dem gemäß wir Präferenzen verwirklichen zur Realisierung und Maximierung unseres persönlichen Glücks, ist diese Erklärung nicht erschöpfend: „satisfaction of *bodily* need-based desires alone falls short of maximizing happiness“ (Brandt 1996, 26). Wir können bspw. in Kauf nehmen, einem starken körperlichen Bedürfnis wie Schlaf zu widerstehen, um etwas zu erreichen, das für uns mit großem Wert verbunden ist, und aus diesem Grund nächtelang Wache im Krankenhaus halten, um die Geburt des eigenen Kindes miterleben zu können. Körperliche Bedürfnisse scheinen zwar eine gewisse Potenz aufzuweisen, können jedoch wiederum „ausgestochen“ werden durch Wünsche, die wir auf andere Weise erworben haben und mit einer gewissen Dringlichkeit vertreten. Auf diesen letzten Punkt geht natürlich die Tatsache zurück, dass wir vermutlich nicht im Krankenhaus einschlafen, wenn die Geburt des eigenen Kindes bevorsteht, jedoch sehr wohl Gefahr laufen einzuschlafen, wenn wir bspw. in der Rolle eines Nachtportiers in einem Lagerhaus tätig sind. An späterer Stelle werde ich zu zeigen versuchen, dass dieser Ansatz scheitert, weil er die Relevanz der vermeintlichen Mediatorfaktoren falsch einschätzt. Tatsächlich sind es nicht „desires“ im Sinne von körperlichen Bedürfnissen, die Werte konstituieren, sondern die damit verbundenen affektiven Qualitäten, die gemäß der hier vorgestellten Position lediglich die Rolle von Mediatorfaktoren spielen würden.

Wie verhält es sich nun mit dem Begriff des „Triebes“, der in der traditionellen psychologischen Forschung (besonders der Psychoanalyse) wohletabliert ist zur Erklärung menschlichen Verhaltens. Wir könnten bspw. „desire“ im Sinne eines Körperzustands als einen „Trieb“ auffassen, z.B. den Sexualtrieb. Triebe wurden lange Zeit für die primitivste und zugleich stärkste motivationale Kraft bei Mensch und Tier gehalten. Wir kennen zahlreiche anekdotische Beispiele, zu welchen Übeltaten sexuelle Lust führen kann, aus Kunst und Literatur ebenso wie aus der Realität. Und auch der Selbstmörder kann seinen Selbsterhaltungstrieb nicht außer Kraft setzen: wird er – vor Ausführung seines eigenen Suizids – unerwartet unter Wasser gedrückt, so wird auch er spontan mit aller Macht darum kämpfen, nicht ertrinken zu müssen. Wenn wir an die absurde Idee denken, durch Anhalten des Atems Selbstmord zu begehen, dann wird schnell deutlich, dass offenbar Triebe (Sexualtrieb oder Selbsterhaltungstrieb) eine sehr hohe Potenz aufweisen und vermeintlich primitiv (im Sinne von elementar) sind: sie lassen sich offenbar nicht auf andere motivationale Kräfte zurückführen und entziehen sich weitestgehend unserer Kontrolle. Die daraus generierten Wertigkeiten (der Attraktion und Repugnanz) sind also nicht nur sehr stark, sondern vermeintlich ontologisch eindeutig auf Trieben fundiert, und nicht auf ein weiter im Hintergrund stehendes Phänomen. Dieser traditionellen Sichtweise hat Tomkins mit seinen Studien zur Affektivität vehement widersprochen (vgl. v.a. Tomkins 1984).

Nach Tomkins bilden Emotionen das primäre Motivationssystem im Menschen, das potenter ist als Triebe. In der Tat benutzt das Triebsystem sogar die verstärkende Wirkung von Emotionen. Emotionen beeinflussen und regulieren Triebe. Triebe unterliegen einer festen raum-zeitlichen Struktur: Hunger bspw. tritt entsprechend einem festen Muster unseres physiologischen Zustands (Zuckergehalt im Blut etc.) und unseres Biorhythmus auf, gleiches gilt auch für andere Triebe wie bspw. Müdigkeit. Weiterhin stehen Triebe mit einem stetigen Objekt in Verbindung: Hunger mit Nahrung, Müdigkeit mit Schlaf, sexuelle Lust mit sexuellen Handlungen. Emotionen hingegen unterliegen weder der einen noch der anderen Beschränkung: sie treten auch jenseits von festen Zyklen auf, und sie lassen prinzipiell jede Art von Objekt zu.⁷¹ Die Zusammenführung der beiden Systeme Emotion und Trieb kann daher eine Komplexität generieren, die weit über derjenigen des Triebsystems alleine liegt. Demnach ist gemäß Tomkins die Freudsche Konzeption der Transformation und Sublimation

⁷¹ Im Rahmen einer kognitivistischen Emotionstheorie würde vermutlich argumentiert, dass auch Emotionen mit stetigen Objekten korrespondieren (mit ihren sog. formalen Objekten). In einem früheren Kapitel habe ich jedoch zu zeigen versucht, dass diese Sichtweise empirisch gesehen unplausibel ist: auch Gefahr, die uns nicht bedroht (z.B. vergangene Ereignisse), kann uns Angst machen.

von Trieben missverständlich: solche Operationen werden nämlich nur dadurch ermöglicht, dass das Triebssystem mit Emotionen verbunden ist. Sexuelle Lust, die auf natürlich Weise mit sexuellen Handlungen korrespondiert (und mit nichts anderem), und die anhält, bis sie befriedigt wird oder aus dem Bewusstsein tritt, könnte niemals auf nachvollziehbare Weise bspw. durch intellektuelle Tätigkeiten „sublimiert“ werden, wenn wir nur das Triebssystem betrachten. Weiterhin ist es offensichtlich, dass Emotionen Triebe verstärken, dämpfen oder sogar aufheben können. Depression oder Kummer kann bspw. unseren Sexualtrieb aufheben oder Müdigkeit unterdrücken. Müdigkeit verschwindet gelegentlich auch, wenn wir frisch verliebt oder sehr aufgeregt sind. Und Verzweiflung und extreme Erregung bzw. Anspannung kann uns fast gleichgültig gegenüber dem Schmerzvermeidungstrieb machen.

Diese theoretischen Überlegungen, die auf Tomkins zurückgehen, können weiterhin durch neuere Ergebnisse aus der experimentellen Hirnforschung unterfüttert werden. Betrachten wir die sogenannten Experimente zum „reward“ und zur „intracranialen Selbst-Stimulation“ (ICSS), die schon eine lange wissenschaftliche Tradition haben. Bei diesem Typ von Experiment werden Elektroden in das Rattenhirn eingepflanzt, u.z. hauptsächlich in monoaminergene Kerne und Faserbahnen. Traditionell wurde der Hypothalamus (als Teil des limbischen Systems) als Lokus der Elektrodenimplantation gewählt, in jüngster Zeit konnte jedoch nachgewiesen werden, dass sich die gleichen Resultate auch im medialen präfrontalen Kortex (mPFC) reproduzieren lassen (Tzschentke 2000).⁷² Die „Belohnung“ („reward“), die über diese implantierten Elektroden verabreicht werden soll, besteht in einer Lusterfahrung, und aus der Beteiligung von emotionsrelevanten Hirnarealen (v.a. des mPFC) können wir folgern, dass es sich um einen hedonisch positiv getönten Zustand *affektiver* Erfahrung handeln muss. In einem bestimmten Typ dieses Experiments werden die präparierten Ratten nun in eine Umgebung gesetzt, in der sie die Option haben, entweder durch Drücken eines Hebels Stromstöße auszulösen, die zu einer solchen hirnhysiologisch induzierten Erfahrung führen, oder aber angebotenes Futter und Wasser zu konsumieren. Ein signifikanter Anteil von Versuchsratten betätigt nun den Hebel bisweilen „bis zum Punkt der Erschöpfung“

⁷² Dies ist insofern wichtig, als wir in einem früheren Kapitel anhand der Studien von Damasio nachvollzogen haben, dass dem präfrontalen Kortex eine vorrangige Stellung als Lokus der Emotionalität zukommt, u.z. genauer gesagt als Lokus der Sekundäremotionen, die auf einer Verbindung mit Gedankenprozessen beruhen bzw. durch diese ausgelöst werden. Das limbische System, genauer die Amygdala, als Lokus der Primäremotionen, reagiert vornehmlich auf primitive und prädeterminierte Umweltreize.

(Jacques 1979), womit nach Christian Elger⁷³ gemeint ist, dass diese Ratten Nahrung und Wasser vollständig vernachlässigen und sich dem „lever-pressing“ so lange hingeben, bis sie schließlich verhungern. Hierin finden wir m.E. ein sehr eindrucksvolles Beispiel, wie emotionaler Affekt letztlich „desires“ im Sinne von Körperzuständen oder Trieben sogar völlig „ausstechen“ kann, u.z. bis zu einem Punkt, wo die vermeintlich homöostatische Funktion, die solche Triebe haben, vollständig unterlaufen wird.

(2.) Im Gegensatz zum vorgestellten „reflektiven“ desire-Ansatz kann entsprechend auch ein „prospektiver“ desire-Ansatz vertreten werden: anstatt das Phänomen „desire“ auf Bedürfnisse, Körperzustände, Triebe, Appetit bzw. „Schmachten“ („craving“) zurückzubeziehen, können wir „desire“ auch vorwärts projizieren auf unsere Lebens- und Handlungs-Ziele und -Zwecke. Innerhalb dieses Paradigmas wird „desire“ mit einer Vielzahl von Begriffen erklärt und bisweilen als „propensity to act“, „goal“, „end“, „preference“ oder „pro-attitude“ (Gaus 1990, 88ff.) verstanden. Der letztere Begriff, „pro-attitude“ oder pro-Einstellung, der auf Davidson (1960) zurückgeht, ist sicherlich der prominenteste und soll im Folgenden auch hauptsächlich verwendet werden. Das Standard-Argument gegen den prospektiven desire-Ansatz scheint mir auch das zugleich einfachste und stärkste Argument zu sein. Denn in der Tat führt der *prospektive* Ansatz zu einem explanatorischen *Regress*. Intuitiv versuchen wir, das Handeln von menschlichen Akteuren auf zu Grunde liegende Wertüberzeugungen zurückzuführen, d.h. auf dasjenige „Gut“, auf das eine Handlung abzielt bzw. das diese Handlung zu realisieren oder befördern geeignet ist. Wir handeln also in „Orientierung am Guten“ (Steinfath 2001), das „Gute“ wiederum, unsere Wertüberzeugungen also, lassen sich auf verschiedene Weise erklären. Wenn wir jedoch gemäß dem prospektiven desire-Ansatz unsere Handlungstendenzen (pro-Einstellungen) als „desires“ identifizieren, und diese wiederum die Menge der von einem Akteur unterhaltenen Werte darstellen sollen, dann stellt sich die Frage, wie diese Werte überhaupt in erster Linie gebildet wurden. Wir glauben schließlich, dass unsere Werte unser Handeln anleiten, und nicht etwa, dass unsere pro-Einstellungen als bereits vorliegende Größen unsere Menge von Wertüberzeugungen bestimmt. Hier handelt es sich m.E. nicht nur um ein Argument aus der intuitiven Kraft des alltäglich Sprachgebrauchs, sondern der prospektive desire-Ansatz scheint schlicht auf einer Fehlanalyse der Begriffe zu beruhen. Der Ansatz scheint zwei fundamental unterschiedliche

⁷³ Ich bedanke mich bei Prof. Dr. Christian E. Elger für diesen Hinweis während eines Vortrags auf der Konferenz „Dimensionen der Person: Genom und Gehirn“, im Wintersemester 2004/05 organisiert vom Institut für Wissenschaft und Ethik, Bonn.

Perspektiven zu vermischen: der prospektive desire-Ansatz ist sicherlich zutreffend aus einer *hermeneutischen* Perspektive. Denn in der Tat ist oft die einzige verlässliche Möglichkeit herauszufinden, welche Werte ein Akteur vertritt, zu beobachten, wie dieser handelt. Die manifesten Handlungstendenzen eines Akteurs informieren also darüber, welche Werte dieser vertritt. Ich denke jedoch, dass dieser Ansatz hingegen nicht die *genetische* Perspektive hinreichend erklären kann, im Rahmen derer zu klären wäre, wie wir bestimmte Wertüberzeugungen überhaupt erwerben. Und gerade diese Frage soll uns hier beschäftigen.

3.3.2. Werte und Glaubenszustände

Wir haben im vorangegangenen Abschnitt gesehen, dass ein Ansatz, der Werte auf desires zurückzuführen versucht, letztlich *prinzipiell* unterdeterminiert bleibt. Die Mediatorfaktoren, darunter v.a. Emotionen und Affektivität, erweisen sich als wesentlich wichtiger als die vermeintlich zu Grunde liegenden Phänomene wie Körperzustände oder Triebe. Die Gleichsetzung von desire mit Zielen und Zwecken kann schon aus begriffslogischer Sicht nicht überzeugen. Es gibt also nach Diskussion des desire-Ansatzes reiche Indizien dafür, dass Emotionen die Grundlage unserer Wünsche („desires“) und Triebe darstellen könnten, und damit qua desire-Ansatz auch die Grundlage unserer Werte.⁷⁴

Stellen wir jedoch erneut die programmatische Frage „valuing: desiring or believing?“, und wenden wir uns dem zweiten Teil dieser Dichotomie zu. Auch der Ansatz, Werte als Glaubenszustände zu verstehen, ist traditionell verbürgt. Hier soll jedoch zunächst darauf aufmerksam gemacht werden, dass mit dieser Sichtweise oft auch eine Position in der Frage um den Zusammenhang von Werten und Emotionen verbunden ist, die genau entgegengesetzt

⁷⁴ Natürlich kann dieses Verhältnis auch in die andere Richtung gelesen werden. So hat Wollheim (1999 und 2003) argumentiert, dass das Erlebnis der Erfüllung oder Frustration eines desires, mit dem entsprechenden Objekt assoziiert, zur Formierung einer mentalen Disposition führen kann, die sich wiederum in einem akuten emotionalen Zustand ausdrücken kann. Auch Lyons (1980, 58) spricht davon, dass emotionale Reaktionen durch „strong preferences“ ausgelöst werden, wobei er selbst diesen Begriff als Teil eines kognitivistischen Paradigmas vorstellt. Wie Marks (1982) jedoch richtig anmerkt, befindet sich Lyons kraft dieses Begriffs schon auf dem Terrain einer desire-Theorie der Emotionen. Nach Wollheim und Lyons würde also ein desire Grundlage und Auslöser für eine emotionale Reaktion liefern, während wir bisher argumentiert haben, dass Emotionen vielmehr eine zentrale Rolle in der Formierung von desires spielen. Wenn man jedoch an beide Ansätze – Wollheim und Lyons – die Frage richtet, wie denn die entsprechenden desire in erster Linie gebildet wurden, so wird die Antwort entweder uninformativ ausfallen (auf die Faktizität von desires abheben), oder aber sich dem oben vorgestellten Regress ausgeliefert sehen, dass auch desires sich nicht ohne Bezugnahme auf Werte – und damit, wie ich denke, auf Emotionen – erklären lassen.

zu meiner hier vertretenen Argumentationslinie verläuft. Nehmen wir an, Werte seien konstituiert durch Glaubenszustände in einem weiten Sinne, also durch Kognitionen. Diese Position verbindet sich nun wie von selbst mit dem emotionstheoretischen Kognitivismus, dem gemäß Emotionen durch Kognitionen ausgelöst werden, u.z. genauer: durch Werturteile und Wertüberzeugungen (vgl. Kap. 2.2.). Oft wird auch die Position vertreten, dass „Werte“ die formalen Objekte von Emotionen seien, denn mit „Angst“ beziehen wir uns z.B. emotional auf Objekte aus der Klasse „das Gefährliche“, womit zweifellos ein Wertprädikat herausgehoben ist. Emotionen würden in diesem Sinne also von Werten (als Glaubenszuständen) begründet. Emotionen wären ontologisch abhängig von Werten, insofern entweder der Begriff des „Wertes“ eine logische Vorbedingung für den Begriff der „Emotion“ darstellt, oder insofern ein unterhaltenes Werturteil sogar eine aktive Ursache in der Auslösung eines emotionalen Zustands wäre.

Im Folgenden werde ich also in zwei Richtungen argumentieren: einerseits denke ich, dass der Ansatz, desires als Glaubenszustände aufzufassen, an sich problematisch ist. Andererseits wird uns diese Kritik schon darauf hinführen, dass Emotionen eine wichtige Rolle in der Konstitution von Werten spielen. Und wenn dies überzeugend gezeigt werden kann, dann verliert die obige These, dass Emotionen durch Werte ausgelöst werden, ihre Plausibilität. In der Tat würde dann das Verhältnis genau in die andere Richtung laufen. Die folgenden Argumente setzen bewusst nicht am emotionstheoretischen Teil der Argumentation an. Natürlich könnten wir auch argumentieren, dass eine Emotionstheorie, die Werte als die formalen Objekte von Emotionen begreift, und die eine Emotionselizitation auf der Grundlage von Werturteilen postuliert, an sich nicht plausibel ist. Dies wurde jedoch schon in Kapitel 2 geleistet, in dem ich versucht habe, erfolgreich Einwände gegen ein streng kognitivistisches (oder „judgmentalistisches“) Paradigma in der Emotionstheorie vorzubringen. An dieser Stelle will ich mich daher bewusst auf den Teil der axiologischen Theorie beschränken.

Eine weitere Anmerkung erscheint mir wichtig: wenn innerhalb der traditionellen Dichotomie „valuing: desiring or believing“ argumentiert wird, dann werden diejenigen Argumente geprüft, die jeweils für oder gegen eine Erklärung von Werten über desires oder über Glaubenszustände sprechen. Beide Kandidaten, desires und Glaubenszustände, befinden sich metatheoretisch betrachtet auf derselben Stufe. Dies lässt sich leicht anhand der eingangs eingeführten Terminologie verdeutlichen. Betrachten wir den desire-Ansatz: die Interpretation von „Werten“ (Nomen) als „desires“ und von „werten“ (Adjektiv) als „desiring“ baut begrifflich offenbar auf denselben substanziellen Grundlagen auf; es wechselt jeweils lediglich der Modus Nomen- bzw. Adjektivform. Ebenso verhält es sich mit dem belie-

Ansatz: die Interpretation von „Werten“ (Nomen) als „Glaubenszuständen“ und von „werten“ (Adjektiv) als „glauben, dass“ baut substantiell auf denselben Voraussetzungen auf. Anders jedoch verhält es sich mit dem von mir vorgeschlagenen Ansatz, Emotionen als Konstituenten von Werten zu verstehen. Hier liegt ein fundamentaler Unterschied vor zwischen der Interpretation von „werten“ (Adjektiv) als „einen gerichteten emotionalen Zustand erleben“ und „Werten“ (Nomen) als „Emotionen“. Letzteres beschreibt ein hedonistisches oder zumindest reduktiv-phänomenologisches Paradigma, dem gemäß jeder Wert allein in einer phänomenalen Erfahrung bestehen würde. Ich denke, dass ein solcher Ansatz sehr plausibel Nozicks berühmten Argument der Lust-Maschine zum Opfer fällt.⁷⁵

In der vorliegenden Arbeit vertrete ich zwar die Position, dass „werten“ (Adjektiv) als „das Erleben eines gerichteten emotionalen Zustands“ zu interpretieren sei, nicht jedoch, dass „Werte“ (Nomen) in „Emotionen“ bestünden. Die drei Kandidaten, desire, Glaubenszustände und Emotionen, scheinen sich offenbar metatheoretisch auf verschiedenen Ebenen zu bewegen, wobei die ersten beiden auf derselben, der letztere auf einer davon verschiedenen Ebene zu verorten wäre. Erfreulicherweise wird damit nebenbei einer von mir gehegten Sorge Genüge getan, dass die Gleichsetzung eines mentalen Phänomens mit einem anderen zum Zwecke der gegen- oder einseitigen Erklärung der Phänomene einen viziösen Reduktionismus beinhaltet. Denn wenn wir tatsächlich über verschiedene Begriffe für bestimmte Phänomene (wie Wert, desire, Glaubenszustand) verfügen, und diese alltagssprachlich unabhängig voneinander benutzen, dann liegt darin m.E. ein gutes Indiz, dass es sich in der Tat um unterschiedliche Phänomene handelt, deren Kern sich gerade nicht durch gegenseitige Gleichsetzung erhellen ließe. Ein Emotions-Ansatz in der Werttheorie scheint diese viziöse Operation zu vermeiden, was sich ganz klar schon daran zeigt, dass eine Gleichsetzung in verschiedenen Modi (nominal und adjektivisch) äquivok ist.

Was lässt sich nun zu Gunsten des belief-Ansatzes vorbringen? Ich möchte in die Diskussion eintreten, indem ich zunächst ein Gedankenexperiment von Griffin (1998, 19ff.)

⁷⁵ In seinem berühmten Gedankenexperiment fordert Nozick (1974, 42ff.) seine Leser dazu auf anzunehmen, eine Lustmaschine könne bei ihrem Probanden einen subjektiven Glückszustand erzeugen, der bis zu dessen Lebensende nicht abnimmt und nicht langweilig wird, aber jegliche Form eines aktiven Lebens ausschließt. Der Proband wäre lediglich ein passives Subjekt dieses Glücksempfindens. Nozick vermutet nun, dass die große Mehrzahl der Menschen sich – vor die Wahl gestellt – gegen ein Leben in der Lustmaschine entscheiden und stattdessen ein bewusstes, aktives Leben in der Realität der Außenwelt vorziehen würde. Und dass die meisten Menschen diese Wahl treffen würden, selbst wenn die hedonistische Bilanz ihres „wirklichen“ Lebens wesentlich schlechter ausfallen würde als diejenige der Lustmaschine. Aus diesem Gedankenexperiment lässt sich demnach schließen, dass Werte nicht ausschließlich in hedonischen Erlebnissen bestehen können.

nachvollziehe. Nehmen wir also an, so Griffin, dass mich der Gedanke beschleicht, das Leben eines Freundes sei in einer Weise erfüllt oder lebenswert, die meinem eigenen Leben abgeht. Stellen wir uns vor, dass mein Freund in seinem Leben Dinge erreicht oder tut, die diesem einen Sinn oder ein Gewicht geben, das mein eigenes Leben ermangelt. Und schließen wir weiterhin absurde Leistungen (wie bspw. einen Guinness-Buch-Rekord im Dauer-Videospielen) aus, sodass das Argument nicht auf dem Prinzip der Leistung an sich beruht. Griffin meint also eine Leistung oder eine Form von Aktivität, die intrinsisch wertvoll wäre und damit einem Leben an sich einen Sinn geben könnte. Was für eine Leistung oder Aktivität könnte dies sein?

Griffins Punkt dabei ist, dass der Leser beim Nachvollziehen dieses Gedankenexperiments dialektische und kognitive Operationen (Nachdenken, Begriffsanalyse, Überlegen und Urteilen) durchführt, verschiedene Aktivitäten in ihrem Wert erwägt und bei als wertvoll beurteilten Aktivitäten den Wert einer besonderen Leistung erwägt etc. Diese Übung in Gedanken stellt sich offenbar als eine Art *Auf-* und *Entdeckungsexpedition* im Geiste dar. Wir erwägen, was geeignet wäre, durch einen eigenen Wert auch einem menschlichen Leben einen intrinsischen Wert geben zu können, und wir erwägen dies weitestgehend *in Absehung* unserer persönlichen Vorlieben, Neigungen und Einstellungen („desires“). So erscheint uns Albert Schweitzers Gründung eines Urwald-Krankenhauses in Lambarene wohl als eine intrinsisch wertvolle Leistung, völlig unabhängig von unseren persönlichen „desires“. Hierin liegt ein intuitiv erfahrbares Argument für den belief-Ansatz in der Werttheorie. Vielmehr als um desires scheint es bei der Konstituierung von Werten um eine Art der Wahrnehmung zu gehen, weswegen Griffin den belief-Ansatz auch ein „aesthetic model“ oder ein „perception model“ nennt. Wir würden also Werte „entdecken“, unabhängig davon, welche desires wir aktuell oder potenziell unterhalten. Wert erschließt sich uns, indem wir Objekte dialektisch definieren, ihre Werthaftigkeit erwägen und nicht zuletzt indem wir ein Phänomen erleben und uns seines Wertes „bewusst“ werden. Es ist offensichtlich, dass der belief-Ansatz leicht zu einem Wertrealismus führen kann: wenn Werte Objekte einer Art von Wahrnehmung oder Überzeugung sind, dann spräche vieles dafür, dass sie Teil der faktischen Wirklichkeit sind, die wir eben schlicht „entdecken“ müssten.

Mein Argument gegen den belief-Ansatz ist in seinem Kern daher von der Humeanischen Unterscheidung zwischen Tatsache und Wert inspiriert und gründet auf der Idee, dass diese Unterscheidung nicht sinnvoll aufrechterhalten werden kann, ohne dass wir uns dazu auf den Einfluss der emotionalen Affektivität berufen würden. Wenn wir also tatsächlich über emotionale Affektivität verfügen müssen, um überhaupt die Unterscheidung zwischen

Tatsachen und Werten verstehen zu können, dann scheint dies ein starkes Indiz für die These zu sein, dass Emotionen in irgendeiner Weise die Konstituenten von Werten sind, und dass gerade nicht Emotionen auf Werten, und diese wiederum auf Glaubenszuständen beruhen.

Wir müssen also im Folgenden prüfen, ob es tatsächlich möglich ist, die Praxis des „Wertens“ erschöpfend dadurch erklären zu können, dass ein Akteur „etwas als wertvoll beurteilt“. Einleitend möchte ich ein Zitat von Mary Midgley anführen, die schreibt: „we cannot really think injustice is bad if it does not at some point sicken us“ (1981, 92). Die Idee, auf der diese lapidare Aussage gründet, ist zwar einerseits, dass eine adäquate Reaktion – und darüber hinaus auch ein adäquates Verständnis – auf eine positiv oder negativ werthafte Tatsache notwendig eine entsprechend adäquate emotionale Reaktion impliziert. Dies wird sich jedoch in differenzierterer Form noch im Verlauf des Arguments ergeben. Denn andererseits beruht Midgleys Aussage auf der Annahme, dass wir im Vorgang zu dieser adäquaten Reaktion auf einen Wert oder Unwert zunächst einmal eine *Tatsache*, ein Objekt oder einen Zustand der Welt als Ungerechtigkeit (in diesem Fall) identifizieren können. D.h. wir können nach Midgley eine Beschreibung der *natürlichen Eigenschaften* jener Tatsache geben, der dann sozusagen als separate Entität ein *Wert* anhaftet. So könnten wir bspw. sagen, dass Ungerechtigkeit sicherlich im Spiel ist, wenn einem Unschuldigen von den Vollzugsbeamten der Staatsgewalt körperliche Gewalt angetan wird. Diese Idee folgt Humes Werttheorie. Hume schreibt in dem vielzitierten Abschnitt der *Treatise of Human Nature*, dass wir zwar einen gegebenen Sachverhalt korrekt bspw. als „vorsätzlichen Mord“ („wilful murder“) identifizieren können, aber:

“[The] vice entirely escapes you, as long as you consider the object. You never can find it, till you turn your reflection into your own breast, and find a *sentiment of disapprobation*, which arises in you, towards this action” (Hume, *Treatise*, III 1,1,36, Hervorh. v. mir, S.H.).

In jüngerer Zeit wurde diese These auch von Ursula Wolf vertreten, die behauptet, dass wir einen Sachverhalt, der mit einem bestimmten Wert in Zusammenhang steht, zwar korrekt identifizieren können, u.z. unabhängig von unseren Affekten, dass jedoch der Wertaspekt dieses selben Sachverhalts uns „nur über unseren Affekt zugänglich gemacht werden kann“ (Wolf 1984, 158). Gemäß Hume, Wolf und Midgley können wir also axiologisch relevante Sachverhalte in völliger Apathie *erkennen* und identifizieren, die *Bewertung* dieses Sachverhalts jedoch wird notwendig auf unseren emotionalen Kapazitäten gründen. Demgegenüber geht Hauskeller einen Schritt weiter und bestreitet, dass wir einen axiologisch relevanten Sachverhalt in Absehung von unseren emotionalen Affekten überhaupt erkennen und identifizieren könnten:

„Solange ich das Leiden des anderen nicht nachfühle, solange ich nicht mit ihm leide, weiß ich letztlich gar nichts von seinem Leid, das heißt, ich weiß nicht einmal wirklich, *dass* er leidet“ (Hauskeller 2001, 230f.)

Wenn wir also über einen Fall von Leid oder Ungerechtigkeit (wie im obigen Zitat von Midgley, oder wie in meinem Beispiel des Übergriffs auf einen Unschuldigen) in einer emotional neutralen Weise reden, dann verwenden wir nach Hauskeller unsere Worte nur gemäß ihres *pragmatischen* Wertes: wir benutzen einen linguistischen token, um auf einen Sachverhalt zu referieren, wie wir dies als kompetentes Mitglied einer Sprachgemeinschaft gelernt haben. Wir verwenden darum jedoch die Worte noch nicht in einer *bedeutungsvollen* Weise, sondern nur gemäß ihres konventionellen strukturellen Wertes im Sprachsystem. Hauskeller illustriert diesen Punkt wie folgt: wenn wir in einer emotional neutralen Weise über Leid oder Ungerechtigkeit reden, dann wissen wir genauso viel über die *Bedeutung* von „Leid“ und „Ungerechtigkeit“, wie wir über das Fahrradfahren wüssten, wenn wir lediglich ein Buch darüber gelesen hätten.

Diese fortschreitenden Positionen einer humaneischen und einer gewissermaßen ultrahumaneischen Sichtweise zielen darauf ab zu bestreiten, dass ein Glaubenszustand das Phänomen des „Wertens“ erschöpfend erklären könne. Ein Glaubenszustand allein liefert lediglich ein unterdeterminiertes Bild vom Phänomen des „Wertens“. Die oben zitierten Autoren vertreten einhellig die Relevanz von Emotionen in dieser Hinsicht. Zu klären bleibt im Folgenden jedoch, in exakt welcher Weise Emotionen nun zur Praxis des „Wertens“ beitragen. Können wir tatsächlich, gemäß der Position von Hume, Midgley und Wolf, eine klare Trennung zwischen einem „Tatsache-Urteil-Aspekt“ und einem „Wert-Affekt-Aspekt“ in dieser Praxis vornehmen? Oder müssen wir, wie Hauskeller nahezulegen scheint, diese Trennung überdenken und evtl. die Grenzen zwischen rationaler Tatsachenerkenntnis einerseits und emotionaler Wertkonstitution andererseits neu definieren?⁷⁶

⁷⁶ Auch Griffin (1998, 9f.) scheint dieser Überzeugung zu sein: „the taste model assumes that we can isolate valued objects in purely natural terms and then, independently, react to them with approval or disapproval. But can we? [...] The act of isolating the objects we value is far from neutral. [...] One cannot here distinguish the identification of the object to be responded to and the response to it.“ Wenngleich Griffin also meinen Zweifel an einer solchen sauberen Trennung in Faktizität und Wert zu teilen scheint, so sind seine daraus gezogenen Schlussfolgerungen doch unterschiedlich von der in dieser Arbeit vorgeschlagenen Lösung. Nach Griffin müssen wir „volitionale Elemente“ in die Analyse einbringen, sodass Werthaftigkeit erklärbar wird auf der Grundlage von allgemeinen menschlichen Antrieben: „There is no adequate explanation of their being desirability features without appeal to certain natural human motivations“. Ich hingegen argumentiere dafür, dass wir emotionalen Affekt in die Analyse einbringen müssen, um Werthaftigkeit erklären zu können.

3.4. Wertrealismus

An dieser Stelle möchte ich den kritischen Durchgang durch Positionen in der Werttheorie fortsetzen. Bisher haben wir die Dichotomie zwischen dem belief- und dem desire-Ansatz diskutiert, zwei Theoriemodelle, die sich dezidiert auf das Phänomen des „Wertens“ (Adjektivform) beziehen, was in anderen Studien gelegentlich leider nicht explizit gemacht wird. Ein klares Verständnis der involvierten Begriffe kann sich jedoch nur dann ergeben, wenn wir konsequent terminologische Differenzen beachten. Im Folgenden werden wir das Modell des Wertrealismus diskutieren und damit auf einen anderen Aspekt des Wertbegriffs abheben, nämlich auf den des „Wertes“ (Nominalform). Auf dieser gewissermaßen eher „objektiven“ Ebene des Wertbegriffs geht es verstärkt um Fragen der Wertontologie, allerdings kann es sich bei dieser Zuordnung nur um graduelle Unterschiede handeln. Positionen in der Debatte um den Begriff des „Wertens“ sind natürlich eng verbunden mit bestimmten Positionen in der Debatte um den Begriff des „Wertes“. Traditionell ist bspw. den Vertretern des „Wahrnehmungsmodells“ bzw. des belief-Ansatzes eher eine wertrealistische Position zuzuordnen: wenn die Praxis des Wertens als die Wahrnehmung von axiologischen Entitäten in der Außenwelt verstanden wird, dann liegt es nahe, im gleichen Zuge die objektive Existenz solcher Entitäten zu postulieren.⁷⁷

Es wurde immer wieder betont, dass sich die vorliegende Arbeit auf den Begriff des Wertens konzentriert, und nicht so sehr auf den des Wertes. Welche Relevanz haben dann die folgenden Ausführungen zum nominalen Aspekt („Wert“) des Begriffs? Eine bestimmte ontologische Position hat keine besondere Relevanz für den Fortgang der Argumentation, wenngleich sich eine solche Position aus meinen Ausführungen ableiten lässt. Ich werde jedoch nicht detailliert darauf eingehen, da es zuvörderst darum gehen soll, zu bestimmen, wie die Praxis des Wertens zu verstehen ist, und wie Werte (als individuelle mentale Zustände) auf die Praxis emotionalen Wertens zu gründen sind. Die folgenden Ausführungen sollen jedoch in aller Kürze zu einem besseren Verständnis der angesprochenen Unterscheidung zwischen Tatsachen und Werten führen. Im Ausgang des letzten Abschnitts hatten wir gefragt, ob diese Unterscheidung im Anschluss an unsere bisherigen Ergebnisse evtl. zu überdenken sei. Die folgenden Ausführungen sollen diese Überlegungen unterfüttern.

Charakteristisch für wertrealistische Positionen ist die Annahme, dass „Werte“ als objektive Entitäten in der Welt existieren. Werte bilden eine distinkte Klasse von

⁷⁷ Hier handelt es sich natürlich um eine metatheoretische Korrelation, und nicht um eine begriffslogische Implikation.

Eigenschaften, zu der wir einen jeweils zu bestimmenden besonderen epistemischen Zugang haben. Schon hier drängt sich die Frage auf: wenn das Sichtbare gesehen werden kann, das Akustische gehört werden kann etc., was könnte dann das geeignete Sensorium des epistemischen Zugangs zu „Werten“ sein? Es scheint kein Objekt zu geben, für das es einen besonderen „Wert-Sinn“ (analog zu Gesichtssinn oder Gehörsinn) geben könnte, etwa so wie Schallwellen mit dem Gehörsinn korrespondieren, und bestimmte chemische Substanzen mit dem Geschmackssinn. Das Fehlen eines besonderen Sinnes-Apparats für Werte stellt – so lautet diese Überlegung – die Existenz distinkter Werteigenschaften in Frage.

Eine mögliche Erwiderung auf diesen Einwand ist die Annahme, dass wir Werte mittels unserer Emotionen wahrnehmen. Emotionalität wäre in diesem Sinne das Sensorium für Werte, das zwar komplexer ist als bestimmte organisch zu fixierende Sinnesapparate und darüber hinaus transkategorial⁷⁸, sich aber dennoch einer Definition nicht entzieht. Ich denke jedoch, wie ich in der Diskussion des belief-Ansatzes zu zeigen versucht habe, dass diese Position fehlerhaft ist, da sie das Subjekt auf der rezeptiven, und nicht auf der produktiven Seite des Prozesses verortet: Emotionen sind nicht Wahrnehmungsmodi für Werte, sondern Subjekte schaffen Werte vermittelt von emotionalen Reaktionen.

Dennoch ist die These der emotionalen Wahrnehmung von Werten intuitiv sehr überzeugend. Im phänomenalen Erleben stellt sich der Prozess, durch den wir „auf einen Wert stoßen“ tatsächlich so dar, dass wir plötzlich involvierter Wert gewahr werden. Dieses Erleben kann uns in die Irre führen, intuitiv eine realistische Position für überzeugend zu halten. Außerdem werden anti-realistische Positionen oft mit einem Werte-Nihilismus gleichgesetzt. Dies erscheint uns schon per se inakzeptabel. Uns liegt als Menschen (und nicht nur als Wert-Theoretiker) mehr an der Debatte um Werte, als an Debatten bspw. in der theoretischen Mathematik, da erstere unser privates und gesellschaftliches Leben betrifft und bestimmte Positionen – wie der Werte-Nihilismus – die Konzeption gesellschaftlichen Lebens und Handelns umwälzen würde. Betrachten wir jedoch die Annahme, es gäbe ein spezifisches Werte-Sensorium, das funktionieren würde wie unsere Sinnesorgane.

⁷⁸ Wahrnehmungsprozesse sind normalerweise als physiologische Reaktionen unserer Sinnesorgane zu erklären, die Erklärung bewegt sich also auf einer theoretischen Ebene, innerhalb einer Kategorie von involvierten Elementen sozusagen. Die Reaktionen des emotionalen Systems ist – wie wir in einem früheren Kapitel gesehen haben – nicht auf lediglich einer Ebene und nicht unter Zuhilfenahme von ausschließlich einer Kategorie von involvierten Elementen zu erklären. Physiologische Erregung, vegetative Hirnprozesse, kognitive Prozesse und phänomenologische Bewusstseinsprozesse wirken zusammen in der Emotionselzitation.

Der bekannteste Vertreter einer Position, die einen Wertrealismus mit der These verbindet, dass Emotionen unseren „Werte-Sinn“ darstellen, ist sicherlich Max Scheler. Scheler (1954, 37f.) glaubte, es gäbe „*echte* und *wahre* Wertqualitäten [...], die ein eigenes [sic] Bereich von *Gegenständen* darstellen, die ihre *besonderen* Verhältnisse und Zusammenhänge haben, und schon als *Wertqualitäten* z.B. höher und niedriger usw. sein können“. Scheler war also auch Vertreter eines Wertatomismus, dem gemäß Werte stets eine bestimmte Valenz haben, unabhängig vom gegebenen Kontext. Werte basieren nicht auf natürlichen Eigenschaften, sondern sind „anschaulich gegeben“ vermittelt von Gefühlen. Nach Scheler lässt sich dies durch folgende Beobachtung illustrieren: fragen wir nach der gemeinsamen Eigenschaft von guten oder schlechten Handlungen, Einstellungen oder Menschen, dann wird die Antwort ebenso tautologisch ausfallen, wie auf die Frage nach der gemeinsamen Eigenschaft von blauen oder roten Gegenständen. Neben einer Vielzahl von Implausibilitäten ist besonders Schelers wertrealistische Annahme, verbunden mit der Atomismus-These problematisch: wir reagieren emotional nicht immer so, wie es adäquat wäre, wie wir also normalerweise würden oder sollten. Ändert dies den ontologischen Status der Werte? Wie jedoch wäre eine solche Varianz zu erklären, wenn Werte stets „anschaulich gegeben“ wären? Eine Trennung in eine objektive, unveränderliche Welt der Werte einerseits und eine veränderliche Welt der Wertewahrnehmung andererseits wäre jedoch unplausibel, da Varianzen in der emotionalen Reaktion prinzipieller Natur sind, und nicht bloß Ausdruck von defizienter Wahrnehmung. Unser epistemisches Sensorium wäre also prinzipiell inadäquat, eine Härte, die Schelers Theorie höchst unplausibel macht.

Auch Kupperman (1999) diskutiert das Paradigma eines „Werte-Sinns“ und stellt einen Katalog von notwendigen Bedingungen für eine solche Sinneskraft auf. Etwas, das ein Kandidat für einen Werte-Sinn sein wollte, müsste folgende Bedingungen erfüllen: (1.) Ein Werte-Sinn müsste eine Person, die über ihn verfügt, in eine bessere Position bringen, Werturteile zu fassen, als wenn diese Person nicht über den Werte-Sinn verfügen würde. Dies impliziert andererseits auch nicht hundertprozentige Reliabilität: auch der Gesichtssinn kann uns täuschen (durch Illusionen oder Halluzinationen). (2.) Ein Werte-Sinn müsste selbstreflexiv sein: wenn wir ein Geräusch hören, dann sind wir uns durch diesen Prozess des Hörens auch darüber bewusst, *dass* wir ein Geräusch hören. Ein solches Bewusstsein müsste auch mit dem Werte-Sinn verbunden sein. (3.) Ein Werte-Sinn muss Urteile aus erster Hand ergeben, und nicht bloß Urteile, die wir nicht selbst gebildet, sondern von dritten Personen erworben haben. (4.) Die Existenz eines Werte-Sinns muss die Annahme implizieren, dass passende Sinnesobjekte für diesen in der Welt existieren. Mit dieser ontologischen Position

hängt die epistemische Position zusammen, dass es mindestens eine (ideale) Person geben muss, die Recht hat in einer Wertedebatte: ein Subjekt mit der „richtigen“ Wahrnehmung (analog zur Sinneswahrnehmung) hätte Recht, alle anderen, die nicht zustimmen, hätten Unrecht.

Nach Kupperman könnten wir einen solchen Werte-Sinn begrifflich analog zu Konzepten wie dem „Sinn für Humor“ oder dem „Sinn für gutes Benehmen“ verstehen. Kupperman verbindet mit dieser Position eine Kritik des rationalen Wahrnehmungsmodells, also des belief-Ansatzes. Wie seine Aufstellung von notwendigen Kriterien zur Bestimmung eines Werte-Sinns ergeben hat, können rein rationale Erwägung oder dialektische Operationen diese Bedingungen nicht erfüllen. Eine Person, die auf der Grundlage von rationaler Argumentation zu der Annahme einer bestimmten Wertüberzeugung gebracht wurde, aber nicht adäquat emotional reagiert, hat den in Frage stehenden Wert eben nicht wirklich „wahrgenommen“. Es kann keinen rationalen Werte-Sinn geben, da dieser das Kriterium (2) nicht erfüllen würde. So sagt Kupperman: „beliefs are rarely ever experienced“ und fährt fort: „Otherwise, the person who says, ‘X has high value’ (without relevant emotional states), is merely *saying* that X has high value, without any sense of an opening to the value of X“ (Kupperman 1999, 64). Jemand, der bspw. im Gespräch angibt, dass er Französische Weine wertschätze, besonders diejenige aus der St. Emilion Region, und diese Wertschätzung mit allerlei Argumenten zu unterfüttern weiß (Sonneneinstrahlung, Bodenbeschaffenheit, feine Säure des Produkts etc.), der jedoch nicht die geringste Reaktion von *Freude* zeigt, wenn man ihm eine Flasche St. Emilion eines guten Jahrgangs schenkt, oder der nicht die geringsten Anzeichen von *Genuss* beim Trinken eines solchen Weines zeigt: eine solche Person würde uns zu Recht vor ein Paradoxon stellen.

Der Wertrealismus steht also vor großen Problemen in der Erklärung des adjektivischen Aspekts des Wertbegriffs: ein realistischer belief-Ansatz lässt die Praxis des Wertens unterdeterminiert, wie wir im vorangegangenen Absatz anhand einer Illustration gesehen haben, und ein realistischer Ansatz, der Werte als die „Erkenntnisobjekte“ von Emotionen auffasst, muss aus folgenden Gründen scheitern: erstens haben wir gesehen, dass emotionale Reaktionen plausibler als eine Praxis des Wertens verstanden werden können, denn als ein epistemisches Sensorium für objektive Werteigenschaften. Zweitens spricht die Tatsache (die ich detailliert in Kapitel 2 diskutiert habe), dass wir in unseren Primäremotionen zunächst auf natürliche Eigenschaften der Außenwelt reagieren, und nicht etwa auf Klassen von Werteigenschaften (formale Objekte) wie „das Gefährliche“ dafür, dass unsere primären emotionalen Reaktionen solche Klassen von Werteigenschaften überhaupt erst generieren. Da

jedoch gerade der Mechanismus der Primäremotionen weitestgehend unbewusst abläuft, verfallen wir in einem späteren Entwicklungsstadium der Illusion, die emotional instanziierten Klassen von Werteigenschaften seien in Wirklichkeit präexistent und gewissermaßen die objektive Voraussetzung für unsere (späteren) emotionalen Reaktionen.

Wie ließe sich eine wertrealistische Position auf ontologischer Basis verteidigen, d.h. in Absehung von den Defiziten dieser Position in der Erklärung der Praxis des „Wertens“? Bekanntlich hat Mackie (1977, Kap. 1.9.) mit seinem berühmten „argument from queerness“ den Realismus attackiert, da er eine Klasse von objektiven Eigenschaften („axiologischen Tatsachen“) postuliert, zu denen wir weder durch Sinneswahrnehmung noch durch wissenschaftliche Methoden Zugang haben. Hat eine Aussage wie „Erdbeereis ist lecker“ tatsächlich denselben ontologischen Status wie „Dieser Hund wiegt 20kg“, „Dieser Metallgegenstand ist spitz“ oder „Dieser Apfel ist grün“? Kupperman (1999, Kap. 5) hingegen, der eine wertrealistische Position verteidigt, merkt an, dass diese Kritik ihre Stoßkraft hauptsächlich aus der missverständlichen und fehlerhaften Gleichsetzung von „axiologischen Tatsachen“ mit „Dingen“ bezieht. Nach Kupperman gibt es jedoch auch nicht-axiologische Tatsachen in der Welt, deren Existenz unkontrovers ist, die sich dennoch nicht in Analogie zu „Dingen“ auffassen lassen. Bede Rundle (1993, 10) gibt das Beispiel der Tatsache, „dass die Inflationsrate gesunken ist“. Hieran wird deutlich, dass Tatsachen erstens keine Dinge sind und daher zweitens als interpretative Konstrukte von Weltzuständen aufgefasst werden müssen. Searle (1995, 211) bezeichnet Tatsachen als „conditions that obtain in the world“, Strawson (1950, 135) postulierte prominent, dass Tatsachen – anders als Dinge – nicht selbst in der Welt existieren, da sie eben Konstrukte von Weltzuständen darstellen, und nicht Weltzustände selbst. Kupperman (1999, 74) bestimmt Werte in eben diesem Sinne als „interpretative constructs related to (less conspicuously interpreted) underlying phenomena“. Die ontologische Debatte über den Wertrealismus zeigt also, dass wir Tatsachen nicht als reine Faktizität interpretieren können: Tatsachen sind nicht „in der Welt“, so wie ein Schreibtisch in der Welt existiert. Die (zweifelloso objektive) Tatsache, dass die Inflationsrate gefallen ist, bedarf interpretativer Konstruktion, um zu bestehen. Erinnern wir uns an dieser Stelle an die im Anschluss an die Position von Hauskeller formulierte Frage, ob wir evtl. die Trennung zwischen Tatsachen und Werten überdenken müssen. Können wir axiologisch relevante Tatsachen überhaupt als wertneutrale Tatsachen wahrnehmen, oder vielmehr verstehen? Da Tatsachen Konstrukte sind, muss in ihre Formierung subjektive Interpretationsarbeit eingehen. Wenn es sich nun um axiologisch relevante Tatsachen handelt, liegt nahe, dass diese Interpretation auch schon eine axiologische Vor-Markierung einschließt.

Da wir gesehen haben, dass der belief-Ansatz Werte unterdeterminiert lässt, spricht also vieles dafür, dass wir axiologische Tatsachen nur dann überhaupt *als Tatsachen* konstruieren und verstehen können, wenn eine emotionale Reaktion oder Disposition in diese Formierung miteingeht. Wir können also mit anderen Worten – in Anlehnung an Hauskeller – sagen: selbst wenn wir „wilful murder“ als wertneutrale Tatsache definieren können, so *verstehen* wir diesen Begriff dennoch nicht adäquat ohne die Zurechnung eines Bewertungsaspekts.

3.5. Axiologischer Anti-Realismus

Anti-realistische Theorieformen teilen zunächst als gemeinsames Merkmal eine Ablehnung des Realismus, d.h. der Annahme, dass objektive axiologische Tatsachen in der Welt existieren, bzw. dass Werteigenschaften eine objektive und von natürlichen Eigenschaften distinkte Klasse darstellen. Eine prominente Spielart des Anti-Realismus finden wir im Emotivismus, den wir schon in der Einleitung ansatzweise diskutiert haben.

Dem Emotivismus gemäß drücken Bewertungen emotionale Zustände oder Einstellungen aus. Damit wird gesagt, dass Bewertungen keine Aussagen, Urteile oder Propositionen sind, die Wahrheitswerte annehmen können und sich auf einen Zustand der Welt beziehen. Gemäß dem Modell des „direction of fit“ müssten solche Werturteile schließlich einen objektiven Zustand der Welt repräsentieren, was unplausibel wäre, wenn wir dadurch zu einer Annahme objektiver Werteigenschaften gezwungen wären. Bewertungen sind also vielmehr Ausdrücke oder Manifestationen unserer emotionalen Reaktionen und Einstellungen, wobei beide unter einem reduktionistischen Blickwinkel gesehen werden: Emotionen sind nicht-rationale Phänomene, die kontingent auftreten auf der Grundlage von nicht-stetigen Mustern biologischer oder psychologischer Reaktionen. Wie bereits erwähnt, wurde der Emotivismus von Axel Hägerström (1953) entwickelt und von Ayer (1936), Stevenson (1937 und 1944), Urmson (1968) und Hare (1952) aufgenommen und populär gemacht. Gegenwärtige prominente Vertreter einer emotivistischen Positionen sind bspw. Blackburn (1984 und 1998) und Gibbard (1990). Emotivistische Positionen stellen sich jedoch in einer großen Bandbreite dar, und der heutige Emotivismus von Blackburn oder Gibbard stellt sich natürlich wesentlich weiter entwickelt dar als die früheren Ansätze.

D’Arms & Jacobson (2000, 724f.) unterscheiden bspw. zwischen einem „simplizistischen Sentimentalismus“ und einem voll ausgebildeten Emotivismus. Im Rahmen des Ersteren stellt sich folgendes Problem: Eine Aussage wie „x ist schlecht“ wird interpretiert als „Ich lehne x ab“. Es könnte also in diesem Rahmen keine substanziell gehaltvollen Wertdebatten geben.

Denn wenn ich aussage, x sei schlecht, und Hans aussagt, x sei gut, dann liefern wir beide jeweils lediglich einen Bericht unserer mentalen Zustände, und in diesem Sinne haben wir beide gleichermaßen Recht. Da eine Bewertung darüber hinaus an kein objektives Kriterium zurückgebunden wird, wäre damit die Debatte zu Ende, bevor sie begonnen hat. Da es jedoch echten Disput über Werte gibt, muss der simplizistische Sentimentalismus in dieser Hinsicht falsch liegen.

Der klassische Emotivismus (von Ayer und Stevenson) stellt eine verfeinerte sentimentalistische Theorieform dar. Die Aussage „x ist schlecht“ wird nicht als Bericht eines mentalen Zustands im Sinne einer *Beschreibung* interpretiert, sondern als ein *Ausdruck* eines mentalen Zustands. Wertdebatten sind demnach möglich, jedoch nicht im klassischen Sinne einer Debatte über objektive Wahrheit, sondern im Sinne einer Debatte über Einstellungen. Da Personen durch Bewertungen Einstellungen zum Ausdruck bringen, die wir befürworten oder ablehnen können, können wir im Rahmen einer Wertdebatte versuchen, Personen mit konkurrierenden Einstellungen zur Modifikation oder Aufgabe bestimmter Einstellungen zu bewegen. Diese Idee liegt Hares „präskriptivistischer“ Theorie zu Grunde. Allerdings denke ich, dass dem Emotivismus darüber hinaus ein grundlegendes Problem anhaftet. Ein Ansatz, der positive Bewertungen als „Befürworten“ und negative Bewertungen als „Ablehnen“ interpretiert, liefert notwendiger Weise eine reduktionistische Analyse unserer axiologischen Sprache und Begriffe. Wenn bspw. alle negativen Bewertungen einheitlich als contra-Einstellung interpretiert werden, dann kann eine solche Theorie keine semantische gehaltvolle oder informative Unterscheidung von Begriffen wie „schlecht“, „hässlich“, „unaufrichtig“ oder „ekelerregend“ liefern. In der sozialen, linguistisch geformten Welt existiert jedoch eine Vielzahl solcher Begriffe negativer wie auch positiver Bewertung, und eine semantisch gehaltvolle Unterscheidung zwischen ihnen scheint eine wichtige Rolle in der sprachlichen Interaktion und persönlichen Selbstreflexion zu spielen.

Häufig fällt in Diskussionen des Emotivismus nicht einmal auf, dass die emotivistische Strategie lediglich eine „rhetorische Strategie“ (Kupperman 1999, 16f.) ist. Der Emotivismus nutzte die seinerzeit verbreitete philosophische Geringschätzung der Emotionen, da diese vermeintlich keine Rationalität zulassen. Der emotivistische Begriff der „Emotion“ beschreibt ein Phänomen, das einerseits nicht-rational, und zugleich andererseits motivational wirksam ist. Ist damit jedoch schon hinreichend ein adäquater Begriff der „Emotion“ erfasst? Es fehlt m.E. zumindest noch der zentrale Aspekt der Affektivität, während die postulierte Nicht-Rationalität zumindest fraglich ist. Schon Urmson (1968) hatte gesehen, dass der Emotivismus nicht wirklich von „Emotionen“ handelt, sondern von „Einstellungen“ in einem

weiten Sinne, der lediglich von den beiden oben zitierten Merkmalen umrissen wird. Eine Person kann nun eine Einstellung (im emotivistischen) Sinne haben, ohne dass sie aus diesem Grund eine Emotion (bzw. emotionale Disposition) hätte, u.z. weder aus alltagssprachlicher noch aus wissenschaftlicher Sicht.

3.6. Naturalismus und Supervenienz

Jede anti-realistische Werttheorie sieht sich vor der Herausforderung, das Verhältnis zwischen Werten und natürlichen Eigenschaften zu erklären, da für diesen Theorietypus schließlich die Strategie, eine distinkte Klasse objektiver axiologischer Eigenschaften zu postulieren, ausgeschlossen ist. Werteigenschaften müssen also in irgendeiner Weise auf natürliche Eigenschaften zurückgeführt werden. Dies kann mit Hilfe einer reduktionistischen Strategie erfolgen. Ein berühmtes Beispiel ist der *begriffliche Naturalismus*, den G.E. Moore mit seinem bekannten „open question argument“ (Moore 1903, § 13) attackiert hatte. Der begriffliche Naturalismus versucht, eine vollständige Beschreibung von Werteigenschaften allein unter Verwendung von natürlichen Eigenschaften zu geben. Nach Moore jedoch kann eine solche Definition niemals analytisch sein, so wie etwa die Definition von „Junggeselle“ als ein „Mann, der noch nie verheiratet war“ analytisch ist. Denn einer Definition von „Mord“ im Rahmen natürlicher Eigenschaften, wie etwa „die vorsätzliche Tötung eines Menschen aus niederen Motiven“, können wir mit der Frage begegnen „ist das gut oder schlecht?“, ohne dass wir damit unser Verständnis der verwendeten Begriffe in Frage stellen müssten, wie dies der Fall wäre, wenn wir fragen würden, ob ein Junggeselle noch nie verheiratet gewesen sei. Ich werde nicht weiter in diese Diskussion eintreten, da meine eigene Argumentation nicht auf eine Position in der Sprach oder Ontologie der Werte abhebt. Betrachten wir jedoch eine andere Form des Naturalismus, die uns zu einem letzten wichtigen Theoriemodell auf dem Feld der anti-realistischen Theorien überleiten soll. Modelle des sog. *substanziellen Naturalismus* („substantive naturalism“) operieren nun nicht mit der Reduktion von Begriffen, sondern argumentieren, dass bestimmte Sachverhalte und Eigenschaften im Bereich der Werte auf stetige Weise mit bestimmten Sachverhalten und Eigenschaften der natürlichen Welt korrespondieren. Verschiedene Erklärungsansätze können im Rahmen eines substanziellen Naturalismus interpretiert werden. Der desire-Ansatz bspw. kann einer solchen Interpretation unterzogen werden, da Werthaftigkeit auf die objektive Tatsache reduziert wird, was eine Person wünscht („desires“) bzw. was diese Person wünschen würde, wenn sie völlig rational wäre (vgl. Brandt 1979, Kap. 1).

In neuerer Zeit ist an die Stelle eines (reduktionistischen) substanziellen Naturalismus die konkurrierende Theoriebegriff der „Supervenienz“ getreten (vgl. Hare 1984). Es muss schließlich ein Verhältnis zwischen natürlichen Eigenschaften und Werteigenschaften bestehen: wenn zwei Sachverhalte G und H identisch sind hinsichtlich ihrer natürlichen Eigenschaften und in vergleichbaren Kontexten auftreten, dann kann – *ceteris paribus* – kein fundamentaler Unterschied zwischen den Werteigenschaften von G und H vorliegen. Der Begriff der Supervenienz, die Annahme also, dass Werte in irgendeiner stetigen Weise auf natürliche Eigenschaften supervenieren, erlaubt im Gegensatz zum Naturalismus eine „Abhängigkeit ohne Reduktionismus“: „Supervenience allows the dependence of the higher level of explanation upon the lower without reduction, and the autonomy of the higher level without dualism“ (Griffin 1996, 44).

Eine schwache Supervenienz-Beziehung umfasst schlicht das Kriterium der *Ununterscheidbarkeit*: damit ist das gerade angeführte Beispiel aufgegriffen, gemäß dem eine Ununterscheidbarkeit auf dem Niveau der natürlichen Eigenschaften – *ceteris paribus* – notwendig eine Ununterscheidbarkeit auf dem Niveau der axiologischen Eigenschaften impliziert. Dieser Ansatz stößt jedoch auf ein Problem: selbst bei weitgehender Diversität auf dem Niveau der natürlichen Eigenschaften können wir immer noch auf Identität auf dem Niveau der axiologischen Eigenschaften stoßen. Es gibt schlicht zahlreiche natürliche Eigenschaften, die zur Beschreibung einer Entität herangezogen werden können, und es ist sehr wahrscheinlich, dass eine Variation in einer Vielzahl dieser Eigenschaften dennoch nichts an den axiologischen Eigenschaften derselben Entität ändern würde. Wir benötigen also ein *Relevanzkriterium*. Dieser Punkt lässt sich an einem anderen, zur Zeit vieldiskutierten Supervenienz-Paradigma erläutern. Nehmen wir an, wir wollten untersuchen, wie mentale Zustände auf physiologische Zustände supervenieren oder gar auf diese zu reduzieren sind; dies macht keinen Unterschied für die Illustration meines Punktes. Bei der Untersuchung dieses Verhältnis schließen wir von vornherein bestimmte physiologische Zustände aus, die nach unserem Vorverständnis nur peripher oder sekundär mit mentalen Zuständen korrespondieren. Die Messung der Hautleitfähigkeit bspw. kann uns zwar Rückschlüsse auf den Grad der psychologischen Erregung liefern, dies jedoch nur peripher: in erster Linie würden wir wohl direkt *Hirnzustände* untersuchen, die mit einem in Frage stehenden mentalen Zustand (z.B. Angst) korrespondieren. Nun kann man einwenden, dass es unkontrovers ist, dass mentale Zustände sich im Hirn ereignen, und dass daher der Ausschluss anderer physiologischer Zustände (wie der Hautleitfähigkeit) auf einem trivialen Kriterium beruht. Wir würden jedoch noch weitere Einschränkungen vornehmen: wenn wir bspw.

Farbwahrnehmungen auf Hirnzustände zurückführen (oder reduzieren) wollen, dann würden wir wahrscheinlich Untersuchung am ARAS (aufsteigenden retikulären Aktivierungssystem) ausschließen, von dem wir wissenschaftlich begründet annehmen, dass es lediglich den Grad der Wachheit eines Subjekts beeinflusst usw.

Die Definition eines Relevanzkriteriums und die Unterscheidung zwischen Basis- und Supervenienz-Niveau baut also notwendig auf einem *Vorverständnis* auf. Ob und inwiefern dies im Falle naturwissenschaftlicher Untersuchung viziöse Auswirkungen haben kann, sei hier dahingestellt. Im Falle der Werte-Supervenienz jedoch wirft diese Frage ein echtes Problem auf: wie definieren wir ein Basis-Niveau von relevanten natürlichen Eigenschaften, ohne bei dieser Relevanz-Entscheidung schon auf in Frage stehende Werteigenschaften des Supervenienz-Niveaus vorzugreifen? Wiederum finden wir die strikte Unterscheidung zwischen natürlichen Tatsachen einerseits und Bewertungen bzw. axiologischen Eigenschaften andererseits in Zweifel gezogen. Griffin (1996, 46) illustriert diesen Punkt anhand eines gegenläufigen Beispiels: können wir uns einen Fall vorstellen, in dem *Differenzen in den axiologischen Eigenschaften bestehen bei gleichzeitiger Identität der (relevanten) natürlichen Eigenschaften?* Stellen wir uns vor, so Griffin, dass sowohl Smith als auch Jones Gedichte schreiben. Smith's Gedichte sind eine echte intellektuelle und ästhetische Leistung, wohingegen Jones' Gedichte, die genauso lang, innovativ, eloquent und variantenreich sind, einen ebenso großen Wortschatz verwenden etc., dies jedoch nicht sind. Betrachten wir die oben gestellte Fragestellung in genau dieser Form (die kursivierte Phrase also): Wie können wir den Unterschied in der Bewertung der Gedichte von Smith und Jones erklären? Lassen wir zunächst das eingeklammerte Kriterium „relevant“ weg. In diesem Fall ist es einfach, stets mit einem (absurden oder trivialen) Unterschied zu argumentieren. Nehmen wir an, Smith und Jones hätten verschiedenfarbige Tinten benutzt, und nehmen wir weiter an, dass Jones' Scheitern, ein ebenso gutes Gedicht zu produzieren, eben an der Farbe seiner Tinte gelegen haben soll.

Diese Lösung kann natürlich weder ernst genommen werden noch zufriedenstellend sein, womit die Notwendigkeit eines Relevanzkriteriums illustriert wäre. Beziehen wir also das eingeklammerte Kriterium „relevant“ aus der obigen Phrase wieder mit ein. Und nehmen wir nun an, dass Smith's Gedichte voller *tiefer Einsicht in wichtige Themen der menschlichen Existenz* sind. Jones' Gedichten fehlt diese Qualität, und genau darum seien nun Smith's Gedichte eine echte Leistung, und Jones' Gedichte eben nicht. In diesem Fall sind die Basis-Eigenschaften „tiefe Einsicht“ und „wichtige Themen der menschlichen Existenz“, auf die die axiologische Eigenschaft („echte Leistung“) von Smith's Gedichten superveniert. Wir können

nun getrost einräumen, dass es sich im Gegensatz zur Farbe der Tinte um *relevante* Basis-Eigenschaften handelt. Doch handelt es sich um *natürliche* Eigenschaften? Weder „Einsicht“ noch „wichtige Themen der menschlichen Existenz“ passen wirklich komfortabel in die Klasse der natürlichen Eigenschaften; beide scheinen vielmehr schon einen weitgehenden Anteil von Bewertungen einzuschließen.

Griffin (loc. cit.) resümiert: „A lot that is natural is not relevant, and a lot that is relevant is not natural“. Bei der Definition von Supervenienz-Beziehungen im axiologischen Bereich tritt offenbar notwendig das Dilemma auf, dass wir entweder keine stetige Supervenienz herstellen können, da wir *alle* erdenklichen natürlichen Eigenschaften mit einbeziehen müssten. Es würde also keine definitorische Identität mehr vorliegen. Oder aber wir vermischen beide Seiten der Supervenienz-Beziehung, indem wir gerade jenes erste Problem beheben wollen. Denn durch Einführung eines Relevanz-Kriteriums schließen wir nicht-relevante natürliche Eigenschaften aus und können nun tatsächlich definitorischen Identität auf der Basis-Ebene herstellen. Es ist jedoch fraglich, ob ein Relevanz-Kriterium auf der *Basis*-Ebene gefunden werden kann, das nicht schon auf Bewertungen zurückgreift, die der *Supervenienz*-Ebene zu Grunde liegen. Wiederum drängt sich der Schluss auf, dass wir keine strikte theoretische Trennung zwischen einerseits Werten und andererseits denjenigen Entitäten vornehmen können, auf denen Werte gründen. Beide Seiten dieser Beziehung scheinen schon im Vorhinein „vermischt“ zu sein, d.h. in einem Verhältnis zueinander zu stehen. Wir benutzen eine evaluative, axiologisch reiche Sprache schon immer auch in der Beschreibung von „natürlichen“ Tatsachen und Weltzuständen.

3.7. Dispositionalismus

Auch der Dispositionalismus, eine weitere Spielart anti-realistischer Werttheorien, kann als eine Bestrebung im Anschluss an Erkenntnisse des Naturalismus bzw. der Supervenienz-Theorie gesehen werden. Wir hatten gesehen, dass der *desire*-Ansatz oft als ein Beispiel eines Modells des substanziellen Naturalismus figuriert. Werthaftigkeit wird auf die objektive („natürliche“) Tatsache reduziert, was eine Person wünscht („*desires*“) bzw. was diese Person wünschen würde, wenn sie völlig rational wäre. Hierbei wird jedoch das Subjekt, das Träger der jeweiligen *desires* ist, als „black box“ behandelt.⁷⁹ Die empirisch zu erhebende oder

⁷⁹ Eine Ausnahme bildet freilich Theorien des „informed-desire“ Ansatzes, der in meiner Ausführung durch den Nebensatz angesprochen ist „was eine Person wünschen würde, wenn sie völlig rational wäre“. Dieser Ansatz teilt

faktisch zu postulierende Tatsache „dass x Objekt eines desires ist“ stellt die Basis-Eigenschaft dafür dar, dass x das Supervenienz-Prädikat „wertvoll“ zugesprochen wird. Auch der Dispositionalismus operiert mit einer Zurückführung von Werten auf natürliche Eigenschaften, allerdings spielt das wertende Subjekt immerhin die theoretische Rolle eines Trägers bestimmter Reiz-Reaktions-Schemata. Ich denke, dass ein Anschluss an diese Theorie durchaus lohnend sein kann; diskutieren wir jedoch zunächst in aller Kürze prägnante Vorzüge und Nachteile.

Gemäß dem Dispositionalismus hat ein Objekt oder Sachverhalt x einen bestimmten *Wert*, wenn x *geeignet* ist, eine bestimmte wertende Reaktion auszulösen (bzw. diese Reaktion mit *hoher Wahrscheinlichkeit* auslöst) bei Subjekten einer bestimmten Spezies, die eine bestimmte *stetige biologische oder natürliche Disposition* aufweisen. Stellen wir uns bspw. vor, wir (als Menschen) trinken Salzwasser. *Normalerweise* werden wir uns unvermittelt übergeben. Der Dispositionalismus gründet nun unsere Bewertung von Salzwasser als „ekelierend“ auf diese „normale“ oder „typische“ Reaktion von Subjekten der Spezies Mensch. Interessant ist hier, dass diese Reaktion auf Salzwasser tatsächlich biologisch adaptiv ist: das Trinken von Salzwasser ist gesundheitsschädlich für Menschen. In diesem Sinne können wir die damit verbundene Bewertung („Salzwasser ist ekelierend/ schlecht“) als „normativ gerechtfertigt“ ansehen. Ist das jedoch immer der Fall? Rein theoretisch wahrscheinlich nicht, da der Dispositionalismus zwei Ebenen (natürlich – axiologisch) miteinander verbindet, die prinzipiell von der Sein-Sollen-Barriere betroffen sind: „natürliche“ (im Sinne von normalen oder typischen) Reaktionen sind nicht *notwendig* auch normativ gerechtfertigt. Steinfath (2002) formuliert daher polemisch, dass gemäß dem Dispositionalismus Werte-Debatten durch empirische Erhebungen beigelegt werden könnten.

D’Arms & Jacobson (2000, 727) liefern eine lebendige Illustration dieses Punktes: Nehmen wir an, Milliarden von „normalen“ Schmeckern („normal tasters“) weltweit finden den BigMac *nicht* ekelierend. Und nehmen wir weiterhin an, dass eine einzige Person den BigMac hingegen ekelierend findet. Können wir nun die Behauptung dieser einen Person verwerfen in Hinblick auf die Milliarden andere Fälle, die sozusagen einen Gegen-Beweis liefern? Intuitiv würden wir wohl antworten, dass nicht, da die Debatte um den BigMac *prinzipiell* Sinn macht, selbst wenn die eine Person am Ende doch widerlegt werden sollte. Und weiterhin schließen wir prinzipiell auch nicht aus, dass in Fragen des Wertes und Geschmacks sogar Milliarden von Menschen letztlich Unrecht haben können. Nach D’Arms

den Vorzug des Dispositionalismus, auf bestimmte Konfigurationen des Subjekts abzuheben: Rationalität im Falle des informed-desire-Ansatzes, bestimmte Reiz-Reaktions-Schemate im Falle des Dispositionalismus.

und Jacobson zeigt dies eben den normativen Aspekt von Wertbegriffen wie „ekelerregend“. Der Dispositionalismus scheint nur die deskriptive Verwendung dieses Begriffs erfassen zu können. Allerdings verwenden wie diesen darüber hinaus auch, um normative Aussagen zu machen: „we use terms like ‘disgusting’ to do such things as criticize, persuade, or simply express disdain for others, and most generally to guide feelings – our own and other people’s“ (loc. cit.).

Wie könnte nun ein Anschluss an den Dispositionalismus erfolgen? Der Dispositionalismus – ebenso wie der Naturalismus und die Supervenienz-Theorie – versucht, Bewertungen auf natürliche Eigenschaften zurückzuführen. Der Vorzug besteht jedoch in der Einführung eines Subjekts als Träger bestimmter Reiz-Reaktions-Dispositionen. Es liegt nahe, emotionale Reaktionen als einen Typus solcher Reiz-Reaktions-Dispositionen zu verstehen. In meiner Kritik (vgl. Kap. 2.6.) der Theorie vom formalen Objekt habe ich argumentiert, dass Emotionen sinnvoller als Reaktionen auf bestimmte natürlich und ästhetische Eigenschaften der Außenwelt zu verstehen sind: Dunkelheit, schnelle Bewegungen, plötzliche Geräusche, bestimmte Konfigurationen eines Gesichts (gefletschte Zähne, in Falten gelegte Stirn etc.), Ausgesetztheit, große Höhe etc. Wir haben weiterhin die Problemhaftigkeit und leztliche Implausibilität einer strengen Trennung zwischen natürlichen und axiologischen Eigenschaften herausgestellt. Wenn wir dies als Lizenz benutzen dürfen, um auch Konstrukte und Begriffe mit minimal axiologischem Anteil als Basis-Eigenschaften für emotionale Reaktionen definieren zu können, dann erweitert sich dieses Spektrum noch einmal: wir können auch grundlegende Themen oder „Salienzmuster“ (DeSousa 1979) hinzurechnen wie Verlassenheit, Ablehnung, Geborgenheit, Zielerreichung und –frustration, Empathie etc. Zwar lehne ich weiterhin eine Definition von emotionalen Reaktionen nach Kategorien ihres „formalen Objekts“ ab. Angst ist zwar normalerweise eine Reaktion auf „das Gefährliche“, jedoch nicht immer.⁸⁰ Wenn wir Emotionen als Dispositionen zu Reaktionen auf bestimmte natürliche Eigenschaften einschließlich bestimmter grundlegender Themen, Schlüsselsituationen oder Salienzmuster verstehen, dann erweitern wir nicht nur die explanatorische Kraft der Theorie (wie im Laufe von Kap. 2 zu zeigen versucht wurde), sondern wir können auch plausibel machen, wie Emotionen Werte *begründen* können.

Es wird noch zu diskutieren sein, wie einerseits das Normativitätsproblem des Dispositionalismus gelöst werden kann, und wie dieses Modell andererseits „höherwertige“, d.h. transpersonale und kulturelle Werte erklären kann. An dieser Stelle will ich nur

⁸⁰ Denken wir wiederum an das Beispiel der „haunting fear“: Angst vor einem schon geschehenen Ereignis, das uns demnach nicht mehr imminently bedrohen kann.

anmerken, dass die Lösung für beides in der kognitiven Überformung von Emotionen liegt. Zwar gründen sich Werte in einer ontogenetischen frühen Entwicklungsphase auf die *Faktizität* von emotionalen Reaktionen. Diese Reaktionen sind jedoch in einer Art „protonormativ“, da Emotionen eine homöostatische und biologisch adaptive Funktion haben. Allerdings ist hier auch mit Atavismen zu rechnen, alten Reaktionsschemata also, die in heutigen Gesellschaften disruptiv wirken. Im frühen Erwachsenenalter jedoch nimmt der Vorrang der sekundären Emotionalität drastisch zu. Und die darin implizierte kognitive Modellierung unserer Emotionen kann ein Adäquanzkriterium zur Verfügung stellen, auf das wir einerseits eine voll ausgebildete Normativität gründen können, und das andererseits einen Erklärungsansatz für „höhere Werte“ liefern kann.

Im Folgenden soll also die Beziehung zwischen Emotionen und Werten genauer untersucht werden, wobei ich einerseits auf Ergebnisse aus Kap. 2 zurückgreife (gelegentlich jedoch wichtige Ergebnisse wiederhole oder in Hinblick auf eine Werttheorie zuspitze), andererseits auf bereits in diesem Kapitel erarbeitete Erkenntnisse des Zusammenhangs von Emotionen und Werten aufbaue.

3.8. Emotionen als die Grundlage von Werten

Wir haben an früherer Stelle ausgiebig die wichtigsten Emotionstheorien der Geschichte und Gegenwart diskutiert. Auch im Rahmen dieses Kapitels kamen diese Theorien gelegentlich zu Wort. So hat sich gezeigt, dass bspw. ein strenger Kognitivismus, eine Urteilstheorie also, eng mit dem belief-Ansatz in der Werttheorie zusammenhängt: Emotionen werden durch Werturteile ausgelöst, d.h. Werte sind die formalen Objekte von Emotionen. Eine solche Position kann zwar erklären, inwiefern Emotionen eine wertindikative Funktion haben, also gewissermaßen ein Sensorium für Werte sind, nicht jedoch, wie Emotionen bereits in die Praxis des Wertens selbst eingehen. Aufgrund dieses explanatorischen Nachteils, wie auch aus Gründen theorie-interner Widersprüche, habe ich an anderer Stelle einen rein kognitivistische Emotionstheorie abgelehnt.

Im Anschluss an Damasio und Zajonc wurde die theoretische Möglichkeit entwickelt, dass Emotionen durch einfache Wahrnehmungen von natürlichen und ästhetischen Eigenschaften ausgelöst werden können. Auf diese Ergebnisse sei verwiesen, wenn im Folgenden eine dispositionale Werttheorie erarbeitet wird, die Werten als eine Fähigkeit zu gerichteten emotionalen Reaktionen versteht. Die ebenfalls erarbeiteten theoretischen Modelle der kognitiven Analyse unserer Emotionsbegriffe wie auch der kognitiven Einflussnahme,

Kontrolle und Modellierung unserer emotionalen Reaktionen via unserer emotionalen Dispositionen wird im späteren Verlauf von Relevanz sein, wenn es darum gehen wird, die Normativität unserer Wertungen mittels der Adäquanz wertender emotionaler Reaktionen zu erklären.

An früherer Stelle in diesem Kapitel ist deutlich geworden, wie emotionaler Affekt den desire-Ansatz in einer Weise durchdringt, die es zweifelhaft erscheinen lässt, dass Wertungen tatsächlich auf desires beruhen oder nicht vielmehr auf der vermeintlichen Mediatorvariable des emotionalen Affekts. Fragen wir uns also zunächst: wenn Emotionen Werte begründen, kraft welches ihren Aspekts tun sie dies? Betrachten wir die verschiedenen Emotionstheorien im Überblick, so lässt sich zusammenfassend sagen, dass stets drei wichtige Aspekte oder Komponenten des Konstrukts oder Syndroms „Emotion“ herausgehoben werden: Kognitionen, Handlungstendenzen (bzw. Motivation oder „desire“) und phänomenal erlebte Affektivität.⁸¹ Wir können nun Kognitionen und desires als die werte-begründenden Komponenten ausschließen, da wir bereits früher in diesem Kapitel gesehen haben, dass Ansätze, die jeweils auf einem dieser Begriffe aufbauen, letztlich unterdeterminiert bleiben müssen. Emotionen hingegen zeichnen sich charakteristisch durch ihre Qualität des emotionalen Affekts aus, schließen jedoch die beiden anderen Komponenten im Sinne eines komplexen psycho-physischen Konstrukts mit ein.

Emotionen sollen also als eine Praxis des Wertens verstanden werden, die Werte kraft ihres emotionalen Affekts begründen. Interessanterweise teilen sowohl Affekt als auch Werteigenschaften jeweils eine für beide typische binäre Valenz-Skala von „positiv“ vs. „negativ“. Zunächst soll jedoch einem möglichen Missverständnis vorgebeugt werden. Einen emotionalen Zustand als Praxis des Wertens zu verstehen, bedeutet nicht, dass wir *überlegt und bewusst* Werturteile mittels unserer Emotionen zum Ausdruck bringen. Auch eine sprachliche Aussage wie „Äpfel sind lecker“ kann als eine Praxis des Wertens aufgefasst werden, insofern eine solche Aussage einen Wert *ausdrückt*. Eine Emotion jedoch ist eine Wertung insofern sie einen Wert *begründet bzw. konstituiert*. Die hier skizzierte Theorie emotionalen Wertens ist also weder emotivistisch noch realistisch zu verstehen: einerseits soll es nicht um die sprachlogische Analyse emotionaler Wertbegriffe gehen, sondern um die psychologische Genese von Werten durch das Erleben emotionaler Reaktionen. Andererseits

⁸¹ Von dem durchaus wichtigen Aspekt der physiologischen Aktivierung kann hier abstrahiert werden, da diese – phänomenal repräsentiert – die emotionstypische Affektivität darstellt. Physiologische Prozesse sind jedoch – wie wir an entsprechender Stelle gesehen haben – immens wichtig für eine exakte Beschreibung des Prozesses der Aktualgenese von Emotionen.

argumentiere ich, dass Werte erst in die Welt kommen durch Wesen, die einer Praxis des Wertens fähig sind. Emotionen konstituieren Werte kraft ihres Affekts, dem eine auch für Werte typische binäre Valenz eignet. Dennoch wird dadurch das Spektrum der Bedeutung unserer Wertprädikate nicht auf eine Befürworten-Ablehnen-Differenz reduziert: Emotionen generieren eine Pluralität von Werten, wie sich schon an der reichen, auf die Emotionsterminologie verweisenden Werte-Begrifflichkeit zeigt.

Betrachten wir im Folgenden einige Fallstudien, die besonders eindringlich illustrieren, in welchem Zusammenhang Werte mit emotionalem Affekt stehen. Erinnern wir uns zunächst an die erwähnte These von Tomkins, dass Emotionen (und damit auch emotionaler Affekt) Triebe verstärken und dämpfen. Dies trifft offenbar auch für verschiedene andere Körperzustände zu, wie wir ebenfalls bereits gesehen haben. Bestimmte Körperzustände wiederum bilden, vermutlich unabhängig vom jeweils vertretenen Ansatz in der Werttheorie, eine vermeintlich elementare Grundlage für Werte. Der negative Wert von Schmerz (und der entsprechen positive Wert der Schmerzvermeidung) scheint interkulturell, ja fast universell zu gelten. Egal welche Wert-Kataloge (z.B. in Kupperman 1999) man konsultiert, „Schmerz“ ist mit großer Wahrscheinlichkeit vertreten. Damasio liefert nun eine anekdotische Fallstudie eines Patienten, der von einer Trigeminus-Neuralgie betroffen war. Diese Krankheit, eine chronische Schmerzreizung innerhalb des Innervationsgebiets des Nervus trigeminus (eines der 12 Hirnnerven), erzeugt demnach chronische und äußerst extreme Schmerzattacken von variierender Dauer im Bereich der Gesichtsmuskeln und Gesichtshaut. Damasio, der zu dem Zeitpunkt in der Ausbildung beim Chirurgen Almeida Lima war, lernte den Patienten auf der Visite am Krankenbett kennen: „crouched in profound suffering, almost immobile“ (Damasio 1994, 265). Lima entschloss sich, eine Präfrontal-Leukotomie durchzuführen, eine Operation, bei der Teile des präfrontalen Kortex entfernt werden. Hierbei ist zu beachten, dass es sich dabei um eines der nach Damasio emotionsrelevanten Hirnareale handelt. Lima hoffte, auf diese Weise die chronischen Schmerzen des Patienten lindern zu können und sah diese radikale Operation durch die extreme Ausprägung der Neuralgie gerechtfertigt. Als Damasio den Patienten nach der Operation wiedersah, spielte er mit seinem Zimmernachbarn Karten und antwortete auf die Frage, wie es ihm gehe: „Oh, the pains are the same, but I feel fine now, thank you“ (ebd., 265f.). Dies könnte Anlass zur Verwirrung geben, wenn wir uns die vermeintliche Universalität von Schmerz als einer negativ werthaften Erfahrung vor Augen führen. Noch verwirrender mag die Aussage des Patienten sein, dass er offenbar den Schmerz noch als Nerveninformation „fühlen“ könne, dass der Schmerz jedoch nicht mehr „wehtue“.

Damasio unterscheidet ausgehend von dieser Beobachtung zwischen „Schmerz“ und „Leiden“, wobei letzteres definiert ist als die Verarbeitung einer „Schmerzlandschaft des Körpers“ *begleitet* von deren Gegenüberstellung mit emotionalen Bildern. Leiden hat stets eine negativ werthafte Tönung, Schmerz hat dies nicht notwendigerweise. Damasio resümiert seine eigenen Überlegungen zu dieser Fallstudie wie folgt:

“Clearly, what the operation seemed to have done, then, was abolish the emotional reaction that is part of what we call pain. It had ended the man’s suffering. [...] But the operation seemed to have done little to the image of local alteration in the body region supplied by the trigeminal nerve, and that is why the patient stated that the pains were the same. While the brain could no longer engender suffering, it was still making ‘images of pain,’ that is, processing normally the somatosensory mapping of a pain landscape.” (Damasio 1994, 266).

Was kann uns diese Fallstudie über einen Mechanismus emotionalen Wertens lehren? Halten wir zunächst fest, dass nicht der Schmerz selbst negativ werthaft getönt ist: der Patient „fühlte sich wohl“ nach seiner Operation, obwohl der Schmerz noch präsent war. Leiden jedoch unterscheidet sich von Schmerz insofern, als es von einer emotionalen Reaktion begleitet wird. Wir können also daraus schließen, dass nicht die negative Wertung auf den Schmerz an sich per neuralem „hard-wiring“ in uns angelegt ist, sondern dass die emotionale Reaktion, die im Normalfall stets die neurale Verarbeitung einer „Schmerzlandschaft“ begleitet, diese negative Wertung begründet. Es könnte eingewendet werden, dass jedoch selbst in Abwesenheit von Leiden, der Ursprung von Schmerzen, nämlich Verletzungen und Krankheit, noch immer als negativ werthaft angesehen würden. Dies ist sicherlich richtig, steht jedoch hier nicht primär zur Frage. Denn *bevor* eine Person in ihrer ontogenetischen Entwicklung überhaupt die kognitiven Fähigkeiten erlangt, um die Konsequenzen von Krankheiten und Verletzungen selbst durch Überlegungen mit einem Wert zu verbinden, hat dieselbe Person längst die Assoziation von Schmerz und Leiden, und damit auch den negativen Wert von Schmerz (per Leiden) gelernt. Die affektiv begründete Wertung ist primär, und diese affektive Grundlage wird selbst dann bestehen bleiben, wenn wir auch auf anderen Wegen (durch Überlegungen oder dialektische Operationen) Werte begründen können. Wahrscheinlich können wir auf ontogenetisch späteren Entwicklungsstufen auch rational begründete Werte sinnvoll vertreten und unterhalten. Wir *erleben* sie jedoch sicherlich nur dann jeweils *als* einen Wert mit akuter Bedeutung, wenn auch eine affektiv getragene Bewertung im Hintergrund steht oder assoziiert ist.

Betrachten wir noch ein weiteres Experiment aus der Hirnforschung, das Aufschlüsse darüber liefert, wie Werte gebildet und verarbeitet werden können, ohne dass bewusste kognitive Prozesse dabei involviert sind. Dennoch scheinen auch solche Wertungen wiederum

in bewusste kognitive Prozesse (der Entscheidungsfindung oder des Überlegens) mit einzugehen. John Marshall und seine Mitarbeiter führten Experimente mit sog. Neglect-Patienten durch. Dabei handelt es sich um Patienten, die an einer besonderen Form der „Blindheit“ in einem der beiden visuellen Felder leiden (vgl. Mashall 1988).⁸² Diese Form von „Blindheit“ wurde durch Weiskrantz (1986), der die ersten Beobachtungen an betroffenen Patienten machte, mit dem Begriff „blindsight“ belegt (vgl. auch Kentridge & Heywood 1999). Blindsight, bzw. Neglect, beruht nicht auf einem organischen Defekt der Sinnesorgane. Die Patienten sind – wie man folgern muss – noch immer in der Lage, Informationen auch im betroffenen visuellen Feld zu registrieren und zu verarbeiten. Weiskrantz führte als Nachweis Wiedererkennungstest von Stimuli (typographische Symbole wie „+“), die in das betroffene visuelle Feld projiziert wurden, durch und konnte zeigen, dass Neglect-Patienten dabei signifikant über Zufallsniveau abschneiden. Die Patienten müssen also in irgendeiner Weise die Stimuli „wahrgenommen“ haben, ohne dass sie sich dessen bewusst wurden.

Marshall führte nun ein ähnliches Experiment durch, in dem er jedoch statt der von Weiskrantz benutzen neutralen typographischen Symbole emotional relevante Stimuli einsetzte. Im Marshall-Experiment werden dem Probanden zwei Blätter Papier vorgelegt, auf denen jeweils die schematische Zeichnung eines Hauses zu erkennen ist. Die Zeichnungen sind so zentriert, dass exakt die Hälften der beiden Häuser in das jeweils von Neglect betroffene visuelle Feld fällt. Die zweite Zeichnung enthält nun auf der Hälfte, die in das von Neglect betroffene visuelle Feld fällt, zusätzlich einen Stimulus: aus dem Dachstuhl des Hauses stoßen auf dieser betreffenden Seite Flammen hervor. Während der Proband nun die beiden Blätter nacheinander fokussiert, soll er gleichzeitig auf einem separaten Blatt zeichnen, was er visuell wahrnimmt. In beiden Fällen zeichnet der Proband jeweils die Hälfte des Hauses, die *nicht* in das von Neglect betroffene visuelle Feld fällt: eine Doppelhaushälfte sozusagen. Da der Feuer-Stimulus von den Experimentatoren bewusst in das von Neglect betroffene visuelle Feld gelegt wurde, taucht er in der Reproduktion der Zeichnung natürlich nicht auf. Der Proband wird nun gefragt, in welchem der beiden Häuser er lieber wohnen würde und antwortet spontan mit dem ersten (d.h. demjenigen *ohne* den Feuer-Stimulus). Auf die Bitte, dafür Gründe anzugeben, kann der Proband nichts Wesentliches antworten.

Durch die Reproduktionsaufgabe (Abzeichnen des Wahrgenommenen), in der der Feuer-Stimulus nicht mehr auftrat, können wir ausschließen, dass dieser Stimulus bewusst

⁸² Ich bedanke mich bei Prof. Dr. Karl Zilles für seinen Hinweis auf diese Studie während eines Vortrags auf der Konferenz „Dimensionen der Person: Genom und Gehirn“, im Wintersemester 2004/05 organisiert vom Institut für Wissenschaft und Ethik, Bonn.

wahrgenommen wurde. Und dennoch scheint er sich signifikant auf die Entscheidungsfindung des Probanden auszuwirken. Nehmen wir an, dass Tendenzen, sich für die eine oder andere Option zu entscheiden, auf zu Grunde liegende Werte verweisen, dann können wir folgern, dass der Proband offenbar einen negativen Wert mit dem brennenden Haus verbunden hat, und einen dementsprechend vergleichbar positiven Wert mit dem anderen Haus. Die verdeckte Prämisse des Experiments ist, dass „Feuer“ eine universelle Wertmarkierung bildet. Und tatsächlich scheint der Proband seine Entscheidung auf Grund der einzigen distinkten Information gefällt zu haben, durch die sich beide Zeichnungen unterscheiden, ohne dass diese – eben jene Wertmarkierung – bewusst wahrgenommen oder verarbeitet wurde.

Menschliche Akteure stellen offenbar kompetente und wirksame Bewertungen an, ohne dass höhere (bewusste, rational reflektierte oder propositionale) kognitive Prozesse daran beteiligt wären. Damit ist Einiges zu Gunsten eines Mechanismus emotionalen Wertens gesagt, der schon bei Primäremotionen aktiv ist, die sich als einfache Reiz-Reaktions-Schemata lediglich durch ihre typische Affektivität auszeichnen. Allerdings ist damit *nicht* zugleich ein Argument dagegen verbunden, dass wir auch auf alternativen Wegen Wertungen anstellen, bewusst und überlegt bspw. Die Möglichkeit, dass wir Werte z.B. auch durch rationale Überlegen bilden, bleibt bestehen. Allerdings sind zwei Anmerkungen zu machen: einerseits könnten wir wissenschaftstheoretisch argumentieren. Wenn wir menschliches Denken und Handeln erschöpfend durch diesen primitiven Mechanismus des Wertens erklären können, gleichzeitig jedoch noch ein komplexeres Erklärungsmodell desselben Phänomens zur Verfügung steht, dann könnte letzteres mit Hinweis auf „Ockhams Rasiermesser“ verworfen werden. Ich würde mich dieser Position jedoch nicht anschließen.

Andererseits könnten wir mit folgender Strategie argumentieren: wenn zwei verschiedene, jedoch nicht notwendig kontradiktorische Erklärungsmodelle eines Phänomens zur Verfügung stehen, von denen eines lediglich primitive mentale Fähigkeiten und Prozesse erfordert, das andere jedoch höhere mentale (bewusste, kognitive, propositional strukturierte) Prozesse, dann könnte daraus gefolgert werden, dass das erste Modell einen ontogenetisch primären Prozess darstellt, wohingegen das zweite Modell diesen Prozess auf einer späteren Entwicklungsstufe eines Individuums abbildet.

Die Individualgenese von „Wert“ lässt sich also als eine Praxis des emotionalen Wertens auffassen, die offenbar keine höheren kognitiven Prozesse oder Kapazitäten erfordert. „Werten“ (in der Adjektivform) wird also als aktuelle oder dispositionale emotionale Reaktivität aufgefasst. Steht jedoch die Beschreibung dieses Mechanismus einer Praxis des emotionalen Wertens in Einklang mit den aktuellen Erkenntnissen zum Mechanismus der

Emotionselizitation? Mit anderen Worten: wenn wir uns „Werten“ als emotionale Reaktivität vorstellen, kann der unterliegende Mechanismus der Emotionselizitation überhaupt die Bedingungen erfüllen, die wir für den Mechanismus emotionalen Wertens postuliert haben? Ich denke, dass ein Blick auf unsere Ergebnisse aus Kap. 2 diese Frage mit ja bescheiden würde. Dennoch möchte ich eine m.E. hochinteressante Illustration aus dem Bereich der Emotionsforschung anführen.

In einer jüngeren Studie (LeDoux 1996, Kap. 1) diskutiert und interpretiert Joseph LeDoux experimentelle Ergebnisse, die zu einem früheren Zeitpunkt mit seinem damaligen Promotionsbetreuer entstanden sind (Gazzaniga & LeDoux 1978; Gazzaniga 1972). Gazzaniga & LeDoux führten eine Studie an sog. „split-brain“ Patienten durch, Patienten also, bei denen die Nervenverbindungen (das corpus callosum) zwischen den beiden Hirnhälften vollständig durchtrennt worden war. Abgesehen von sehr alten Strukturen bspw. des Hirnstamms trennt (räumlich) und verbindet (synaptisch) das corpus callosum das menschliche Gehirn in einem Längsschnitt entlang des Scheitels von frontal nach okzipital. LeDoux beschreibt eine Studie an P.S., einem der Patienten aus der split-brain Gruppe. Gazzaniga & LeDoux boten emotional relevante Stimuli (wie „Mom“ und „Devil“) abwechselnd in das rechte und linke visuelle Feld von P.S. dar. Die verdeckte (aber empirisch bestätigte) Prämisse, die dem Experiment zu Grunde liegt, ist, dass bei split-brain Patienten Informationen, die nur in einer der beiden Hirnhälften repräsentiert sind (bspw. durch Darbietung eines Stimulus ausschließlich in einem der beiden visuellen Felder), nicht in die jeweils andere übergehen kann. Diejenigen Strukturen (die Großhirnhälften), die durch das corpus callosum räumlich getrennt und synaptisch verbunden werden, können nach Durchtrennung des corpus callosum nicht mehr miteinander kommunizieren. Erinnern wir uns nun, dass die Sehnerven über Kreuz verlaufen und im sog. chiasma opticum die Seite wechseln: sie sind jeweils in der zu ihrem Ursprung im Auge contralateralen Hirnhälfte verschaltet. Ein Reiz, der im linken visuellen Feld (LVF) dargeboten wird, wird also in der rechten Hemisphäre repräsentiert, ein Reiz, der im rechten visuellen Feld (RVF) dargeboten wird, wird demnach in der linken Hemisphäre verschaltet.

Wenn Gazzaniga & LeDoux einen Reiz im RVF darboten, war P.S. in der Lage, auf Aufforderung sowohl zu verbalisieren, was er wahrnahm, als auch anzugeben, welche Art von emotionalem Gefühl (auf einer positiv-negativ-Skala) der Reiz in ihm auslöste. Dieses Ergebnis entspricht den Erwartungen: das Sprachzentrum befindet sich in der linken Hemisphäre, also dort, wo visuelle Reize aus dem RVF repräsentiert werden. P.S. sollte also in der Tat in der Lage sein, die entsprechenden Reize zu verbalisieren. Allerdings würden wir

aus demselben Grund erwarten, dass dies nicht mit Reizen funktioniert, die im LVF dargeboten werden, da diese in der rechten Hemisphäre repräsentiert werden – außerhalb der Reichweite des Sprachzentrums sozusagen.⁸³ Tatsächlich bestätigten die Ergebnisse auch in diesem Fall die Erwartungen: wenn Reize im LVF dargeboten wurden, war P.S. nicht in der Lage, zu verbalisieren, was er wahrnahm. Erstaunlicherweise jedoch war er sehr wohl in der Lage, eine *Bewertung* dieses Reizes anzustellen *und* dieselbe kompetent zu *verbalisieren*, obwohl er keinerlei Bewusstsein darüber hatte, um *was* es sich bei dem entsprechenden Reiz überhaupt handelte. Ich denke, dass durch diese neuerlichen empirischen Ergebnisse die Richtigkeit meiner an früherer Stelle angestellten theoretischen Überlegungen bestätigt wird. Die Art des Prozesses der Emotionselizitation, die wir für das hier zu erarbeitende Modell emotionalen Wertens unterstellen müssen, scheint realistisch möglich zu sein, evtl. sogar einen Standardweg der Emotionselizitation darzustellen. Der Mechanismus der Aktualgenese der Emotionen erfordert offenbar nicht notwendig propositional strukturiertes Wissen oder die überlegte und bewusste Verarbeitung von Informationen. Alle diese Aspekte können zwar – und sind de facto – involviert sein im Mechanismus der Sekundäremotionen. Der Mechanismus der Primäremotionen jedoch läuft – wie wir an früherer Stelle gesehen haben – spontan ab, oder mit Damasio (2001, 102) Worten: „The devices which produce emotions [...] can be engaged automatically, without conscious deliberation“.

3.9. Individuelle emotionale Reaktionen als eine Praxis des Wertens

Wir haben versucht, ein Modell der Wertgenese durch individuelle emotionale Reaktionen bzw. durch die Disposition zu solchen Reaktionen zu entwickeln. Ich stimme also Steinfaths (2002) Vermutung zu, dass Werte durch Emotionen begründet werden, u.z. durch die Praxis individueller emotionaler Reaktionen. Wir haben weiterhin gesehen, dass emotionaler Affekt die zentrale Rolle in diesem Modell spielt: kraft seiner erlangen Werte die – auch für Emotionen typische – positiv-negativ-Valenz. Weiterhin wurde argumentiert, u.z. im Sinne eines *a-fortiori*-Arguments, dass selbst primitive (d.h. primäre) emotionale Reaktionen Werte begründen können, und dass dieser Mechanismus die ontogenetische Grundlage für evtl. komplementäre Mechanismen der Wertgenese durch rationale Überlegung darstellt. Dabei

⁸³ Es ist zwar bekannt, dass das Gehirn hochadaptiv ist und daher gelegentlich funktionale Bereiche, die lädiert und deaktiviert wurden, in anderen anatomischen Arealen des Gehirns neu anlegt. Diese Adaptivität ist im jungen Alter am höchsten und nimmt im späteren Alter ab. Offenbar wurde das Sprachzentrum von P.S. nicht funktional in der rechten Hemisphäre redupliziert.

bleibt jedoch Emotionalität – im Sinne eines angemessenen Modus, Werte authentisch zu vertreten und zu unterhalten – stets konstitutiv beteiligt: rationale Wertungen allein lassen, wie wir gesehen haben, den Wertbegriff unterdeterminiert. Ich denke, dass diese verschiedenen Grade, in denen jeweils emotionale und kognitive Anteile am Mechanismus der Wertgenese beteiligt sind, anschaulich durch Damasio's Modell verschiedener „Stufen“ der Emotionalität abgebildet werden: von primären Emotionen in einem frühen Lebensalter, über sekundäre Emotionen im Erwachsenenalter bis hin zu „als-ob“-Emotionen als eine Art Derivatform der Emotionen oder „Quasi-Emotionalität“. Die verschiedenen Konzepte sollen hier nicht mehr dargestellt werden, für weitergehende Informationen sei auf das entsprechende Kapitel dieser Arbeit verwiesen. Betrachten wir jedoch diese Ergebnisse in einem Überblick, so wird deutlich, dass es sich bei dem erarbeiteten Modell um einen *dispositionalistischen* und einen *subjektivistischen* Ansatz in der Erklärung des Wertens handelt.

Der Ansatz ist *dispositionalistisch*, da wir nach Damasio mit Primäremotionen dispositional auf bestimmte natürliche Merkmale von Dingen oder Sachverhalten reagieren. Diese unterliegenden Dispositionen sind entweder angeboren („hard-wired“) oder durch eine Konditionierung erworben. Die verhältnismäßige Invarianz des Wertens auf dieser primitiven Stufe lässt sich vermutlich einerseits durch die zerebrale und physiologische Architektur erklären, die wir als Menschen mit allen anderen Menschen teilen, andererseits durch ähnliche Sozialisationserfahrungen.⁸⁴ Und in der Tat sind die natürlichen Merkmale, auf die wir mit Angst, Freude, Trauer oder Zorn reagieren, relativ konstant (dank vergleichbarer kognitiver und physiologischer Charakteristika); leichte Varianzen ergeben sich wiederum im Fall von Ekel (der ebenfalls als eine Primäremotion gilt, vgl. Damasio 2001, 102), der offenbar stärker kulturell modelliert ist. Diese Varianzen jedoch betreffen Randbereiche. Der Kernbereich der zu Grunde liegenden natürlichen Eigenschaften wiederum ist relativ konstant: ich möchte behaupten, dass alle Menschen mit Ekel auf das Trinken von Salzwasser oder das Essen von verdorbenen Speisen reagieren.

Ein dispositionalistischer Ansatz hat – wie wir gesehen haben – zahlreiche Vorzüge: einerseits werden die Fallstricke des Realismus vermieden, andererseits kann – im

⁸⁴ Die Tatsache, dass durch den Verweis auf die Vergleichbarkeit von Sozialisationserfahrungen gleichzeitig auch bestimmte regionale oder sozio-geografische Varianzen in der emotionalen Wertung Platz in der Theorie finden, halte ich eher für einen Vorzug als einen Nachteil des Theoriemodells: Es ist weithin bekannt, dass solche interkulturellen Unterschiede in der Emotionalität bestehen, sodass eine Theorie emotionalen Wertens m.E. auch die Aufgabe hätte, diese theoretisch adäquat abbilden zu können.

Unterschied zum Naturalismus – die Supervenienz von Wert-Eigenschaften auf natürliche Eigenschaften im Sinne einer „Abhängigkeit ohne Reduktionismus“ erklärt werden. Weiterhin bezieht der Dispositionalismus – und dies halte ich für einen gewissen Vorzug – das wertende Subjekt ein, insofern keine *begriffliche* Supervenienz konstatiert, sondern die Supervenienzbeziehung auf ein *Tun* des wertenden Subjekts gegründet wird. Die Beziehung von Wert-Eigenschaften zu bestimmten natürlichen Eigenschaften ergibt sich nicht einfach begriffslogisch, sondern wird durch eine Praxis bzw. ein *Tun* von *wertenden Lebewesen* überhaupt erst begründet. Dem großen Defizit des Dispositionalismus, der zwar die Faktizität, nicht jedoch die Normativität von Werten erklären kann, ist wie folgt zu begegnen: der dispositionalistische Mechanismus des Wertens wird auf einer ontogenetisch späteren Entwicklungsstufe durch alternative Formen des Wertens *ergänzt* (*nicht* ersetzt). Sobald wir als Individuen anfangen, unsere emotionalen Reaktionen kritisch zu reflektieren, werden damit per se auch unsere emotionalen Wertungen einer kritischen Prüfung unterzogen. Normativität kann nun gewährleistet werden, indem wir annehmen können, dass geprüfte und akzeptierte Wertungen als solche belassen werden. Wertungen wiederum, die uns als inadäquat erscheinen, würden modifiziert oder unterdrückt. Es ist also anzunehmen, dass eine Kontrolle unserer Emotionalität – im Sinne einer dispositionalen Verantwortung für unsere Emotionen – besteht, wie wir einerseits bereits in Kap. 2 gesehen haben, andererseits im Verlaufe des vorliegenden Kapitels noch eingehender diskutieren werden.

Der Erklärungsansatz ist wiederum *subjektivistisch* in dreierlei Hinsicht (vgl. Steinfath 2001a, 163): (1.) ob etwas als angenehmen, langweilig, fürchterlich etc. bewertet wird (d.h. mit einem spezifischen Wertprädikat belegt wird), hängt *ontologisch* von unseren Emotionen ab. Ohne die Emotion „Furcht“ gäbe es demnach auch nichts „Fürchterliches“ in der Welt. (2.) Emotionen verweisen auf ein erlebendes Subjekt, auf dessen soziales Leben und weltliches Dasein. Agnes Heller betont in ihrer Studie diese Tatsache, dass Emotionen eine Weise sind, in der ein Subjekt persönlich in die Welt – v.a. als soziale Welt – *involviert* ist: Fühlen bedeutet „In-etwas-Involviertsein“ (Heller 1981, 16). Diesen Punkt hatte wir auch in der empirischen Psychologie vorgefunden: erinnern wir uns an Lazarus' Theorie, dass der Ursprung von Emotionen in sozialen „Transaktionen“ bestehe (Lazarus 1984). Diese Transaktionen schließen sechs sog. „appraisal dimensions“ ein, die sich auf die Ziele des entsprechenden Individuums beziehen, auf sein soziales Leben, coping-Optionen und Perspektiven für Zukunftsentwicklungen (Lazarus 1991). Eine emotionale Wertung bewertet also zunächst etwas als gut oder schlecht *für mich*. (3.) Emotionale Wertungen sind subjektiv

insofern sie *individuell* sind: sie gründen auf den emotionalen Reaktionen eines Individuums. Damit ist zugleich auch gesagt, dass eine gegebene Wertung nicht notwendig auch für andere Personen zutrifft. Dies wird gelegentlich vorschnell als viziöser Aspekt ausgezeichnet. Ich denke jedoch, dass diese Ansicht auf einer begrifflichen Amalgamierung von „Wert“ mit „moralischem Wert“ oder „objektivem Wert“ beruht. Steinfath (2001a) unterscheidet zwischen „subjektivem“ und „objektivem“ Wert, in analoger Weise unterscheidet Helm (1994 und 2002) zwischen „import“ („Bedeutung“) und „value“ („Wert“). Hier stehen wir vor dem gleichen Einwand, dem sich der Dispositionalismus ausgesetzt sieht, nämlich dass das erarbeitete Modell nicht die Normativität von Wertungen (hier aufgefasst im Sinne einer transpersonalen oder objektiven Geltung) erklären könne. Im Falle des Dispositionalismus-Vorwurfs waren wir dem Einwand mit dem Mechanismus eines „reflective endorsement“ begegnet. An späterer Stelle wollen wir diskutieren, wie wir von der Subjektivität von Werten zu einer normativen transpersonalen Geltung gelangen können. An dieser Stelle sei lediglich gesagt, dass der erarbeitete Mechanismus emotionalen Wertens, wie er hier dargestellt wird, einen ontogenetisch frühen Mechanismus darstellt. Die bereits angesprochene Ergänzung dieses Mechanismus soll an späterer Stelle diskutiert werden.

Halten wir an dieser Stelle nur fest, dass emotionaler Affekt auch in Fällen transpersonaler Wertungen eine zentrale Rolle spielen wird. Ein möglicher theoretischer Erklärungsmechanismus steht dazu mit Damasio's „als-ob“ Emotionen bereit. In diesem Stadium der emotionalen Entwicklung eines Individuums wird emotionaler Affekt unabhängig von der peripher-physiologischen Grundlage des Körpers und kann sich mit jedem beliebigen mentalen Phänomen assoziieren. Die Affektivität von „als-ob“ Emotionen kann sich auf Gedanken, Überzeugungen, Urteile, Pläne, Projekte, Absichten, Überlegungsstrategien, moralische Werte und Normen beziehen. Anhand seiner „somatic marker hypothesis“ (Damasio 1994, Kap. 8) hat Damasio nachgewiesen, dass wir diese affektive Tönung unseres Denkens und Bewusstseins voraussetzen müssen, um Aufmerksamkeit, zielgerichtetes Denken und Überlegen und Handeln von menschlichen Akteuren überhaupt erklären zu können.

Ich möchte jedoch noch in aller Kürze illustrieren, warum ich die Subjektivität emotionalen Wertens zunächst nicht für einen viziösen Aspekt halte, ja vielmehr denke, dass diese Subjektivität die explanatorische Kraft einer Werttheorie sogar erweitern kann. Nach Steinfath (2001, 121) sind Emotionen irreduzibel private Phänomene. Ideen, Überzeugungen und Meinungen können kommuniziert werden, wir können sie „aus zweiter Hand“ erhalten, ohne dass sich dadurch ihr ontologischer Status ändert, bzw. ohne dass sie dadurch weniger

„authentisch“ würden. Wenn Hans mir mitteilt, „dass x zutrifft“, und ich ihm glaube kraft eines überzeugenden Beweises oder kraft Hans' unleugbarer Autorität, dann wird die Überzeugung „dass x zutrifft“ zu meiner eigenen und unterscheidet sich hinsichtlich ihres ontologischen Status als in meinem Geiste subsistierend nicht von anderen Überzeugungen, die ich vertrete: sie ist und bleibt eine *Überzeugung*, wenngleich ich evtl. eher bereit bin, eine auf Autorität begründete Überzeugung aufzugeben, als eine auf Augenzeugenschaft gegründete. Emotionen können nicht auf diese Weise kommuniziert oder „weitergegeben“ werden, ja sie können oft kaum adäquat verbalisiert werden. Eine verbalisierte oder kommunizierte Emotion hat einen fundamental anderen ontologischen Status als eine *erlebte* Emotion.

Gaus (1996) weist auf eine Parallele zwischen diesem Erlebnisaspekt und Werten hin. Stellen wir uns vor, dass Hans von einer kompetenten Autorität darüber informiert wurde, dass der Witz J lustig ist. Hans vertritt nun dieses Werturteil („dass J lustig ist“) auf einer zwar nicht idealen, aber empirisch akzeptablen Grundlage. Hermann hingegen hat J angehört und über den Witz gelacht. Auch Hermann denkt, „dass J lustig ist“. Wenn wir dieses Beispiel betrachten, dann scheint Hans Wertüberzeugung defizient zu sein. Hans hat zwar eine Überzeugung, aber nicht den damit zusammenhängenden mentalen Zustand, Hermann hat beides. Offenbar scheint eine subjektivistische Theorie den Mechanismus des Wertens sogar wesentlich adäquater erklären zu können als eine nicht-subjektivistische. Erinnern wir uns an Midgleys Aussage, wir könnten nicht wirklich denken, dass Ungerechtigkeit schlecht sei, wenn sie uns nicht auch irgendwie „krank machen“ würde. Wir denken tatsächlich auch intuitiv, dass wir uns Wertungen in irgendeiner Weise „zu eigen machen“ müssen, damit wir sie authentisch vertreten. Ein Modell emotionaler Wertungen hat kein Problem mit der Erklärung dieser Tatsache.

3.10. Die Intentionalität emotionalen Wertens

Bevor wir das Kapitel mit einer Diskussion der Möglichkeit schließen, wie transpersonaler Wert als begründet auf emotionalen Wertungen verstanden werden kann, möchte ich mich einem Problem zuwenden, das bisher noch nicht zur Sprache gekommen ist. „Werten“ ist eine *intentionale* Praxis: wir bewerten etwas *als* etwas (einen Apfel als lecker). Inwiefern kommt Emotionen – v.a. primären Emotionen – Intentionalität zu? Im Falle von sekundären Emotionen ist diese Frage nicht allzu virulent, da wir im Rahmen eines weitgehend kognitivistischen Paradigmas lediglich auf den kognitiven Gehalt solcher Emotionen

hinweisen müssten. Es kann jedoch nicht befriedigend sein, die Praxis des Wertens auf emotionale Reaktionen insgesamt zu gründen, ohne dass die Möglichkeit der Intentionalität auch für primäre Emotionen nachgewiesen werden kann. Ich denke, dass diese Möglichkeit besteht. Wir haben im Laufe dieser Arbeit festgestellt, dass primäre Emotionen lediglich durch Sinneswahrnehmungen ausgelöst werden können. Unsere Aufgabe bei der intentionalen Begründung der Praxis emotionalen Wertens liegt also darin, nachzuweisen, dass Sinneswahrnehmungen über eine zumindest primitive Intentionalität verfügen. Gelingt uns dies, so können wir *a fortiori* schließen, dass auch primäre Emotionen intentional in diesem Sinne sind. Daraus wiederum ließe sich schließen, dass auch komplexere oder höherstufige Emotionen intentional sein müssen: zumindest schon auf Grund ihrer ontologischen Abhängigkeit von primären Emotionen, in jedem Fall aber kraft ihres zusätzlichen kognitiven Anteils.

„Werten“ bezieht sich erstens auf ein binäres Spektrum positiver oder negativer Valenz und beschreibt zweitens eine intentionale Praxis. Steinfath (2001, 152ff.) fragt sich nun, welche Phänomene die beiden obigen Bedingungen erfüllen und antwortet: (a) Sinneswahrnehmung (deren Valenz etwas als *angenehm* oder *unangenehm* auszeichnet), (b) Stimmungen (die uns die Welt als *erfüllt* oder *leer* erfahren lassen) und (c) Emotionen (die eine *positive* oder *negative* affektive Qualität einschließen). Lassen wir also entsprechend unserer Eingangsbemerkung aus dem letzten Abschnitt (b) und (c) bei Seite und betrachten wir (a).

Sinneswahrnehmungen – insofern diese eine hedonische Tönung einschließen – involvieren die Bewertung von etwas als entweder angenehm oder unangenehm (womit wir uns auf das Kriterium der binären Valenz beziehen). Wir können „angenehm“ und „unangenehm“ als *weite* (oder allgemeine) axiologische Prädikate auffassen, analog zu „gut“ und „schlecht“. Dieser hedonisch Aspekt wird schon durch die Erfahrung einer Sinneswahrnehmung selbst gegeben und ist daher nicht auf Wertungen durch rationale Lebewesen beschränkt, auch Tiere können vermutlich solche Wertungen anstellen. Der Unterschied ist hierbei, dass Menschen diese Erfahrungen verbalisieren und über sie reflektieren können. Damit werden wir bereits zum zentralen Problem dieses Typs von Wertungen geführt. Denn die Verbalisierungen der Wertungen durch Sinneserfahrungen sind lediglich *expressive* Aussagen. Diese Aussagen können strenggenommen keine Behauptung hinsichtlich des entsprechenden *Objekts* machen. Sie können bspw. nicht behaupten, dass das Objekt bestimmte Eigenschaften habe, oder dass es bestimmte Empfindungen unter bestimmten Bedingungen auslöst. Solche Wertungen sind allein auf die entsprechende

Erfahrung gegründet, und um es genauer zu sagen: auf den hedonischen Aspekt dieser Erfahrung. Daher kann die Verbalisierung einer solchen Erfahrung selbst keine Bewertung sein, sondern lediglich eine linguistische Referenz auf einen Erlebniszustand. Mit den Worten Steinfaths (ebd.) gesagt: solche Verbalisierungen sind „Artikulationen der Wertungen, die schon in den jeweiligen sinnlichen Empfindungen selbst liegen“.

Die Intentionalität einer solchen Wertung – im Sinne der Beziehung zwischen der Sinneserfahrung und ihrem Objekt – scheint kompromittiert zu sein. Tatsächlich scheint die *Sinneserfahrung selbst* das Objekt der entsprechenden Wertung zu sein, und nicht etwa das *Objekt dieser Sinneserfahrung*. Und selbst wenn die Objekte von Sinneserfahrungen natürlich kausal mit den entsprechenden Sinneserfahrungen verbunden sind, so können wir diese Objekte selbst offenbar nur in einem *abgeleiteten* Sinne als „angenehm“ oder „unangenehm“ bezeichnen. Ist damit jedoch jegliche intentionale Beziehung zum externen (kausalen) Objekt ausgeschlossen? Ich denke, es gibt zwei Wege, tatsächlich eine Unterscheidung zu treffen zwischen der Haltung des Bewertens und dem bewerteten Objekt.

(1.) Wir können behaupten, dass Sinneserfahrungen intentional sind in dem Sinne, in dem Brentano (1874, Bd. I, 120-137) von der Intentionalität als dem „Zeichen des Psychologischen“ gesprochen hat (vgl. in diesem Sinne auch Crane 1998 und Goldie 2002). Die Klasse der psychologischen (d.h. mentalen) Prozesse und Zustände zeichnet sich demnach gerade dadurch aus, dass ihr Intentionalität eignet. Auch Steinfath nimmt an, dass Sinneswahrnehmungen intentional sind, insofern sie unseren erlebten Körperzustand repräsentieren. Eine ähnliche Idee hatten wir mit Damasio's Theorie der mentalen Verarbeitung von „Körperlandschaften“ angetroffen. Der übliche Einwand seitens kognitivistischer Theorien weist darauf hin, dass bspw. „Furcht“ (als ein intentionaler Zustand) sich *auf etwas* bezieht, so wie Furcht vor der Dunkelheit sich eben auf die Dunkelheit bezieht. Schmerz jedoch (als ein nicht-intentionaler Zustand) bezieht sich angeblich *nicht* auf irgendetwas. Ich denke jedoch, dass diese Ansicht leicht durch einen Hinweis auf medizinische Phänomene wie *Phantomschmerzen* widerlegt werden kann. Dieser Begriff bezeichnet die Erfahrung von Schmerzen in einer nicht mehr vorhandenen Extremität. Vergessen wir dabei nicht – und dies macht der analoge englische Begriff „phantom limbs“ deutlich – dass kraft einer Schmerzempfindung in dieser Erfahrung natürlich auch ein intentionales Objekt mit repräsentiert sein muss, nämlich das intentionale Objekt der jeweils betroffenen Extremität. Stellen wir uns einen Patienten mit einer amputierten Hand vor, der Schmerzen in dieser (fehlenden) Hand empfindet. Damit scheint ein starker empirischer Beweis verbunden zu sein, dass selbst einfach Sinnesempfindungen (wie Schmerz) die

intentionale Repräsentation ihres Objekts (hier: der Hand) mit einschließen. In dieser Hinsicht ist wiederum interessant, dass es aus der physiologischen Sicht der Innervierung (d.h. der Lage von Nervenknötchen und des Verlaufs von Nervenbahnen) überhaupt kein *reales Objekt* „Hand“ auszumachen ist.⁸⁵ Da wir jedoch unseren eigenen Körper nur durch nervös vermittelte Körperbilder erfahren können, ist also fraglich, warum wir offenbar wie selbstverständlich die anatomische Einheit „Hand“, die jedoch keinen eigenen, speziell innervierten Bereich darstellt, *als* eine eigene Einheit *erleben*. Wir müssen also annehmen, dass wir im Laufe unserer ontogenetischen Entwicklung ein intentionales Objekt der Hand *konstruiert* haben, damit wir überhaupt in der Lage sind, Repräsentationen einer Hand (bspw. die Repräsentation eines haptischen Stimulus auf der Hand oder die Repräsentation eines Schmerzes in der Hand) bilden können. In dieser Linie hat auch Deigh (1994) für ein weiteres Verständnis des Intentionalitätsbegriffs im Sinne von „Gerichtetheit“ argumentiert.

(2.) Wir können darüber hinaus auch eine Unterscheidung treffen zwischen einem Aspekt des „Wertens“ und einem Aspekt des „Bewertet-Werdens“ innerhalb einer gegebenen Sinneserfahrung selbst. Am Beispiel von „Schmerz“ illustriert, könnten wir zwischen der Qualität der Schmerzhaftigkeit einerseits und der negativen Bewertung dieser Schmerzhaftigkeit andererseits unterscheiden. Letztere wäre dann dafür verantwortlich, dass uns Schmerzen Ungemach bereiten, und dass letztlich „Schmerz“ für uns einen negativen Wert darstellt. Damit wäre jedoch die Frage aufgeworfen, ob diese negative affektive Reaktion auf Schmerz der Erfahrung von Schmerz analytisch intrinsisch ist, oder ob sie nur kontingent (wenn auch mit hoher Korrelation) mit dieser assoziiert ist. Hier möchte ich auf Damasio's Fallstudie des bereits angesprochenen Trigeminal-Neuralgie-Patienten hinweisen. Erinnern wir uns, dass dieser nach seiner Leukotomie-Operation zwar noch denselben Schmerz empfinden konnte, dass dieser jedoch nicht mehr „wehtat“. Es muss also prinzipiell möglich sein, in der Erfahrung von Schmerz die Empfindung von Schmerzhaftigkeit von der negativen Bewertung derselben zu dissoziieren. Wenn dies prinzipiell *empirisch* möglich ist (wenngleich unter Normalbedingungen nicht auftritt), dann spricht dies dafür, dass die entsprechende *begriffliche* Analyse (die m.E. begrifflich in jedem Fall konsistent ist) nicht fehlgeleitet sein kann. Ich denke also, dass sowohl (1) als auch (2) jeweils plausible Wege darstellen, wie wir die Möglichkeit der Intentionalität von Sinneswahrnehmungen einräumen können.

⁸⁵ Ich danke Herrn A/Prof. Philip Cam für diesen Hinweis.

3.11. Transpersonale Geltung von emotionalen Wertungen

Abschließend wird im vorliegenden Abschnitt noch zu klären sein, wie die transpersonale (oder objektive) Geltung von Werten im Rahmen eines Modells emotionaler Wertungen zu erklären ist. Wir haben weiter oben gesehen, dass diesem Modell ein charakteristischer Subjektivismus eignet, insofern Wertungen als individuelle emotionale Reaktionen aufgefasst werden. Im Rahmen meiner Ausführungen zum dispositionalistischen Aspekt dieses Modells habe ich zwar darauf hingewiesen, dass wir als wertende Subjekte ausreichend biologische und kognitive Eigenschaften teilen, um eine grundlegende Invarianz von Wertungen sicherzustellen. Dennoch wird ein stärkerer Rektifikationsmechanismus wünschenswert sein. Allerdings möchte ich darauf hinweisen, dass meiner eigentlichen Argumentation schon Genüge getan sein dürfte, indem ich nachgewiesen habe, dass Werte *ontologisch* auf Emotionen gründen, d.h. dass Prozesse emotionalen Wertens ontogenetisch primär sind, und dass daher emotionaler Affekt auch bei Instanzen höherer (kognitiver, transpersonaler) Wertungen konstitutiv bleibt, damit ein ausreichend determiniertes Bild vom Wertbegriff gegeben werden kann. Der Nachweis, dass *allen* Werten qua ihrer ontogenetischen Entstehungsgeschichte aus Instanzen emotionalen Wertens Affektivität eignet, genügt bereits, um in den letzten Schritt meiner Argumentation einzutreten, in dem die zentrale Fragestellung dieser Arbeit wieder aufgenommen werden und eine Lösung von Michael Smith's „moral problem“ versucht werden soll. Dennoch möchte ich kurz ausführen, wie ich denke, dass auch emotionale Wertungen transpersonale Geltung haben können.

Wenn wir nach einem Rektifikationsmechanismus fragen, so fragen wir damit nach einem *Angemessenheitskriterium* emotionaler Wertungen. Damit ein Wert transpersonale Geltung beanspruchen kann, muss er angemessen sein. Diese Bedingung kann in der Tat ein Problem für eine Theorie emotionalen Wertens darstellen. Stellen wir uns vor, Hans habe eine Spinnenphobie und halte aus diesem Grund Spinnen für sehr gefährliche Lebewesen, die es um jeden Preis zu vermeiden gilt. Die damit zum Ausdruck gebrachte Bewertung kann keine transpersonale Geltung beanspruchen, weil sie ihrem Objekt nicht angemessen ist. Ebenso könnten wir uns eine Person vorstellen, die durch schwere Fehler in ihrer Erziehung Freude an Gewalttaten empfindet und diesen dementsprechend einen bestimmten Wert beimisst. Auch in diesem Fall ist die Bewertung den gesellschaftlichen Konventionen und moralischen Normen nicht angemessen.

Auf dem Feld der Emotionstheorien wird ein Angemessenheitskriterium über die Kategorie des formalen Objekts eines Emotionstyps definiert. Allerdings halte ich dieses Kriterium – dezidiert auf den Bereich einer Theorie der Emotionen bezogen – für fehlerhaft:

zwar können wir sagen, dass Furcht vor Spinnen „unangemessen“ ist, da sich in diesem Fall die Emotion „Furcht“ auf ein materiales Objekt (Spinne) bezieht, das nicht als Instanziierung ihres intentionalen Objekts („das Gefährliche“) gelten kann. Ich habe an früherer Stelle argumentiert, dass diese Position wenig zur Erhellung der Genese und Analyse von emotionalen Zuständen beiträgt. Denn solche „unangemessenen“ Emotionen treten de facto auf. In der Fragestellung der Beziehung von Emotionen und Werten wiederum kann dieses Kriterium sinnvoll sein. Eine schlüssige Formulierung eines solchen Kriteriums finden wir bei Mulligan (1998), der in seinem Artikel ebenfalls den Grundriss einer Theorie emotionaler Wertungen entwickelt. Ich stimme Mulligans Meinung zu, dass die Emotionselizitation auf natürlichen Eigenschaften aufbauen muss, damit Emotionen wiederum die Grundlage für Bewertungen bilden können. Allerdings versteht Mulligan nur *angemessene emotionale Zustände* als Instanzen von Wertungen. Ich möchte einwenden, dass emotionale Zustände insgesamt Wertungen sind, dass jedoch Mulligans Formulierung eine korrekte Beschreibung *transpersonal gültiger Werte* darstellt. Ein Angemessenheitskriterium lässt sich dann relativ leicht formulieren, Mulligan liefert eine sehr präzise Fassung:

“x is an emotion appropriate to y and its natural properties F-H iff (1) b is a true or veridical (re)presentation of y and of its properties F-H, and (2) b is justified if it is a conceptual representation of y (or fits y if it is a non-conceptual presentation of y), and (3) b justifies x, and (4) no relation of justification or fit mentioned is defeated.”
(Mulligan 1998, 179)

Mulligans Angemessenheitskriterium beschreibt eine Beziehungen zwischen emotionalen Zuständen und ihren Objekten. Wenn sich diese Beziehung als angemessen darstellt, kann die entsprechende Emotion als die Instanz einer transpersonal gültigen Wertungen gelten. Helm (2001; vgl. auch 2001a) schlägt eine alternative Möglichkeit vor, die Angemessenheit von emotionalen Zuständen durch das Konzept von „Reaktionsmustern“ („patterned responses“) zu verstehen. Auf diese Weise lässt sich ein Angemessenheitskriterium zusätzlich über die Beziehungen mehrerer emotionaler Zustände mit gleichem Fokus untereinander definieren. Nach Helm sind emotionale Zustände auf zwei Arten miteinander verbunden:

(1.) „Transitional commitments“: Helm unterscheidet zwischen vorwärtsgerichteten und rückwärtsgerichteten Emotionen, je nach chronologischer Bestimmung der entsprechenden Objekte. „Hoffnung“ wäre demnach ein Beispiel für eine vorwärtsgerichtete Emotion, „Trauer“ eines für eine rückwärtsgerichtete. Die Bedingung des „transitional commitment“ postuliert nun, dass – im Verlauf der Zeit – eine gegebene vorwärtsgerichtete Emotion das Auftreten einer entsprechenden rückwärtsgerichteten Emotion mit gleichem Fokus einschließt. Wenn wir uns vorstellen, dass Hans sich auf seinen Sommerurlaub freut, dann können wir kraft dieser Tatsache annehmen, dass Hans traurig sein wird, wenn der Urlaub

vorbei ist. Mit Helms Worten: „Consequently, as the relevant things happen in the world, forward-looking emotions rationally ought to become the corresponding backward-looking emotions“ (Helm 2001, 193).

(2.) „Tonal commitments“: Helm greift mit diesem Kriterium die bekannte binäre Valenz emotionalen Affekts auf. Wir erleben Emotionen als positiv oder negativ getönt. Daraus ergibt sich nun ein kontrafaktisches oder spekulative Kriterium über die affektive Qualität von emotionalen Zuständen in Bezug auf ihr auslösendes Objekt:

“To say that emotions involve *tonal commitments* is to say that if one experiences a positive emotion in response to something good that has happened or might happen, then, other things being equal, one rationally ought to have experienced the corresponding negative emotion if instead what happened (or conspicuously might happen) were something bad; not to experience this emotion would be rationally inappropriate.” (Helm 2001, 193f.)

Zunächst könnten Helms Ausführungen missverständlich sein. Ich betone noch einmal, dass nicht primär auf die Beziehung Emotion-Objekt abgehoben wird, sondern auf die Beziehung Emotion – Emotion, in diesem Fall jedoch Emotion – kontrafaktische Emotion. Gemeint ist hier nicht, dass positive Emotionen sich auf „positive“, d.h. „gute“ Objekte beziehen. Diese Annahme wäre ein „moralistischer Fehlschluss“ (D’Arms & Jacobson 2000a). Gemeint ist hingegen eine kontrafaktische Bedingung, ausgehend von einem bestehenden emotionalen Zustand. Wenn also Person H mit der Emotion e(positiv) auf die Tatsache x reagiert, dann können wir kontrafaktisch annehmen, dass wäre nicht-x eingetreten, hätte H mit e(negativ) reagiert. Wenn also Hans sich freut, dass seine Freundin Susi zu Besuch kommt, dann können wir kontrafaktisch annehmen, dass er wohl traurig gewesen wäre, wenn Susi nicht zu Besuch gekommen wäre. Da es jedoch nicht um die Beziehung Emotion – Objekt geht, trifft die Bedingung des „tonal commitment“ ebenso gut auf den folgenden Fall zu: Hans ist traurig, weil seine Freundin Susi zu Besuch kommt, woraus wir schließen können, dass Hans sich gefreut hätte, wenn Susi nicht zu Besuch gekommen wäre.

Helms kontrafaktische Bedingung wird von ihm selbst als eine spekulative Kategorie eingeführt. Wir können von einem eingetretenen Ereignis bzw. vorliegenden Objekt und einer gegebenen emotionalen Reaktion darauf lediglich spekulativ schließen, dass im Falle eines gegenteiligen Ereignisses die contra-tonale Emotion aufgetreten *wäre*. Ich denke jedoch, dass wir diese Bedingung auch sinnvoll und ohne Mühe erweitern können, indem wir annehmen, dass tonale Beziehungen auch zwischen *hinreichend ähnlichen Situationen* bestehen. Wenn Hans sich also über den Besuch von Susi gefreut hat, können wir *voraussagen*, dass ihre Absage am nächsten Wochenende – *ceteris paribus* – ihn traurig machen wird.

Die beiden Bedingungen beschreiben also rationale Beziehungen zwischen mehreren Emotionen mit demselben „Fokus“, verstanden als „the background object of the subject’s concern that makes intelligible the evaluation implicit in the emotion“. Dabei beschreibt (1) diachrone und (2) synchrone Beziehungen zwischen Emotionen. An die Stelle des Angemessenheitskriteriums setzt Helm nun den Begriff der Rationalität eines emotionalen Musters, wobei diese Rationalität als die *Konstanz* und *Integrität* des Musters verstanden wird:

“In this way, the mutual commitments among emotions define a pattern with a common focus. This pattern is *rational* in the sense that the pattern is partly definitive of the rationality of its elements: it is normal and appropriate to feel emotions that fit into such patterns, and it is abnormal and inappropriate either to fail to feel emotions that fit into such patterns or to feel emotions that do not.” (Helm 2001, 194)

Helms Theorie der „patterned responses“ kann somit auch die Angemessenheit von subjektiven emotionalen Wertungen (die Helm bekanntlich „import“ nennt) erklären. Nehmen wir an, dass ich eine phobische Angst vor der Dunkelheit habe, und aus diesem Grund mein Schlafzimmer bei Nacht als einen bedrohlichen Ort bewerte. Aus Mulligans Sicht könnten wir nachweisen, dass meine Angst fehlgeleitet, weil unangemessen ist. Diese Erkenntnis setzt jedoch meistens eine dritte-Person-Perspektive voraus. Mit Helm können wir alternativ argumentieren, dass ich erkennen kann, dass es sich bei meiner Angst vor der Dunkelheit um eine fehlgeleitete Wertung handelt, da dieser emotional bewertende Bezug auf den entsprechenden Fokus *nicht* Teil eines rationalen Musters ist. Die anderen, vermutlich nicht fehlgeleiteten Emotionen desselben Musters fungieren somit als Verifikations- und Falsifikationsmechanismen für meine phobische Angst, die ich wiederum sozusagen durch eine intrapersonale Musterüberprüfung als „fehlgeleitet“ anerkennen und als Grundlage für das von mir vertretene System transpersonaler Wertungen dadurch ausschließen kann.

Rufen wir uns abschließend noch einmal ins Gedächtnis, dass es verschiedene Ebenen von Wertungen gibt. So kennen wir sicherlich die Unterscheidung zwischen subjektivem und objektivem Wert, evtl. können wir Letzteren noch weiter aufschlüsseln, z.B. in universellen oder moralischen Wert. Diese Kategorien schließen sich nicht gegenseitig aus, sondern bilden Überlappungen. Da die ontogenetisch primären Wertungen eines Individuums subjektiv im oben ausgeführten Sinne sind, müssen wir annehmen, dass durch einen heuristischen oder Filtermechanismus subjektive Werte in den Rang höherer Werte mit objektiver Geltung erhoben werden können. Wir haben anhand von Mulligans Angemessenheitskriterium oder Helms „patterned-response“-These zwei mögliche Beispiele solcher Rektifikationsmechanismen kennen gelernt. Umgekehrt können wir auch einräumen, dass wir

auch nicht-emotionale Wertungen anstellen können, die wiederum als rationale Kritik unserer emotionalen Wertungen dienen können. Diese jedoch werden wir nur dann auch authentisch vertreten und unterhalten, wenn wir sie in einem gewissen Sinne zu subjektiven Werten machen, sie uns als zu eigen machen. Auch hier möchte ich noch einmal betonen, dass das von mir erarbeitete Modell emotionalen Wertens nicht als die einzige Erklärung der Wertgenese angesehen werden soll. Ich kann mir die Möglichkeit von alternativen Modellen vorstellen, glaube aber, dass das hier vorgestellte repräsentativ für den ontogenetisch frühen Mechanismus des Wertens ist. Dadurch jedoch werden Werte insgesamt auf das ontologische Fundament von Emotionen gestellt. Im folgenden Abschnitt werden wir nun noch zu zeigen versuchen, inwiefern Emotionen einer aktiven Kontrolle des Akteurs unterliegen. Denn die rationale Einsicht, dass eine gegebene emotionale Wertung unangemessen ist, muss sich letztlich auch in der emotionalen Reaktivität des Akteurs niederschlagen, um eine wirkliche Modifikation der vertretenen Werte zu ermöglichen. Wenn dieser Nachweis erfolgreich erbracht wird, dann können wir zu Recht annehmen, dass auch ein Modell emotionalen Wertens transpersonal gültige Werte begründen kann. Im letzten Schritt der Argumentation sollen daraufhin die gewonnen Erkenntnisse für einen Lösungsversuch des „moral problem“ eingesetzt werden.

3.12. Verantwortung für Emotionen und emotionale Wertungen

Emotionen – so die prominente These von DeSousa (1979, 137-141) – liegen nicht völlig außerhalb unserer Verantwortungsbereichs. Dies wird zwar von bestimmten Strömungen in der Emotionstheorie bestritten⁸⁶, aber normalerweise machen wir auch alltagssprachlich Personen zu einem gewissen Grad für ihre Gefühle verantwortlich. De Sousas Emotionstheorie kann diesen bestimmten, aber begrenzten Grad von Verantwortung erklären. Wenn Emotionen Aufmerksamkeitsmuster sind, dann können wir unsere Emotionen beeinflussen, indem wir unsere Aufmerksamkeit beeinflussen. Und unsere Aufmerksamkeit liegt in unserer Macht. Allerdings ist es sehr viel einfacher, willentlich Aufmerksamkeit auf

⁸⁶ Die Zuordnung zu den gängigen Paradigmen in der Emotionstheorie ist hier interessanter Weise jedoch nicht einheitlich vorzunehmen, besonders nicht innerhalb des kognitivistischen Paradigmas: Das Konzept der Verantwortlichkeit für Emotionen ist gewöhnlich integraler Teil eines (neo-)stoischen (mithin des rein-kognitivistischen Paradigmas) und/ oder eines existenzphilosophischen Ansatzes (bspw. Nussbaum 1996 oder Solomon 1980 u.a.), bereitet jedoch Theoretikern der weit verbreiteten gemischt-kognitivistischen Ansätze größere Probleme.

etwas zu richten, als sie willentlich von etwas abzuwenden. Wir geraten sogar in einen logischen Widerspruch: ich kann nicht durch willentlichen Entzug der Aufmerksamkeit auf x eine Nicht-Aufmerksamkeit auf x (im Sinne eines Nicht-Wahrnehmens) erreichen, da x in meinem Willensakt des Aufmerksamkeitsentzugs immer noch repräsentiert ist. Wir können uns also sehr viel eher in eine Emotion versetzen, als wir eine unerwünschte Emotion abschalten könnten.

Auch an dieser Stelle sollten wir versuchen, eine a-fortiori-Argumentation zur Verantwortung für Emotionen zu erarbeiten. Wenn wir annehmen, dass emotionale Reaktionen (und damit auch emotionale Wertungen) spontan ablaufen (also ohne aktives Zutun des Akteurs und evtl. sogar ohne aktive Kontrollmöglichkeit seitens des Akteurs) und bisweilen nicht einmal einen kognitiven Aspekt einschließen, der einen rationalen Zugang ermöglichen könnte, dann sollte unser Argument auch unter diesen Voraussetzungen Erfolg haben können, um damit die ganze Bandbreite des Phänomens „Emotion“ erfassen zu können. Der Schlüssel zur Lösung dieses Problems liegt – wie wir sehen werden – in der Dichotomie von *Disposition und Aktualität*, unter der sich Emotionen beschreiben lassen. Unser Fazit wird lauten: wir sind verantwortlich für unsere Emotionen insofern diese feste Dispositionen – analog zu Habitus oder Charaktereigenschaften – darstellen. Zwar mögen aktuelle Auslösungsbedingungen nicht in unserer Verantwortung stehen, die Entwicklung und Festigung der zu Grunde liegenden Disposition verantworten wir hingegen sehr wohl.

Charaktereigenschaften können wir als feste, habituierte Handlungsmuster von Akteuren auffassen (wobei wir Charaktereigenschaften im Gegensatz zur natürlichen Persönlichkeit verstehen sollten), deren Implementierung in der praktischen Akteursidentität meistens zurückzuführen ist auf einige wenige Entscheidungen, die wiederum nicht innerhalb der Dynamik eines Habitus getroffen wurden, sondern sozusagen in einer originären Position („In Zukunft will ich ehrlich sein!“). „Verantwortung“ für Charaktereigenschaften ist also im Sinne von „Zuschreibbarkeit“ zu verstehen. Analog dazu soll ein Begriff der Verantwortung für Emotionen verstanden werden: Emotionen sind nicht rein passive „Widerfahrnisse“, sondern sind in einer bestimmten Hinsicht dem entsprechenden Akteur zuschreibbar.

Der status quo des „Charakters“ (und damit auch des emotionalen Charakters im Sinne emotionaler Dispositionen) einer erwachsenen Person scheint zu großen Teilen abhängig zu sein von Einflüssen, die sich zur Zeit der entsprechenden Prägung der Verantwortung dieser Person entzogen. Nehmen wir an, Hans (als Erwachsener) ist aggressiv und löst seine Probleme mit seinen Mitmenschen durch physische Gewaltanwendung. Und nehmen wir an, dass diese Gewohnheit dadurch zu erklären ist, dass Hans (als Kind) in seiner Familie keine

anderen Konfliktlösungsstrategien kennengelernt hat. Natürlich machen wir Hans verantwortlich für die gebrochene Nase, die er als erwachsener Mann Hermann zugefügt hat. Verantwortung für Handlungen ist zwar ebenfalls ein kontroverses Thema, wenn jedoch Hans den Fausthieb ausgeführt hat und nicht unter erheblichem Einfluss von berauschenden Substanzen stand, dann ist der Fall zumindest juristisch eindeutig. Interessant ist jedoch die Frage, ob wir Hans für seine gewalttätige Gewohnheit verantwortlich machen. Denn aktuell vorhandene Gewohnheiten sind sehr mächtig, und zur Zeit der Etablierung dieser Gewohnheit stand Hans in völliger Abhängigkeit zu seinem Elternhaus, hatte also selbst keinen Einfluss auf den Modus seiner Erziehung.

Charakter wird verständlich gemacht, u.z. für uns selbst wie für andere, durch unsere „Geschichte“, wobei hier bei dem Begriff der Aspekt des Narrativen den Vorrang hat vor dem Aspekt des historisch-genealogischen. Einen solchen narrativen Ansatz finden wir prominent bei Jonathan Jacobs (2001). Meistens, so Jacobs, wird eine Geschichte zugänglich über ein Schlüsselement, z.B. die Tatsache, dass Hans' Vater alkoholsüchtig war, und die Art, wie dieses Element narrativ in Hans' Charakter eingebettet ist. Auch Temperament und andere Persönlichkeitsmerkmale wie Schüchternheit, Aggressivität, Geduld, Ängstlichkeit sind Aspekte des Charakters und fließen damit auch in das Handeln eines Akteurs mit ein. Weiterhin wird die Geschichte einer Person immer auch Elemente von Glück und Unglück aufweisen, Ereignisse, die außerhalb der Kontrolle eines Akteurs liegen, wie Gesundheit und Krankheit, soziale Kontakte und persönliche Begegnungen, Katastrophen, Erfolg und Misserfolg, oder auch Talente, persönliches Geschick und Begabung für bestimmte Dinge. All dies wird notwendig auch – zumindest als Hintergrundebene – in der Bestimmung eines Charakters eine Rolle spielen. Jacobs Argument ist nun einerseits, dass je besser wir solche Elemente aus der Lebensgeschichte einer Person kennen – und je mehr wir davon kennen – desto besser werden wir den Charakter dieser Person bewerten, verstehen und beschreiben können. Dennoch, und darauf insistiert Jacobs, scheint es wichtig, demgegenüber dem Begriff der Verantwortung für Charakter eine bestimmende Rolle einzuräumen. Warum?

Beispiel für willentliches bzw. absichtliches Handeln sind breit gefächert. Der kleine Hans, der seinen Luftballon loslässt und ihm daraufhin nachläuft, handelt sicher absichtlich in seinem Versuch, den verlorenen Ballon wieder einzufangen. Ebenso handeln seine Eltern absichtlich, wenn sie sich – *nach reiflicher Überlegung* – entscheiden, Hans zur Grundschule in der Marktstraße zu schicken, und nicht zur Grundschule in der Bahnhofsstraße. Der

Unterschied ist hier nicht im Alter zu suchen,⁸⁷ denn, wie Jacobs es formuliert, „even for adults there is a great deal of voluntary motion that does not involve much in the way of explicit judgment or deliberation. Turning over on the sofa to get more comfortable is voluntary and so is rubbing one’s eyes after taking off glasses“ (ebd., 13). Wir sprechen also von Verantwortlichkeit für Handeln – im Sinne von Zuschreibbarkeit auf der Grundlage von Absichtlichkeit oder Willentlichkeit – auch bei Handlungen, die ohne Überlegung ausgeführt werden: nach dem Ballon greifen oder sich auf dem Sofa umdrehen. Und da wir also bei einer ganzen Bandbreite von Handlungen von Verantwortung sprechen – wobei dieses Spektrum die Opposition zwischen absichtlich und unabsichtlich (bzw. willentlich und spontan) transzendiert –, liegt die Vermutung nahe, dass der Begriff der Verantwortung sich auch beim Charakter anlegen lässt. Dennoch ist dies eine stärkere Behauptung, die eine wesentlich ausgefeiltere Argumentation verlangt. Die Tatsache, dass wir auch bei spontanen Handlungen von Verantwortung sprechen ist ein Indiz, kein Beweis dafür, dass wir auch sinnvoll von Verantwortung für Charakter sprechen können.

„We are accustomed to calibrating attributions of responsibility on the basis of the extent to which it is appropriate to expect agents to be aware of what they are doing and the extent to which reason, decisions, and understanding can be expected to inform their actions. Thus, with respect to an adolescent or a grown-up, but not with respect to a very young child, it makes sense to investigate whether negligence, mitigating, or aggravating conditions were involved. With maturation we expect an increasing knowledge of particulars and an enlarged general awareness of what is acceptable, what invites recrimination, anger, gratitude, or admiration, and so forth.“ (Jacobs 2001, 11)

In der Tat setzt ein voll habituerter Charakter den gegenwärtigen Handlungsoptionen eines Akteurs bestimmte Grenzen, und evtl. ist es tatsächlich schwieriger, vorhandene Charaktereigenschaften zu durchbrechen und zu ändern, als bestimmte erwünschte Eigenschaften aufzubauen. Weiterhin ist offenbar, dass Unglück und Misserfolgserfahrungen große Einschränkungen auf das menschliche Handeln bewirken, und dass Glück und

⁸⁷ Wenngleich wir ganz eindeutig Altersunterschiede geltend machen bei der Zuschreibung von Verantwortung. Wir beziehen also jenseits der Dichotomie von „absichtlich“ und „unabsichtlich“ auch bestimmte Fähigkeiten, Vermögen und Kapazitäten des Akteurs mit ein, z.B. Wahrnehmungsvermögen und Bewusstsein sowie die Fähigkeit zum abstrakten, modellhaften Durchdenken von Situationen. Nehmen wir an, ein Kleinkind macht seine ersten Schritte und wirft dabei eine Stehlampe um (vgl. Jacobson 2001, 12f.). Wir würden das Kind nicht für Letzteres verantwortlich machen, weil diese Konsequenz gewissermaßen außerhalb seines Wahrnehmungshorizonts liegt. Zugegebenermaßen ist jedoch dieses Kriterium graduell altersabhängig: wollte man solche Kriterien für eine erwachsene Person formulieren, dann würden wohl nur sehr hochspezialisierte und „enge“ Lebensbereiche in Frage kommen, um einen erwachsenen Akteur in analoger Weise von Verantwortung auszuschließen.

Vorhandensein von Glücksgütern ein großer Vorteil ist. Dennoch – und diese Meinung teilt auch Jacobs – macht es Sinn, von Verantwortung für Charakter zu sprechen, und dies scheint gestützt zu werden durch die aristotelische Philosophie. Denn auch Aristoteles geht davon aus, dass Absicht und Willentlichkeit nicht auf Handlungen zu beschränkt ist, die Resultat von expliziter Überlegung oder Entscheidung wären (vgl. EN 1111a26 und 1112a15). Wir sind also auch verantwortlich für gewisse Typen spontaner Handlungen, nämlich solche, die einem *habitus* entsprechen. Im Zusammenhang mit dem menschlichen Charakter spricht Aristoteles (in EN 1114b1 und 1114b23) von „aitios“, was soviel bedeuten kann wie: wir sind *ursächlich* verantwortlich für unseren Charakter. Dies ist auch der Verantwortungsbegriff, dem ich in der vorliegenden Studie nachgehen will. Dabei ist jedoch wichtig anzumerken, dass beide Stellen von „aitios“ jeweils mit „pos“ qualifiziert sind, d.h. wir sind „irgendwie“ kausal verantwortlich für unseren Charakter. An anderer Stelle (EN 1114b24) heißt es wiederum: wir sind „teilweise“ verantwortlich für unseren Charakter („synaitioi“, wörtlich: mit-verantwortlich). In diesem Sinne formuliert auch Kosman (1980, 113):

„But it is nonetheless possible to engage in a certain range of conduct deliberately designed to make one the kind of person who will characteristically feel in appropriate ways, at appropriate times, and so on. And in this sense, feelings are deliberate and chosen, since the hexeis from which these feelings emanate are deliberate and chosen, since (in turn) the actions that lead to these hexeis are deliberate and chosen, and deliberately chosen to make one the kind of person who characteristically will have the appropriate feelings.“

Jacobs liefert zwei Fallbeispiele, die er einander gegenüberstellt: einerseits stellt er eine fiktive Person vor, die die Angewohnheit hat, in Klatsch und Tratsch involviert zu sein. Nehmen wir nun an, dass diese Person die Angewohnheit ablegen will, dann wird dieser Prozess wohl in mehreren Stadien ablaufen, beginnend mit der bewussten und jeweils situational neu zu treffenden Willensbildung, sich nicht auf Klatsch einzulassen. Als zweites Stadium postuliert Jacobs die *Anerkennung* einer neuen Rolle durch das soziale Umfeld: dies ist ebenso wichtig wie die aktive Bemühung des Akteurs selbst, weil es das Selbstverständnis mitkonstituiert. Schließlich wird der neugewonnene Charakter ebenso zur zweiten Natur des Akteurs wie es zuvor seine Identität als „Waschweib“ oder „Klatschtante“ war, d.h. Abstinenz von Klatsch und Tratsch erfordert keine jeweils neue Bemühung mehr, sondern erfolgt mühelos aus der neugewonnenen praktischen Identität. Das zweite Fallbeispiel ist nun ganz analog zum ersten, jedoch mit einer fiktiven Person, die ihre (quasi-)pathologische Ungeduldigkeit überwinden will und einen eher entspannten, ruhigen Charakter anzunehmen bestrebt ist. Diese zweite Person arbeitet, anders als die Person aus dem ersten Fallbeispiel, an

einem habituellen Merkmal ihres Temperaments. Nach Jacobs sieht hier das letzte Stadium anders aus: möglicherweise wird für diese Person geduldiges, entspanntes Agieren und Reagieren niemals zur zweiten Natur, sondern die je neue situationale Willensbildung bleibt auch nach langer Habituation immer noch erforderlich. Bestimmte Persönlichkeitsmerkmale, wie z.B. Temperament, Grundstimmungen etc., teilen zwar ihre habituelle Natur mit anderen Charaktermerkmalen wie z.B. Gewohnheiten, sind jedoch viel schwieriger zu verändern als letztere.

„The efforts one makes to be more honest, for example, are different from the efforts one makes to be less prone to fear or pessimism. It is a different sort of undertaking. The reasons for this have to do with the different ‘materials’ that are being worked with and on. Still, across that range and the various degrees of plasticity there is a significant role for the agent’s own voluntariness and responsibility.“ (Jacobs 2001, 19)

Auch Charles Taylor (1976) spricht von „Verantwortung für unser Selbst“ im folgenden Sinne: nehmen wir an, dass Verantwortung für Handlungen ein allgemein akzeptiertes Konzept ist. Worauf würde dann eine Konzeption von „Verantwortung für unser Selbst“ zielen?

„Naturally we think of the agent as responsible, in part, for what he does; and since he is an evaluator, we think of him as responsible in part for the degree to which he acts in line with his evaluations. But we are also inclined to think of him as responsible in some sense for these evaluations themselves.“ (Taylor 1976, 289)

Für Taylor bedeutet „Verantwortung für unser Selbst“, ein Begriff, den ich im Sinne von „Verantwortung für unseren Charakter“ verstehe, verantwortlich zu sein für die „starken Bewertungen“ („strong evaluations“), die ein Akteur macht. Diese Idee baut auf Taylors Unterscheidung zwischen qualitativen (i.e. starken) und quantitativen Bewertungen auf. Letzteres lässt sich am besten dadurch illustrieren, dass für Taylor ein Akteur, der lediglich quantitative Bewertungen anstellt ein „simple weigher of alternatives“ (ebd., 287) ist: jemand also, der lediglich versucht, konfligierende Wünsche zu koordinieren oder in anderer Weise effizient zu arrangieren. Nehmen wir an, um ein Beispiel von Taylor selbst zu variieren, Hans ist hungrig, möchte aber mit Susi schwimmen gehen. Die Zeit reicht nicht mehr für ein ausgiebiges Mittagessen, also entscheidet sich Hans dafür, auf der Busfahrt zum Schwimmbad ein Butterbrot zu essen. Damit hat Hans beide Wünsche befriedigt, zwar nicht jeweils maximal für sich, aber doch maximal in Kombination. Nach Taylor betreffen qualitative Bewertungen nun ein reicheres Vokabular, demgemäß eine Person bestimmte Phänomene (auch die eigenen Wünsche z.B.) als gut oder schlecht, edel oder niedrig, schön oder hässlich beurteilt, oder im Rahmen von dichten Wertprädikaten wie „feige, großzügig,

etc.“. Für Taylor können wir Personen zu einem gewissen Grad verantwortlich machen für ihre starken Bewertungen. Der Begriff der Bewertung, so Taylor, geht schließlich auf das Verb „(be-)werten“ zurück und impliziert mithin die Idee, dass wir als Personen aktiv beteiligt sind daran, Bewertungen hervorzubringen. Ohne Personen, die bewerten, gäbe es keine Bewertungen in der Welt. Wir können also verantwortlich sein für unsere Bewertungen insofern diese ontologisch von uns abhängig sind.⁸⁸ Offenbar versteht Taylor das „Selbst“, oder wie wir wohl ersetzen dürfen den „Charakter“, als eine Menge von unterhaltenen Bewertungen. Ich weise schon an dieser Stelle darauf hin, dass das hier erarbeitete Modell emotionaler Wertungen gerade diesen Aspekt von Verantwortlichkeit per ontologischer Abhängigkeit im Sinne von Taylor problemlos erklären kann.

Für Charles Taylor gründet sich Verantwortung für das Selbst also auf unsere Praxis des Wertens, durch die wir uns mit bestimmten Zielen, Projekten und Zuständen in der Welt als assoziiert verstehen. In einer ähnlichen Perspektive versteht Harry Frankfurt (1987) im Rahmen seines berühmten Modells eine Person als konstituiert durch diejenigen „first-order desires“, denen sie ihre Zustimmung gibt. Nach Frankfurt haben wir erste-Stufe-Wünsche („first-order desires“), und welchen dieser Wünsche wir auch immer auführen, diesen nennen wir begrifflich unseren „Willen“. Weiterhin haben wir zweite-Stufe-Wünsche, d.h. Wünsche, die sich darauf beziehen, *welche* erste-Stufe-Wünsche wir haben *wollen*. Letztlich haben wir zweite-Stufe-Volitionen („second-order volitions“), also Willensinstanzen, die sich darauf beziehen, welche erste-Stufe-Wünsche wir als Ausdruck unseres Willens haben wollen. Frankfurt merkt weiterhin an, dass auch weitere Stufen möglich wären. Eine Entscheidung terminiert nun den fortlaufenden Prozess der Abwägung zwischen verschiedenen Wünschen und Präferenzen steigender Ordnung: „The decision determines what the person really wants by making the desire upon which he decides fully his own. To this extent the person, in making a decision by which he identifies with a desire, *constitutes himself*“ (ebd., 38). In Anlehnung an Aristoteles denkt Frankfurt nun, dass Handlungen willentlich sind, wenn das bewegende Prinzip „innerhalb des Akteurs“ („inside the agent“) zu finden ist. Dies ist jedoch nicht als räumliche Lokalisation zu verstehen: das auslösende Prinzip eines epileptischen Anfalls ist räumlich im Subjekt zu verorten, und dennoch ist es kein Prinzip willentlichen Handelns. Die Frage, ob das bewegende Prinzip für eine Handlung „innerhalb“ des Akteurs zu verorten ist, wird vielmehr dadurch entschieden, ob der Akteur dieses Prinzip durch eine

⁸⁸ Ich möchte an dieser Stelle auch explizit auf die augenfällige Parallele zu der vom Verfasser der vorliegenden Arbeit entworfenen Position hinweisen.

Entscheidung in sein „psychisches Leben“ integriert hat⁸⁹, d.h. ob der Akteur sich selbst als Person in einer Weise konstituiert hat, dass das in Frage stehende Prinzip Teil dieser Konstitution ist.

Entgegen Aristoteles will jedoch Frankfurt sein Konzept der Verantwortung nicht auf kausale Autorschaft basieren (ebd., 39f.). Nach Aristoteles ist eine Person verantwortlich für ihren Charakter, insofern sie bestimmte Maßnahmen ergriffen hat (oder es unterlassen hat, bestimmte Maßnahmen zu ergreifen), durch die bestimmte Habitus und Dispositionen entstanden sind. Nach dem aristotelischen Modell sind also das Tun und Lassen einer Person kausaler Ursprung bestimmter Dispositionen, Charaktereigenschaften oder Wünsche. Nach Frankfurt kann dies jedoch nur erklären, inwiefern eine Person verantwortlich ist für bestimmte einzelne *Charaktereigenschaften*. Demgegenüber betrifft die Frage, ob eine Person verantwortlich sei für ihren *Charakter* jedoch vielmehr, ob sie Verantwortung für ihre Charaktereigenschaften übernommen hat. Frankfurt betrachtet also verschiedene Charaktereigenschaften unabhängig von der Frage, ob ihre Existenz der kausalen Autorschaft der Person direkt (durch ihr Tun) oder indirekt (durch ihr Lassen) zuzuschreiben ist. Verantwortung für den Charakter hängt davon ab, ob eine Person für eben jene schlicht vorhandenen Charaktereigenschaften Verantwortung übernimmt, d.h. sich mit ihnen identifiziert und sie Teil dessen macht, wodurch sich diese Person als Person konstituiert. Diese Konzeption erlaubt es Frankfurt, zu unterscheiden zwischen zwei verschiedenen Formen des „inneren Konflikts“: (1.) der Konflikt einer Person mit ihren widerstreitenden Wünschen, und (2.) der Konflikt einer Person mit sich selbst, d.h. mit einer ambivalenten Perspektive zu einer gegebenen Frage, wo bestimmte höherstufige Wünsche und Volitionen sich immer noch widersprechen. Diese zweite Form des Konflikts wird nun durch Frankfurts Modell gelöst: die Entscheidung, sich in einer bestimmten Weise zu konstituieren verschafft einer Person wieder eine einheitliche Perspektive, indem diese bestimmte Wünsche und Volitionen als konstitutiv für ihr Personsein ansieht und sich mit ihnen identifiziert. Konflikte mit widerstreitenden Wünschen werden dadurch zwar nicht per se gelöst, aber sie werden nicht mehr als Konflikt mit sich selbst wahrgenommen, sondern in ähnlicher Weise, wie wir wahrscheinlich versuchen würden, einen epileptischen Anfall zu unterdrücken, als etwas, mit dem wir uns nicht identifizieren.⁹⁰ Wenngleich Frankfurts Ansatz eine interessante neue

⁸⁹ „What counts, even with respect to a moving principle that operates as an element of his psychic life, is whether or not the agent has constituted himself to include it.“

⁹⁰ Ich halte Frankfurts Beispiel hier für unglücklich, denn man könnte ja argumentieren, dass eine solch fundamentale Erkrankung, die notwendig auf unser Selbstverständnis als Person einwirkt, besser ertragen

Perspektive in die Frage der Verantwortung einführt, halte ich seine globale Kritik an Aristoteles nicht für stichhaltig. Frankfurt geht von der Perspektive faktisch vorhandener Wünsche und Dispositionen aus. Das aristotelische Modell jedoch kann erklären, wie diese überhaupt entstanden sind, bzw. sich aus kontingenten Affektionen zu festen Dispositionen entwickeln konnten. Und wenn in diesem Fall ein Tun oder Lassen des Subjekts die Ursache war, dann machen wir Personen sehr wohl dafür verantwortlich.

Wie kann man also Verantwortung für Charakter, und im Besonderen Verantwortung für dispositionale Emotionen, aus einer aristotelischen Perspektive verstehen? Interessant ist, dass wir Emotionen intuitiv als Teil des Charakters einer Person betrachten, und dass wir eine Person in dieser Hinsicht auch beurteilen. Wir bevorzugen Menschen in unserem sozialen Umfeld, denen wir einen guten Charakter zusprechen, dies auch in Hinsicht auf ihre Emotionalität. So würden wir, um ein triviales Beispiel zu geben, lieber mit einer Person Tennis spielen, die Niederlagen mit Witz und Souveränität hinnimmt, als mit jemandem, der bei jeder Niederlage in cholerische Wutanfälle verfällt. Wir bevorzugen es also, dass die Personen, mit denen wir es in unserem Alltagsleben zu tun haben, die „richtigen“ Emotionen haben. Und dies erstreckt sich, wenn wir immer noch bei unserem intuitiven Verständnis der Dinge bleiben, sogar (entgegen der üblicherweise Kant zugeschriebenen These) in den Bereich der Moral. Wir bevorzugen Personen, die Gutes tun und dabei positive emotionale Erlebnisse haben, z.B. Gefühle der Sympathie und Erfüllung, gegenüber Personen, die Gutes tun mit Widerwillen und Verbitterung. Soweit trägt jedenfalls eine Erläuterung unseres Alltagsverständnisses. Zwar erfasst Kant einen ebenfalls wichtigen Punkt, wenn er von emotionalen Einstellungen absieht, oder sogar denjenigen Akteur für den Edleren hält, der die „stiefmütterliche Natur“ (vgl. Kant, KpV AA 146) überwindet und dennoch Gutes tut. Was Kant hier erfasst, können wir auch intuitiv nachvollziehen, als eine Kategorie des *moralischen Verdienstes*. Und dennoch können wir wohl festhalten, dass wir – *ceteris paribus* – den gutgestimmten Wohltäter bevorzugen gegenüber dem Misanthropen, der sich durch Selbstbeherrschung zum Wohltäter aufschwingt.

Die weitere Frage ist natürlich, ob wir Emotionen lediglich die Rolle von zusätzlichen Perfektionen des moralischen Handelns zusprechen wollen, oder ob wir richtige Emotionen als dessen notwendige Komponente oder Grundlage verstehen wollen. Wenn wir für eine Sichtweise optieren, die sich in Richtung von Letzterem bewegt, ergibt sich ein Problem, das auch schon dann auftritt, wenn wir Personen hinsichtlich ihres (emotionalen) Charakters

werden kann, wenn man sich schicksalhaft mit ihr identifiziert. Wären wir dann nach Frankfurts Verständnis in diesem Sinne „verantwortlich“ für die Erkrankung als Teil unseres Selbstverständnisses als Person?

bewerten, auch in moralisch irrelevanten Fragestellungen: „A central notion of our moral lives is responsibility, and responsibility presupposes choice; emotion is unchosen“. (Sabini & Silver 1987, 169)

In bestimmten Spielarten der Emotionstheorie ist die Postulierung von Verantwortung für Emotionen nicht weiter fragwürdig. Solomon (1973 u.f.) und Sartre (1964) z.B. halten einen Entscheidungsmoment für inhärent im Erleben von Emotionen. Gemäß einer solchen Theorie wären wir ohne Weiteres verantwortlich für unsere Emotionen. Eine mögliche Lösung in der aristotelischen Tradition liegt in der Unterscheidung zwischen aktueller Entscheidung und vorausgehenden Entscheidungen. Wir können z.B. verneinen, dass Emotionen das Objekt einer aktuellen Entscheidung darstellen. Wenn wir den cholerischen Tennisspieler synchron betrachten, dann hat dieser aktuell in der Situation, in der er ein Match verliert, wahrscheinlich keine Wahl, wütend zu werden oder nicht, d.h. sein akuter emotionaler Zustand ist nicht Gegenstand einer Entscheidung *in diesem Moment*. Das schließt jedoch wiederum nicht aus, dass Emotionen, oder genauer: emotionale Dispositionen (als habitualisierte Charaktermerkmale)⁹¹, durch rationale Einflussnahme modelliert werden können, somit also Gegenstand einer langen Serie von Entscheidungen sind, die schließlich in einer festen Disposition resultiert. Mit dieser Sichtweise befinden wir uns schon im Themenbereich des Charakterbegriffs. Indem wir bestimmte emotionale Dispositionen einüben, bestimmen wir zu einem Teil, was für Personen wir werden, was für einen Charakter wir uns schaffen.

Die aristotelische Sichtweise ist zwar tragfähig für einige Fälle, nicht jedoch im Allgemeinen. Sabini & Silver illustrieren dieses Problem mit dem Beispiel der Drogenabhängigkeit: wir nehmen an, dass der Erstkonsum von z.B. Heroin in den meisten Fällen Gegenstand einer freien Entscheidung ist, d.h. nicht unter Umständen hoch-affektiver Emotionalität, psychischen Drucks, oder einer verengten Urteils- und Bewertungsperspektive stattfindet. Alle weiteren Instanzen des Heroinkonsums nach Eintreten der Abhängigkeit jedoch sind äußerst schwer zu kontrollieren, eben wegen der oben angeführten und noch weiteren situationalen Bedingungen. Dennoch halten wir diese exemplarisch imaginierte, heroinabhängige Person in gewissem Sinne für verantwortlich für ihre Abhängigkeit, u.z. im Hinblick auf genau jene Instanz des Erstkonsums, die Gegenstand einer freien Entscheidung

⁹¹ Sabini & Silver sprechen durchgängig von „emotions“. Diese Redeweise ist natürlich ungenau, denn wir beurteilen niemanden im Hinblick auf seine akuten Emotionen, sondern auf emotionale Muster und Dispositionen. Die Ausführungen von Sabini & Silver bleiben jedoch tragfähig, wenn wir davon ausgehen, dass Letzteres eigentlich gemeint ist.

war, und die direkt in die Passivität, d.h. Heteronymie der Abhängigkeit geführt hat. Andererseits müssen wir uns vergegenwärtigen, dass wir eine Person, die eine Opiatabhängigkeit entwickelt hat als Resultat von medizinisch indizierten Morphiumgaben, nicht, oder nicht im gleichen Ausmaß, für diese Abhängigkeit für verantwortlich halten.

Sabini & Silver geben drei Bedingungen, die zutreffen müssen, um Personen Verantwortung für ihre Emotionen zuzusprechen:

„[...] to ground claims that people are responsible for their emotions, we must show:

1. that people know, or should know, that their failures to control their passions will result in their becoming the sort of person who will later be unable to control that emotion;
2. that people have, and know they have, techniques to accomplish this – Aristotle’s claim that moderation begets moderation is empirical and in need of empirical support; and
3. that people were not (blamelessly) in an uncontrollable emotional state when they failed to do that which would have led to more properly formed emotions (this must be shown to avoid regress).” (Sabini & Silver 1987, 171)

Ein auf diesen Bedingungen aufbauender Begriff der Verantwortung für Emotionen wäre also zumindest für moralisch relevante Zuschreibungen von Lob und Tadel zu beachten. Bleibt jedoch noch die Tatsache, dass wir Personen hinsichtlich ihrer charakteristischen Emotionen beurteilen, selbst wenn dieser Begriff der Verantwortung nicht greift, also selbst dann, wenn die emotionalen Charakteristika dieser Personen völlig außerhalb ihrer Kontrolle liegen.

Sabini & Silver unterscheiden zwischen drei Typen von Beurteilung: (1.) *pragmatisch*: wir beurteilen etwas in Hinblick auf unsere Wünsche, Ziele, Absichten etc., (2.) *moralisch*: wir beurteilen etwas mit Bezug auf moralische Werte und Normen, unabhängig von unserer eigenen Person. Wir haben oben also gezeigt, inwiefern wir eine quasi-moralische Beurteilung der emotionalen Charakteristika einer Person rechtfertigen können. Diese Rechtfertigung beruhte jedoch auf dem Begriff der Verantwortung für das, was Gegenstand der Beurteilung ist. In welchem Sinne können wir nun intersubjektiv nachvollziehbare und rechtfertigbare Beurteilungen anstellen über jemandes Charakter, wenn die in Frage stehenden Charakterzüge, wie z.B. bestimmte Emotionen, nicht im Bereich der Verantwortung dieser Person liegen? (3.) In *ästhetischer* Hinsicht: wir beurteilen bestimmte emotionale Dispositionen als negativ, weil sie „unschön“ sind. Eine cholerische Disposition „verunstaltet“ den Charakter. Andere Dispositionen beurteilen wir in ästhetischer Hinsicht als positiv: Ausdrücke von Sympathie, Zuneigung oder Freude sind – da, wo sie angemessen sind – eine „Zierde“ für den Charakter. Wir können also jemandes Charakter mit Hinblick auf

ästhetische Kategorien bewerten. Sabini & Silver merken an: „these grounds do not include responsibility, but they are objective, what anyone should see, impartial, independent of the judge’s interests“ (ebd., 173).

Der Ansatz von Sabini & Silver weist m.E. in die richtige Richtung, insofern bei der Frage nach der Verantwortung für Emotionen das Kriterium der Voraussicht und möglichen Abwendbarkeit in Anschlag gebracht wird. Denn in der Tat kann eine solche Verantwortung nicht im Sinne einer kausalen Verursachung begriffen werden, wie es etwa bei den meisten Handlungen möglich und sinnvoll ist. Wie in einem früheren Beispiel illustriert: wenn Hans den Fausthieb gegen Hermann ausgeführt hat, dann ist er auch für Hermanns gebrochene Nase verantwortlich. Wir „verursachen“ jedoch Emotionen nicht, sondern wir „erleben“ sie. Diesen Widerfahrnis-Charakter erfasst das traditionelle Verständnis von Emotionen als „passiones“. Dennoch sind wir unseren Emotionen nicht passiv ausgeliefert. Es ist sicherlich nicht möglich, eine Konzeption von Verantwortung für akute, nicht voraussehbare emotionale Zustände zu konstruieren. Dies ist aber auch nicht das Ziel dieses Kapitels, da nicht solche akuten Zustände, sondern allein emotionale Dispositionen als Konstituenten des Charakters aufgefasst werden können. Da es sich hier gerade um Dispositionen handelt, die also eine vorausgehende Phase der Habituation einschließen, greift das Kriterium der Abwendbarkeit. Die wohl ausgefeilteste Theorie dieser Art hat Justin Oakley (1992) vorgelegt.

Oakley (ebd, Kap. 4) denkt, dass wir, um zu zeigen, wie wir verantwortlich sein können für unsere Emotionen, zwei Punkte nachweisen müssen: (1.) dass und wie wir unsere Emotionen kultivieren können und (2.) dass und wie wir Kontrolle über unsere Emotionen haben können. Weiterhin unterscheidet er zur Klärung, was gemeint sein soll mit „Verantwortung für Emotionen“, zwischen 3 Bedeutungen von „A is responsible for X“:

- (1.) A caused X. This meaning need not carry any implication of credit or blame, and is usually associated with actions and events, rather than agents.
- (2.) A is (2.1.) legally or (2.2.) morally blameworthy of X.
- (3.) A is open to creditworthiness or blameworthiness for X.

Wir können uns in der folgenden Diskussion auf Punkt (1) „A caused X“ beschränken. Denn wir sind auf der Suche nach einem Kriterium der Verantwortung im Sinne der Zurechenbarkeit; wir fragen mit anderen Worten danach, inwiefern wir eine Person als „Akteur“ ihrer Emotionen auffassen können, und nicht lediglich als deren passives Subjekt. Dabei geht es nicht in erster Linie um die substanzielle Ausgestaltung der konkreten

Emotionen. Die Frage soll also nicht sein, ob Hans sich qua Zuschreibung von Verantwortung einem moralischen Vorwurf dafür ausgesetzt sehen muss, dass er etwa *neidisch* auf seinen Nachbarn ist. Wir wollen also nicht untersuchen, inwiefern *bestimmte* Emotionen Anlass zu moralischem Lob oder Tadel des Akteurs geben könnten. Der zu erbringende Nachweis der Verantwortung im Sinne von Zurechenbarkeit soll sich insgesamt auf alle Emotionen (oder genauer: auf alle emotionalen Dispositionen) eines Akteurs beziehen, mithin also einen Nachweis der Verantwortung für dispositionale Emotionalität *tout court* darstellen. Auf diese Weise könnte in einem neuerlichen Schritt sichergestellt werden, dass das in diesem Kapitel erarbeitete Modell emotionaler Wertungen nicht ein Phänomen zu validen Wertungen erhebt, das in Wirklichkeit kontingent ist und sich in dem wertenden Subjekt als „Widerfahrnis“, also passiv, ereignet. Wenn wir zeigen können, dass ein Akteur prinzipiell verantwortlich für seine emotionalen Dispositionen ist (gleich welchen Inhalts und welchen substanziellen Emotionstyps), dann können wir es mithin als bewiesen betrachten, dass Akteure auch für die von ihnen angestellten emotionalen Wertungen verantwortlich sind.

Betrachten wir also Oakleys Kriterium (1) näher. Gelegentlich wird die Ansicht vertreten, dass die Fähigkeit, eine bestimmte Handlung willentlich auszuführen, ein notwendiges Kriterium dafür ist, verantwortlich für diese Handlung zu sein. In dieser Weise kann man Hans „verantwortlich“ dafür nennen, dass er aufsteht, seinen Arm hebt oder seine Luft anhält, nicht jedoch dafür, dass er niest. Nach dieser Formulierung wäre Verantwortung für Emotionen ausgeschlossen, da wir offensichtlich nicht fähig sind, auf Willensentschluss eine bestimmte Emotion zu erleben. Handelt es sich hier jedoch um einen tragfähigen Begriff von Verantwortung? In der Tat gibt es viele Handlungen, die nicht allein auf Grund eines Willensentschlusses ausgeführt werden können, und dennoch sprechen wir auch dort von Verantwortung. Wie Oakley sagt, sind wir verantwortlich dafür, ein Spiel zu gewinnen oder ein Auto zu fahren, und dennoch erfordern beide Handlungen einen wesentlichen Anteil externer Faktoren zu ihrem Gelingen, weit über den reinen Willensentschluss hinaus. Wir könnten evtl. eine Modifikation dieser Definition formulieren, etwa dass ich verantwortlich für etwas bin, wenn ich es willentlich tun oder haben kann, vorausgesetzt dass die richtigen Rahmenbedingungen gegeben sind. Allerdings versagt auch diese Modifikation für die Frage nach der Verantwortung für Emotionen, da selbst die detaillierteste Beschreibung des Kontexts noch keine hinreichende Bedingung für die Genese einer bestimmten Emotion liefert. Eine Situation kann strukturell ein perfektes Paradigma für „Zorn“ sein, und dennoch bin ich aktuell einfach nicht wütend.

Das konkurrierende Modell ist nun, Verantwortung auf die Frage der Abwendbarkeit zu gründen. In diesem Sinne wäre die Fähigkeit, eine Handlung oder ein Ereignis zu beenden bzw. einen Zustand nicht mehr zu haben, eine notwendige Bedingung für die Zuschreibung von Verantwortung. Wir sind also nicht verantwortlich für etwas, das wir nicht beenden können, indem wir uns dazu entscheiden, es zu beenden. Allerdings existieren auch hier *hard cases*: Wir können Hans nicht dafür verantwortlich machen, dass er mit dem Auto von einer Klippe fährt, weil seine Bremsen versagen. Hans kann dieses Ereignis schließlich nicht einfach dadurch beenden, dass er sich entsprechend entscheidet. Trotzdem können wir uns Fälle vorstellen, wo dennoch eine Verantwortung besteht. Nehmen wir an, Hans ist ein erwachsener Mann, der in seiner Urteilskraft nicht pathologisch eingeschränkt ist. Hans müsste also wissen, dass Autos regelmäßig gewartet werden müssen, damit keine Funktionsstörungen auftreten. Er wäre also in einem gewissen Sinne verantwortlich für den Unfall, *genau dann wenn* er es z.B. seit Jahren versäumt hätte, die Bremsen kontrollieren zu lassen. Deren Versagen ist dann zwar aktuell nicht mehr zu verhindern, hätte aber evtl. durch entsprechende vorausgehende Entscheidungen und Handlungen verhindert werden können. Wenn wir also entscheiden wollen, ob eine Person verantwortlich ist für etwas, das diese Person aktuell nicht mehr beenden oder abwenden kann, sollten wir betrachten, wie diese Person in eben diejenige Situation kam, dass der aktuelle Zustand nun unvermeidbar ist. Dieser Ansatz scheint auch für Emotionen vielversprechend zu sein.

In diesem Zuge wären auch Thomas Nagels Beobachtungen im Rahmen seines berühmten Artikels „Moral Luck“ zur Rolle des Glücks bei der Feststellung von juristischer Verantwortlichkeit zu nennen (Nagel 1976). Er schlägt folgendes Beispiel vor: zwei betrunkene Autofahrer fahren aus dem Pub nach Hause, wobei der eine einen Passanten anfährt, der andere ereignislos nach Hause kommt. Der erste Fahrer ist natürlich verantwortlich dafür, den Passanten angefahren zu haben, obwohl dessen Anwesenheit auf der Straße außerhalb der Kontrolle des Fahrers lag und damit – im Vergleich zum anderen, der eben „Glück“ hatte – ein Fall von Pech für diese Fahrer ist. Nach Oakley gilt dies auch analog für moralische Verantwortung: wenn Hans alle Ersparnisse seiner Familie in Gold investiert, und ein unvorhersehbarer Crash des Goldpreises an den Börsen eintritt, dann ist Hans dennoch verantwortlich für den Ruin seiner Familie, selbst dann, wenn dieser Verlust nicht auf Inkompetenz oder schlechte Planung zurückzuführen ist, sondern in erster Linie auf außerordentliches Pech.

Es wird also deutlich, dass wir das Modell der Abwendbarkeit noch weiter modifizieren müssen. Denn einerseits wollen wir auch Fälle erfassen, wie Hans' Versäumnis, die Bremsen

warten zu lassen, und andererseits wollen wir Fälle erklären können, wo Verantwortung trotz Einflusses von Glück zu bestehen scheint. Oakley setzt an mit einer Konzeption von „Voraussicht“. Voraussicht zu haben oder schuldhaft zu ermangeln kann als ein notwendiges Kriterium für Verantwortung gelten. Allerdings haben wir offenbar keine volle Voraussicht über unsere Emotionen. Oakley schlägt darum eine Differenzierung von zwei Begriffen vor: *Handlungs-Voraussicht* („act foresight“) und *schematische Voraussicht* („schematic foresight“). Der zweite Begriff soll erklären, wie ich durch Haben oder Ermangeln von Voraussicht auch dann verantwortlich für etwas sein kann, selbst wenn der Ausgang meines Handelns nicht völlig deutlich ist für mich, da entweder notwendig Glück hinzutreten muss (oder Pech hinzutreten kann), oder aber wenn mein Handeln einen so hohen Grad der Komplexität aufweist, dass nicht alle Einflüsse abgewägt werden können. In der Tat können wir oft schematische Voraussicht über unsere Emotionen haben: so können wir unser Emotionsleben kultivieren, indem wir herauszufinden versuchen, welche Schlüsselsituationen (DeSousa 1979: „paradigm scenarios“) relativ reliabel mit bestimmten Emotionen korrelieren. Wir können daraufhin versuchen, bestimmte Elemente zu isolieren, die mit bestimmten Emotionen verknüpft zu sein scheinen, um diese dann entweder generalisiert zu entwickeln oder aber zu unterdrücken. Dies können wir z.B. durch einfach Habituation erreichen, indem wir Tätigkeiten häufig ausführen, die mit erwünschten Emotionen zusammenhängen, bzw. solche vermeiden, die mit unerwünschten Emotionen gekoppelt sind. Auch Oakley hält für diese Habituation ein reflektives Zutimmen für nötig: wir müssen bestimmen, welche unserer Emotionen den Charakter widerspiegeln, den wir haben möchten, um genau diese Emotionen dann zu verstärken.

Ein solches Unternehmen nennt Oakley „erworbene Spontaneität“ („learned spontaneity“). Der Begriff ist durchaus kein Paradoxon. Gemeint ist damit, eine Fähigkeit oder Disposition, die in ihrer perfekten Ausprägung spontan abläuft, zunächst einmal durch Training zu erwerben, sie dann aber zu kultivieren, in bestimmte Bahnen zu lenken und unter einen gewissen Grad von Kontrolle zu stellen. Ganz analog verhält es sich z.B. mit dem Erlernen einer Fremdsprache oder eines Musikinstruments. Wir lernen dabei zunächst einzelne Elemente, z.B. morphologische Wortveränderungen, syntaktische Regeln oder aber einzelne Töne, Tonleitern und Phrasen. Danach erlernen wir durch Transferaufgaben, wie regelgerechte Generalisierungen ausgeführt werden, sodass wir schließlich kompetent „improvisieren“ können. Am Ende ist das Sprechen einer Fremdsprache oder das Spielen des Saxophons Spontaneität: wenn wir vor dem Aussprechen eines neuen Satzes jedes Mal überlegen müssten, oder wenn wir beim Saxophon-Solo jede Phrase zuerst gedanklich

vorkonstruieren müssten, dann haben wir eigentlich die jeweilige Fertigkeit noch nicht wirklich gelernt. Wir sind noch immer Schüler, keine Meister.

Wie können wir nun dieses Modell der Abwendbarkeit auf emotionale Zustände anwenden? Offenbar gibt es hier einige Problemfälle. Oakley diskutiert das Beispiel einer Frau, die an prämenstrueller Depression (PMD) leidet. Nehmen wir an, die PMD wäre abwendbar durch entsprechende Medikation. Wäre diese Frau demnach verantwortlich für ihre PMD, wenn sie sich entscheidet, kein entsprechendes Mittel einzunehmen? Oder nehmen wir an, Hermann leidet an einer Depression, weil seine Freundin ihn verlassen hat. Ist er auch verantwortlich zu machen für seinen Zustand, da er ihn ja hätte abwenden können, indem er schlicht keine Liebesbeziehung eingegangen wäre? Diese Beispiele sind polemisch und provokativ. Oakley schlägt eine Lösung auf zwei Analyseebenen vor:

- (1.) Wir müssen die Frage stellen: Gab es eine vorausgehende Handlung, die wir vernünftigerweise hätten ausführen (unterlassen) können, um unseren jetzigen emotionalen Zustand abzuwenden? Wenn wir also nur mit Hinblick auf etwas versagt haben, das man nicht vernünftigerweise von uns hätte erwarten können, dann liegt keine Verantwortung vor.
- (2.) Angenommen, wir sind verantwortlich (im obigen Sinne) für ein vorausgehendes Tun oder Lassen. Hätten wir vernünftigerweise die emotionalen Konsequenzen dieses Tun oder Lassens vorhersehen können? Falls nicht, liegt keine Verantwortung vor.

Oakley Fallbeispiele illustrieren jeweils einen der beiden Punkte:

(Ad 1): Handelt es sich im Falle der Frau mit PMD-Syndrom um die Unterlassung einer Handlung, die man vernünftigerweise von ihr hätte erwarten können? Nehmen wir an, sie sei der Meinung, dass Arzneimittel nur für besonders wichtige Zwecke eingesetzt werden sollten. Und bisher hat sie sich immer mit ihren Symptomen abgefunden. In diesem Fall liegt wohl keine verantwortungsrelevante Unterlassung vor. Nehmen wir jedoch an, dass die Frau unter schweren PMD-Symptomen leidet und normalerweise ein Mittel dagegen einnimmt. In dem speziellen Fall jedoch hat sie es aus nicht direkt relevanten Gründen (Unlust: „she couldn't be bothered“) versäumt, das notwendige Mittel zu besorgen. In diesem Fall können wir sagen, dass sie durch das Versäumnis einer Handlung, die man vernünftigerweise hätte erwarten können (und die sie de facto normalerweise auch ausführt), verantwortlich für ihren emotionalen Zustand ist.

(Ad 2): Vorausgesetzt, dass ein verantwortliches Tun oder Lassen stattgefunden hat: Hans hat eine Liebesbeziehung begonnen, was letztlich seine eigene Entscheidung war. Ob Hans nun verantwortlich für seine Depression nach der Trennung ist, hängt davon ab,

ob er vernünftigerweise die Konsequenzen hätte voraussehen können. In den allermeisten Fällen würden wir diese Frage mit nein beantworten, womit Hans natürlich nicht verantwortlich für seine Depression wäre. In einigen wenigen Fällen kann dennoch Verantwortung vorliegen, was Hans evtl. durch die weniger charmanten Mitglieder seines Freundeskreis erfahren müsste, die ihm z.B. sagen würden: „Selbst schuld, ich hab’s dir ja gesagt...“. Der Punkt ist hier, dass evtl. bestimmte Konstellationen des Charakters, der Persönlichkeit, des sozialen Umgangs etc. vorliegen könnten, durch die vorauszusehen wäre, dass die Liebesbeziehung von Hans sehr problematisch verlaufen und sehr unangenehm enden würde. Nehmen wir z.B. an, dass Hans, der fest an eheliche Treue glaubt, sich mit Susi zu verheiraten wünscht, deren polygamer Lebensstil allseits bekannt ist (und nur von Hans selbst geflissentlich übersehen wird). Bei diesem Kriterium ist jedoch m.E. darauf zu achten, dass man nicht in einen unbilligen Paternalismus gerät. Indem wir als selbständige Akteure unser Leben leben, öffnen wir uns notwendig der Möglichkeit, unangenehme Erfahrungen zu machen, die aller Voraussicht nach mit einer bestimmten Wahrscheinlichkeit eintreten können. Dennoch generieren solche Erwägungen nicht immer Gründe für oder gegen entsprechende Handlungen.

Oakley fasst sein Programm der Kultivierung der Emotionalität wie folgt zusammen:

I have been suggesting that we can and ought to cultivate our emotional capacities as traits of character, which involves not just knowing how to produce certain emotions, but involves the development of a whole emotional outlook, an appreciation of situations in terms of one’s own and others’ emotions, and a capacity to respond naturally with and act out of emotion. In other words, I am suggesting that we can and ought to become emotionally sensitive people. (Oakley 1992, 143)

Abschließend werde ich einige mögliche Einwände gegen Oakleys Modell behandeln. Zunächst könnten wir entgegen, dass wir gemäß diesem Modell der Abwendbarkeit gar nicht für unsere Emotionen verantwortlich sind, sondern lediglich für die Handlungen, die wir willentlich ausführen, um eben unsere Emotionalität zu beeinflussen. Dazu ist zu sagen, dass diese Handlungen ja gerade intentional auf die Kultivierung unserer Emotionalität bezogen sind. Das zutreffende Paradigma einer solchen Handlung ist weniger das des „Spazierengehens“ bspw., sondern vielmehr das des zielgerichteten Handelns, wie etwa der erwähnte Fausthieb, durch den Hans auch für die Verletzung seines Gegners verantwortlich ist. Es sei zugestanden, dass das Verhältnis zwischen solchen Handlungen der Emotionskultivierung und den entsprechenden Emotionen selbst nicht ein streng kausales Verhältnis wie das zwischen dem Fausthieb und der gebrochenen Nase ist. Dennoch ist dieses Verhältnis weit diessseitig von reiner Zufälligkeit zu verorten, es besteht eine starke Korrelation zwischen vorausgehenden Bemühungen und dem emotionalen Erleben, auf das

diese zielen. Nur darum, nebenbei bemerkt, zeitigt schließlich auch die Praxis der modernen Psychotherapie eine Wirkung in der Veränderung (pathologischer) emotionaler Dispositionen ihrer Patienten.

Betrachten wir nun also Oakleys (1992, Kap. 4.7) Auseinandersetzung mit dem Verhältnis zwischen seinem Ansatz und einem anderen gängigen philosophischen Modell der Verantwortung, nämlich dem „Kausalansatz“: „I am responsible for Z if and only if my beliefs and desires, as reasons, cause my doing or having Z.“ Dieser Ansatz bringt selbst bei Standardfällen Probleme mit sich. Stellen wir uns einen Bahnwärter vor, der vergisst, eine Weiche zu stellen, und dadurch ein Zugunglück auslöst. Es ist eindeutig, dass er – *ceteris paribus* – verantwortlich ist für sein Versäumnis. Der Kausalansatz kann dies jedoch nicht erfassen. In diesem Fall wäre er gewissermaßen deflationär (kein Zusprechen von Verantwortung, wo diese aber geboten wäre). Andersherum können die in der Definition genannten Wünsche und Glaubenszustände ihrerseits etwas sein, das nicht in der Verantwortung des Akteurs liegt. Denken wir an das Beispiel der Person, die eine Opiatabhängigkeit entwickelt auf Grund von langfristigen, medizinisch indizierten Morphiumgaben. Die daraus entstandenen Wünsche („desires“), das Verlangen nach Morphium also, liegen ihrerseits nicht in der Verantwortung dieser Person. In diesem Fall wäre der Kausalansatz inflationär (Zusprechen von Verantwortung, wo diese nicht geboten ist). Schließlich stellt sich das Problem der devianten Kausalkette. Nehmen wir an, ich gehe zu Hermanns Party mit der alleinigen Absicht, ihn *coram publico* zu beleidigen. Bevor ich jedoch überhaupt mit Hermann sprechen kann, erfahre ich in aller Vertraulichkeit von Susi, dass Hermann tödlich beleidigt ist durch meinen unangemessenen Kleidungsstil, wobei ich meine Kleidung tatsächlich völlig zufällig und ohne die Absicht, Hermann zu beleidigen, ausgewählt habe. Ich hätte also meine Absicht erreicht, jedoch nicht auf dem beabsichtigten Wege. Bin ich in diesem Fall verantwortlich dafür, dass Hermann beleidigt ist?

Abgesehen von den aufgeführten Problemen dieses Ansatzes mit einigen Standardsituationen, scheint es zweifelhaft, ob er überhaupt zur Erklärung der Verantwortung für Emotionen tauglich ist. Denn der Definition nach ergibt sich Verantwortung für Z nur dann, wenn meine Glaubenszustände und Wünsche – *als Gründe* – verursachen, dass ich Z tue oder habe. Die Beispiele haben deutlich gemacht, dass dies im Sinne von Gründen zum absichtsvollen Handeln gemeint ist. Zwar können Emotionen auf Grund von Glaubenszuständen und Wünschen – als Auslösebedingungen – entstehen, nicht jedoch im Sinne einer kausalen Handlungserklärung durch Gründe. Denn in letzterem Fall sind die vorliegenden mentalen Zustände intentional auf die entsprechende Handlung gerichtet,

Emotionen jedoch können auf der Grundlage anderer mentaler Zustände entstehen, ohne dass diese intentional auf die Elizitation von emotionalen Zuständen ausgerichtet sind. Gerade darin liegt das gewisse Maß an Passivität, dem das Subjekt hier unterworfen ist. Ein Ansatz, der lediglich aktuelle Kausalität erfassen kann, wird also Verantwortung für Emotionen nicht erklären können.

Anhand von Oakleys Modell der Abwendbarkeit („avoidability“) und der von ihm definierten Bedingungen der möglichen Voraussicht lassen sich jedoch emotionale Dispositionen sehr überzeugend unter die Verantwortung des Akteurs stellen. Und gerade diese Phänomene – im Sinne von emotionalen Einstellungen als Charaktermerkmale oder „habitus“ – hatte ich als die Basis einer Praxis individuellen emotionalen Wertens identifiziert. Die Implementierung einer Verantwortung für Charaktermerkmale war also notwendig, um die Verantwortung für emotionale Dispositionen erklären zu können, welche wiederum notwendig ist, um ein Modell der Verantwortung – im Sinne von *Zuschreibbarkeit* oder *Zurechenbarkeit* – für emotionale Wertungen begründen zu können. Auf der Grundlage dieses Modells in Verbindung mit den Überlegungen zur transpersonalen Geltung emotionaler Wertungen nach Mulligan und Helm kann m.E. überzeugend erklärt werden, inwiefern emotionalen Dispositionen Gültigkeit, Reliabilität und Zuschreibbarkeit zukommt, die jeweils analog zu der von bspw. Urteilen und Überzeugungen zu verstehen ist.

4. Handeln auf Grund von emotionalen Einstellungen: metaethische Situierung des Modells

4.1. Einleitung

Im vorliegenden Kapitel sollen nun die Fäden der bisherigen Argumentation zusammengenommen werden. Wie in der Einleitung dargestellt, ist das Ziel der vorliegenden Arbeit, eine Lösung des von Michael Smith diagnostizierten „moral problem“ zu versuchen, u.z. entlang der Linien eines emotiv-kognitivistischen Ansatzes. Sowohl dieser Ansatz als auch Smith's Problemdiagnose wurden in der Einleitung kurz vorgestellt. Der hier vertretene Lösungsansatz beinhaltet zumindest dem Titel nach eine konzeptionelle Paradoxie: Emotivismus und Kognitivismus sind schließlich zwei konkurrierende Paradigmen in der Handlungstheorie. Smith's Problem beruht gerade auf der Tatsache, dass wir über beide Paradigmen verfügen, und dass beide mit einiger Plausibilität aufwarten können. Ein Ansatz also, der eine Lösung des Problems gerade in der harmonischen Verbindung dieser beiden Paradigmen suchen würde, erschiene zu Recht als argumentativ zirkulär: denn wenn sich beide Paradigmen verbinden ließen, würde das Problem überhaupt nicht bestehen.

Die Paradoxie soll jedoch allein auf den Titel beschränkt bleiben. Ich verwende die beiden Begriffe, weil ich denke, dass wir sie mit einer revisionistischen Bedeutung unterlegen sollten, ausgehend von den bisher dargestellten Forschungsergebnissen, sodass das Problem m.E. tatsächlich gelöst werden könnte, da wir nachweisen würden, dass die beiden Begriffe in der Diagnose desselben nicht adäquat benutzt wurden. Im Kapitel zur Emotionstheorie habe ich zu zeigen versucht, dass Emotionen keine nicht-rationalen Phänomene sind, und dass insofern also ein kruder Emotivismus ein inadäquates, weil übermäßig stark vereinfachtes Bild dieses Phänomens zeichnet. Im Kapitel zur emotionalen Begründung von Werten habe ich wiederum zu zeigen versucht, dass wir die „Werte“ und „desires“ nicht unabhängig von Emotionalität verstehen können, ja dass sogar letztere eine konstitutive Rolle in der Erklärung von Wertungen und Wünschen spielen muss. Insofern können wir eine rein-kognitivistische Perspektive in der Handlungserklärung als relativiert ansehen, als Emotionen in der Handlungsbegründung durch die zentralen Motive „Werte“ und „desires“ mit eingehen.

Im vorliegenden Kapitel wollen wir zunächst das „moral problem“ von Michael Smith detailliert darstellen, und insofern einen ausführlicheren Anschluss an die Einleitung versuchen. Mein eigener Lösungsansatz lässt sich dann abschließend relativ kurz gefasst

formulieren in Kap. 4.9., wobei auf die relevanten Ergebnisse früherer Kapitel verwiesen sei. Zur angemessenen Situierung meiner Ausführungen sollen im Laufe der Diskussion des Problems von Smith auch allgemein Erläuterungen zu Begriffen und Modellen der Metaethik einfließen bzw. konkurrierende Positionen zu Wort kommen. Zuletzt sei noch angemerkt, dass es sich bei diesem Ergebnis lediglich um eine Stellungnahme zu Michael Smith's Ergebnissen handelt, wobei nicht zur Frage steht, ob seine Analyse per se zutreffend ist. Sicherlich kann auch Smith's im weitesten Sinne Humeanisch geprägter Theorie-Rahmen zur Disposition gestellt werden, etwa durch Kantianische Ansätze in der Erklärung menschlichen Handelns. Obwohl ich denke, dass meine bisherigen Ergebnisse auch eine sinnvolle Stellungnahme in diesem weiteren Feld der Debatte ermöglichen würden, soll diese prinzipientheoretische Fragestellung der Handlungstheorie hier ausgespart werden, um den Rahmen der Arbeit nicht zu sprengen.

4.2. Michael Smith's „moral problem“

Smith (1994) spricht von einem „concern“ im Bereich der normativen Moralphilosophie, also auch im Bereich des Nachdenkens über moralische Fragen, das jeder anstellt, und nicht allein der Theoretiker. Diese „Sorge“ dreht sich darum, die „richtigen“ Antworten zu finden, und „richtig“ zu handeln. Wir scheinen also zu denken, dass moralische *Fragestellungen* prinzipiell eine „korrekte Antwort“ zulassen. Daraus lässt sich dann weiter ableiten, dass diese Fragestellungen bestimmten (moralischen?) *Tatsachen* entsprechen, und dass sich eine Antwort also aus der korrekten Beschreibung der bestehenden *Umstände* ergibt. Wir spezifizieren die Umstände und erörtern die Tatsachen im Rahmen der Fragestellungen mittels Argumentationen und moralischem Diskurs (beides können wir unter „kognitive“ Phänomene fassen, also solche, die an den Gebrauch des Denkens und der Vernunft appellieren). Smith (1994, 6) nennt diese Beobachtung die „*objectivity of moral judgment*“.

Weiterhin diagnostiziert Smith ein zweites Prinzip von Moralität (aus metaethischer Perspektive): nämlich die „*practical implications of moral judgment*“, das er anhand eines vielzitierten Fallbeispiels illustriert:

„Suppose we are sitting together one Sunday afternoon. World Vision is out collecting money for famine relief, so we are waiting to hear a knock on the door. I am wondering whether I should give to this particular appeal. We debate the pros and cons of contributing and, let's suppose, after some discussion, you convince me that I should contribute. There is a knock on the door. What would you expect? I take it that you would expect me to answer the door and give the collector my donation. But suppose I say instead ‚But wait! I know I should give to famine relief. But what I

haven't been convinced of is that I have any reason to do so!' And let's suppose that I therefore refuse to donate.“ (ebd.)

Nach Smith beläuft sich diese zweite Merkmal auf folgende Feststellung: moralische Überzeugungen scheinen Meinungen (Glaubenszustände oder Überzeugungen) darzustellen darüber, welche Gründe wir haben, auf bestimmte Weise zu handeln. Das Haben solcher Meinungen wiederum scheint – *ceteris paribus* – mit dem Haben eines entsprechenden Handlungsmotivs einherzugehen.

Wenn wir in einem weiteren Schritt zusätzliche Prämissen hinzuziehen, die von Smith weitestgehend geteilt werden (vgl. auch Smith 1987 und Davidson 1963), wird ein vermeintliches Paradoxon offenbar. Den Bezugsrahmen für diese zusätzlichen Prämissen liefert die Humeanische Interpretation der menschlichen Psychologie. Nach Hume gibt es zwei grundsätzliche psychologische Zustände. *Glaubenszustände* stellen die Welt dar, wie sie ist. Sie sind wahrheitsfähig und tragen einen Wahrheitswert je nachdem, ob sie in der Tat die Welt adäquat repräsentieren, d.h. ob die Welt tatsächlich so ist, wie wir sie zu sein glauben. „*Desires*“ wiederum repräsentieren einen Zustand, wie die Welt sein soll.⁹² Darum unterliegen sie nicht dem Kriterium der Wahrheitsfunktionalität (wenngleich evtl. einer wahrheitsfähigen Semantik von „Erfüllungsbedingungen“, vgl. Scarano 2001).

Hume geht also davon aus, dass beide – „beliefs“ und „desires“ – distinkte psychologische Zustände sind, die sich auch jeweils auseinanderhalten lassen. Mit anderen Worten: es gibt keine notwendige Verknüpfung zwischen einem bestimmten Glaubenszustand und einem bestimmten desire. Der Beweis wird angetreten, indem die grundsätzliche Möglichkeit der folgenden Konstellation intuitiv zu erfassen sei: Wir können uns immer einen Akteur P vorstellen, der zwar einen Glaubenszustand x hat, jedoch nicht den dementsprechenden Wunsch y. Wenn das so wäre, d.h. wenn es eine notwendige Verbindung zwischen beiden gäbe, dann müsste man auch Wünsche – zumindest in abgeleiteter Form – wahrheitsfunktional beurteilen können. Hume geht davon aus, dass unsere Wünsche nicht nur nicht wahrheitsfunktional sind, sondern weiterhin, dass die von uns vertretenen Wünsche auch nicht Objekt rationaler Kritik sein können. Wünsche zeichnet eine besondere Faktizität aus: dass wir eine bestimmte Menge von Wünschen vertreten, ist schlicht eine Tatsache über uns als Person (oder Akteur), die anzuerkennen ist, auf die jedoch nicht normativ Bezug genommen werden kann. Smith resümiert dieses „standard picture“ (also die Humeanische Interpretation) der menschlichen Psychologie wie folgt:

⁹² Vgl. auch den damit zusammenhängenden Begriff des „direction of fit“, der „Passensrichtung“ also, in Anscombe 1957, überblickshaft bspw. in Scarano 2001.

„The standard picture of human psychology is important because it provides us with a model for explaining human action. Crudely, our beliefs tell us how the world is, and thus how it has to be changed, so as to make it the way our desires tell us it is to be.“ (Smith 1994, 8f.).

An dieser Stelle sollten wir einen Blick zurück werfen auf Smith's Darstellung der zwei Prinzipien der Moralität: zunächst wurde im Rahmen des *Objektivitätsprinzips* diagnostiziert, dass moralische Fragestellungen die Möglichkeit von „korrekten Antworten“ zulassen. Dieses Prinzip können wir auch einen „Kognitivismus“ in der metaethischen Analyse nennen: moralische Urteile zu fällen bedeutet, auszudrücken, wie die Sachlage der relevanten moralischen Tatsachen ist. Moralische Überzeugungen beziehen sich demnach auf den Zustand der Welt, u.z. bezogen auf eben den Zustand der moralischen Tatsachen in der Welt. Die daraus resultierende Implikation für eine Humeanische Handlungstheorie wäre, dass die Frage, welche moralischen Überzeugungen ein Akteur vertritt, völlig abgekoppelt wird von der Frage, wie dieser Akteur handeln soll. Auf der Grundlage einer gegebenen moralischen Überzeugung wäre es demnach kontingent (d.h. nicht abhängig von dieser moralischen Überzeugung), ob ein Akteur dementsprechend handelt. Denn der Glaubenszustand allein würde nicht zum Ausführen einer Handlung führen, wenn nicht ein entsprechender Wunsch (bzw. Pro-Einstellung) zugleich vorläge, und dies wiederum ist keine Tatsache, die normativer Betrachtung zugänglich ist, sondern einfach „a further fact about a person's psychology“ (ebd., 9).

Das *Prinzip der praktischen Implikationen* wiederum postuliert, dass das Haben einer bestimmten moralischen Überzeugung – *ceteris paribus* – mit dem Haben einer entsprechenden Motivation einhergeht. Wiederum ergibt sich eine Implikation für die psychologische Erklärung menschlichen Handelns, wenn wir dieses Prinzip in Beziehung zu den dargestellten Grundlinien der Humeanischen Handlungstheorie setzen: nach Hume ergibt sich durch eine rationale Betrachtung des Zustands der Welt keinerlei Zwang oder Verpflichtung, aus diesem Grund eine bestimmte Menge von Wünschen (Pro-Einstellungen, „desires“) zu haben, oder einen Wunsch x zu haben und aus dem gleichen Grund einen Wunsch y nicht. Auch hier liegt also ein Problem: das Prinzip der praktischen Implikationen und die Humeanische Handlungstheorie führen zu konfligierenden Ergebnissen. Die einzige Lösung sieht Smith vordergründig darin, moralischen Überzeugungen ihren kognitiven Gehalt abzusprechen und anzunehmen, dass moralische Überzeugungen tatsächlich expressive Äußerungen unser Pro-Einstellungen (bzw. Dispositionen zu Pro-Einstellungen) sind. Diese Strategie entspricht also dem Projekt einer expressivistischen oder emotivistischen

Moralphilosophie, wie es in ähnlicher Form vor allem von Ayer (1936), Stevenson (1937 und 1944) und Hare (1952) jeweils respektive unternommen wurde.

Dieses Ergebnis jedoch konfliktiert wiederum mit dem Objektivitätsprinzip. Wenn moralische Überzeugungen tatsächlich lediglich expressive Äußerungen unserer Pro-Einstellungen sind, dann wären sie mithin auch keiner rationalen Kritik zugänglich. Auch die Tatsache, welche moralischen Überzeugungen ein Akteur vertritt, wären dann mit Smith's Worten schlicht „a further fact about a person's psychology“, und damit eben nicht normativ zu bewerten. Dies scheint jedoch schon intuitiv nicht tragfähig zu sein und das Projekt der Moralphilosophie insgesamt in Frage zu stellen. Rücken wir jedoch von dem Postulat des Non-Kognitivismus von moralischen Überzeugungen ab, geraten wir – wie oben dargestellt – wiederum in einen Externalismus in der Erklärung von motiviertem Handeln auf der Grundlage von Überzeugungen. Tatsächlich ergibt sich also vor dem Hintergrund der dargestellten Sachlage – den zwei Prinzipien der Moralphilosophie und der Humeanischen Handlungstheorie – ein viziöser Zirkel.

Smith nennt dies „the moral problem“ (vgl. auch Smith 1989 und 1994a) und fährt fort mit der folgenden Einschätzung der Problemlage: „As I see it, the moral problem is in fact the central organizing problem in contemporary meta-ethics“ (1994, 11). Sayre-McCord (1997, 55), der die von Smith im Rahmen des von diesem diagnostizierten „moral problem“ diskutierten Thesen prüft, weist darauf hin, dass die Bezeichnung des Problems bzw. Paradoxons – und mithin auch der Titel des Buches („The Moral Problem“) – irreführend sein mag. Wenn wir das moralische Problem apostrophieren, was gewöhnlich innerhalb eines bestimmten Kontextes – zeitgeschichtlich oder gesellschaftspolitisch z.B. – geschieht, so meinen wir damit Phänomene wie den „Verfall der Werte“, institutionelle Ungerechtigkeiten oder internationales politisches Unrecht etc. Michael Smith's „moral problem“ ist jedoch – wie Sayre-McCord richtig anmerkt – gar kein moralisches, sondern ein metaethisches Problem. Dies wird natürlich auch an Smith's Ausführungen (vgl. den eingangs in diesem Abschnitt zitierten Satz) deutlich, sei jedoch nur als klärend Anmerkung unserer Untersuchung zur Seite gestellt.⁹³

⁹³ Sayre-McCord denkt weiterhin, dass Smith Analyse nicht direkt etwas über unsere moralische *Praxis* aussagt, sondern nur über unser *Verständnis* dieser Praxis: „[I]f, as I believe, our thought is not after all that incoherent, then the problem is a metaethical one that plagues our understanding of our own practice but does not itself constitute a fundamental problem within the practice“ (1997, 55 FN 1). Ich halte diesen Einwand für nicht zutreffend, da bei der Moral, wie bei anderen kulturellen Praktiken, stets von einer Wechselwirkung zwischen der Praxis selbst und unserem Selbstverständnis als daran teilnehmenden Akteuren sowie unserem Verständnis

Smith bringt das Moral Problem in eine Formulierung von drei miteinander unvereinbaren Aussagen (vgl. McNaughton 1988, 23):

- „1. Moral judgments of the form ‚It is right that I Φ ‘ express a subject’s beliefs about an objective matter of fact, a fact about what it is right for her to do.
2. If someone judges that it is right that she Φ s then, ceteris paribus, she is motivated to Φ .
3. An agent is motivated to act in a certain way just in case she has an appropriate desire and a means-end belief, where belief and desire are, in Hume’s terms, distinct existences.“⁹⁴ (Smith 1994, 11)

Nach Smith liegt die Unvereinbarkeit in folgender in den drei Aussagen enthaltenen Argumentationskette: nach (1) wird eine moralischen Überzeugung als Glaubenszustand identifiziert, der entsprechend (2) notwendig mit einer Handlungsmotivation in Verbindung steht. (3) sagt nun aus, dass eine solche Motivation im Haben einer bestimmten Pro-Einstellung (desire) besteht. Die Argumentation (1, 2, 3) führt also zu dem Postulat, dass eine notwendige Verbindung zwischen Glaubenszuständen und Pro-Einstellungen besteht. Dies jedoch wird von (3) bestritten.

Nach Smith bieten sich vordergründig folgende theoretische Rahmenbedingungen für eine Lösung des „moral problem“:

- (I.) Theoretiker, die der Humeanischen Handlungstheorie treu bleiben wollen, haben die Option entweder (Ia) das *Objektivitätsprinzip* und damit die Aussage (1) zu verwerfen. Hierbei wird angenommen, dass moralische Überzeugungen eben keine Glaubenszustände seien. (Diese Strategie wird u.a. verfolgt von Ayer 1936, Hare 1952, Blackburn 1984, 1986 und 1987, und Gibbard 1990). Die andere Option in diesem Zusammenhang besteht darin, (Ib) das *Prinzip der praktischen Implikationen* und damit die Aussage (2) zu verwerfen. Bestritten wird hier also, dass eine notwendige Verbindung zwischen moralischen Überzeugungen und entsprechender Motivation besteht. (Diese Strategie – bekannt als „Externalismus“

dieser Praxis selbst auszugehen ist. Probleme, die sich auf unser *Verständnis* einer Praxis beziehen, betreffen somit auch die Praxis *selbst*.

⁹⁴ Sayre-McCord bringt dieselben drei Prämissen auf einen leicht abgewandelten Wortlaut und bezieht diese dabei auf die unabhängig von Smith’s eigener Terminologie populären Begrifflichkeiten „Kognitivismus“, „Internalismus“ und „Humeanismus“, „1. Moral judgments express beliefs about matters of fact (call this ‚cognitivism‘). 2. If someone makes a moral judgment then she is motivated appropriately, ceterus [sic] paribus (call this ‚internalism‘). and 3. Beliefs can motivate only thanks to their interaction with conceptually independent desires (call this ‚Humean psychology‘)“ (Sayre-McCord 1997, 56)

– wird verfolgt von Frankena 1958, Foot 1972, Scanlon 1982, Railton 1986 und Brink 1986 und 1989).

- (II.) Theoretiker, die insgesamt mit der Humeanischen Handlungstheorie brechen, werden hingegen eine einheitliche Option verfolgen, nämlich die Aussage (3) zu verwerfen, d.h. mit der Annahme brechen, dass Motivation zu erklären sei durch Pro-Einstellungen und zweckrationale Glaubenszuständen, wobei beide gemäß Hume als distinkte Phänomene angesehen werden. (Diese Strategie, die traditionell eher einer Kantianischen Position entspricht, wird u.a. verfolgt von Nagel 1970, McDowell 1978, Platts 1981, McNaughton 1988 und Dancy 1993.)

Wie bereits eingangs in diesem Kapitel angemerkt, will ich im Folgenden keine breit angelegte Abwägung des Humeanischen gegen den Kantianischen Rahmens unternehmen, wenngleich eine solche Auseinandersetzung fruchtbare Ergebnisse auch für die vorliegende Debatte liefern kann. Es liegt eine große Zahl entsprechende Studien von herausragender Qualität vor. Wie wir später sehen werden, liegt Smith's eigene Lösung weitestgehend selbst im Humeanischen Rahmen, weswegen auch ich auf Grund des „principle of charity“ diesen Rahmen nicht verlassen und eine mit diesem kompatible alternative Lösung vorschlagen möchte. Wir werden also an späterer Stelle Smith's Ausführung weiter verfolgen, bevor wir schließlich selbst Stellung zur Lösungsstrategie nehmen. Zunächst jedoch scheint es ratsam, einige Grundbegrifflichkeiten und Theoriepositionen in der Metaethik, die bei Smith bereits angeklungen sind, zu klären.

4.3. Internalismus und Externalismus

Praktische Rationalität muss zwei Bedingungen erfüllen: Gründe zum Handeln müssen einerseits Motive sein (also Handlungen erklären), andererseits müssen sie normative Gründe sein (also Handlungen rechtfertigen). Ein beträchtlicher Teil der metaethischen Forschungsdebatte wird um die Frage geführt, welche theoretische Position beide Bedingungen erfüllen könne. Die Humeanische (Gosepath 2002: „internalistische“) Position kann sehr plausibel die motivationale Wirkung von Gründen erklären, ohne jedoch eine zufriedenstellende Erklärung ihrer Normativität liefern zu können. Demgegenüber können anti-Humeanische (Gosepath: „externalistische“) Positionen die normative Dimension von Gründen erklären, geraten aber in Schwierigkeiten, wenn es um die Erklärung ihrer motivationalen Wirkung geht, da ja der Externalismus geradezu programmatisch vom Akteur und dessen desires und Motivlage absieht.

Wir haben bereits in Smith's Worten die beiden Desiderata einer Moraltheorie aus metaethischer Perspektive kennengelernt: das Objektivitätsprinzip einerseits und das Prinzip praktischer Implikationen andererseits. In andere Worte gefasst finden wir diese Beobachtung auch bei Stefan Gosepath:

„An action is rational only if it is carried out for the right kind of reason, one that is motive and obligation at once. A theory of reasons should show both that and how a reason justifying an action is also capable of contributing to the explanation of the action, and likewise, that a reason explaining an action is also capable of justifying an action.“ (Gosepath 2002, 230)

Die zentrale Frage dieser Arbeit (und natürlich auch von Smith's eigener Studie) ist natürlich, wie eine Koordinierung beider Prinzipien möglich sei. Und genau diese Fragestellung bildet die Internalismus-Externalismus-Debatte ab. Im Rahmen des Humeanischen Modells menschlicher Motivation sind Gründe diejenigen Verbindungen von Überzeugungen und desires, die in einem gegebenen Fall handlungsrelevant sind. Die Humeanische Position können wir damit als *internalistisch* bezeichnen (vgl. prominent Williams 1980), wobei dieser Begriff wie folgt zu verstehen ist:

“Internalism is the view that the presence of a motivation for acting morally is guaranteed by the truth of ethical propositions themselves. On this view the motivation must be so tied to the truth, or meaning, of ethical statements that when in a particular case someone is (or perhaps merely believes that he is) morally required to do something, it follows that he has a motivation for doing it” (Nagel 1970, 7)

Praktisches Überlegen, welche Handlungsoptionen verwirklicht werden sollen, geschieht im Rahmen der Humeanischen Perspektive auf der Grundlage der desires und der relevanten Überzeugungen des Akteurs. Nach Gosepath (2002, 230) erfüllt das Modell sowohl die Anforderung, Handlungen erklären zu können, als auch sie rechtfertigen zu können.

Überzeugungen und desires, die in einem gegebenen Fall handlungsrelevant werden, *rechtfertigen* die entsprechende Handlung, wenn gezeigt werden kann, dass der Akteur die Handlungsoption gewählt hat, von der er glaubte, dass sie am besten seine vorliegenden desires erfüllen oder befördern könne. Gosepath spricht von drei Schritten in der Rechtfertigung von Handlungen nach dem belief-desire Modell: „Instrumentelle Rationalität“ ist die Bezeichnung für die Art von Handlungsrechtfertigung, die nur mit dem Ausweis einer Zweck-Mittel-Beziehung zwischen der Handlung und dem relevanten desire befasst ist. „Prudentielle Rationalität“ schließt darüber hinaus noch ein, dass die relevanten desires und Überzeugungen durch Überlegen kohärent gemacht wurden, und dass der Akteur sein eigenes Wohlergehen hinzugezählt hat. Ein weiterer Schritt, und wie ich denke: eine andere Kategorie rationaler Rechtfertigung, wird erreicht, wenn darüber hinaus nachgewiesen werden kann, dass auch die vorliegenden desires und relevanten Überzeugungen selbst rational, d.h. wohl-

begründet, -überlegt und -durchdacht, sind. In diesem Fall spricht Gosepath von „substanzieller Rationalität“.⁹⁵

Überzeugungen und desires, die in einem gegebenen Fall handlungsrelevant werden, *erklären* die entsprechende Handlung (d.h. sind motivational wirksame Faktoren in der Darbietung der Handlung) mit Bezug auf die Humeanische Motivationstheorie (Hume, Treatise II 3,3). Denn nach Hume besteht Motivation aus einem relevanten desire und einer Zweck-Mittel-Überzeugung. Hier setzt auch Williams' berühmte Internalismus-Externalismus These an (1980, vgl. auch 1995a). Denn nach Williams zeigt die Humesche Motivationstheorie, dass die Behauptung, es bestehe ein externer Grund (ein normativer Grund) für Hans, die Tür zu öffnen, nicht die Existenz eines Motivs (eines internen bzw. motivierenden Grundes) für Hans impliziert, die Tür zu öffnen. Williams schlussfolgert also, dass externe (normative) Gründe nicht zur Erklärung von Handlungen taugen.

„In Williams' view, if they are to be motivational, justifying reasons have to exist relative to the agent's ‚subjective motivational set‘. [...] In order for external grounds to exist, rationality must be in the position of producing a completely new motive without proceeding from subjectively given motives by means of a sound deliberative route. It is unclear how this can work.“ (Gosepath 2002, 230f.)

Die Neo-Humeaner stimmen daher darin überein, dass Motivation nur dann möglich ist, wenn ein entsprechendes desire vorliegt (Smith 1987, Blackburn 1995, Dreier 1997). Die (Neo-)Humeanische Konzeption von Handlungsgründen, die eine *internalistische* Theorie darstellt, steht wiederum den verschiedenen *externalistischen* Konzeptionen praktischer Rationalität gegenüber. Diese lassen sich nach Gosepath wie folgt charakterisieren:

„The various externalist models are united in disputing the need of rational reasons for action to be based consistently on motives of the acting person. In contrast, they maintain that justificatory grounds exist that are valid independently of a person's contingent motivational constitution.“ (Gosepath 2002, 231)

Damit wird die Existenz *kategorischer* Gründe postuliert, neben den aus dem Internalismus bekannten *hypothetischen* (d.h. zweckrationalen) Gründen. Thomas Scanlon (1999, Kap. 1) bspw. sieht die Einsicht in normative Weltansprüche als ein Beispiel, das eine kategorische Verpflichtung jenseits der faktischen Motivlage des Akteurs darstellt.

Als solche externalistischen Gegenpositionen zum (internalistischen) Humeanismus figurieren zwei Hauptschulen, der Kantianismus (O'Neill 1989, Korsgaard 1996 und 1996a,

⁹⁵ Gerade in Hinsicht auf die Erklärung dieser substanziellen Rationalität steht der Humeanismus vor einem enormen Problem, das ich mit dem Begriff der Wahrheitsfähigkeit zu fassen versucht habe. „Desires“ zeichnen sich durch ihre Faktizität aus, die jedoch nicht wahrheitsfähig in einem normativen Sinne ist. An diesem Punkt muss eine alternative Theorie rechtfertigender (also normativer) Gründe ansetzen.

Hampton 1998) und der Aristotelismus (McDowell 1979 und 1985, Foot 2001). Im Rahmen externalistischer Theorieformen ist auch der Einwand erhoben worden, dass praktisches Überlegen nicht nur – wie im Humeanismus – Zweckrationalität ausdrückt, sondern auch im Sinne von Gosepaths „substanzieller Rationalität“ ein Überlegen hinsichtlich der Handlungsziele und –zwecke sein kann. Mit den Worten eines programmatischen Aufsatztitels (Kolnai 1977) könnte man diesen Einwand formulieren als: „deliberation is of ends“ (vgl. auch Wiggins 1987, Richardson 1994).

Es lassen sich sechs typische Einwände gegen die internalistische, neohumeanische Position aufzeigen (vgl. Gosepath 2002)

(1.) *Inbetrachtziehung der eigenen Zukunft*: Intuitiv sehen wir es als irrational an, wenn ein Akteur durch die mögliche Vermeidung eines zukünftigen Übels oder die mögliche Erreichung eines zukünftigen Guts völlig unbeeindruckt bleibt. In der Tat stellen jedoch solche Formen von zukunftsorientierter Planung oder Kooperation mit anderen schwierige Fälle für das belief-desire Modell dar. Denn wie wäre es in diesem Rahmen zu erklären, dass Akteure Vorkehrungen für eine ungewisse Zukunft treffen oder sich mit den Kontingenzen einer Kooperation mit anderen abgeben, wenn sie dadurch momentan bestehende desires frustrieren würden? Zu diesem Punkt werden gemeinhin drei Gegenargumente vertreten (vgl. Gosepath 2002, 232): (a) in eben diesen Fällen wird ein latentes, dispositional vorliegendes desire seitens des so handelnden Akteurs angenommen, sich um seine Zukunft zu kümmern. Aus Erfahrung wissen Akteure, dass bestimmte Wünsche in zukünftigen Situationen sich phänomenal aufdrängen werden und haben aus diesem Grund auch aktuell ein (latentes, dispositionales) desire, sich um diese zukünftigen, lediglich zu erwartenden desires zu kümmern. (b) Auch ein Modell „struktureller Rationalität“ (Nida-Rümelin 2001) wird als Lösungsvorschlag bemüht. Im Sinne von Odysseus' selbstaufgelegter Fesselung an den Mast sollen zukunftsorientierte Handlungen verstanden werden als Teil einer globalen Struktur von Handlungen, die ein Akteur im Ganzen vertritt und durch sein Handeln befördert. Handlungen, die also punktuell (zeitlich oder kontextuell) als irrational gelten müssten, da sie faktisch vorliegende desires frustrieren, können dennoch im Rahmen einer solchen selbstgewählten, globalen Handlungsstruktur als rational gelten. Es handelt sich also um das Prinzip der momentanen Selbstbeschränkung zu Gunsten von zukünftigen Vorteilen (Elster 1979, 2000). Bratman (1983) gibt jedoch zu bedenken, ob der Humeanische Theorierahmen, der in seiner aktuellsten Form höchstens den Begriff der pro-Einstellung zur Unterfütterung des desire-Begriffs anbieten kann, ein deskriptiv ausreichend starkes Modell darstellt, um eben diese Formen von Zukunftsplanung als langfristige Nutzenabwägung theoretisch adäquat

fassen zu können. Hier handelt es sich aber evtl. nicht um eine grundlegende Inkompatibilität des (Neo-) Humeanismus, als vielmehr um eine konzeptionelle Schwäche. (c) Schließlich können auch im Sinne von Nagel (1970) zukünftige desires schlicht als per se schon rationale Gründe verstanden werden. Auch wenn wir zukünftige desires momentan noch nicht phänomenal erleben, so haben diese dennoch den Effekt von veritablen Handlungsgründen und können daher als solche gezählt werden, obwohl sie nicht aus der aktuell faktischen Motivlage des Akteurs heraus zu erklären sind.

(2.) *Die Quellen der Motivation*: Ein viel grundlegenderer Einwand zielt auf die Stellung von „desire“ zwischen Handlungserklärung und normativen Gründen ab. Nach Ansicht verschiedener Gegner des Neo-Humeanismus muss die Erklärung von Handlungsmotivation nicht auf bereits vorliegende desires rekurrieren (Nagel 1970, Kap. 5; McDowell 1978; Wallace 1990; Quinn 1993; Raz 1999a). Gemäß diesem Ansatz sind desires nicht die Grundlage für Gründe, sondern ergeben sich selbst wiederum aus rechtfertigenden und erklärenden Gründen, d.h. desires können selbst „motiviert“ sein (vgl. Steinfath 2001). Sowohl Nagel (1970) als auch Korsgaard (1986) akzeptieren die internalistische Forderung, dass Handlungen durch ein Motiv erklärt werden müssen, d.h. letztlich durch das Vorliegen eines desires. Allerdings – und dies entgegen den Humeanismus – sind sie davon überzeugt, dass desires sich wiederum selbst oft aus Überlegung ergeben, d.h. aus der Erwägung von normativen Gründen. Damit wären nicht sie die eigentlich Quelle der Motivation, sondern diese Position käme in letzter Instanz den normativen Gründen und dem praktischen Überlegen zu.

(3.) *Der Standpunkt praktischen Überlegens*: weiterhin wird dem belief-desire Modell vorgeworfen, dass es mit einem inkorrekten Begriff des Willens operiert. Nicht die in uns vorliegenden desires bestimmen den Kurs unseres Handelns, sondern unser Wille: wir wollen bestimmte Handlungsoptionen einschlagen gegenüber anderen Optionen, weil wir gute Gründe für diese Optionen erwogen haben und gute Gründe gegen andere, konkurrierende Optionen. Wir überlegen uns also, was in einer gegebenen Situation Inhalt unseres Willens sein sollte in Ansehung der bestmöglichen Handlungsoption für uns. Dieser Standpunkt zeichnet also ein Bild von uns als Akteuren im Sinne von „autonomen Deliberatoren“. Willensentscheidungen werden in den meisten Fällen aktiv getroffen. Wie jedoch hängt dieses Bild zusammen mit demjenigen, das der Humeanismus von uns zeichnet, als Akteuren, die der Handlungsmotivation durch bestehende desires und Überzeugungen gewissermaßen passiv unterliegen? Verschiedentlich ist darauf hingewiesen worden, dass die letztere Auffassung schon unserem Selbstverständnis als Personen widerspricht (Quinn 1993; Williams

1995). Der Konflikt läuft evtl. letztlich darauf hinaus, welchem Gesichtspunkt der Vorrang zukommt: entweder dem hermeneutisch rekonstruierten Selbstverständnis unserer moralischen Praxis (d.h. wir handeln aus Gründen allein, weil wir uns schließlich als autonome Personen *verstehen*), oder der empirisch plausibleren Handlungserklärung durch motivationale mentale Zustände? Schließlich könnte diese quasi-empirische Analyse menschlicher Motivation im Sinne Humes doch letztgültig zutreffend sein, selbst wenn es uns aus der inneren (erste-Person-) Perspektive nicht so erscheint (Tugendhat 1993): unsere vermeintlich Autonomie als Personen könnte ja nur eine Illusion sein, gewissermaßen ein hermeneutischer Überbau der psychologistischen Handlungserklärung. Zur Klärung dieser wichtigen Frage fordert Gosepath (2002, 233) – wie ich denke zu Recht – eine theoretisch reflektierte Epistemologie von „desire“.⁹⁶ Schließlich hängt eine bestimmte Konzeption von praktischer Rationalität direkt von dem implementierten Begriff des „desire“ ab und *vice versa*. Die Frage ist also nicht zu entscheiden in Absehung der konkreten inhaltlichen Ausgestaltung der desire-Konzeption: „Theories of practical reasoning, mental ontologies, and moral psychology are very tightly linked. That means in turn, that it is far from clear what the direction of explanation is – from the former to the latter or the reverse“ (ebd.).

(4.) *Rechtfertigung*: Einer der zentralen Schwachpunkte des belief-desire Modells – zumindest aus der Sicht seiner Gegner – ist, dass es die Rechtfertigung von Handlungen nicht erklären kann, sofern desires als faktisch vorliegende, dem Akteur nicht zugängliche Handlungsdispositionen verstanden werden. Desires rechtfertigen Handlungen zwar aus einem begriffsanalytischen Standpunkt heraus (vgl. die oben gemachten Unterscheidungen zwischen verschiedenen Begriffen praktischer Rationalität), nicht aber vom Standpunkt praktischen Überlegens aus. Eine Entität, die Handlungen von diesem Standpunkt aus (also im Sinne einer „substanziellen Rationalität“) zu rechtfertigen geeignet wäre, muss wahrheitsfunktional sein. Gerade diese Eigenschaft wird desires gewöhnlich abgesprochen. Nach Ansicht der Kritiker des Humeanismus (Quinn 1993; Velleman 1992; Hampton 1998; Korsgaard 1997; Raz 1999b und 1999c; Dancy 2000a) können desires als faktisch vorliegende Zustände (im Sinne eines „Seins“) keine normativen Ansprüche (im Sinne eines „Sollens“) begründen: Normativität muss auf Wertüberzeugungen, -urteilen und –vorstellungen aufbauen, und nicht auf einer Faktizität. Der auf die Rechtfertigung abhebende Einwand gegen den Humeanismus tritt in unterschiedlicher inhaltlicher Ausgestaltung auf.

⁹⁶ Ich denke, dass meine Ausführungen in Kap. 3 eine interessante Grundlegung für ein solches Unternehmen liefern; leider werden wir jedoch im Rahmen dieser Arbeit nicht vertiefend in eine solche Diskussion eintreten können.

Elijah Millgram bspw. (1997) vertritt einen sog. „praktischen Empirizismus“, dem gemäß Ziele, pro-Einstellungen und Bewertungen nur darum eine sinnvolle Handlungsanleitung geben können, weil sie durch praktische Lebenserfahrung informiert sind. Wir lernen aus Erfahrung, welche Phänomene für ein menschliches Leben von zentraler Bedeutung sind, über einen Prozess der „praktischen Induktion“ (ebd.). Im Verlauf dieses Prozesses sammeln wir Daten für diese Induktion anhand der Leitlinie von hedonisch positiv getönten (also „lustvollen“) Erfahrungen: der Lust kommt also die Funktion eines Indikators von „desirability“ zu (Millgram 1993). Oder wie Gosepath sehr treffend den Unterschied dieser Konzeption zu einem humanischen Hedonismus herausstellt: „Pleasure’s role in cognition is not that of a goal but that of a guide“ (Gosepath 2002, 234).

Weiterhin besteht die Möglichkeit nachzuweisen, dass desires durchaus eine evaluative Funktion bzw. einen evaluativen Aspekt haben, bspw. dann, wenn wir uns selbst bestimmen, indem wir uns zu unseren eigenen (first-order) desires verhalten, ihnen gegenüber eine wertende Stellung einnehmen (Wallace 1999a). In diesem Sinne können wir Frankfurts Theorie von höherstufigen Wünschen und Volitionen verstehen. Normativität entsteht im Rahmen dieser Konzeption aus der Tatsache, dass wir uns zunächst wertend gegenüber unseren niedrig-stufigen Wünschen verhalten, und uns daraufhin eine Selbstbestimmung geben, was für eine Person wir zu sein wünschen. Dieser Wunsch wiederum – was für eine Person wir sein wollen, und damit welche first-order desires wir wollen bzw. ablehnen – ist ein höherstufiger Wunsch bzw. Volition, der zwar selbst wieder nach dem desire-Paradigma verstanden wird, aber kraft der wertenden Stellungnahme mit Normativität versehen ist. „Wünschen“ und „Wollen“ bei Frankfurt sind also im Vergleich zu klassisch Humeanischen Konzeptionen weiterentwickelte Begrifflichkeiten, die nicht nur faktisch vorliegende („naturalistische“) Handlungsdeterminantien beschreiben, sondern evaluativ aus praktischem Überlegen und unserer Konzeption der eigenen Person hervorgehen. Diese Ganzheitlichkeit der Perspektive einer Identifikation mit einer „Person, die wir zu sein wünschen“ bringt eine normativ gehaltvolle Bewertung auf der Ebene höher-stufiger desires hervor (Frankfurt 1998).

Im Sinne Korsgaards entsteht die Normativität von Handlungszielen aus dem Willen, u.z. genauer aus der autonomen Selbstbestimmung unseres Willens im Sinne Kants (Korsgaard 1996, § 3.3.1; 1997). Normativität ist damit notwendig an unsere praktische Identität als selbstbestimmte Akteure gebunden, weil dieses Selbstverständnis als autonome Akteure uns den Willen als eine Instanz der Selbstbestimmung erfahren lässt, die Ziele unabhängig von vorliegenden Neigungen, Wünschen oder Dispositionen souverän bestimmen kann.

In Verteidigung der internalistischen Position haben sich wiederum Vertreter einer (neo-)humeanischen Linie zu Wort gemeldet und angemerkt, dass auch der von ihnen vertretene desire-Begriff als evaluativ gehaltvoll verstanden werden könne. So weist Davidson (1978) darauf hin, dass desires konstitutiv von Wertüberzeugungen abhängen, insofern wir Objekte stets nur *sub specie boni*, also im Ansehen eines Gutes, wünschen oder erstreben („we desire“). Velleman (1992) und Stocker (1979) haben jedoch Gegenbeispiele vorgebracht, dass es nicht unplausibel ist, dass gerade die „Schlechtigkeit“ eines Objekts oder Handlungsziels eben dieses für den Akteur „desirable“ macht. Die Formel, dass desires stets auf einem positiven Wert aufbauen, trifft also nicht uneingeschränkt zu. Dennoch wird man fragen müssen, ob die erwähnten Gegenbeispiele die These *in toto* widerlegen können, oder ob es sich nicht vielmehr um kuriose Ausnahmefälle von der leitenden Theorie handelt.

Eine besonders harte Linie in der Verteidigung der Normativität von desire wird von Stampe (1987) vertreten, dem gemäß desires schlicht per se normative Gründe darstellen. Denn desires implizieren nach Stampe begriffslogisch, dass das propositionale Objekt dieses mentalen Zustands als ein „Gut“ angesehen wird, wonach einem desire wiederum kraft dieses Wertungscharakters Normativität zukommt.

(5.) *Objektive Sachverhalte*: Eine besonders radikale Kritik des belief-desire Modells finden wir im Rahmen von realistischer Interpretation des Begriffs „Grund zum Handeln“. Es handelt sich insofern um einen besonders radikalen Bruch mit dem belief-desire Modell, als dieses Gründe als mentale Zustände auffasst („Psychologismus“). Eine realistische Position, im Sinne Bittners oder Dancys, postuliert hingegen, dass Gründe zum Handeln vielmehr objektive Sachverhalte darstellen (wie z.B. „dass das Wetter schön ist“, also abstrakte, nicht-dingliche Sachverhalte). Bittner (2001, Kap. 4) vertritt die These, dass der narrative Charakter von Handlungserklärung ein Indiz dafür sei, dass es sich bei Gründen zum Handeln eben um die Stellungnahme zu einem objektiven Sachverhalt handelt, und nicht um einen mentalen Zustand. Allerdings spricht er aus diesem Grunde (ebd., Kap. 9) Gründen auch ihren normativen Stellenwert ab. Demgegenüber will Dancy (1995 und 2000, Kap. 5) die normative Dimension von motivierenden Gründen erhalten.

(6.) *Der Status von Normen von Rationalität*: Rechtfertigungsverhältnisse können auch über Rationalitätsnormen bestimmt werden. Was rational geboten ist, ist geboten auf der Grundlage von bestimmten Rationalitätsnormen. Prominent lässt sich hier Kants Bestimmung instrumenteller Rationalität anführen, dass wer den Zweck will, darum per se auch das Mittel wolle (Kant, Grundlegung, AA IV 417). Dieser Zusammenhang soll nach Kant begriffsanalytisch bestehen. Die Schwäche der Kantischen Argumentation liegt jedoch darin,

dass sie Irrationalität *a priori* ausschließt. Diese Konzeption, in den Sozialwissenschaften als „Revealed-Preference-Account“ bekannt geworden, läuft darauf hinaus, dass Handlungszwecke nur empirisch bestimmt werden können. Denn wer ein Mittel – etwa als fragwürdig – ablehnt und aus diesem Grund den Handlungszweck nicht realisiert, dem würde eben dieser Zweck abgesprochen. Irrationalität ist damit faktisch ausgeschlossen.

Mit Blick auf diese Kategorisierung der Einwände gegen den Humeanischen Internalismus durch Gosepath, möchte ich festhalten, dass der Punkt (4), nämlich die Frage der *Rechtfertigung*, zentral zu sein scheint. Nicht weniger wichtig sind sicherlich diejenigen Fragen, die das Selbstverständnis menschlicher Akteure angehen: der Humeanismus scheint fundamental unseren Intuitionen zu widersprechen, wenn er bspw. postuliert, dass der Vernunft lediglich eine zweckrationale (oder instrumentalistische) Funktion zukomme. Tatsächlich denken wir jedoch immer auch über unsere Handlungsziele nach. Solche Erwägungen sind jedoch durch einen eher tugendethisch orientierten Ansatz zu klären. Im Folgenden wollen wir uns verstärkt auf die Untersuchung von Handlungsgründen konzentrieren. In dieser Hinsicht ist deutlich geworden, dass der Humeanismus sich großen Problemen gegenüber sieht, Rechtfertigung zu erklären, u.z. in verschiedener Hinsicht: Erstens scheint die Faktizität des desire-Begriffs keine normativ gehaltvollen Aussagen zu erlauben, zweitens scheinen rechtfertigende Gründe in Bezug zu objektiven Tatsachen der Außenwelt zu stehen, und nicht allein in einem mentalen Zustand (desire) des Akteurs zu bestehen. Und drittens scheint unklar, inwiefern Rationalität, als Garant einer normativen Aufladung der Humeanischen Motivationstheorie, in die Formierung von desires eingehen könnte.

Warum sollten wir jedoch überhaupt an einem Humeanischen Rahmen festhalten? Ich denke, dafür spricht die schlichte und überzeugende Beobachtung von Smith (vgl. o.), dass die prinzipielle Möglichkeit einer Dissoziation von Überzeugungen und desires beweise, dass es sich um zwei distinkte Entitäten handle. Wir kennen empirisch Fälle von „Akrasie“, in denen wir zwar eine starke praktische Überzeugung vertreten, jedoch mangels Motivation (oder „Lust“) nicht die entsprechende Handlung realisieren. Handlungsgründe, zumindest insofern sie eine Handlung motivational erklären, sollten offenbar als mentale bzw. psychologische Größen aufgefasst werden. Eventuell wird der Widerspruch, der zwischen dieser Annahme, und derjenigen, dass Handlungsgründe (insofern sie eine Handlung normativ rechtfertigen) objektive Tatsachen darstellen, zu lösen sein, indem wir zwischen zwei Kategorien von Gründen unterscheiden. Diese vordergründige Lösung wiederum bereitet den

Boden vor für das eigentliche in dieser Arbeit diskutierte Problem, nämlich wie die beiden Kategorien von Handlungsgründen zu koordinieren seien.

Der Begriff „Internalismus“ jedoch – als Verbindung zwischen moralischen Tatsachen bzw. moralischen Überzeugungen einerseits und Motivation andererseits – wird in der philosophischen Debatte auf verschiedene Art gebraucht (Smith 1994, 60f.). Eine Form sieht Smith in der These „If an agent judges that it is right for her to Φ in circumstances C, then she is motivated to Φ in C.“ Oder, wie er sagt, in anderen Worten bringen moralische Urteile Motivation mit sich *simpliciter*. Diese These bezeichnet Smith als *starken Internalismus*, der z.B. auch Phänomene wie Willensschwäche nicht erfassen kann, und damit in Bezug auf einige empirische Tatsachen unplausibel erscheinen mag. Smith schlägt eine plausiblere Variante der These vor, die er „defeasible internalism“ nennt (vgl. auch Pettit & Smith 1993a): „If an agent judges that it is right for her to Φ in circumstances C, then either she is motivated to Φ in C or she is practically irrational.“⁹⁷ Wir sehen hieran, dass das Kriterium des Motivationsinternalismus zwar ein wichtiges Desiderat jeder Moraltheorie sein muss (und die Implementierung desselben ist ja gerade der Gegenstand unserer Untersuchung). Andererseits ist ein zu starker Internalismus auch wiederum intuitiv (und empirisch) nicht akzeptabel. Wir wollen zwar in der Lage sein erklären zu können, warum wir auf Grund

⁹⁷ Smith identifiziert weiterhin eine Variante der Internalismus-These, die nicht so sehr auf die praktischen Implikationen moralischer Überzeugungen abhebt, sondern auf die Rationalität moralischer Überzeugungen. Diese These lässt sich nach Smith wie folgt formulieren: „If it is right for agents to Φ in circumstances C, then there is a reason for those agents to Φ in C“ (Smith 1994, 62). Eine solche These wird vertreten von Korsgaard (1986). Die Rationalismus-These schließt den „defeasible internalism“ ein: Ein Akteur hat einen Grund, auf bestimmte Weise zu handeln, wenn er motiviert wäre so zu handeln, falls er rational ist (Korsgaard 1986). Daraus folgt, dass ein Akteur, der glaubt, einen Grund zu haben, auf bestimmte Weise zu handeln (d.h. der annimmt, er wäre dazu motiviert, falls er rational wäre), dann praktisch irrational ist, wenn er eben nicht dazu motiviert ist. Wie Smith (1992) sagt, wäre der Akteur praktisch irrational nach seinen eigenen Standards. Daher schließt die Rationalismus-These den „defeasible internalism“ ein: denn gemäß ersterer entspricht die Überzeugung, dass es richtig ist, auf bestimmte Weise zu handeln, der Überzeugung, dass es einen Grund gibt, auf diese Weise zu handeln. Nach Smith ist dieses Implikationsverhältnis jedoch nicht umkehrbar: der „defeasible internalism“ schließt nicht die Rationalitäts-These ein. Man könnte bspw. im Sinne des Expressivismus argumentieren, dass das Vertreten einer Überzeugung mit einer entsprechenden Motivation einhergeht, bzw. falls nicht, dass der betreffende Akteur in diesem Falle praktisch irrational wäre. Dennoch würde der Expressivismus in Abrede stellen, dass diese Überzeugungen – moralische Anforderungen also – Vernunftforderungen seien, und daher die Rationalitätsthese ablehnen. Zwar gilt die These vom Motivationsinternalismus (nämlich in Bezug auf die expressiv geäußerten Präferenzen), aber die zu Grunde liegende Kategorie ist in diesem Fall keine rationale (im Sinne eines objektiven *Grundes* zu handeln).

unserer praktischen Überzeugungen handeln. Wir können jedoch wiederum nicht theoretisch postulieren, dass wir *stets* auf Grund unserer Überzeugungen handeln, denn das ist empirisch gesehen nicht der Fall. Das Kriterium des Motivationsinternalismus stellt also ein Rationalitätskriterium dar: wir akzeptieren zwar, dass es empirisch möglich ist, dass ein Akteur nicht auf Grund seiner praktischen Überzeugungen handelt, denken jedoch weiterhin, dass wenn er dies nicht tut, dass dies erklärungsbedürftig ist. Wir halten einen solchen Akteur (bzw. einen Akteur in einer solchen Situation) für praktisch irrational.

Auch Roskies (2003) untersucht in Ihrer Studie die These des „belief-internalism“⁹⁸. Ihrer Meinung nach steht diese These vor einem Dilemma: die Annahme, dass moralische Überzeugungen eine intrinsische motivationale Wirkung haben, ist entweder als schwache These zu verstehen und damit nicht mehr substantiell informativ. Oder diese Annahme ist als starke These zu verstehen, in welchem Fall Roskies sie für falsch hält. Interessanterweise glaubt Roskies, dass gerade VM-Patienten, also typische Damasio Patienten (vgl. Kap. 2.8. der vorliegenden Arbeit), „lebende Gegenbeispiele“ sind, die Indizien gegen das Zutreffen einer starken Internalismus-These liefern.

Roskies (2003, 52) identifiziert drei Prinzipien des Internalismus, nämlich (1.) *Notwendigkeit*: moralische Überzeugungen sind notwendig mit Motivation zum Handeln verbunden. (2.) *Intrinsizität*: die internalistische Verbindung zwischen moralischen Überzeugungen und Handlungsmotivation besteht kraft des Inhalts der Überzeugungen, und nicht auf Grund anderer, externer Begleitumstände. Wie Alfred Mele (1996, 730) es formuliert, sind moralische Überzeugungen „motivation constituting“.⁹⁹ Und (3.) *Spezifität*: Ursprünglich war der Internalismus stets mit der These der Spezifität von moralischen Überzeugungen verbunden. Dieser gemäß kommt moralischen Überzeugungen (gegenüber nicht-moralisch relevanten Überzeugungen) eine besondere Eigenschaft zu, nämlich gerade diejenige, motivational wirksam zu sein. Andere Überzeugungen („Schnee ist weiß“) haben dagegen nicht diese charakteristische (internalistische) Verbindung mit Handlungsmotivation (vgl. Mele 1996, 732).

⁹⁸ Also die Annahme, dass moralische Überzeugungen eine intrinsische motivationale Wirkung haben. Diese These haben wir kurz „Internalismus“ genannt.

⁹⁹ Vgl. Mele (1996, p. 732): „Believing oneself to be morally required to A metaphysically or conceptually (hence internally, in a formal sense) guarantees that one has a motivation for Aing; second, what is guaranteed, more precisely, is that motivation for Aing is built into any belief that one is (oneself) morally required to A and is internal to belief of that kind in this sense.“

Roskies geht in ihrer Studie weiterhin von einem moralischen Kognitivismus aus (also der Annahme, dass moralische Überzeugungen Glaubenszustände sind), und spezifiziert diese Position wie folgt: moralische Überzeugungen sind Glaubenszustände mit einem „sollen“-Inhalt in der ersten Person. Es wird folgende Internalismus-These zu Grunde gelegt (Roskies 2003, 55), die wir unschwer als Smith's „starke“ Internalismus-These identifizieren werden: „If an agent believes that it is right to Φ in circumstances C, then he is motivated to Φ in C“. Diese These glaubt Roskies als empirisch falsch nachweisen zu können anhand des Beispiels von VM-Patienten (vgl. Damasio 1994). Die bezeichneten Patienten leiden an einer Läsion des ventromedialen präfrontalen Kortex und wurden von Damasio als Träger einer Defizienz sowohl in der Emotionalität als auch in der praktischen Entscheidungsfindung diagnostiziert.

„Ventromedial (VM) prefrontal cortex is anatomically connected to a wide variety of brain areas, including those associated with perception, reasoning, declarative knowledge, and with emotion and visceral control. VM cortex is therefore uniquely functionally situated to mediate between the neural systems for arousal and emotion, and those supporting linguistic cognition.“ (Roskies 2003, 55)

Rufen wir uns noch einmal ein Beispiel eines VM-Patienten vor Augen. Wir haben kursorisch den von Damasio (1994) studierten Patienten Elliot kennengelernt, ebenso wie die historische Figur des Phineas Gage, über den als medizinhistorische Kuriosität die gleiche Symptomlage überliefert ist. Betrachten wir einen kurzen Steckbrief des Patienten „EVR“:

„By age 35, in 1975, EVR was a successful professional, happily married, and the father of two. He led an impeccable social life, and was a role model to younger siblings. In that year, an orbitofrontal meningioma was diagnosed and, in order to achieve its successful surgical resection a bilateral excision of orbital and lower medial cortices was necessary. EVR's basic intelligence and standard memory were not compromised by the ablation. His performances on standardized IQ and memory tests are uniformly in the superior range [97–99th percentile]. He passes all formal neuropsychological probes [...]. Standing in sharp contrast to this pattern of neuropsychological test performance, EVR's social conduct was profoundly affected by his brain injury. Over a brief period of time, he entered disastrous business ventures (one of which led to predictable bankruptcy), and was divorced twice (the second marriage, which was to a prostitute, only lasted 6 months). He has been unable to hold any paying job since the time of the surgery, and his plans for future activity are defective.“ (Damasio et al. 1990)

Das für VM-Patienten typische Syndrom nennt Damasio „acquired sociopathy“. Roskies fasst die Befundlage wie folgt zusammen:

„1. VM patients are able to make appropriate moral and social judgments when queried. When presented with hypothetical situations, the conclusions they reach about moral questions concur with those which normals typically reach. Psychological

evaluation shows that some VM subjects attain the highest level of abstract moral reasoning.¹⁰⁰

2. Clinical histories and observation suggest that VM patients are impaired in their ability to act effectively in many moral situations.

3. Normal subjects produce a skin-conductance response (SCR) to emotionally charged or value-laden stimuli. In contrast, VM patients do not generally produce SCRs when presented with such stimuli. However, other tests produce normal SCRs in VM patients, demonstrating that the autonomic nervous system itself is undamaged.

4. VM patients display and report attenuated or absence of affect when faced with situations that reliably elicit emotions in normals.“ (Roskies 2003, 56f.)

Interessanterweise lieferte der Patient EVR keine signifikanten Reaktionen auf passiv wahrgenommene emotionale Stimuli in einem Hautleitfähigkeits-Test. Wurde EVR jedoch aufgefordert, seine Reaktion zu den Stimuli zu verbalisieren, war gleichzeitig seine Reaktion der Hautleitfähigkeit normal (d.h. signifikant messbar, wie man es bei normalen Probanden erwarten würde, da sie durch die emotionalen Stimuli in Erregung versetzt würden, vgl. Damasio et al. 1990, 88).

Roskies bezieht sich auf Ihre Zusammenfassung der VM-Befunde (s.o., 1-4) und postuliert, dass das Fehlen einer Reaktion im Hautleitfähigkeits-Test auf das Fehlen von Motivation hindeutet bzw. mit diesem gleichzusetzen ist:

„For the purposes of this paper, I take a measurable SCR [skin-conductance response, S.H.] to be evidence of the presence of motivation, and lack thereof to be indicative of absence of motivation. This simplification is warranted: the presence of the SCR is reliably correlated with cases in which action is consistent with judgment, and its absence is correlated with occasions in which the VM patient fails to act in accord with his or her judgments. Thus, the SCR is a reliable indicator of motivation for action.“ (Roskies 2003, 57)

Bezogen auf die handlungstheoretische Fragestellung in der Internalismus-Externalismus Debatte zieht Roskies folgendes Resümee bezüglich dem Befund bei VM-Patienten: VM Patienten verfügen über moralisch relevantes *deklaratives Wissen* (in dem Maße, wie dieses vor der Läsion vorlag) und erweisen sich in Tests (wie dem Kohlberg-Test, vgl. o.) als fähig, praktische Überlegungen anzustellen. VM-Patienten bieten jedoch *disruptives Handeln* in moralisch oder prudenziell relevanten Situationen dar und zeigen in diesen Situationen offenbar *nicht* die übliche *motivationale und emotionale Reaktion*, sowohl hinsichtlich ihres

¹⁰⁰ Interessant ist, dass nach Saver & Damasio 1991 der Patient EVR in Kohlbergs „moral reasoning test“ ein Resultat von „late conventional/ early post-conventional“ im moralischen Denken erreichte. Zu beachten ist, dass nur 11% der us-amerikanischen Männer im Alter von 36 Jahren (also dem Patienten EVR vergleichbare Probanden einer Kontrollbedingung) diesen Status erreichen.

subjektiven, phänomenalen Erlebens, als auch hinsichtlich empirisch messbarer Parameter wie der Hautleitfähigkeit.

Wie ist das motivationale Defizit von VM-Patienten zu erklären? Roskies stellt die interessante Beobachtung an, dass das motivationale Defizit von VM-Patienten eine hohe Spezifität aufweist: sie behalten „appetitive Motivation“ in dem Sinne, als sie Nahrung aufsuchen, wenn sie hungrig sind, die Gesellschaft von anderen Menschen suchen, und offenbar nach Belohnung oder Erfolg streben, was sich daran zeigt, dass sie immerhin aktiv an Damasio's berühmten Iowa Gambling Test (vgl. Bechara et al. 2005) teilnehmen. Zwar erweisen sich Pläne zur Erreichung von Erfolg und Belohnung (im finanziellen Sinne) oft als inadäquat oder höchst defizitär, doch die zu Grunde liegende Motivationskraft scheint bestehen zu bleiben. Roskies resümiert treffend:

„Thus, while certain kinds of motivation are impaired, others, often those with explicit and immediate rewards, remain intact. Ethical motivations are a subset of the impaired motivations. [...] There is evidence that their impairment involves action as it relates to the expectation of delayed reward or punishment. This may indicate a deficit in prudential motivation, which is often involved in motivation in ethical situations.“ (Roskies 2003, 58).

Eine interessante Frage im Sinne eines Exkurses drängt sich auf: warum bieten VM-Patienten, als soziopathische, „amoralisch“ Akteure, verhältnismäßig selten besonders boshafte, gewalttätiges oder gar grausames Verhalten dar? Roskies vermutet einerseits, dass die hier vorliegende Form von Amoralität eher in einer Untätigkeit besteht, da ja offenbar moralische Motivation fehlt. Für Handlungen von besonderer Boshaftigkeit jedoch – so könnte man vermuten – wäre hingegen eine besonders starke Motivation, amoralisch zu sein, nötig. Andererseits liegt ein interessanter Befund vor, dem gemäß diejenigen VM-Patienten, die ihre Läsion in einem frühen Lebensalter erlitten, eher zu Gewalttätigkeit neigen, als diejenigen, deren Läsion in einem späteren Lebensalter auftrat (Anderson et al. 1999). Wenngleich also bewusste moralische Motivation dafür, gewaltabstinent zu sein, fehlt, bzw. Indifferenz gegenüber den Optionen „Gewalt“ vs. „Friedfertigkeit“ besteht, so ist doch anzunehmen, dass die Mehrzahl derjenigen Patienten, deren Läsion in einem späteren Lebensalter auftrat, zuvor schon typische per Konvention oder Moral geforderte Verhaltensweisen habituell gelernt haben. Diese Habitus wirken auch nach der Läsion, wenngleich – wie gesagt – kein Bewusstsein hinsichtlich der motivationalen Grundlage mehr besteht. Andersherum lässt sich schlüssig annehmen, dass Patienten, die schon in jungem Alter eine VM-Läsion erlitten, keine entsprechenden Habitus zur Gewaltabstinenz gelernt haben, und dem gemäß eher zu Gewalttätigkeit tendieren. Dies jedoch wiederum nicht aus einer starken (wenngleich niederträchtigen) Motivation heraus, sondern vielmehr kontingent

per Indifferenz: Gewaltverzicht bzw. die Achtung der Unversehrtheit Anderer wird nicht als ein in besonderem Maße schützenswertes Gut wahrgenommen und kann daher je nach Gelegenheit und Belieben zur Disposition gestellt werden.

Kehren wir jedoch nun zu der von Roskies zu Grunde gelegten Internalismus-These zurück, nämlich:

„If an agent believes that it is right to Φ in circumstances C, then he is motivated to Φ in C“.

Die Anwendung der Befunde von VM-Patienten auf diese These ergibt ein erstaunlich erhellendes Ergebnis: VM-Patienten haben offenbar moralische Überzeugungen. Dies kann durch Tests (z.B. Kohlberg, s.o.) nachgewiesen werden. Allerdings zeigen VM-Patienten nicht die üblichen Hinweise auf Motivation, die mit entsprechenden moralischen Überzeugungen korreliert. Die These eines starken Internalismus, nämlich dass moralische Überzeugungen stets eine entsprechende Motivation einschließen bzw. mit sich führen, kann also durch diesen Befund schlicht widerlegt werden. Soweit können wir Roskies Schlussfolgerung mit tragen. Allerdings überrascht uns das Ergebnis nicht sonderlich, nachdem wir anhand von Smith's Ausführungen schon per begrifflicher Analyse intuitiv gesehen hatten, dass ein starker Internalismus nicht haltbar sein dürfte. Viel aufschlussreicher muss daher die Frage sein, was die Hinzuziehung des Befundes der VM-Patienten über die Schlüssigkeit eines „defeasible internalism“ aussagen könnte. Erinnern wir uns, dass Smith diese These wie folgt definiert: „If an agent judges that it is right for her to Φ in circumstances C, then either she is motivated to Φ in C or she is practically irrational“. Wir müssen uns also fragen: würden wir das Verhalten der VM-Patienten (die Dissoziation von praktischen Überzeugungen und Handlungsmotivation) als *irrationales* Verhalten betrachten? Ich denke ja. Erinnern wir uns weiterhin, dass VM-Patienten nicht in erster Linie durch defiziente Rationalität auffallen, sondern vielmehr durch defiziente Emotionalität, dann drängt sich eine alternative Interpretation auf. Wir können einräumen, dass die besagten Probanden über praktische Überzeugungen verfügen. Die mangelnde Handlungsmotivation ließe sich daher v.a. mit Blick auf die Kap. 3 herausgearbeitete Verbindung zwischen Emotionalität und Werten bzw. desires als ein Fehlen eines relevanten desires bzw. einer relevanten Wertung verstehen. Und dies wiederum deutet abermals in die Richtung einer Lösung des „moral problem“, die wohl an der substanziellen Ausgestaltung des desire-Begriffs wird ansetzen müssen.

4.4. Noch einmal: die Humeanische Motivationstheorie

Wir haben bereits kurz angerissen, dass die Humeanische Motivationstheorie in der These besteht, ein motivierender Grund zum Handeln bestehe aus einem desire (bzw. einer pro-Einstellung) und einer relevanten Zweck-Mittel-Überzeugung. Stellen wir uns vor, Hans verspürt Durst (desire) und ist der Überzeugung, dass ein Glas Apfelsaft seinen Durst stillen wird. Humes Theorie wurde von einigen zeitgenössischen Theoretikern als „Dogma“ bezeichnet und abgelehnt (Nagel 1970, McDowell 1978, 1979 und 1981, und Platts 1979 und 1981). Das Humeanische Dogma besteht jedoch nach Smith (1994) in zwei Prinzipien, bei denen unklar ist, ob die Kritiker des Humeanismus beide gleichermaßen ablehnen, oder nur das stärkere der beiden.

Die *stärkere* These besagt, dass Motivation ihren *Ursprung* im Zusammenwirken einer Pro-Einstellung und eines zweckrationalen Glaubenszustands hat (vgl. Davidson 1963). Smith formuliert diese These wie folgt:

„P1: R at t constitutes a motivating reason of agent A to Φ iff there is some ψ such that R at t consists of an appropriately related desire of A to ψ and a belief that were she to Φ she would ψ .“ (Smith 1994, 92)

Diese Formelsprache lautet in Anwendung auf unser obiges Beispiel: Hans (A) hat jetzt (t) einen motivierenden Grund (R) [also einen, in dessen Licht seine Handlung erklärt werden oder verständlich gemacht werden kann] zu trinken (Φ), wenn er seinen Durst stillen will (ψ), sodass sein Grund jetzt aus einem angemessen in Beziehung gesetzten Paar einer Pro-Einstellung seinen Durst zu löschen (ψ) und einem Glaubenszustand besteht, dass wenn Hans trinken würde (Φ), dann wird er seinen Durst stillen (ψ).

Eine schwächere These besagt nun, dass Motivation lediglich die *Anwesenheit* einer relevanten Pro-Einstellung und eines zweckrationalen Glaubenszustandes erfordert:

„P2: Agent A at t has a motivating reason to Φ only if there is some ψ such that, at t, A desires to ψ and believes that were she to Φ she would ψ .“ (93)

Nach Smith sind sich Kritiker der Humeanischen Motivationstheorie einig in der Ablehnung von P1, während einige unter ihnen jedoch die schwächere These P2 gelten lassen, da dieser gemäß nur die *Anwesenheit* des desire-belief Paares erforderlich ist, was jedoch gleichzeitig konsistent ist mit der Ablehnung der These P1, dass Motivation ihren *Ursprung* in diesem desire-belief Paar habe (in diesem Sinne Nagel 1970, 29 und McDowell 1978, 15).

P1 verbindet motivierende Gründe mit Pro-Einstellungen und Glaubenszuständen. Smith führt i.F. Beispiele gegen die Tatsache an, dass P1 notwendige bzw. hinreichende Gründe dafür liefert, einen Zustand als motivierenden Grund zu bestimmen:

Folgende Beispiele sind aufgestellt gegen die These, dass P1 eine *notwendige* Bedingung dafür liefert, einen Zustand als motivierenden Grund zu bestimmen:

„(i) Suppose I now desire to purchase an original Picasso, but I do not now believe that were I to purchase the painting before me I would do so – suppose I don’t believe that it is a Picasso. If the painting before me is indeed a Picasso, then surely it would be appropriate for an outsider to say I have a reason to purchase the painting before me. But I lack the relevant belief.“ (Smith 1994, 94)

„(ii) Suppose I knowingly stand on someone’s foot so causing that person pain. Surely we can imagine its being appropriate for an outsider to say I have a reason to get off the person’s foot even though I lack the relevant desire, and, indeed, even if I desire to cause that person pain.“ (ebd.)

In (i) finden wir also einen Fall, in dem der Akteur einen *Grund* hätte, das Gemälde vor ihm zu kaufen, und dies sogar in einem schlichten zweckrationalen Rahmen: der Akteur hat den Wunsch (desire), einen Picasso zu erwerben, hätte er auch die Überzeugung, dass das Gemälde vor ihm ein Picasso ist, so würde er damit einen im Humeanischen Paradigma verstandenen motivierenden Grund haben, das Gemälde vor sich zu kaufen. Der Akteur ist aber in Unkenntnis über die Autorschaft des Gemäldes vor ihm. Ähnlich gelagert, da aus der Außensicht der dritten-Person-Perspektive konstruiert, ist (ii). Auch hier hat der Akteur einen *Grund*, seinen Fuß von dem seines Opfers zu nehmen, der jedoch – anders als in (i) – nicht aus Unkenntnis über die relevanten Tatsachen der Situation motivational unwirksam ist, sondern weil dieser Grund nicht durch die Wünsche und Überzeugungen des Akteurs getragen ist.

Folgendes Beispiel ist aufgestellt gegen die These, dass P1 eine *hinreichende* Bedingung dafür liefert, einen Zustand als motivierenden Grund zu bestimmen:

„(iii) Suppose I now desire to drink a gin and tonic and believe I can do so by mixing the stuff before me with tonic and drinking it¹⁰¹. Suppose further that this belief is false; that the stuff before me is petrol rather than gin. Surely it would be appropriate for an outsider to say I have no reason to mix this stuff with tonic and drink it. Yet I have both the relevant belief and desire.“ (ebd.)

In (iii) finden wir also ein Beispiel, in dem der Akteur einen *Grund* hat, die Mischung zu trinken, obwohl er auch einen (rechtfertigenden) *Grund* hat, diese Mischung nicht zu trinken. Wenn wir sagen, er hat einen Grund dafür, dann heben wir auf eine Erklärung im Rahmen seiner vorliegenden, relevanten mentalen Zustände ab: der Akteur ist in dem Glauben, bei dem Getränk handle es sich um einen Gin and Tonic, und diese Annahme zusammen mit seinem koordinaten desire *motiviert* ihn zum Trinken und *erklärt* sein Trinken zugleich. Wenn wir sagen, der Akteur hat keinen Grund, die Mischung zu trinken, dann heben wir auf die

¹⁰¹ Dieses berühmte Beispiel des Gin-Trinkers stammt von Williams (1980).

Wahrheit und Falschheit von relevanten Tatsachen und Sachverhalten der Welt ab, die evtl. dem Akteur nicht zugänglich sind, oder über die er nur qua einer falschen Repräsentation verfügt. Ein Verweis auf solche Zustände der Welt *rechtfertigt* eine Handlung, oder – wie in diesem Fall – rechtfertigt sie nicht.

Wir sprechen von Gründen in zwei Hinsichten: im Sinne von *erklärenden (motivierenden)* Gründen oder im Sinne von *rechtfertigenden (normativen)* Gründen. Beides sind Aspekte am Begriff des Grundes, die unterschiedlich stark ausgeprägt sein können.¹⁰² Theoretisch kann man daher unterscheiden zwischen motivierenden und normativen Gründen. In den obigen Gegen-Beispielen wird diese Ambiguität ausgenutzt, um gegen das Prinzip P1 zu argumentieren. Smith resümiert die Fallbeispiele wie folgt:

„By contrast with normative reasons, then, which seem to be *truths* of the form ‚It is desirable or required that I Φ ‘, motivating reasons would seem to be *psychological states*, states that play a certain explanatory role in producing action.“ (Smith 1994, 96)

Offenbar sind motivierende und normative Gründe also sogar unterschiedliche Kategorien: während erstere mentale Zustände sind, sind letztere Propositionen der allgemeinen Form „Es ist wünschenswert oder erforderlich, dass A Φ -t“. Und weiterhin scheint vordergründig keine Beziehung zwischen beiden Kategorien zu bestehen: ein Akteur kann einen motivierenden Grund haben, ohne einen normativen Grund zu haben, und *vice versa*. P1 bezieht sich nur auf motivierende Gründe.

4.5. Motivierende Gründe

Normative Gründe sind Gründe, die für eine Handlung sprechen, während motivierende Gründe diejenigen Gründe sind, die eine Person tatsächlich zum Handeln veranlassen.¹⁰³ Die Unterscheidung zwischen normativen und motivierenden Gründen kann als die Grundlage für die Debatte zum praktischen Überlegen überhaupt angesehen werden (Halbig 2002), ein ähnliches Ergebnis hatten wir bereits im Rahmen der Diskussion der Internalismus-

¹⁰² Vgl. Michael Woods (1972, 189): „the concept of a reason for an action stands at the point of intersection, so to speak, between the theory of the explanation of actions and the theory of their justification“.

¹⁰³ Normative Gründe lassen sich darüberhinaus in subjektive und transpersonale Gründe gliedern (Audi 2000, 119): erstere haben normative Kraft nur für eine bestimmte Person in einer bestimmten Situation, während letzteren eben transpersonale Normativität zukommt. Halbig (2002) illustriert dies durch folgendes Beispiel: „lustvolle ästhetische Erfahrungen zu suchen“ stellt – *ceteris paribus* – einen transpersonalen normativen Grund dar, während die Option, dies durch Betrachtung spanischer Barockmalerei zu tun, nur für bestimmte Personen mit bestimmten Wünschen, Interessen, Überzeugungen, Geschmacksneigungen etc. einen guten (normativen) Grund darstellen würde.

Externalismus-Debatte erarbeitet. Anzumerken ist, dass „motivierende Gründe“ oft synonym zu verstehen sind mit „explanatory reasons“, also „erklärenden Gründen“¹⁰⁴, ein Begriff, den manche Autoren eben anstatt des ersteren benutzen. Was die zur Verfügung stehenden Theoriemodelle motivierender Gründe anbetrifft, so scheint der humesche Instrumentalismus auch in der zeitgenössischen Debatte zur praktischen Rationalität noch als „Rückfallposition“ zu fungieren. Dieser Theorie zu Folge kann die Vernunft lediglich Mittel zur Erreichung vorgegebener Ziele aufsuchen, jedoch nicht beitragen zur Findung und Setzung unserer Ziele selbst.

Nach Halbig (2002) stellen normative und motivierende Gründe nicht zwei verschiedene Arten von Gründen dar, sondern beide konvergieren. Dies trifft zumindest für die Erste-Person-Perspektive zu. Denken wir noch einmal an das (leicht variierte) klassische Beispiel von Williams (1980): nehmen wir an, ich bin sehr durstig und glaube, vor mir ein Glas Wasser zu haben. Ich hätte einen guten (normativen) Grund, das Glas „Wasser“ zu trinken, und derselbe Grund ist auch das, was mich dazu motiviert. Aus der Perspektive einer dritten Person, die – so nehmen wir an – weiß, dass das Glas Spiritus („petrol“) enthält, gäbe es keinen normativen Grund, der diese Handlung rechtfertigt, wenngleich ich offenbar einen motivierenden Grund habe, der die Handlung erklären kann (mein Irrtum bezüglich des Inhalts des Glases). Ich würde jedoch Halbig widersprechen, insofern es zwar sicher ein Desiderat des praktischen Überlegens ist, normative und motivierende Gründe zur Deckung zu bringen, dies aber in bestimmten Fällen auch in der Erste-Person-Perspektive nicht der Fall ist. Denken wir an klassische Beispiele quasi-akratischen Handelns, oder Handelns wider besseren Wissens. Nehmen wir an, dass ich bereits zwei Schokotörtchen zum Nachtsch verspeist habe, und dass aus diesem Grund nichts mehr (aus der Perspektive normativer Gründe) für die Verspeisung eines dritten spricht. Obwohl ich mir dessen bewusst bin, esse ich noch ein drittes Törtchen, und offenbar bestand ein motivierender Grund für mich, so zu handeln, angesichts meines besseren Wissens um einen widersprechenden normativen Grund. Ich denke, wenn man wie Halbig die prinzipielle Konvergenz von motivierenden und normativen Gründen postuliert, verliert die Debatte um dieses Thema viel von ihrer Relevanz. Wir wollen ja eben nicht nur solch „paternalistische“ Fälle erklären können, wo ein naiver

¹⁰⁴ Halbig merkt jedoch an, dass in gewisser Hinsicht „erklärende Gründe“ weiter gefasst werden können als „motivierende Gründe“. Er nennt folgendes Beispiel: „Fragt man nach der Erklärung dafür, warum sich jemand unmöglich aufführt, kann die Antwort auf einen nur kausal wirksamen Faktor (zum Beispiel eine psychische Störung) verweisen, der selbst aber keinen motivierenden Grund bildet, in dessen Licht die Person entsprechend handelt.“ (2002, 962 FN 4)

Akteur über seine Irrationalität belehrt wird, sondern eben auch solche „hard cases“, wo normativ richtiges Handeln nicht dargeboten wird, trotz Einsicht in die widersprechenden Gründe. Meine Einwände in dieser Hinsicht lassen sich relativ bequem auf die Postulate des bereits angesprochenen „defeasible internalism“ einschränken.

4.5.1. Der Humeanismus

Wie allgemein bekannt, besteht der Humeanischen Theorie motivierender Gründe zu Folge ein motivierender Grund aus einem Wunsch „dass p“ und einer Überzeugung, „dass Handlung h zur Erreichung, Realisierung, oder Förderung von p dient“ (vgl. Hume, Treatise II 3,3). Die Humeanische Theorie – wie wir gesehen haben, eine internalistische Theorie – lässt sich mithin weiter unter die psychologischen Theorieformen kategorisieren, da ein motivierender Grund als die koordinierte Verbindung von *mental*en Zuständen des handelnden Subjekts angesehen wird. Anzumerken ist jedoch noch, dass zwischen beiden Komponenten eine Asymmetrie angenommen wird, mit einer Gewichtung zum Wunsch hin. Hume erklärte dies noch (ebd.) mit der Unterscheidung zwischen der kausal wirksamen Kraft des Wunsches gegenüber der lediglich kanalisierenden Funktion der Überzeugung. Im neo-Humeanischen Diskurs (vgl. Smith 1987) wird üblicherweise mit dem Begriff der „Passensrichtung“ („direction of fit“, Anscombe 1957, 56f.; vgl. zum Begriff auch Platts 1979, 256f.) gearbeitet: dem Wunsch kommt demnach das Schwergewicht zu, weil seine Passensrichtung „Welt-auf-Geist“ ist, und nicht umgekehrt, wie im Falle der Überzeugung. Kritik am Humeanismus kann nach Halbig (2002) nun zwei verschiedene Ziele angreifen: (1.) entweder man attackiert die Asymmetrie-These, oder (2.) man versucht, den Humeanismus in seiner Ganzheit zu widerlegen, z.B. durch Vertreten eines reinen Kognitivismus.

(ad 1): Nagel (1970) kritisiert die Asymmetrie-These durch Unterscheidung von „motivierenden“ und „motivierten“ Wünschen (ebd., 29f.). Nagel räumt zwar ein, dass Wünsche stets notwendiger Bestandteil praktischen Überlegens sind, dass aber einige Wünsche (nämlich „motivierte“) erst auf Grund einer Überzeugung gebildet werden. In diesem Fall hat also die für den Wunsch grundlegende Überzeugung den Vorrang, da sie jenen eben kausal bedingt. Ein ähnlicher Ansatz ist in jüngerer Zeit von Steinfath (2001) anhand seines Modells „neigungsförmiger Wünsche“ herausgearbeitet worden (vgl. ebd., Kap. 3).

Dancy (1993, Kap. 2.1; und erneut 2000, 81) äußert wiederum kritische Vorbehalte gegenüber Nagels Vorschlag. Nach Dancy hat eine sog. „hybride“ Theorie wie die von Nagel ein Problem damit, erklären zu können, warum einige Überzeugungen stets das Komplement

eines Wunsches benötigen, um eine Handlung zu motivieren, während andere einen solchen Wunsch selbst kausal hervorbringen können. Können wir also verschiedene Klassen von Überzeugungen benennen, um diesem Unterschied gerecht werden zu können? Ein solches Projekt erscheint vordergründig als schwierig. Tatsächlich jedoch wird sich zeigen, dass die von mir favorisierte Lösung nicht allzu weit von Dancys Vorschlag entfernt sein dürfte. Zwar denke ich, dass eine substanzielle Neufassung des in der Theorie praktischen Überlegens eingesetzten desire-Begriffs nötig sein wird. Dennoch wäre Dancys Vorschlag u.U. zu plausibilisieren, indem wir annehmen würden, dass eben diejenigen Überzeugungen, die unter bestimmten Bedingungen Wünsche generieren können, als „emotionale Einstellungen“ aufzufassen sind. Ein ähnlicher Lösungsansatz soll schließlich am Ende des vorliegenden Kapitels unternommen werden.

McDowell (1979) hatte zu dieser Frage schon Stellung genommen, indem er den Vorschlag machte anzunehmen, dass bei im aristotelischen Sinne tugendhaften Akteuren eine Überzeugung stets den relevanten Wunsch hervorbringen könne, während andere (auch bei den nach Aristoteles sog. „selbstbeherrschten“, „enkratês“) nur dann aufgrund einer Überzeugung handeln würden, sofern sie schon von voneherein mit einem entsprechenden Wunsch (bzw. Pro-Einstellung) ausgestattet sind. Anstatt davon auszugehen, dass Überzeugungen manchmal (und unter nicht vorhersagbaren oder theoretisch greifbaren Bedingungen, so Dancy) Wünsche hervorbringen können, schlägt Dancy nun vor, im Rahmen einer rein kognitivistischen Theorie davon auszugehen, dass Überzeugungen stets zur Motivation zu einer Handlung ausreichen.

4.5.2. Der Kognitivismus

Dem reinen Kognitivismus nach sind es einzig und allein Überzeugungen, die uns zu Handlungen motivieren. Eine solche Theorieform wird vertreten z.B. von Dancy (1993, Kap. 1-3). Gemäß Dancy besteht ein motivationaler Zustand aus einem Überzeugungspaar: die eine Überzeugung sagt aus, wie die Welt in ihrem status quo ist, die andere, wie sie aussehen wird nach erfolgreichem Abschluss der Handlung. Dancy identifiziert „Wunsch“ nicht als ein eigenständiges mentales Phänomen, sondern sagt, dass Wünsche in dem Zustand bestehen, „motiviert zu sein“. Da diese Beschreibung auf den von ihm analysierten motivationalen Zustand zutrifft, bestehend aus einem Überzeugungspaar, meint Dancy, dass man über einen auf diese Weise motivierten Akteur sagen könne, dass er so zu handeln „wünscht“ („desires“). Ein Wunsch ist also kein Bestandteil des motivationalen Zustands, sondern dieser Zustand selbst, wobei dieser aus einem Paar von Überzeugungen besteht. Auch Raz (1999) scheint

einen reinen Kognitivismus zu vertreten: ihm gemäß kann die Überzeugung, dass eine gewisse Handlung unbedingt geboten ist, motivational wirksam werden, selbst entgegen einer widerstreitenden Neigung.

Beide Formen von Kognitivismus begegnen jedoch einem ähnlichen Problem wie die Theorie von Nagel (1970): wie lässt sich erklären, dass offenbar nur bestimmte Überzeugungen motivational wirksam sind, andere jedoch nicht? Raz' (1999, 25) Lösungsvorschlag ist folgender: es gibt nicht eine bestimmte *Klassen* von Überzeugungen, die motivational wirksam sind, sondern es gibt Überzeugungen, die motivational wirksam werden kraft ihres *Inhalts*, der sich in diesen Fällen eben auf einen normativen Aspekt der Wirklichkeit bezieht.

Auch der Kognitivismus muss das Problem akkratischen Handelns oder Nicht-Handelns lösen können. Nach Raz jedoch zeigen diese Fälle nicht, dass das Modell zur Erklärung rationalen Handelns falsch sei, sondern sie zeigen vielmehr, dass wir nicht immer in der Lage sind, rational zu handeln, oder mit Audis Worten, dass unsere Fähigkeit zur Rationalität (Audi 2000, 113: „capacity-rationality“) nicht zur Anwendung kommt. Ein ähnlicher Einwand ist bereits von Velleman (1989; 1992 und 1996) her bekannt. Velleman akzeptiert die Unterscheidung zwischen „Motiven“ und „Gründen“ (motivierenden und normativen Gründen), glaubt jedoch, dass die Verbindung zwischen beiden Entitäten wiederum ein Motiv sei, das konstitutiv für menschliches Handeln („agency“) ist: ein Akteur zu sein, bedeutet u.a., den Ansprüchen normativer Gründe zu folgen. In Fällen von Akrasie bspw. müssten wir demnach annehmen, dass dieses „agency-constituting motive“ abwesend ist. Tatsächlich denkt Velleman, dass wir in diesen Fällen temporär unsere autonome Handlungsfähigkeit einbüßen: wir handeln zwar noch auf der Grundlage instrumentaler Motivation, allerdings ohne dass wir uns die entsprechenden desires zu eigen machen würden. Nach Velleman ist jedoch das akteurs-konstituierende Verbindungsmotiv nicht unter unserer Kontrolle, weswegen wir dessen Verlust und Wiedergewinn jeweils passiv ausgeliefert sind.

Auch Thomas Scanlon (1999) vertritt eine rein kognitivistische Theorie. Wir werden zu einer Handlung allein dadurch motiviert, dass wir bestimmte Überzeugungen haben. Nach Scanlon stellt es kein Problem für den Kognitivismus dar, dass Überzeugungen unterschiedlich starke motivierende Kraft haben, und dass damit akkratisches Handeln oder Nicht-Handeln auftreten kann. Denn meine Überzeugung, dass ich z.B. eine bestimmte Verpflichtung gegenüber einer anderen Person habe, kann unterschiedlich starke motivationale Kraft haben, abhängig davon, wie ich persönlich zu dieser anderen Person stehe, und in welchem Kontext meine praktische Überlegung stattfindet. Scanlon denkt nicht

nur, dass Wünsche keine Relevanz in der Erklärung von praktischem Überlegen haben, sondern leugnet sogar, dass Wünsche ein eigenständiges mentales Phänomen sind (ebd., Kap. 1.8). Wünsche sind nach Scanlon eine bestimmte Spezies von Gründen. Der Wunsch z.B. nach Sydney zu reisen, wäre im Sinne von Scanlons Begriff der „directed attention“ (ebd., 39) schlicht damit gleichzusetzen, dass sich bestimmte Gründe, oder ein bestimmter Grund, mir in besonderem Maße aufdrängen. Halbig nennt dies die „Anmutung“ eines Grundes (2002, 969).

Wir hatten bereits gesehen, dass eine psychologistische Theorieform sehr gut in der Lage ist, Gründe im Sinne von Motiven zu erklären, d.h. denjenigen Aspekt am Begriff des Grundes zu erfassen, mit dem wir auf die individuelle Handlungsmotivation des Akteurs abheben. Wie wir jedoch ebenfalls – anhand der diskutierten Beispiele von Smith – gesehen haben, lassen sich normative Gründe nicht leicht im Rahmen dieses Paradigmas erklären.

So vertritt bspw. auch Dancy (2000, 103-108) die Auffassung, dass der sog. „normative constraint“ im Rahmen psychologischer Theorien nicht zu erklären ist. Halbig (2002, 970) resümiert: „Fragt man sich jedoch nach dem ontologischen Status von guten Gründen, wird klar, dass Sachverhalte, nicht psychologische Zustände normative Gründe konstituieren.“ Halbig gibt folgendes Beispiel. Es ist der *Sachverhalt*, dass die Gemälde von Velázquez eine außerordentliche ästhetische Qualität besitzen, der einen normativen Grund darstellt, sie zu betrachten, und nicht etwa ein mentaler Zustand des Akteurs, wie z.B. dessen *Überzeugung*, dass die Bilder künstlerisch wertvoll sind. Eine realistische (also anti-psychologistische) Theorie kann also normative Gründe wesentlich einfacher beherbergen als der Psychologismus. Dennoch denke ich, dass wir dieses Argument bzgl. des ontologischen Status von Gründen weiter verfolgen können. Denn wie ich in meinen Ausführungen zur emotionalen Wertung diskutiert habe, sind axiologische Sachverhalte nicht etwa Sachverhalte, die unabhängig von einer bestimmten wertenden Praxis in der Welt bestünden. Axiologische Aspekte der Wirklichkeit entstehen überhaupt erst durch unsere Praxis des Wertens, die ich als emotionale Reaktionen oder Dispositionen identifiziert habe. Letztlich müssten wir also die Formulierung von Halbig überarbeiten: Fragt man nach dem ontologischen Status von normativen Gründen, wird klar, dass die zugrundeliegenden axiologischen Sachverhalte wiederum durch psychologische Zustände oder Dispositionen von Akteuren konstituiert werden, dass also letztlich normative Gründe ebenfalls ontologisch auf psychologischen Zuständen fußen.

4.6. Normative Gründe

Betrachten wir somit die andere „Kategorie“ von Gründen zu handeln: normative Gründe. Wie wir gesehen haben, sind motivierende Gründe diejenigen Gründe, „in deren Licht wir handeln“ (Halbig 2003), während normative Gründen diejenigen Gründe sind, die tatsächlich für eine bestimmte Handlung sprechen. Die beiden Arten von Gründen können prinzipiell auseinandertreten: wir können bspw. aus einem (motivierenden) Grund handeln, der sich als (normativ) falsch herausstellen kann, u.z. entweder aus der Perspektive einer dritten Person, oder aus der Sicht in Retrospektive. Denken wir noch einmal an Williams' (1980, 102) berühmtes Beispiel vom Gintrinker: hat dieser einen Grund, sich Spiritus in sein Tonic Water zu gießen, weil er es fälschlicherweise für Gin hält?¹⁰⁵ Andererseits können bisweilen gute normative Gründe darin versagen, motivational wirksam zu werden, etwa in Fällen von Willensschwäche, oder weil dem Akteur bestimmte Informationen fehlen, er bestimmte Verbindungen nicht herstellt oder Schlussfolgerungen nicht zieht etc. Selbst wenn beide Arten von Gründen auseinandertreten können, so ist es doch ein Desiderat jeder überzeugenden Theorie praktischer Rationalität, beide zusammenzubringen: wir müssen nachweisen, dass normative Gründe auch motivational wirksam sein können – dass wir aus ihnen heraus auch handeln können. So formuliert Halbig (2003, 133):

„Der Internalist nimmt an, dass für P nur dann ein guter Grund, p zu tun, besteht, wenn P auf der Grundlage seines Wissens um alle für sein Handeln relevanten Gesichtspunkte und rationalem praktischem Überlegen motiviert wäre, p zu tun. Der Externalist bestreitet dies.“

Im Klartext heißt das: im Sinne des Internalismus besteht für einen Akteur nur dann ein (normativer) Grund zu handeln, wenn dieser im Zustand optimaler Information auch motiviert ist, entsprechend zu handeln. Der Externalismus postuliert, dass ein normativer Grund zu handeln auch dann besteht, wenn dieser sich nicht mit der Motivlage des Akteurs deckt.

Nach Dancy (2000) formuliert Williams mit seiner These lediglich eine notwendige Bedingung an die Konzeption normativer Gründe: nämlich dass sie motivationale wirksam sein müssen, u.z. in dem Sinne, als ein Akteur in einem optimal informierten Zustand – in einer kontrafaktischen Situation zur gegebenen Situation also – auf Grund eines (guten) normativen Grundes in dieser kontrafaktischen Situation auch motiviert werden sollte. Williams These sagt jedoch nichts aus über die Quellen – Halbig (2003, 134): die „metaphysischen Konstitutionsbedingungen“ – normativer Gründe. Auch Halbig ist der

¹⁰⁵ Wedgwood (2002, 348) hält die Handlungsmotivation des Gintrinkers für „subjektiv rational“, aber „objektiv irrational“, womit er eine Unterscheidung von Brandt (1979, 72f.) aufnimmt.

Ansicht dass das Internalismus-Externalismus Problem einerseits und die Frage nach den Quellen normativer Gründe andererseits „distinkte Problembereiche, die systematisch quer zueinander stehen“ darstellen.

4.6.1. Die Quellen normativer Gründe

Halbig (2003) diskutiert (1.) Wünsche (desires), (2.) Werte und (3.) natürliche Eigenschaften als Kandidaten für die Quellen normativer Gründe. Wir wollen im Folgenden seine Kategorisierung mitvollziehen, jedoch auf (3) wiederum verzichten. Die Supervenienz von desires und Werten auf natürliche Eigenschaften wurde bereits in Kapitel 3.6. dieser Arbeit diskutiert, wir können also an dieser Stelle darüber hinweggehen. Interessanter scheint daher der Zusammenhang von normativen Gründen mit eben jenen gegenüber natürlichen Eigenschaften „höherstufigen“ Entitäten wie desires und Werten. Allerdings schickt Halbig ein beachtenswertes *caveat* voraus: bisweilen werden Wünsche und Werte in einem sehr pauschalen Sinn verwendet, um eine Art abgekürzte Argumentation für den entsprechenden Kandidaten zu führen, die nach Halbig jedoch nicht überzeugen kann. Betrachten wir zunächst also diese Überlegungen zu unzulässigen „Abkürzungen“ in der Argumentation:

Da wir uns im Rahmen einer Theorie praktischer Rationalität mit intentionalem Handeln beschäftigen (und nicht bspw. mit Zwangshandlungen, die per se schon eher einen Verhaltens- als einen Handlungscharakter haben), können wir unkontrovers behaupten, dass jede Handlung in irgendeinem Sinne „gewollt“ wird. Bedeutet dies jedoch zugleich, dass wir also handeln, um unsere Wünsche zu realisieren? Scanlon (1999, 37) unterscheidet zwischen pro-Einstellungen und Wünschen im engeren Sinne.¹⁰⁶ Nach Scanlon soll diese Unterscheidung klar machen, dass jede intentionale Handlung in einer trivialen Weise eine pro-Einstellung voraussetzt. Denn nur dies kann erklären, warum wir überhaupt handeln. Selbst der Kantianer, der sich für einen Handlungskurs entgegen seinen Neigungen (also Wünschen im engeren Sinne) entscheidet, muss eine gewisse pro-Einstellung gegenüber seiner Handlung haben, sonst wäre sein Handeln nicht zu erklären. Bei Wünschen im engeren Sinne jedoch muss es sich im Gegensatz zu dieser trivialen Bedeutung von pro-Einstellung um „eine gegenüber Überzeugungen eigenständige Klasse mentaler Zustände handeln, die eine von ihnen unabhängige Quelle der Handlungsmotivation und –rechtfertigung bilden

¹⁰⁶ Diese Terminologie von Scanlon wird hier nur beispielhaft wiedergegeben, jedoch nicht aufgenommen. In der großen Mehrheit der Literatur wird „pro-Einstellung“ als ein quasi-Synonym von „Wunsch“ verwendet, und dieser Begrifflichkeit folge auch ich.

kann“ (Halbig 2003, 135). Es muss ein „attachment not required by reason“ (Raz 1999, 111) vorliegen. Wir können also nicht auf pauschale und zugleich triviale Weise Wünsche als Quelle normativer Gründe bestimmen, ohne dass die Theorie entscheidend an Informations- und Beschreibungskraft verliert.

Ebenso verhält es sich mit einem trivialen Begriff von Werten. Man könnte argumentieren, dass Werte die Quelle normativer Gründe seien, da wir immer um eines Gutes willen – also *sub specie boni* – handeln. Die Argumentation stellt sich also strukturell analog zu der These dar, Handeln wäre sowieso immer an einer pro-Einstellung orientiert. Auch im Fall von Werten jedoch muss diese Strategie misslingen, zumindest wenn etwas Substanzielles über die Quelle normativer Gründe gesagt werden soll. Denn in der Tat zielt jede intentionale Handlung immer auf ein Gut (einen Wert) ab. Dieser jedoch ergibt sich schon begriffslogisch aus dem Begriff des Handelns: da ein Akteur das Ziel einer Handlung zu realisieren versucht, kann dieses *ex hoc* als das „Gut“ dieser Handlung gesetzt werden. Ob es sich um einen substanziellen Wert handelt, ist damit jedoch noch nicht gesagt. In diesem Sinne müsste nachgewiesen werden, dass sich normative Gründe nicht (nur) auf empirisch ablesbare Handlungsziele beziehen, sondern auf Werte, die unabhängig von den Zielen und Wünschen eines Akteurs bestehen.

4.6.2. Wünsche (desires) als Quellen normativer Gründe

Es ist unstrittig, dass Wünsche (desires) eine Quelle praktischer Gründe darstellen: ein Akteur hat sicherlich einen Grund, p zu tun, wenn dadurch die Erfüllung eines seiner Wünsche realisiert oder befördert wird. Strittig ist hingegen die Frage, ob Wünsche möglicherweise die *einzig*e Quelle praktischer Gründe bilden. Intuitiv würden wir wohl antworten, dass auch andere Erwägungen, u.z. sowohl moralische als auch prudentielle Gründe, gegen die Verfolgung gegebener Wünsche sprechen können. Scanlon (1999, 43) und Raz (1999, Kap. 3) hingegen gehen einen Schritt weiter und behaupten, dass Wünsche *niemals* eine Quelle praktischer Gründe darstellen.

Aus einer instrumentalistischen Perspektive könnte argumentiert werden, dass Wünsche, falls sie auftreten, normative Gründe zum Handeln liefern. Das wohlbekanntes Problem (analog zum Emotivismus) ist jedoch, dass sich nicht weiter rational argumentieren lässt, ob Wünsche auftreten *sollen*. Die Frage der Normativität kann nicht beantwortet werden. Praktisches Überlegen hat im instrumentalistischen Paradigma allein zweckrationale Funktion: Wünsche bilden den Ausgangspunkt solchen Überlegens, sie können jedoch nicht dessen Gegenstand sein, d.h. praktisches Überlegen kann nicht kritisch zu den Wünschen

selbst Stellung beziehen. Scanlon (1999, Kap. 1.9) wendet jedoch ein, dass erstens Wünsche als eine einheitliche Klasse mentaler Zustände nicht existieren und zweitens auch keine andere Klasse mentaler Zustände zur Verfügung steht, die die Rolle von Wünschen im Rahmen des instrumentalistischen Paradigmas übernehmen könnten.

Betrachten wir ein Beispiel von Halbig (2003, 136): nehmen wir an, ich wolle den Prado besuchen und werde gefragt, warum ich einen guten Grund dazu habe. Die tautologische Antwort, „weil dies meinen Wunsch erfüllt, den Prado zu besuchen“ erscheint nicht zufriedenstellend. Vielmehr wird von mir erwartet, Gründe anzugeben wie: ich erwarte mir davon einen ästhetischen Genuss, oder ich halte es per se für einen guten Grund, sich anhand von Kunst zu bilden, oder ich habe mir vorgenommen, alle Gemälde von Velázquez zu sehen. Natürlich kann eingewendet werden, dass hinter diesen Erwägungen evtl. elementare „Wünsche“ stehen. Halten wir dennoch zunächst fest, dass es sich bei diesen Erwägungen um Konstruktionen unseres praktischen Überlegens handelt, und nicht um einheitliche Mitglieder einer Klasse mentaler Zustände, die wir „Wunsch“ nennen. Was könnte jedoch ein homogener mentaler Zustand sein, der die Rolle von Wünschen im Sinne des Instrumentalismus einnehmen könnte? In der Tat könnten wir hier auf die hintergründigen „elementaren“ Wünsche verweisen. Da jedoch keinerlei Mediation durch praktisches Überlegen vorliegen dürfte, wollte man nicht aus dem instrumentalistischen Paradigma herausfallen, müsste es sich bei solchen elementaren Wünschen jedoch um „Triebe“ oder appetitive Körperzustände handeln (Durst, Müdigkeit etc.). Einerseits ist fraglich, ob solche naturalistischen Phänomene noch in der Lage wären, als normative Gründe zu fungieren. Ich denke, dass man mit nein antworten müsste (vgl. Quinn 1993, 246f.). Andererseits habe ich an früherer Stelle gezeigt, dass solche Phänomene nicht so „roh“ oder „naturalistisch“ sind, wie mit strategischer Überzeichnung oft angenommen wird. Auch Triebe unterliegen gewissen Mediatorvariablen, v.a. der Emotionalität. Emotionen wiederum sind Konstrukte, die zwar auf elementaren affektiven Strukturen aufbauen, jedoch beim erwachsenen Menschen stets kognitiv überformt sind. Es ist letztlich also fraglich, ob es überhaupt einen mentalen Zustand des Menschen gibt, der die Eigenschaften erfüllen könnte, die der Instrumentalismus seinem Phantom-Begriff des „Wunsches“ zuschreibt.

Welches Bild würde sich jedoch ergeben, wenn wir zugestehen würden, dass der Instrumentalismus zutrifft, und dass Wünsche Ausgangspunkt unseres praktischen Überlegens sind? In diesem Fall hätte praktisches Überlegen die Aufgabe, konkurrierende Wünsche gegeneinander abzuwägen und evtl. prinzipiell miteinander harmonisierende Wünsche hinsichtlich ihrer Erfüllungsbedingungen zu koordinieren (Scanlon 1999, 50ff.). Das Bild des

Akteurs gemäß dem Instrumentalismus ist das eines „simple weigher of alternatives“ (Taylor 1976, 287).

Raz (1999, 50ff.) jedoch macht auf eine grundsätzliche Implausibilität dieser Position aufmerksam: wenn sich praktisches Überlegen tatsächlich nur darauf beschränken würde, den intensivsten oder am meisten salienten Wunsch („desire“) herauszufinden und geeignete Mittel zu seiner Erfüllung zu finden, dann wäre aus dieser Perspektive unverständlich, warum wir überhaupt darüber nachdenken, *was* wir wollen (im Sinne von „what we desire“). Und das tun wir de facto. Gemäß dem Instrumentalismus sollte sich jedoch unser Nachdenken nur auf die geeigneten Mittel zu vorgegebenen Zwecken richten. Wie Halbig prägnant formuliert: „Es würde also reichen, uns selbst beim Handeln zuzuschauen – die Handlung, die wir vollziehen, ist eben genau die, die wir am meisten gewollt haben und für die nach der instrumentalistischen Konzeption normativer Gründe eben deshalb auch der beste Grund bestand, sie zu vollziehen“ (Halbig 2003, 136f.). Auch Audi (2000, 52ff.) weist darauf hin, dass wir nicht lediglich über die Möglichkeiten zur Realisierung vorgefundener Wünsche nachdenken, sondern vielmehr darüber, welche *Gründe* es für oder gegen bestimmte Handlungsoptionen geben könne. Der Instrumentalismus scheint also ein zu unseren Intuitionen quer stehendes Bild praktischen Überlegens zu zeichnen. Nach Halbig (2003) könnte der Instrumentalismus diese Härte nur vermeiden, indem die Normativität von Wünschen nicht an ihre de facto zu erwartende motivationale Kraft (d.h. ihrer phänomenalen Intensität) zurückgebunden würde, sondern indem auf ein eigenständiges, substanzielles Kriterium zur Gewichtung von Wünschen rekuriert würde. Damit jedoch ist die Normativität von Wünschen gleichsam per Redundanz aufgehoben, da im Hintergrund ein von den bei einem Akteur vorliegenden (und gefühlten) Wünschen unabhängiges Kriterium stünde. Wir könnten bspw. postulieren, dass per Rekurs auf ein Kriterium des menschlichen Wohlergehens und der Selbsterhaltung bestimmte selbstzerstörerische Wünsche, die dennoch phänomenal salient und motivational stark sind, ausgeschlossen würden.

Bei der Erklärung normativer Gründe zum Handeln scheint der Instrumentalismus jedoch in eine weitere Implausibilität zu geraten. Der Instrumentalist müsste annehmen, dass jeder Wunsch per se einen normativen Grund zum Handeln darstellt, da ja außer dem faktischen Bestehen oder Nichtbestehen kein weiteres Kriterium zu Entscheidung über unsere Handlungsziele hinzugezogen werden soll. Scanlon (1999, Kap. 1.10) jedoch weist darauf hin, dass bestimmte Handlungsziele oder –optionen nicht erst im Laufe des gegeneinander Abwägens fallengelassen werden sollen, sondern dass wir intuitiv glauben, bestimmte Handlungsziele oder –optionen sollten von vornherein aus dem Prozess des praktischen

Überlegens ausgeschlossen sein. Wenn wir uns bspw. die Person eines Arztes (mit den aus dieser Rolle erwachsenden Pflichten) vorstellen, so nehmen wir an, dass ein Arzt im Notfall verpflichtet ist, einem Patienten Hilfe zu leisten, u.z. in Absehung von bestimmten Aspekten, die die Hilfeleistung nicht beeinträchtigen würden. Sollte dieser Arzt nun einen Patienten mit Herzstillstand beatmen, der jedoch zufällig äußerst ungepflegt ist und schlechten Atem hat, so würden wir erwarten, dass ein Arzt den Gedanken an die Unterlassung dieser Maßnahme mit Hinblick auf die zu erleidenden Unannehmlichkeiten *nicht* erst nach Abwägen bestimmter Gründe (bzw. gemäß dem instrumentalistischen Paradigma: Wünsche) verwirft, sondern diesen Gedanken *schon von Vornherein* nicht als Kandidaten im Prozess des praktischen Überlegens zulässt. Halbig (2003, 137 FN 6) vergleicht diese Beobachtung von Scanlon mit dem von McDowell (1978, 26) geprägten Konzept des „silencing“. Nach McDowell gibt es bestimmte Erwägungen oder Wünsche, die nicht erst im Laufe des praktischen Überlegens „zum Schweigen gebracht“ werden, sondern gar nicht erst zu diesem Prozess zugelassen werden sollen. Dieser Position unterliegt eine aus der Tugendethik stammende Sicht auf die Phänomene der Akteursidentität und des praktischen Überlegens: der tugendhafte Akteur („spoudaios“ oder „phronimos“; bspw. der „Gerechte“) wird bestimmte Wünsche oder Handlungsoptionen (bspw. zu betrügen, um ein Ziel zu erreichen) gar nicht erst aufkommen lassen. In dieser Art soll einem Arzt eben der kompromisslose Wille zu helfen und zu heilen gewissermaßen eingeschrieben sein. Handlungsoptionen, die das Wohl des Patienten grundlos kompromittieren, sollen daher von Vornherein ausgeschlossen sein. In der Tat scheint dies auch unser intuitives (Alltags-) Verständnis einer moralisch gefestigten Persönlichkeit zu sein. Gerade dieser Intuition kann jedoch der Instrumentalismus nicht gerecht werden, da alle faktisch vorliegenden Wünsche zunächst einen guten (normativen) Grund zum Handeln darstellen. Erst im Laufe des Prozesses des praktischen Überlegens würden dann bestimmte Handlungsoptionen sozusagen „aussortiert“. Wie Halbig es formuliert:

„Wünsche können zwar *zu schwach* sein, um ausschlaggebende Handlungsgründe zu fundieren, sie können aber innerhalb des instrumentalistischen Modells nicht *von vornherein* als unzulässig ausgeschlossen werden“ (Halbig 2003, 138).

Audi (2000) entwickelt hingegen eine Theorie praktischer Gründe auf der Basis von Wünschen, die diesen Vorwürfen gegenüber dem Instrumentalismus zu entgehen versucht. Audi folgt weitgehend der These, dass Handeln aus einem motivationalen Element und dem Vorliegen einer Überzeugung zu erklären sei. Allerdings hält er Wünsche – wenn auch für die wichtigsten – jedoch nicht für die einzigen motivationalen Elemente der menschlichen Psyche. Außerdem räumt er ein, dass Wünsche wiederum rationaler Prüfung und Bewertung zugänglich seien (2000, 62ff.).

Zur Behebung der Schwächen des Humeanischen Modells auf dem Feld einer Theorie normativer Gründe hat auch Audi (2000) jüngst einen Lösungsvorschlag vorgelegt, an den im Rahmen dieser Arbeit angeschlossen werden kann. Audi schließt sich der Humeschen These an, dass ein Handlungsgrund sich aus dem Vorliegen eines Wunsches sowie dem Vorliegen einer Überzeugung, dass eine bestimmte Handlung die Erfüllung dieses Wunsches realisiert oder befördert, besteht. Allerdings glaubt Audi, dass wir in zweierlei Hinsicht Einfluss auf unsere Wünsche haben, die gemäß dem klassischen Humeanismus lediglich als faktisch vorhanden oder nicht-vorhanden beschrieben, nicht jedoch rational kritisiert oder modifiziert werden können. Einerseits denkt Audi (ebd., 112f.), dass Überzeugungen die Rationalität von „intrinsischen Wünschen“ aufheben können. Wir können bspw. erkennen, dass das Objekt unseres Wunsches sich in Wirklichkeit anders darstellt als zunächst geglaubt. Oder wir können durch einen Prozess des praktischen Überlegens unserer Handlungsziele zu der Überzeugung gelangen, dass das Objekt unseres Wunsches in Wirklichkeit kein würdiges Objekt unseres Bemühens durch Handeln darstellt. In beiden Fällen führe das Kommen zu eben jener Überzeugung zugleich auch zur Aufgabe des entsprechenden Wunsches, außer in Fällen von Irrationalität. Andererseits räumt Audi (ebd., 118f.) die Möglichkeit einer Formierung von Wünschen auf der Grundlagen von Wertüberzeugungen ein. Wir können also nicht nur durch rationale Einsicht Wünsche fallen lassen, die sich als Quelle schlechter Handlungsgründe erwiesen haben, sondern wir bilden kraft unserer Wertüberzeugungen auch aktiv Wünsche aus. Audi positioniert sich also auch in dieser Hinsicht in Gegensatz zum klassischen Humeanismus, der bekanntlich annimmt, dass Wünsche schlicht vorliegen oder nicht vorliegen. Audi räumt also – abschließend betrachtet – sowohl eine schwache Kontrolle über unsere Wünsche im Rahmen einer Möglichkeit zur Modifikation oder Repression von Wünschen ein, als auch eine starke Kontrolle über unsere Wünsche im Rahmen einer Möglichkeit des aktiven Formierens von Wünschen.

4.6.3. Werte als Quellen normativer Gründe

Raz definiert praktische Vernunft als eine „faculty for pursuing value“ (1999, 68). Werte stellen Gründe für intentionales Handeln dar. Raz hält seine Theorie damit für eine „klassische Herangehensweise“ (ebd., 22), wie sie etwa von Platon und Aristoteles überliefert ist. Damit legt er eine Theorie praktischer Rationalität als Handeln um des Guten willen vor (vgl. Steinfath 2001).

Eine solche Theorie setzt sich dem möglichen Einwand aus, dass wir offenbar bisweilen Handlungsoptionen wählen können, nicht etwa, weil wir dadurch einen Wert realisiert oder

befördert glauben, sondern *gerade* weil sie dem von uns als ein Gut erkanntem *widersprechen*. Halbig (2003, 142) führt etwas das vieldiskutierte Beispiel des Satan aus John Miltons „Paradise Lost“ an, der in Rebellion gegen Gott das Böse zu seinem Guten erwählt. Auf Anscombe (1957, 75) geht der bekannte Lösungsversuch zurück anzunehmen, dass ein solcher Akteur dennoch versucht, irgendeinen Wert zu realisieren, wenn es sich auch um einen falsch bestimmten oder äußerst verquerten Wert handeln sollte. Velleman (1999, 119) wendet ein, dass jedoch damit gerade das „Satanische“ an seinem Handeln missverstanden werde. Die Wahl gerade dieses Beispiels jedoch trägt viel zur Kontroverse bei. Stellen wir uns hingegen einen menschlichen (und keinen quasi-göttlichen oder mythischen) Akteur vor, so erscheint Anscombes These recht plausibel. Nehmen wir an, dass Hans „die Schule schmeißt“, nachdem er von den Lehrern als Autoritätspersonen insgesamt enttäuscht wurde und gesellschaftliche Autorität nun per se in Frage stellt. Wenn er auch offenbar seine beruflichen Karrierechancen kompromittiert, so scheint er dennoch im Lichte eines Wertes oder Gutes zu handeln: des selbstbestimmten Lebens oder der Freiheit bspw.

An dieser Stelle sei jedoch angemerkt, dass der Schwerpunkt der vorliegenden Arbeit *nicht* in einer Suche nach der Quelle von normativen Gründen liegt. Vielmehr dienen meine Ausführungen, sowie auch die gewonnen Erkenntnisse zur emotionalen Begründung von Werten, gerade dazu, die *Verbindung* zwischen normativen und motivierenden Gründen deutlich zu machen. Mein Ansatz kann eine Verbindung zur Verfügung stellen, die erklärt, warum eine gegebene praktische Überzeugung (als normativer Grund) mit einem Handlungsmotiv verbunden ist (als motivierendem Grund).

4.7. Wie ist das „moral problem“ zu lösen?

Oder: das „reconciliation problem“

Rufen wir uns noch einmal die Problemlage vor Augen, die Smith im Rahmen der Inkonsistenz der bereits oben zitierten drei Propositionen gefasst hatte:

- „1. Moral judgments of the form ‚It is right that I Φ ‘ express a subject’s beliefs about an objective matter of fact, a fact about what it is right for her to do.
2. If someone judges that it is right that she Φ s then, ceteris paribus, she is motivated to Φ .
3. An agent is motivated to act in a certain way just in case she has an appropriate desire and a means-end belief, where belief and desire are, in Hume’s terms, distinct existences.“ (Smith 1994, 11)

Wir hatten weiterhin gesehen, dass sich das Problem – die Frage, wie Handeln aus Gründen zu erklären sei – herunterbrechen lässt auf die Frage, welcher Erklärungsansatz Richtigkeit beanspruchen kann: eine Humeanische (psychologistische oder emotivistische) Herangehensweise, oder eine Kantianische (kognitivistische) Herangehensweise. Dabei ist deutlich geworden, dass wir offenbar weiter unterscheiden müssen zwischen Gründen, die Handeln motivieren, und Gründen, die Handeln rechtfertigen. Prinzipiell kann für beide Kategorien ein individuelles Theorieparadigma zur Erklärung gewählt werden. Die zentrale Frage lautet dann jedoch, wie beide Kategorien wieder miteinander zu verbinden seien, damit erklärbar wird, wie ein Akteur durch *denselben* Grund zum Handeln motiviert wird, der dieses Handeln auch normativ rechtfertigt. Wir können einige Punkte formulieren, in denen unsere Intuitionen uns nahelegen, was eine adäquate Theorie des Handelns aus Gründen erklären können müsste. Wiland (2002, 451f.) formuliert drei Kriterien, die ich für relativ unkontrovers halte.

- (1.) Praktisches Überlegen hat eine *explanatorische* Dimension: wenn wir aus einem Grund heraus handeln, dann kann das Anführen eben dieses Grundes erklären, warum wir so handelten, wie wir handelten. Eine adäquate Theorie praktischen Überlegens sollte zeigen können, wie Gründe Handlungen erklären können. Wir finden hier Smith's Begriff des Prinzips der praktischen Implikationen wieder.
- (2.) Praktisches Überlegen hat eine *normative* Dimension: in Fällen, in denen wir nicht so handeln, wie wir Grund hätten zu handeln, scheint unserem Handeln Rechtfertigung zu fehlen. Mit anderen Worten können Gründe unser Handeln rechtfertigen. Hier finden wir wiederum Smith's Objektivitätsprinzip.
- (3.) Praktisches Überlegen unterliegt der „Erklärungs- und Rechtfertigungs-Bedingung“ (Dancy 2000, 101ff.: „*explanatory/normative constraint*“). Diese Bedingung verlangt, dass eine Theorie praktischen Überlegens Gründe nach (1) und Gründe nach (2) zur Deckung bringen kann. D.h. Gründe, die eine gegebene Handlung *h* rechtfertigen (2), sollen möglichst auch im tatsächlichen explanatorischen Schema (1) von *h* vorkommen. Umgekehrt verlangt (3), dass derselbe Grund, der meine Handlung *h* erklären kann, auch in der Lage ist, *h* zu rechtfertigen. In Kürze gefasst kann man sagen: Eine Theorie praktischen Überlegens soll zeigen können, wie Gründe, die eine Handlung rechtfertigen, dieselbe Handlung gleichzeitig erklären können, und umgekehrt. Und mit diesem Punkt (3) erfasst Wiland die Problemlage, die Smith eben „moral problem“ nennt.

Weiterhin hatten wir angekündigt, genau wie Smith auch im Rahmen der Humeanischen Theorie motivierender Gründe zu bleiben. Der Humeanische Psychologismus hat keine Probleme damit, (1) zu erfüllen. Wenn wir nach einer Erklärung dafür suchen, warum Hans h ausführte, lässt sich sehr plausibel auf einige von Hans' mentalen Zuständen verweisen. Diese mentalen Zustände (wie etwa pro-Einstellungen) stellen dann die Erklärung dafür dar, warum Hans h ausführte. Allerdings stellt (2) Probleme dar für eine solche Theorieform. Eine solche quasi-faktische Erklärung von Handeln kann evtl. die Dimension der Normativität nicht erfassen. Hans könnte z.B. einen Grund haben, h auszuführen, obwohl er keinen mentalen Zustand hat, der ihn dazu motiviert. Nehmen wir an, Hans (als wohlzogener Junge) hätte einen Grund, der alten Dame seinen Sitzplatz anzubieten, ohne jedoch einen entsprechenden mentalen Zustand zu haben (er hat keine Lust, oder es kümmert ihn nicht etc.). In diesem Fall hätte Hans einen rechtfertigenden Grund (2), h auszuführen, ohne jedoch einen explanatorisch relevanten Grund zu haben.

Demgegenüber hat der Kognitivismus keine Probleme, gerade den normativen Aspekt des praktischen Überlegens zu erfassen, indem er Gründe zum Handeln als nicht-mentale, gewissermaßen objektive Tatsachen versteht bzw. als universalisierbare Vernunftforderungen (Scarano 2001). Ein rechtfertigender Grund, h auszuführen, der sich aus einer bestimmten objektiven Tatsachenkonstellation ergibt, würde bestehen, auch wenn Hans kein Motiv hat, h auszuführen, bzw. h nicht ausführt. Wie kann jedoch der Kognitivismus die Bedingung (1) erfassen, d.h. wie kann eine Tatsache, die völlig unabhängig von Hans' mentalem Zustand ist, erklären, warum Hans h ausführte? Wiland (2002, 453) führt eine überzeugende Illustration an, warum nicht-mentale Tatsachen keine erklärenden Gründe zum Handeln liefern können:

Nehmen wir an, Hans nimmt im Winter jeden Abend eine Tablette eines immunstärkenden Mittels aus Echinacea-Extrakt ein, weil er glaubt, dass sich dadurch Grippeerkrankungen vermeiden lassen, und weil Hans eine pro-Einstellung gegenüber der Vermeidung von Grippen hat.¹⁰⁷ Wir erklären Hans' Handlung durch Anführen von mentalen Zuständen (seine Wünsche, Absichten, Glaubenszustände etc.). Nehmen wir nun an, Hermann teilt Hans' Überzeugung zur Wirksamkeit von Echinacea-Extrakt, hat jedoch eine andere Konstellation seiner mentalen Zustände, sodass diese mit der Tatsache vereinbar sind, dass

¹⁰⁷ In diesem Schema von Gründen, die eine Handlung erklären, habe ich einem davidsonischem Modell gemäß einen Glaubenszustand und eine pro-Einstellung aufgeführt. Welchem der beiden Phänomene nun welcher Rang zukommt, soll an dieser Stelle nicht diskutiert werden. Wichtig ist hier nur, dass die Bedingung des Psychologismus, nämlich das Vorhandensein eines mentalen Zustands, in jedem Fall erfüllt ist, sei es, dass man diesen Zustand als einen Glaubenszustand oder als eine pro-Einstellung versteht.

Hermann nicht täglich Echinacea-Extrakt nimmt. Vielleicht glaubt Hermann, dass er sowieso kerngesund ist, oder er hält die Vermeidung von Grippe nicht für besonders wichtig, oder er hat schlicht keine Lust, jeden Tag eine Tablette einzunehmen. Der Psychologismus kann also ein aktorsrelatives Schema von Handlungserklärungen liefern. Nehmen wir jedoch mit dem Kognitivismus nun an, dass Hans die Tabletten einnimmt, nicht weil er entsprechende mentale Zustände aufweist, sondern weil eine objektive Tatsache seine Handlung erklärt. Hier ergibt sich das folgende Problem in Bezug auf Hermanns Unterlassung dieser Handlung: wenn es eine Tatsache in der Welt gibt, die Hans' Handlung hinreichend erklärt, dann müssen wir annehmen, dass dieselbe Tatsache auch im Fall von Hermann besteht. Dann wäre jedoch das Bestehen dieser Tatsache (die Hans' Handlung notwendig erklärt oder determiniert) nicht vereinbar mit Hermanns Unterlassung derselben Handlung, da der Kognitivismus eben nicht auf die individuelle Konstellation mentaler Zustände verweisen kann. Im Übrigen scheint es schwierig zu sein, sich vorzustellen, was eine solche objektive Tatsache sein könnte (im Sinne eines kognitivistischen Realismus).

Welche objektive Tatsache könnte eine Handlung hinreichend erklären bzw. determinieren, in allen analogen Fällen (d.h. für vergleichbare Akteure in vergleichbaren Situationen etc.)? Interessanterweise sind die plausibelsten Beispiele für eine solche Tatsache stets mit nicht-reflektierten oder instinkthaften Handlungen befasst. Wir können z.B. sagen dass der Braunbär in den Winterschlaf geht, weil die entsprechende Jahreszeit gekommen ist, d.h. weil ein bestimmter kalendarischer Zeitpunkt eingetreten ist. Der Verweis auf diese Tatsache kann die Handlung des Braunbärs hinreichend erklären, bzw. determiniert diese Handlung. Und in diesem Fall können wir auch behaupten, dass wenn das Bestehen dieser Tatsache einen Grund zum Schlafen für Puh den Bären darstellt, dann stellt das Bestehen derselben Tatsache auch einen Grund für Balu den Bären dar. Balus Unterlassung wäre in diesem Fall nicht kompatibel mit dem Bestehen dieser Tatsache. Wichtig anzumerken bei diesem Beispiel ist, dass es um erklärende Gründe geht, nicht um rechtfertigende. Offenbar kann der Realismus ein adäquates Modell erklärender Gründe zum Handeln liefern, wo sich die Frage des Normativen nicht stellt: wie etwa bei instinktiven oder triebhaften Reaktionen in der Tierwelt. Sobald wir jedoch die Möglichkeit der Unterscheidung von erklärenden und rechtfertigenden Gründen annehmen, scheint der Realismus zu versagen, wenn es darum geht, ein Modell der hinreichenden Handlungserklärung zu liefern, ohne sich auf normative Gründe zu beschränken.

Wir haben oben jeweils die Problemlage in der Erklärung von motiverenden und normativen Gründen diskutiert. Mit Wilands Punkt (3) tritt nun eben jenes Problem hinzu, wie

beide Kategorien von Gründen im Rahmen einer „kompatibilistischen“ Theorie (Wiland 2002) zu koordinieren seien. Das damit neu hinzutretende Problem, Smith's „moral problem“, nennt Wiland (ebd., 454) treffend „the reconciliation problem“:

„Suppose that one and the same reason could not both explain and justify Agnes's Φ -ing. This means that the thing that explains why she Φ -ed is necessarily different from the thing that makes her Φ -ing the appropriate thing to do. Try as she might, Agnes cannot Φ for the reason that makes Φ -ing appropriate. On this view, a person never acts for the reason that makes that action reasonable. But it seems as though we sometimes act for the reason that justifies our action [...]. On the compatibilist view, however, this is impossible.” (Wiland 2002, 455f.)

Jenseits von Wilands treffender Analyse der Problemlage teile ich jedoch nicht seine abschließende Beurteilung: dass eine kompatibilistische Lösung nicht möglich sei. Im Folgenden soll daher dargestellt werden, wie Michael Smith das „moral problem“ bzw. das „reconciliation problem“ zu lösen versucht, und im Anschluss daran mein eigener Lösungsvorschlag mit Blick auf die in früheren Kapiteln erarbeiteten Ergebnisse gegeben werden.

4.8. Michael Smith's Lösungsansatz

Smith (1994, Kap. 5) geht davon aus, dass die Humeanische Handlungstheorie adäquat zur Erklärung von motivierenden Gründen geeignet ist (zur Verteidigung des Humeanischen Psychologismus vgl. auch Wiland 2003), und dass konkurrierende Anti-Humeanische Ansätze sogar defizitär sind. Sein Lösungsansatz zum „moral problem“ im Sinne des von Wiland (2002) verorteten „reconciliation problem“ ist jedoch eine Revision der Humeanischen Theorie normativer Gründe. Letztere scheint nämlich verantwortlich zu sein für das Scheitern des Humeanismus am „moral problem“. Das Problem am Humeanismus ist, dass die desires eines Akteurs nicht rational zugänglich oder kritikfähig sind:

„The distinctive Humean view of normative reasons, then, is that the rational thing for an agent to do is simply to act so as maximally to satisfy her desires, whatever the content of those desires.“ (Smith 1994, 130).

Verfolgen wir nun Smith's Lösungsansatz. Bekanntlich unterscheidet Smith – wie auch Philip Pettit – zwischen der *intentionalen* und der *deliberativen* Erklärung von Handlungen (Pettit & Smith 1990; Pettit 1993, 20-24). Erstere befasst sich mit den psychologischen Zuständen, die eine Handlung eines Akteurs erklären, letztere mit den Propositionen (d.h. vom Akteur erwägten Gründen und Gedanken), die eine Handlung erklären oder rechtfertigen. Nach Smith (1994, 132) besteht eine intentionale Erklärung aus den Elementen

„desiring“ und „believing“ (einem klassischen Humeanischen Paradigma also), während eine deliberative Erklärung aus den Elementen „valuing“ und „believing“ besteht.¹⁰⁸

Smith baut sein Modell von normativen Gründen damit auf dem Wert-Begriff auf: wir haben einen (normativen) Grund etwas zu tun, wenn wir damit einen Wert herbeiführen oder befördern. Daher stellt sich für ihn die Frage danach, wie normative mit motivierenden Gründen verknüpft seien, als Frage danach, wie ‚Wert‘ zu erklären sei: im Sinne von „desiring“, oder im Sinne von „believing“? Und somit ergibt sich für die Verknüpfung von normativen Gründen mit motivierenden Gründen folgende Situation: wenn Werte durch Glaubenszustände erklärt werden, dann kann im Rahmen eines Humeanischen Paradigmas wiederum nicht erklärt werden, warum wir auf Grund unserer normativen Gründe jemals handeln. Wenn Werte durch desires erklärt werden, dann kann umgekehrt nicht erklärt werden, warum wir jemals Handlungen unterlassen, zu denen wir einen (normativen) Grund haben. Beides ist nicht konsistent mit der empirischen Sachlage.

Wir haben bereits festgehalten, dass Smith's Lösung des moral problem einerseits eine Beibehaltung des Humeanischen Paradigmas für motivierende Gründe beinhaltet. Diesen Teil werden wir im Folgenden übergehen, da ich Smith's Argumentation für unkontrovers halte: vieles spricht für einen Humeanismus in der Erklärung motivierender Gründe, und sollte der Leser Anstoß an diesem Erklärungsparadigma nehmen, so wollen wir diesen Teil der Lösung dennoch im Sinne von Smith mittragen, allein um des *principle of charity* in der Argumentation willen. Der zweite Teil der Smith'schen Lösung besteht nun in einer Revision der Humeanischen Theorie normativer Gründe. Wir finden dazu zwei Formulierungen bei Smith, u.z. einmal mit explizitem Bezug auf ein Modell „normativer Gründe“:

„To say that we have a normative reason to Φ in certain circumstances C is to say that we would want ourselves to Φ in C if we were fully rational. If there is a normative reason for some agent to Φ in certain circumstances C then there is a like normative

¹⁰⁸ Damit ist die schon von Watson (1975, 101) herausgehobene Diskrepanz zwischen einer Handlungserklärung durch desires oder durch Werte angesprochene (da sich beide Begriffe chiastisch jeweils in seinem Modell motivierender bzw. normativer Gründe wiederfinden): „Consider the case of a woman who has a sudden urge to drown her bawling child in the bath; or the case of a squash player who, while suffering an ignominious defeat, desires to smash his opponent in the face with the racquet. It is just false that the mother values her child's being drowned or that the player values the injury and suffering of his opponent. But they desire these things none the less. They desire them in spite of themselves. It is not that they assign to these actions an initial value which is then outweighed by other considerations. These activities are not even represented by a positive entry, however small, on the initial ‚desirability matrix‘.“

reason for all those who find themselves in circumstances C to Φ .” (Smith 1994, 182)¹⁰⁹

Darin enthalten ist eine bereits früher erwähnte Prämisse, nämlich:

„C2: If an agent believes that she has a normative reason to Φ , then she rationally should desire to Φ .” (Ebd., 148)

Smith entwirft zur Lösung des von ihm eigens diagnostizierten „moral problem“ bzw. später so apostrophierten „reconciliation problem“ also gewissermaßen eine „Ideal Advisor Theory“ (Sayre-McCord 1997, 58) normativer Gründe. In Kapitel 6 seines Buches (insbes. Smith 1994, 185) buchstabiert Smith schließlich seine Lösung aus, u.z. mit Bezug auf die oben zitierten drei Propositionen (an denen das „moral problem“ im Sinne eines Inkonsistenz-Problems demonstriert wurde, vgl. Kap. 4.2. der vorliegenden Arbeit), von denen zu zeigen war, dass sie wahr und miteinander konsistent seien.

(Ad 1. „Objektivitätsprinzip“) Smith’s Begriff von „richtig“ ist die Konzeption von dem, was wir zu tun wünschen würden, wenn wir vollständig rational wären, wobei dieser Wunsch sich substantziell auf ein angemessenes Objekt beziehen muss. Daher können wir in der Tat unsere moralischen Überzeugungen als Ausdruck unserer Glaubenszustände über objektive Tatsachen verstehen. Denn unsere moralischen Überzeugungen drücken aus, welche normativen Gründe für Handlungen wir zu haben glauben, wobei Smith diese Gründe – im Sinne eines Anti-Humeanischen Modells – als kategorische Vernunftpflichten versteht. Aus dem gerade Ausgeführten geht hervor, dass diese Überzeugungen auch im Sinne von (2. „Prinzip der praktischen Implikationen“) mit Motivation verknüpft sind. Und Smith hat weiterhin gezeigt, dass diese motivationale Wirksamkeit im Rahmen einer Humeanischen Handlungstheorie von Glaubenszuständen und desires als distinkte Phänomene zu erklären ist (ad 3. „Humescher Psychologismus“). Das „moral problem“ scheint damit im Sinne von Smith gelöst.

„My handing back a wallet I found in the street in such and such circumstances is right, for example, only if, under conditions of full rationality, we would all want that if we find a wallet in the street in such and such circumstances, then we hand it back. Of course, if this is indeed true, then it is an *a priori* truth. The fact that we would have

¹⁰⁹ Wedgwood (2002, 349) hält Smiths Formulierung für missverständlich: „Michael Smith interprets [the internalism requirement] as the claim that it is rational for A to Φ in circumstances C just in case A would want herself to Φ in C if she were fully rational. But to say that it is rational for A to Φ is just to say that it is rationally *permissible* for A to Φ , not that A is rationally *required* to Φ : even if it is rational for A to Φ , it might also be rational for A not to Φ but to do something else instead. So, even if it is rational for one to Φ , this does not entail that one *would* want oneself to Φ if one were fully rational; at most, it entails that wanting oneself to Φ is *compatible* with being fully rational“.

such a desire under conditions of full rationality will be a consequence of the theory that systematically justifies our desires, and this task of theory construction is itself a relatively *a priori* enterprise [...]“ (Smith 1994, 187)

Smith vertritt also einen moralischen Rationalismus *normativer* Gründe, nämlich die These, dass Überzeugungen darüber, *was wir moralisch zu tun verpflichtet sind*, gleich Überzeugungen darüber sind, *was wir zu tun einen Grund haben* (ebd., 58).¹¹⁰ Weiterhin vertritt Smith die These, dass Überzeugungen darüber, *was wir zu tun einen Grund haben*, gleich Überzeugungen darüber sind, *was wir zu tun wünschen* würden, wenn wir vollständig rational wären.¹¹¹ Smith's Lösung des moral problem ist also eine dezidiert anti-Humeanische Theorie von normativen Gründen (eben jene „Ideal Advisor Theory“), die jedoch kompatibel mit einer Humeanischen Theorie von Motivation (d.h. von motivierenden Gründen) ist, dass Akteure nur dann handeln, wenn zusätzlich zu einer Überzeugung auch das entsprechende desire vorliegt. Smith lehnt jedoch einen strengen Internalismus ab, dem gemäß Akteure stets entsprechend ihrer Überzeugung darüber handeln, *was sie zu tun einen Grund haben*. Diese Art von Internalismus ist empirisch widerlegbar. Stattdessen vertritt er einen „defeasible internalism“, also die These, dass die „*ceteris-paribus*“ Bedingung in der Internalismus-Proposition „If someone makes a moral judgment then – *ceteris paribus* – she is motivated appropriately“ eine *Rationalitätsbedingung* ist. D.h. wir handeln aufgrund unserer Überzeugungen, *sofern* wir vollständig rational sind. Gegenbeispiele wie der Zustand schwerer Depression oder Willensschwäche zeigen nach Smith eben, dass die direkte Verbindung von Überzeugungen und Handlungsmotivation aufgehoben ist, weil unsere Rationalität als Akteur kompromittiert ist. Damit positioniert er sich entgegen dem strengen Internalismus, dem gemäß die Verbindung von „*einen Grund haben*“ zu „*motiviert sein zum*

¹¹⁰ Der Unterschied zwischen moralischen Überzeugungen einerseits und anderen (praktischen) Überzeugungen darüber, *was wir zu tun einen Grund haben* andererseits muss natürlich herausgearbeitet werden. Dies ist jedoch nicht Gegenstand der vorliegenden Arbeit. Es sei lediglich darauf hingewiesen, dass bspw. im Sinne Kants das Kriterium der Kategorizität angewendet werden kann (so Scarano 2001). Bei Smith selbst (1994, 183f.) scheint sich auch die Ansicht zu finden, wonach moralische Überzeugungen – im Gegensatz zu nicht-moralisch relevanten praktischen Überzeugungen – sich auf fundamentale menschliche Werte beziehen, wie z.B. „*human flourishing*“, *Gerechtigkeit* etc.

¹¹¹ Dieser Zustand des „vollständig rational sein“ wird nach Sayre-McCord (1997, 58) inhaltlich gefüllt mit folgenden Bedingungen: „Thus, as Smith would have it, a person has reason to Φ in C if and only if were she fully rational she would want herself, as she actually is, to do it – where her being fully rational would be a matter of her having no false beliefs, her having all relevant true beliefs, and her deliberating correctly. It is morally right for a person to Φ in C, Smith then holds, if and only if she has reason (of the appropriate sort and strength) to do it.“

Handeln“ nicht durch ein Mediatorkriterium wie „praktische Rationalität“ hergestellt wird, sondern diese Verbindung schon begriffsanalytisch, sozusagen *a priori*, besteht.¹¹²

Smith's Lösung des „moral problem“ hängt entscheidend von der Interpretation seiner Proposition „C2“ ab, nämlich: „If an agent believes that she has a normative reason to Φ , then she rationally should desire to Φ .“ (Smith 1994, 148). Genau mit dieser Proposition soll schließlich die gesuchte Verbindung zwischen normativen und motivierenden Gründen hergestellt werden, wobei „desiring to Φ “ (unter der Voraussetzung, dass ein gegebener Zustand der Welt eine entsprechende Zweck-Mittel-Überzeugung rechtfertigt, sozusagen im Sinne einer „triggering condition“) als ein motivierender Grund aufgefasst wird. Aus einer Perspektive der metaethischen Semantik und der moralischen Epistemologie (Scarano 2001, 16) könnte diese Proposition nun unter folgendem Gesichtspunkt untersucht werden:

„C2.1.: A believes that she has a normative reason to Φ “,
ergo: „C2.2.: A rationally should desire to Φ .“

Es ginge also um die (zentrale) Frage in Smith's Lösungsansatz, wie zu erklären sei, dass die Überzeugung, man habe einen normativen Grund zu handeln, gleichbedeutend sei damit, dass man auf derselben Grundlage auch einen motivierenden Grund hat zu handeln, sofern man vollständig rational ist. Es ginge um eine Untersuchung der (logischen) Schlüssigkeit dieser komplexen Proposition. Das Feld dieser Fragestellung soll jedoch im Folgenden nicht weiter bearbeitet werden. Natürlich ließe sich auf der Grundlage eines fundamentalen Skeptizismus, wie ihn bspw. Lewis Carroll für das logische Schließen anhand seines Gleichnisses mit Achilles und der Schildkröte illustriert hat (Carroll 1895, vgl. Engel 1998), diese Schlüssigkeit stets anzweifeln. Wir können immer auch wieder von neuem bezweifeln, dass aus einer Überzeugung von einem normativen Grund irgendein Handlungsmotiv hervorgehen könne.

Betrachten wir jedoch vielmehr die moralpsychologischen Implikationen dieser Smith'schen These. Das von Sayre-McCord (1997) erarbeitete Fazit dazu liefert uns in dieser Hinsicht eine sinnvolle Hinführung. Sayre-McCord diagnostiziert, dass Smith's Analyse unserer moralischen Begriffe (v.a. normativer und motivierender Gründe) zwar zutreffend sei,

¹¹² Vgl. Sayre-McCords (1997, 64 FN 15) Illustration: „When Huck Finn, for instance, finds that he can't bring himself to turn Jim in, even though he believes he has a duty to, the problem is not with Huck's motives, but with his evaluative belief, and it is the belief he has reason to change rather than the desire. It may well be irrational not to desire to do what one believes one has reason to do or what one believes would be morally right to do, but it may be one's beliefs, rather than one's desires, that have gone wrong.“

dass wir aber dadurch noch immer nicht das Prinzip praktischer Implikationen („practicality requirement“) moralischer Überzeugungen erklären können. Mit Sayre-McCords Worten (ebd. 83): „the approach we [Smith und Sayre-McCord, S.H.] share still has nothing close to an adequate explanation of the practicality of evaluative thought“ und fährt (ebd., FN 32) fort:

„My own view is that we will find an explanation of C2 not in an account of our concept of what it is to have a reason but instead in an account of our concept of what it is to value something (where this is, contra Smith, not merely a matter of having a belief with a certain content)“.

Ich möchte entgegen Sayre-McCord festhalten, dass ich sowohl denke, dass Smith's Analyse motivierender und normativer Gründe richtig sei, als auch, dass anhand dieser Analyse das gesuchte Prinzip praktischer Implikationen (unter Wahrung des Objektivitätsprinzips) zu erklären sei. Die dezidiert moralpsychologische Erklärung, auf die auch Sayre-McCord in seinem Fazit abhebt, lässt sich – so räume ich ein – noch nicht allein aus Smith's Analyse gewinnen. Und in eben dieser Richtung weist Sayre-McCords Anmerkung den Weg, den wir nun auf der Grundlage der in der vorliegenden Arbeit gewonnenen Erkenntnisse beschließen können. Wie genau lässt es sich moralpsychologisch erklären, dass wir ein aktuales, in einer durch einen motivierenden Grund getragenen Handlungserklärung einsetzbares *desire* (im Sinne von C2.2.) gewinnen können aus einer *Überzeugung* (im Sinne von C2.1.), dass wir einen entsprechenden normativen Grund haben?

4.9. Die Lösung des „moral problem“ aus moralpsychologischer Perspektive

Es scheint also offensichtlich, dass Smith's Lösungsvorschlag zentral von dem Dreh- und Angelpunkt eines *desire*-Modells abhängt. Es muss schließlich gezeigt werden, wie ein *desire* aus einer Überzeugung zu generieren sei. So stellt auch Smith daher die Frage, was denn nun ein „*desire*“ sei (Smith 1994, Kap. 4.5). Dieser Frage sind wir an früherer Stelle ausführlich nachgegangen. Hier soll nur in aller Kürze Smith's Gedankenlinie aufgenommen werden, um vor diesem Hintergrund eine alternative Theorie entwickeln zu können. Smith (ebd., 104) betrachtet zunächst den phänomenologischen Ansatz, „*desire*“ als „*feeling*“ zu interpretieren. Hume selbst scheint diesen phänomenologischen *desire*-Begriff zu haben. Im „*Treatise*“ sagt er, dass *desires* eine Art von „*passions*“ seien, und „*passions*“ wiederum eine Art von „*feeling*“. ¹¹³ Wir müssten uns unserer *desires* also ständig bewusst sein, oder wie Barry

¹¹³ Hume, *Treatise* II, 3,3: „It is obvious, that when we have the prospect of pain or pleasure from any object, we feel a consequent emotion of aversion or propensity, and are carried to avoid or embrace what will give us this

Stroud sagt: unsere desires wären stets „phänomenologisch salient“ (1977, 163). Es ist schon intuitiv nicht plausibel, dass diese Sichtweise zutreffen kann, da die Mehrzahl unserer desires tatsächlich über weite Strecken dormant sein dürften, und nur selektiv oder situational salient werden. Auf diesen Einwand, der auf die Dichotomie von Aktualität und Disposition abhebt, werden wir im Rahmen meines eigenen Lösungsvorschlags noch zurückkommen. Weiterhin gibt Platts (1981, 74-77) zu bedenken, dass desires zwar einerseits wie Empfindungen (also „feelings“, also phänomenologisch salient) seien, insofern sie einen *phänomenologischen* Gehalt haben; sie jedoch andererseits wiederum verschieden von Empfindungen seien, insofern sie auch einen *propositionalen* Gehalt haben (für weitere Einwände vgl. Smith 1994, 105f.). Die in der Einleitung zu dieser Arbeit bereits angesprochene emotivistische Konzeption ist geradezu dezidiert im Gegensatz zum Geiste der Platts'schen Konzeption aufgestellt, insofern „desires“, verstanden als emotive Pro- oder Contra-Einstellungen, eben rein motivationale Zustände ohne kognitiven Inhalt sein sollen.¹¹⁴ Platts (1979, 256f.) bezieht sich auf Anscombe (1957), um seine These zu formulieren, dass desires die für sie typische Passensrichtung haben, die gerade in umgekehrter Richtung verläuft wie bei Glaubenszuständen. Auch Davidson (1963, 4) versteht den desire-Begriff in diesem Sinne und schlägt daher bekanntlich vor, statt „desire“ den Begriff „Pro-Einstellung“ zu verwenden im Sinne von „psychological state with which the world must fit“. Platts räumt jedoch bekanntlich ein, dass auch desires Elemente von Glaubenszuständen mit einschließen, dass also ihre

uneasiness or satisfaction. It is also obvious, that this emotion rests not here, but making us cast our view on every side, comprehends whatever objects are connected with its original one by the relation of cause and effect. Here then reasoning takes place to discover this relation; and according as our reasoning varies, our actions receive a subsequent variation. But it is evident in this case that the impulse arises not from reason, but is only directed by it. It is from the prospect of pain or pleasure that the aversion or propensity arises towards any object: And these emotions extend themselves to the causes and effects of that object, as they are pointed out to us by reason and experience.“

¹¹⁴ In neuerer Zeit ist vermehrt auf die Implausibilität hingewiesen worden, eine solche Konzeption unter den Begriff des „Emotivismus“ zu fassen, sofern damit eine ernst gemeinte Anlehnung an den Emotionsbegriff impliziert ist: „Instead of exploring the specific ethical roles of various emotions, reactive attitudes, desires, and conceptions of worth, these were generically lumped under headings such as ‘pro-attitudes’ (or the opposite), which hardly revealed the textures of the issues“ (Jacobs 2001, 89). In diesem Sinne nennt auch Kupperman (1999, 16f.) den Emotivismus eine „rhetorical strategy“, insofern dieser zwar die Geringschätzung von „Emotion“ als vermeintlich nicht-rationales Phänomen ausnutzt, ohne jedoch einen substanziellen Emotionsbegriff liefern zu können. Schon Urmson (1968) merkte scharfsichtig an, dass der Emotivismus gar nicht über „Emotionen“ in einem substanziellen Sinne handle, sondern über „Einstellungen“ („attitudes“). Und wir können zahlreiche Einstellungen vertreten, ohne zugleich eine entsprechende Emotion zu erleben.

Passensrichtung nicht ausschließlich in die ihnen typische Richtung weist. Dies ist aber problematisch, wenn gerade die Passensrichtung – nach Anscombe – als klassifizierendes Kriterium hergenommen wird. Platts zweifelt, ob es solche Zustände gibt, die zwar in ihrem Aufbau hybrid, in ihrer Passensrichtung einheitlich seien.

Nun gibt es Lösungsversuche zu beiden vorgestellten Problemfeldern, die hier nur angerissen werden sollen. Hume selber schlägt eine Lösung des Problems der phänomenologischen Salienz vor, indem er weiterhin die Existenz von sog. „calm desires“ annimmt:

„Now it is certain, there are certain calm desires and tendencies, which, though they be real passions, produce little emotion in the mind, and are more known by their effects than by the immediate feeling or sensation.“ (Treatise II, 3, 3, 8).

Können wir – als Entgegnung auf den zweiten Einwand – nun annehmen, dass Pro-Einstellungen sowohl motivational als auch kognitiv gehaltvoll sind? Auch Smith (1994, 118ff.) diskutiert die Überlegung, ob es evtl. einen Zustand geben könnte, der beide Passensrichtungen aufweist. Ein solcher Zustand wurde von Altham (1986) hypothetisch „besire“ genannt (als Wortprägung aus den beiden Begriffen „belief“ und „desire“). Ein „besire“, dass es richtig ist zu Φ -en wäre charakterisiert dadurch, dass es ein Zustand ist, der auf die Welt passen muss, weil er sich auflöst, sobald der Akteur mit einer kognitiven Information konfrontiert wird, dass es nicht richtig ist, zu Φ -en. Andererseits muss die Welt auf den Zustand passen, weil dieses „besire“ dem Akteur eine Disposition zu Φ -en zuschreibt. Smith nennt dies eine anti-Humeanische Perspektive, die das „moral problem“ von vorneherein aufhebt. Wir können darin auch die Kantische Perspektive finden. Smith konzediert nun, dass ein Humeaner nicht leugnen muss, dass ein Akteur zugleich in einem Zustand sein kann, zu glauben, dass es richtig ist zu Φ -en und zu Φ -en zu wünschen (desire). Allerdings ist das Humeanische Argument von desire und belief als distinkten Phänomenen anders zu verstehen: Aus der Tatsache, dass ein Akteur prinzipiell in einem Zustand sein kann, in dem er überzeugt ist, dass es richtig ist, eine bestimmte Handlung auszuführen, und dennoch nicht das entsprechende desire hat, lässt sich schließen, dass „desire“ und „belief“ eben distinkte Phänomene sind.

Smith (1994, 177) analysiert „desirability“ wie folgt: Wünschenswert („desirable“) ist das, was wir wünschen („desire“) würden, wenn wir vollständig rational wären. Dies sieht Smith als ein Modell von Werten als Glaubenszuständen (im Gegensatz zu Werten als desires). Smith's Formel lautet bekanntlich: „C2: If an agent believes that she has a normative reason to Φ , then she rationally should desire to Φ “. Wenn wir jedoch die oben diskutierten Einwände betrachten (und umso mehr, wenn wir die in Kap. 3 dieser Arbeit gewonnen

Erkenntnisse hinzuziehen), so wird deutlich, dass Smith's Lösung nicht in letzter Instanz überzeugen kann. Wir können seine These zwar metaethisch deskriptiv verstehen, insofern damit gesagt ist: „Ein Akteur, der der Überzeugung ist, er habe einen normativen Grund zu Φ -en, sollte auch das entsprechende desire haben. Anderenfalls ist diese Akteur irrational zu nennen.“ Wir hatten uns jedoch gerade zum Ziel gesetzt, den *moralpsychologischen* Mechanismus dieses mysteriösen Übergangs von einem normativen Grund (den Smith nach dem rationalistischen Paradigma versteht) zu einem aktuellen desire (das Smith im Rahmen des Humeschen Paradigmas versteht) aufzuklären. Und gerade dies kann Smith's Lösungsvorschlag nicht auf der Basis des bestehenden desire-Modells leisten. Der Humesche desire-Begriff, dem auch Smith folgt, postuliert schließlich „kontingente desires“ in dem Sinne von „a desire that the agent is not in any way rationally required to have“ (Wedgwood 2002, 345). Eine solche Konzeption löst gerade nicht das Problem der Verantwortlichkeit oder des „endorsements“ von desires, das bspw. über die Zuschreibung eines kognitiven Aspekts gelöst werden könnte, was jedoch im Rahmen des Humeanischen Paradigmas nicht möglich ist. Es ist also nicht einsichtig, warum eine Überzeugung ein desire generieren können sollte. Wenn dies nämlich der Fall wäre, stellte sich das „moral problem“ überhaupt nicht in der Form, wie Smith es diagnostiziert. Schon Williams (1980, 109) betont, dass falls ein Prozess des Überlegens ein neues desire hervorbringen sollte, so müssten wir dennoch von einem primären desire ausgehen, von dem *ausgehend* der Akteur überhaupt erst praktische Überlegungen angestellt hat, um seine neue Handlungsmotivation zu erreichen. Wallace (1990, 370) nennt dies das „desire-out, desire-in“-Prinzip: „processes of thought which give rise to a desire (as ‚output‘) can always be traced back to a further desire (as ‚input‘)“.

Ich möchte daher im Folgenden eine eigene Lösungsalternative dieses moralpsychologischen Problems vorschlagen, u.z. auf der Grundlage der in früheren Kapiteln erarbeiteten Erkenntnisse zu „Emotion“, „Wert“ und „desire“. Eine zentrale Rolle in diesem Lösungsvorschlag nimmt der Begriff der „emotionalen Einstellung“ ein, zu dessen Definition ich auf in Kap. 2.11. von mir formulierte Thesen zurückgreife, u.z. insbesondere *T7* („Disposition vs. Aktualität“) und *T6* („Manifestation von Emotionen in Handlungstendenzen“). Als „emotionale Einstellung“ verstehe ich zunächst eine emotionale Disposition in der Art, wie dieser Begriff bereits in der Einleitung von akuten emotionalen Zuständen („occurrent emotional states“) unterschieden wurde. Weiterhin jedoch handelt es sich bei einer emotionalen Einstellung entweder (1.) um die dispositionale Form einer Sekundäremotion (wie Scham, Empörung etc., die also per se schon einen starken kognitiven Anteil hat), oder aber (2.) um die dispositionale Form einer Primäremotion, verbunden mit

einer entsprechenden „heißen Kognition“ (Goller 1992) desselben Fokus (Helm 2002) bzw. demselben Schlüsselszenario (DeSousa 1979) zugehörig. Diese zweite Form lässt sich in Anlehnung an Damasio früher diskutierte „als-ob-Emotionen“ verstehen, die die emotionale Tönung von Bewusstseinsinhalten beschreiben, unabhängig von der für akute emotionale Zustände so typischen peripher-physiologischen Aktivierung. In einer Weise wie wir sie anhand der Theorien von Damasio, DeSousa, Helm und Heller kennengelernt haben, ist davon auszugehen, dass bestimmte „heiße“ Kognitionen, also solche mit besonderer Adressatenqualität wie bspw. praktische Überzeugungen, Handlungsmaximen oder Wertüberzeugungen, als emotionale Einstellungen vorliegen, die eine Verbindung auch mit einer „als-ob“ Form von Primäremotionen darstellen können. In beiden Fällen also (1,2) soll der Begriff der „emotionalen Einstellung“ eine kognitiv gehaltvolle emotionale Disposition beschreiben, die eine individuelle Wertung einschließt ($desire_2$), welche sich wiederum in einem aktuellen „desire“ ($desire_1$) manifestieren kann. Daher unterscheide ich folgende drei Ebenen von Handlungsgründen, die ich tabellarisch wiedergebe:

(a)	(b)	(c)
motivierende Gründe	mentale Repräsentationen von normativen Gründen	normative Gründe
<ul style="list-style-type: none"> - desire₁ („of A to Ψ“) - Zweck-Mittel-Überzeugung („were A to Φ, she would Ψ“) 	<u>emotionale Einstellung =</u> <ul style="list-style-type: none"> - individuelle Wertung (desire₂) - praktische Überzeugung Handlungsmaxime heiße Kognition - emotionale Disposition 	Propositionen mit praktischer Implikation („sollen“-Sätze) der Form <ul style="list-style-type: none"> - „Ich sollte/ bin verpflichtet zu Φ-en.“ - „Φ-en ist geboten/ verboten.“ - „Es ist besser zu Φ-en.“
desire ₁ = aktual, phänomenal salient („Humescher desire-Begriff“)	desire ₂ = dispositional, phänomenal nicht-salient, kraft der affektiven Fundierung ein Aktualisierungspotenzial, das in der Wertung enthalten ist	
	← Aktualisierung durch passende Schlüsselszenarien	← Prozesse rationaler Habituation <ul style="list-style-type: none"> - kognitives Lernen - Argumentation, gesellschaftlicher Diskus - reflektierte Erfahrung

Der hier unternommene Lösungsvorschlag erlaubt m.E. eine Erklärung des „moral problem“ aus moralpsychologischer Sicht im Rahmen eines „defeasible internalism“, einer Theorie also, die zwar intrinsische motivationale Wirksamkeit moralischer Überzeugungen annimmt, jedoch zu einem Grad, der nicht apriorisch oder analytisch zu verstehen ist, wodurch wiederum faktisch mögliche Fälle bspw. der Willensschwäche ausgeschlossen werden. Der vorliegende Ansatz buchstabiert Smith's „Rationalitätskriterium“ des

Internalismus weiter aus. Erinnern wir uns an seine Prämisse „C2: If an agent believes that she has a normative reason to Φ , then she rationally should desire to Φ “. Wir sind nun in der Lage, nicht nur deklarativ zu postulieren, dass praktische Irrationalität vorliegen muss, wenn ein Akteur nicht auf Grund vertretener normativer Gründe handelt, sondern können auch normativ und prognostisch explizieren, unter welchen Umständen normative Gründe mit motivierenden Gründen koordiniert sind, und unter welchen nicht, sodass Handeln letztlich ausbleibt. Wir können als Fazit festhalten, dass alle jenen vertretenen normativen Gründe das Aktualisierungspotenzial zu motivierenden Gründen haben, deren mentale Repräsentation des Akteurs als „emotionale Einstellung“ darzustellen ist. Das abschließende Erklärungsniveau der motivierenden Gründe wiederum entspricht einer unkontroversen Variante des Humeanismus, bspw. bei Smith (1994) selbst. Ich hatte darauf hingewiesen, dass ich mich nicht argumentativ auf diese Theorieebene beziehe, sondern sie kommentarlos von Smith übernehme. Das Modell der emotionalen Einstellung wiederum basiert zentral auf zwei wichtigen Ergebnissen unserer früheren Untersuchungen. Zum einen haben wir gezeigt, dass desires und individuelle Wertungen auf Emotionen basieren. Dadurch wird einerseits eine dispositional-motivationale Komponente implementiert, andererseits konnten wir in gewisser Hinsicht auch eine Verantwortung von Akteuren für ihre Emotionen nachweisen, u.z. über einen rationalen Zugang auf dispositionaler Ebene. Die Verbindung mit geeigneten Kognitionen (Handlungsmaximen bspw.) wiederum ist auf zwei Weisen vorstellbar: Entweder insofern diese direkt als Komponente der emotionalen Einstellung figurieren (sofern diese eine Sekundäremotion darstellt), oder insofern diese sich mit einer emotionalen Wertung zu einer emotionalen Disposition verbinden, die aktuell als „als-ob-Emotion“ erlebt wird oder bisweilen auch phänomenologisch nicht-salient ist.

Wir können also aus moralspsychologischer Sicht eine überarbeitete Fassung von Smith's Prämisse C2 vorschlagen, nämlich:

„C2': Wenn ein Akteur A glaubt, einen normativen Grund zu haben zu Φ -en, dann hat A auch einen motivierenden Grund zu Φ -en, sofern eine entsprechende emotionale Einstellung bei A vorliegt, d.h. sofern A den entsprechenden normativen Grund als eine koordinierte Verbindung einer emotionalen Wertung und einer entsprechenden praktischen Überzeugung repräsentiert.“

Normative Gründe, die in ihrer mentalen Repräsentation nicht als emotionale Einstellungen darzustellen sind, liefern lediglich kontingente Motivationsbedingungen. Auf diesem Wege erreichen wir einen „defeasible internalism“. Das „Prinzip praktischer Implikationen“ bzw. die Frage des Motivationsinternalismus scheint also durch den vorliegenden Lösungsansatz gewahrt bzw. affirmativ beantwortet zu werden.

Kann das vorliegende Modell jedoch auch die Frage nach der Wahrheitsfunktionalität moralischer Überzeugungen zufriedenstellend beantworten bzw. wahrt es das Smith'sche „Objektivitätsprinzip“? Hier möchte ich zunächst auf meine Anmerkung verweisen, dass die vorliegende Arbeit allerhöchstens die Grundlagen für eine emotivistisch-kognitive Moralpsychologie liefern kann, nicht jedoch die Grundlagen für eine Theorie der Moral selbst. Prinzipiell könnte offen gelassen werden, wie das Phänomen eines „normativen Grundes“ zu interpretieren sei. Allerdings war mein Anliegen in dieser Arbeit, einen Lösungsansatz des „moral problem“ zu erarbeiten, der die Wahrheitsfunktionalität von moralischen Überzeugungen wahren kann (bei gleichzeitiger Wahrung des Prinzips der praktischen Implikationen). Darum soll natürlich auch diese Frage noch abschließend diskutiert werden, mit dem lediglich einschränkenden Hinweis, dass es an dieser Stelle primär um die Frage der Moralpsychologie gehen soll. Die Frage lässt sich relativ einfach bescheiden, indem wir annehmen, dass normative Gründe in Überzeugungen über Zustände der Welt bestehen. Dieser (realistisch-)rationalistische Ansatz ist durchaus in das vorliegende Modell implementierbar. Die mentalen Repräsentation von normativen Gründen wiederum (emotionale Einstellungen also) sind wahrheitsfähig in dreierlei Hinsicht: (1.) schließen sie, sofern es sich um die dispositionale Form von Sekundäremotionen handelt, schon per se einen kognitiven Aspekt ein. (2.) Sind sie wahrheitsfähig im Rahmen der bereits diskutierten Theorien zur Angemessenheit von Emotionen (Mulligan 1998) bzw. zur Musterkonsistenz (Helm 2002) von emotionalen Reaktionen. Und (3.) sind sie wahrheitsfähig im Sinne einer normativen Bewertung insofern sie – als emotionale Dispositionen – der Verantwortung eines Akteurs für seine Charaktereigenschaften unterliegen und daher Gegenstand einer reflektierten Aneignung (Korsgaard 1997: „reflective endorsement“) bzw. Ablehnung und in geringem Grade sogar einer habituellen Gestaltung, Modellierung und Konstituierung sind.

5. Bibliografie

- Abelson, Robert P. (1963): Computer simulation of „hot“ cognition, in: S.S. Tomkins & S. Messick (Hgg.): Computer simulation of personality, New York: Wiley, S. 277-298.
- Adamos, Maria Magoula (2001): Aristotle on Emotions and Contemporary Psychology, in: Sfondoni-Mentzou, Demetra (Hg.) (2001): Aristotle and Contemporary Science, Bd. 2, New York: Peter Lang, S. 226-235.
- Adolphs, R./ Tranel, D./ Bechara, A./ Damasio, H./ Damasio, A.R. (1996): Neuropsychological approaches to reasoning and decision-making, in Damasio, A.R./ Christen, Y./ Damasio, H. (Hgg.): Neurobiology of decision-making, Heidelberg: Springer, S. 157-179.
- Adolphs, R./ Tranel, D./ Damasio, H./ Damasio, A.R. (1995): Fear and the human amygdala, in: Journal of Neuroscience 15, S. 5879-5891.
- Alston, William P. (1967): Emotion and Feeling, in: The Encyclopaedia of Philosophy, hrsg. v. Paul Edwards, Bd. 2, New York: Macmillan.
- Altham, J.E.J. (1986): The Legacy of Emotivism, in: Graham Macdonald & Crispin Wright (Hgg.): Fact, Science and Morality: Essays on A.J. Ayer's Language, Truth and Logic, Oxford, S. 275-288.
- American Psychiatric Association, The (Hg.) (DSM-IV-TR): Diagnostic and Statistical Manual of Mental Disorders, 4., überarb. Ausg., Washington (DC) 1994.
- Anderson, Elizabeth (1993): Value in Ethics and Economics, Cambridge (Mass.).
- Anderson, John R. (1996): Kognitive Psychologie, 2. Aufl., Heidelberg/ Berlin/ Oxford: Spektrum.
- Anderson, S.W./ Bechara, A./ Damasio, H./ Tranel, D./ Damasio, A.R. (1999): Impairment of social and moral behavior related to early damage in human prefrontal cortex, in: Nature Neuroscience 2, S. 1032-1037.
- Annas, Julia (1995): The Morality of Happiness, Oxford.
- Annas, Julia (1996): Aristotle and Kant on Morality and Practical Reasoning, in: Engstrom & Whiting (Hg.) (1996): Aristotle, Kant, and the Stoics, S. 237-258.
- Anscombe, G. E. M. (1958): Modern Moral Philosophy, in: Philosophy 33, S. 1-19 (wiederabgedr. in: Anscombe, G.E.M. (1981): Collected Philosophical Papers, Bd. 3, Minneapolis: University of Minnesota Press, S. 26-42).
- Anscombe, G.E.M. (1957): Intention, Oxford: Basil Blackwell.

- Antonaccio, Maria (2001): *Picturing the Soul: Moral Psychology and the Recovery of the Emotions*, in: *Ethical Theory and Moral Practice* 4, S. 127-141.
- Aquila, Richard E. (1975): *Causes and Constituents of Occurrent Emotion*, in: *Philosophical Quarterly* 25, S. 346-349.
- Aristoteles (EE): *Eudemische Ethik*, übers. v. Franz Dirlmeier, 2. Aufl., Darmstadt: WBG 1969 (=Aristoteles, *Werke*, begr. v. E. Grumbach, hrsg. v. H. Flashar, Bd. 7).
- Aristoteles (EN): *Nikomachische Ethik*, übers. v. Franz Dirlmeier, 4. Aufl., Darmstadt: WBG 1967 (=Aristoteles, *Werke*, begr. v. E. Grumbach, hrsg. v. H. Flashar, Bd. 6).
- Aristoteles (EN): *Nikomachische Ethik*, übers. v. Olof Gigon, 3. Aufl., München 1998.
- Aristoteles (Rh.): *Rhetorik*, übers. u. erläutert v. Christof Rapp, Berlin: Akademie Verlag 2002 (=Aristoteles, *Werke*, begr. v. E. Grumbach, hrsg. v. H. Flashar, Bd. 4).
- Aristotelis Opera Omnia ex rec. Immanuelis Bekkeri, 2. Aufl., hrsg. v. Olof Gigon, Berlin 1960ff.
- Arnold, Magda B. (1950): *An excitatory theory of emotion*, in: M.L. Reymert (Hg.): *Feelings and emotions*, New York: McGraw-Hill, S. 11-33.
- Arnold, Magda B. (1960): *Emotion and Personality*, New York.
- Arpaly, Nomi & Schroeder, Timothy (1999): *Praise, Blame and the Whole Self*, in: *Philosophical Studies* 93, S. 161-188.
- Arpaly, Nomi (2000): *On Acting Rationally against one's Best Judgement*, in: *Ethics* 110, S. 488-513.
- Arregui, Jorge V. (1996): *Descartes and Wittgenstein on Emotions*, in: *International Philosophical Quarterly* 36, 3, S. 319-334.
- Audi, Robert (1993): *Action, Intention, and Reason*, Ithaca (NY).
- Audi, Robert (1997): *Moral Knowledge and Ethical Character*, New York: OUP.
- Audi, Robert (2000): *The Architecture of Reason. The Structure and Substance of Rationality*, Oxford.
- Audi, Robert (2004): *Reasons, Practical Reason, and Practical Reasoning*, in: *Ratio* (new series) 17, S. 119-149.
- Averill, James R. (1982): *Anger and Aggression. An Essay on Emotion*, New York/ Heidelberg/ Berlin: Springer.
- Ayer, Alfred J. (1936): *Language, Truth, and Logic*, New York.
- Badhwar, Neera K. (1996): *The Limited Unity of Virtue*, in: *Noûs* 30, 3, S. 306-329.
- Baeyer-Katte, W. (1973): *Angst*, 2., erw. Aufl., Frankfurt a.M.: Suhrkamp.
- Baier, Annette (1978): *Hume's Analysis of Pride*, in: *Journal of Philosophy* 75, S. 27-40.

- Baier, Annette (1990): What Emotions Are About, in: J.E. Tomberlin (Hg.): *Philosophical Perspectives 4: Action Theory and Philosophy of Mind*, Atascadero: Ridgeview, S. 1-29.
- Baier, Kurt (1991) Types of Responsibility, in: *The Spectrum of Responsibility*, hg. v. Peter French, New York: St. Martin's Press, S. 17-31.
- Barbalet, J.M. (1999): William James' Theory of Emotions: Filling in the Picture, in: *Journal for the Theory of Social Behaviour* 29, 3, S. 251-266.
- Baron, R.M./ Kenny, D.A. (1986): The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations, in: *Journal of Personality and Social Psychology* 51, S. 1173-1182.
- Bateson, Gregory (1982): *Geist und Natur*, Frankfurt a.M.: Suhrkamp.
- Baugh, Bruce (1990): Sartre and James on the Role of the Body in Emotion, in: *Dialogue: Canadian Philosophical Review* 29, S. 357-373.
- Bechara, A./ Damasio, H./ Damasio, A.R. (2000): Emotion, decision-making and the orbitofrontal cortex, in: *Cerebral Cortex* 10, S. 295-307.
- Bechara, A./ Damasio, H./ Tranel, D./ Damasio, A.R. (1997): Deciding advantageously before knowing the advantageous strategy, in: *Science* 275, S. 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, in: *Trends in Cognitive Sciences* 9, 4, S. 159-162.
- Bechara, A./ Tranel, D./ Damasio, H./ Adolphs, R./ Rockland, C./ Damasio, A.R. (1995): Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans, in: *Science* 269, S. 1115-1118.
- Beck, A.T. (1979): *Wahrnehmung der Wirklichkeit und Neurose. Kognitive Psychotherapie emotionaler Störungen*, München.
- Bedford, Erol (1961): Pleasure and Belief, in: *Proceedings of the Aristotelian Society* 33, supplementary volume, S. 73-92.
- Bedford, Erol (1962): Emotions, in: Chappell, V. (Hg.): *The Philosophy of Mind*, Englewood Cliffs (NJ): Prentice-Hall, S. 111-113.
- Bem, D.J. & Funder, D.C. (1978): Predicting more of the people more of the time: Assessing the personality of situations, in: *Psychological Review* 85, S. 485-501.
- Bem, D.J. (1972): Self-perception theory, in: L. Berkowitz (Hg.): *Advances in Experimental Social Psychology*, New York: Academic Press, Bd. 6, S. 1-62.
- Bennett, Jonathan (1980): Accountability, in: *Philosophical Subjects. Essays Presented to P.F. Strawson*, hg. v. Zak Van Straaten, New York: OUP, S. 14-47.

- Benthien, C./ Feig, A./ Kasten, I. (Hgg.) (2000): Emotionalität. Zur Geschichte der Gefühle, Köln/ Weimar/ Wien: Böhlau.
- Ben-Ze'ev, Aaron (1997): Appraisal Theories of Emotions, in: Journal of Philosophical Research 22, S. 129-143.
- Ben-Ze'ev, Aaron (2000): The Subtlety of Emotions, Cambridge: MIT Press.
- Bermond, B./ Nieuwenhuysse, B./ Fasotti, L./ Schuerman, J. (1991): Spinal cord lesions, peripheral feedback, and intensities of emotional feelings, in: Cognition and Emotion 5, S. 201-220.
- Birnbacher, Dieter (1998): Gegenstand und Ursache der Emotion in Spinozas Affektenlehre, in: Stefan Hübsch & Dominic Kaegi (Hgg.): Affekte. Philosophische Beiträge zur Theorie der Emotionen, Heidelberg: C. Winter, S. 101-116.
- Bittner, Rüdiger (2001): Doing Things for Reasons, Oxford: OUP.
- Blackburn, Simon (1984): Spreading the Word. Groundings in the Philosophy of Language, Oxford.
- Blackburn, Simon (1985): Errors and the Phenomenology of Value, in: Ted Hondrich (Hg.): Morality and Objectivity, London: Routledge & Kegan Paul, S. 1-22.
- Blackburn, Simon (1986): Morals and Modals, in: Graham Macdonald & Crispin Wright (Hgg.): Fact, Science and Morality: Essays on A.J. Ayer's Language, Truth and Logic, Oxford: Basil Blackwell, S. 119-142.
- Blackburn, Simon (1987): How to Be an Ethical Antirealist, in: Peter A. French/ Theodore E. Uehling, Jr./ Howard K. Wettstein (Hgg.): Midwest Studies in Philosophy XII: Realism and Anti-Realism, University of Notre Dame Press 1987, S. 361-375.
- Blackburn, Simon (1995): Practical Tortoise Raising, in: Mind 104, S. 695-711. (wiederabgedr. in: Wallace (1999): Reason, Emotion and Will).
- Blackburn, Simon (1998): Ruling Passions. A Theory of Practical Reasoning, Oxford.
- Blackburn, Simon (2002): How Emotional is the Virtuous Person?, in: Goldie, Peter (Hg.): Understanding Emotions. Mind and Morals, Aldershot u.a.: Ashgate, S. 81-96.
- Blankenburg, Wolfgang (1993): Affektivität und Personsein aus psychiatrischer Sicht. Am Beispiel von Angst und Hoffnung im menschlichen Dasein, in: Fink-Eitel & Lohmann (1993), 307-333.
- Blankenburg, Wolfgang (1995): Das Sich-Befinden zwischen Leiblichkeit und Gefühl, in: Großheim, Michael (Hg.): Leib und Gefühl. Beiträge zur Anthropologie, Berlin 1995, 193-214.

- Blasi, Augusto (1999): Emotions and Moral Motivation, in: *Journal for the Theory of Social Behaviour* 29, 1, S. 1-19.
- Bless, Herbert (1997): *Stimmung und Denken. Ein Modell zum Einfluss von Stimmungen auf Denkprozesse*, Bern u.a.: Huber.
- Böhme, Gernot (1989): *Für eine ökologische Naturästhetik*, Frankfurt a.M.: Suhrkamp.
- Böhme, Gernot (1992): Atmosphäre als Grundbegriff einer neuen Ästhetik, in: *Kunstforum* 120, 247-255.
- Bollnow, Otto F. (1941): *Das Wesen der Stimmungen*, 7. Aufl., Frankfurt a.M.: Klostermann 1988.
- Bormann, Karl (1981): Zur stoischen Affektenlehre, in: Ingrid Craemer-Ruegenberg (Hg.): *Pathos, Affekt, Gefühl*, Freiburg i.Br./ München 1981, 79-102.
- Boyd, Richard (1988): How to be a Moral Realist, in: Geoffrey Sayre-McCord (Hg.): *Essays on Moral Realism*, Cornell University Press, S. 181-228.
- Brady, Michael S. (1998): Reasons and Rational Motivational Access, in: *Pacific Philosophical Quarterly* 79, S. 99-114.
- Brady, Michael S. (2000): How to Understand Internalism, in: *The Philosophical Quarterly* 50, S. 91-97.
- Brady, Michael S. (2002): Skepticism, Normativity, and Practical Identity, in: *The Journal of Value Inquiry* 36, S. 403-412.
- Brandt, Richard B. (1979): *A Theory of the Good and the Right*, Oxford.
- Brandt, Richard B. (1988): The Structure of Virtue, in *Midwest Studies in Philosophy* 13, S. 64-82.
- Brandt, Richard B. (1946): Moral Valuation, in: *Ethics* 56, S. 106-121.
- Brandt, Richard B. (1996): *Facts, Values, and Morality*, Cambridge/ New York: Cambridge University Press.
- Bratman, Michael E. (1983): Taking Plans Seriously, in: *Social Theory and Practice* 9, S. 271-287 (wiederabgedr. in Wallace (1999)).
- Breggin, P.R. (1964): The psychophysiology of anxiety with a review of the literature concerning adrenaline, in: *Journal of Nervous and Mental Disease* 139, S. 558-568.
- Brentano, Franz (1955): *Psychologie vom empirischen Standpunkt, mit Einl., Anm. u. Reg.* hrsg. v. Oskar Kraus, 2 Bde, Hamburg: Meiner.
- Brewer, Talbot (2001): Rethinking Our Maxims: Perceptual Salience and Practical Judgment in Kantian Ethics, in: *Ethical Theory and Moral Practice* 4, S. 219-230.

- Brink, David O. (1984): Moral Realism and the Skeptical Arguments from Disagreement and Queerness, in: *Australasian Journal of Philosophy* 62, S. 111-125.
- Brink, David O. (1986): Externalist Moral Realism, in: *Southern Journal of Philosophy Supplement* 24, S. 23-42.
- Brink, David O. (1989): *Moral Realism and the Foundations of Ethics*, Cambridge: CUP.
- Broad, C.D. (1954): Emotion and Sentiment, in: *Broad's Critical Essays in Moral Philosophy*, hrsg. v. David Cheney, London: Allen & Unwin 1971.
- Broome, John (1997): Reasons and Motivation, in: *Proceedings of the Aristotelian Society Supplementary Volume* 77, S. 131-146 (wiederabgedr. in: Wallace (1999): *Reason, Emotion and Will*).
- Brunschwig, Jacques/ Nussbaum, Martha (1993): *Passions and Perceptions. Studies in Hellenistic Philosophy of Mind*, Cambridge.
- Buck, R. (1980): Nonverbal behavior and the theory of emotion: the facial feedback hypothesis, in: *Journal of Personality and Social Psychology* 38, 5, S. 811-824.
- Burrington, Dale (1999): Blameworthiness, in: *Journal of Philosophical Research* 24, S. 505-527.
- Busemeyer, J.R./ Townsend, J.T./ Stout, J.C. (2002): Motivational underpinnings of utility in decision making. Decision field theory analysis of deprivation and satiation, in: Moore, Simon & Oaksford, Mike (Hgg.) (2002): *Emotional Cognition: From brain to behaviour*, Amsterdam/ Philadelphia: John Benjamins (= *Advances in Consciousness Research* 44), S. 197- 219.
- Cacioppo, J.T./ Uchino, B.N./ Crites, S.L./ Snyder-Smith, M.A./ Smith, G./ Berntson, G.G./ Lang, P.J. (1992): Relationship between facial expressiveness and sympathetic activation in emotion: a critical review, with emphasis on modelling underlying mechanisms and individual differences, in: *Journal of Personality and Social Psychology* 62, S. 110-128.
- Calhoun, Cheshire (1984): *Cognitive Emotions?*, in: Calhoun, C./ Solomon, R. (Hgg.): *What is an emotion?*, OUP.
- Calhoun, Cheshire/ Solomon, Robert (1984): *What is an Emotion?*, New York.
- Cam, Philip (2002): Fact, Value and Philosophy Education, in: *Critical & Creative Thinking*, 10, 1, S. 21-28.
- Cannon, W.B. (1927): The James-Lange theory of emotions: A critical examination and an alternative theory, in: *American Journal of Psychology* 39, S. 106-124.

- Cannon, W.B. (1929): *Bodily changes in pain, hunger, fear and rage: an account of recent researches into the function of emotional excitement*, New York: Appleton-Century Company.
- Card, Claudia/ Hunt, Lester (Hgg.) (1991): *Character. Essays in Moral Psychology*, Ithaca (NY).
- Carr, David (2003): *Character and Moral Choice in the Cultivation of Virtue*, in: *Philosophy* 78, S. 219-232.
- Carroll, Lewis (1895): *What the Tortoise said to Achilles*, in: *Mind, New Series*, vol. N, S. 278-280.
- Carroll, Noel (1998): *Art, Narrative, and Moral Understanding*, in: Jerrold Levinson (Hg.): *Aesthetics and Ethics. Essays at the Intersection*, Cambridge 1998, 126-160.
- Cates, Diana F. (1997): *Choosing to Feel. Virtue, Friendship, and Compassion for Friends*, Notre Dame (Indiana).
- Chapman, John W./ Galston, William A. (Hgg.) (1992): *Virtue*, New York/ London (= *Nomos* 34).
- Charland, Louis C. (1997): *Reconciling Cognitive and Perceptual Theories of Emotion: A Representational Proposal*, in: *Philosophy of Science* 64, 4, S. 555-579.
- Charland, Louis C. (2001): *In Defence of „Emotion“: Critical Notice of Paul E. Griffiths' What Emotions Really Are: The Problem of Psychological Categories*, in: *Canadian Journal of Philosophy* 31, 1, S. 133-154.
- Christianson, S.A. & Loftus, E.F. (1991): *Remembering emotional events: the fate of detailed information*, in: *Cognition and Emotion* 5, S. 81-108.
- Christianson, S.A. (1992): *Emotional stress and eyewitness memory: a critical review*, in: *Psychological Bulletin* 112, S. 284-309.
- Churchland, Paul M. (1984): *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind*, Cambridge (MA): MIT Press.
- Churchland, Paul M. (1995): *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain*, Cambridge (MA): MIT Press.
- Cicero, Marcus Tullius (Tusc.): *Disputationes Tusculanae*, hrsg. mit einer engl. Übers. v. J.E. King, London 1960 (=Loeb Classical Library 141).
- Ciampi, Luc (1982): *Affektlogik*, Stuttgart: Klett-Cotta.
- Ciampi, Luc (1999): *Die emotionalen Grundlagen des Denkens. Entwurf einer fraktalen Affektlogik*, 2., durchges. Aufl., Göttingen: Vandenhoeck & Ruprecht.

- Claparède, Edouard (1911): Recognition and „me-ness“, in: D. Rapaport (Hg.): *Organization and pathology of thought*, New York: Columbia University Press 1951, S. 58-75.
- Clarke, Stanley G. (1986): Emotions: Rationality without Cognitivism, in: *Dialogue* 25, S. 663-674.
- Cléro, Jean-Pierre (1985): *La philosophie des passions chez David Hume*, Paris.
- Cohen, N.J. & Eichenbaum, H. (1993): *Memory, amnesia, and the hippocampal system*, Cambridge (MA): MIT Press.
- Cohon, Rachel (1993): Internalism about Reasons for Action, in: *Pacific Philosophical Quarterly* 74, S. 265-288.
- Cole, Jonathan (1999): *Über das Gesicht*, München.
- Collins, A. (1997): The Psychological Reality of Reasons, in: *Ratio* 10, S. 8-23.
- Cooper, John M. (1999): *Reason and Emotion: Essays on Ancient Moral Psychology and Ethical Theory*, Princeton: Princeton University Press.
- Coriando, Paola-Ludovika (2002): *Affektenlehre und Phänomenologie der Stimmungen. Wege einer Ontologie und Ethik des Emotionalen*, Frankfurt a.M. (=Philosophische Abhandlungen Bd. 85).
- Cornelius, Randolph R. (1991): Gregorio Marañón's two-factor theory of emotion, in: *Personality and Social Psychology Bulletin* 17, S. 65-69.
- Cornelius, Randolph R. (1996): *The Science of Emotion. Research and Tradition in the Psychology of Emotions*, Upper Saddle River (NJ) u.a.: Prentice Hall.
- Cottingham, John (1998): *Philosophy and the good Life. Reason and the Passions in Greek, Cartesian, and Psychoanalytic Ethics*, Cambridge.
- Craemer-Ruegenberg, Ingrid (1981): Biologisch-anthropologische Auskünfte über Gefühle, in: *Ibid.* (Hg.): *Pathos, Affekt, Gefühl*, Freiburg i.Br./ München 1981, 307-326.
- Craemer-Ruegenberg, Ingrid (1993): Begrifflich-systematische Bestimmung von Gefühlen. Beiträge aus der antiken Tradition, in: Fink-Eitel & Lohmann (1993), 20-32.
- Craemer-Ruegenberg, Ingrid (Hg.) (1981): *Pathos, Affekt, Gefühl*, Freiburg i.Br./ München.
- Crane, T. (1998): Intentionality as the Mark of the Mental, in: A. O'Hear (Hg.): *Current Issues in Philosophy of Mind*, Cambridge, S. 229-251.
- Crisp, Roger & Slote, Michael (Hgg.) (1997): *Virtue Ethics*, Oxford.
- Critchley, H.D./ Mathias, C.J./ Dolan, R.J. (2001): Neural Correlates of first and second-order Representation of Bodily States, in: *Nature Neuroscience* 4, S. 207-212.
- Cullity, Garrett/ Gaut, Berys (Hgg.) (1997): *Ethics and Practical Reason*, Oxford.
- D'Arms, Justin & Jacobson, Daniel (2000): Sentiment and Value, in: *Ethics* 110, S. 722-748.

- D'Arms, Justin & Jacobson, Daniel (2000a): The Moralistic Fallacy: On the „Appropriateness“ of Emotions, in: *Philosophy and Phenomenological Research* 61, 1, S. 65-90.
- Damasio, A.R./ Grabowski, T.J./ Bechara, A./ Damasio, H./ Ponto, L.L.B/ Parvizi, J./ Hichwa, R.D. (2000): Subcortical and Cortical Brain Activity during the Feeling of Self-generated Emotions, in: *Nature Neuroscience* 3, S. 1049-56.
- Damasio, A.R./ Tranel, D./ Damasio, H. (1990): Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli, in: *Behavioral Brain Research* 41, S. 81-94.
- Damasio, Antonio R. (1994): *Descartes' Error. Emotion, Reason, and the Human Brain*, New York.
- Damasio, Antonio R. (1996): The somatic marker hypothesis and the possible functions of the prefrontal cortex, in: *Philosophical Transactions of the Royal Society London B: Biological Sciences* 351, S. 1413-1420.
- Damasio, Antonio R. (1999): *The Feeling of what happens. Body and Emotion in the Making of Consciousness*, New York.
- Damasio, Antonio R. (2001): Reflections on the Neurobiology of Emotion and Feeling, in: João Branquinho (Hg.), *The Foundations of Cognitive Science*, Oxford: Clarendon, S. 99-108.
- 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, in: *Science* 264, S. 1102-1105.
- Dana, C.L. (1921): The anatomic seat of the emotions: A discussion of the James-Lange theory, in: *Archives of Neurology and Psychiatry* 6, S. 634-639.
- Dancy, Jonathan (1993): *Moral Reasons*, Oxford: Basil Blackwell.
- Dancy, Jonathan (1995): Why There is Really No Such Thing as the Theory of Motivation, in: *Proceedings of the Aristotelian Society* 95, S. 1-18 (wiederabgedr. in: Wallace (1999): *Reason, Emotion and Will*).
- Dancy, Jonathan (2000): *Practical Reality*, Oxford: OUP.
- Dancy, Jonathan (Hg.) (2000a): *Normativity*, Oxford: Blackwell.
- Darley, J.M. & Batson, C.D. (1973): 'From Jerusalem to Jericho': A Study of Situational and Dispositional Variables in Helping Behavior, in: *Journal of Personality and Social Psychology* 27, S. 100-108.
- Darwall, Stephen (1983): *Impartial Reason*, Cornell University Press.

- Darwall, Stephen (1983): *Impartial Reason*, Cornell.
- Darwall, Stephen (1998): *Philosophical Ethics*, Boulder (CO): Westview Press.
- Darwall, Stephen (Hg.) (2003): *Virtue ethics*, Oxford/ Malden (MA): Blackwell (= Blackwell readings in philosophy, Bd. 10).
- Davidson, Donald (1963): *Actions, Reasons, and Causes*, wiederabgedr. in: Davidson (1980), S. 21-42.
- Davidson, Donald (1970): *Mental Events*, wiederabgedr. in: Davidson (1980), S. 207-225.
- Davidson, Donald (1977): *Hume's Cognitive Theory of Pride*, wiederabgedr. in: Davidson (1980), S. 277-290.
- Davidson, Donald (1978): *Intending*, wiederabgedr. in: Davidson (1980), S. 83-102.
- Davidson, Donald (1980): *Essays on Actions and Events*, Oxford: OUP.
- Davis, Wayne A. (1987): *The Varieties of Fear*, in: *Philosophical Studies* 51, 287-310.
- De Sousa, Ronald (1979): *The Rationality of Emotions*, in: Rorty (1980), 127-152. (zuerst in: *Dialogue* 18).
- De Sousa, Ronald (1987): *The Rationality of Emotion*, Cambridge (Mass.)/ London.
- De Sousa, Ronald (2002): *Emotional Truth*, in: *Proceedings of the Aristotelian Society Supplementary Volume* 76, S. 247-263.
- Deigh, John (1994): *Cognitivism in the Theory of Emotions*, in: *Ethics* 104, 824-854.
- Dennett, Daniel (1995): *Our Vegetative Soul: The Search for a Reliable Model of the Human Self*, in: *Times Literary Supplement*, 25. August, S. 3f.
- Descartes, René (1649): *Die Leidenschaften der Seele*, franz.-dt. hg. v. Klaus Hammacher, 2., durchges. Aufl., Hamburg: Meiner 1996.
- Di Muzio, Gianluca (2000): *Aristotle on Improving One's Character*, in: *Phronesis* 45, 3, S. 205-219.
- Döring, Sabine A. (2002): *Die Moralität der Gefühle. Eine Art Einleitung*, in: *Die Moralität der Gefühle*, *Deutsche Zeitschrift für Philosophie*, Sonderband 4, hrsg. v. Sabine A. Döring u. Verena Mayer, Berlin, S. 15-35.
- Döring, Sabine A./ Peacocke, Christopher (2002): *Handlungen, Gründe und Emotionen*, in: *Die Moralität der Gefühle*, *Deutsche Zeitschrift für Philosophie*, Sonderband 4, hrsg. v. Sabine A. Döring u. Verena Mayer, Berlin, S. 81-103.
- Doris, John (1998): *Persons, Situations, and Virtue Ethics*, in: *Nous* 32, S. 504-530.
- Dorschel, Andreas (1997): *Über die Intentionalität von Emotionen*, in: *International Studies in Philosophy* 29, 4, S. 11-21.
- Dorschel, Andreas (1999): *Emotion und Leib*, in: *Philosophia Naturalis* 36, 1, S. 35-52.

- Dorschel, Andreas (1999): Emotion und Verstand, in: *Philosophisches Jahrbuch* 106, 1, S. 18-40.
- Downing, George (2001): Emotion Theory Revisited, in: *Heidegger, Coping, and Cognitive Science: A Festschrift for Hubert Dreyfus*, MIT Press, Bd. 2, S. 245-270.
- Dreier, James (1997): Humean Doubts about the Practical Justification of Morality, in: *Cullity & Gaut (1997)*, S. 81-99.
- Dretske, F. (1981): *Knowledge and the Flow of Information*, Cambridge (MA): MIT Press.
- Dreyfus, Hubert L. & Dreyfus, Stuart E. (1987): *Künstliche Intelligenz. Von den Grenzen der Denkmachine und dem Wert der Intuition*, Reinbek: Rowohlt.
- Eccles, John C. & Popper, Karl (1982): *Das Ich und sein Gehirn*, München/ Zürich.
- Ekman, P./ Davidson, R.J./ Friesen, W.V. (1990): The Duchenne smile: emotional expression and brain physiology II, in: *Journal of Personality and Social Psychology* 58, S. 342-353.
- Ekman, P./ Friesen, W.V./ Ancoli, S. (1980): Facial signs of emotional experience, in: *Journal of Personality and Social Psychology* 39, S. 1125-1134.
- Ekman, P./ Friesen, W.V./ Ellsworth, P. (1972): *Emotion in the human face: guidelines for research and integration of findings*, New York: Pergamon Press.
- Ekman, P./ Friesen, W.V./ O'Sullivan, M./ Chan, A./ Diacoyanni-Tarlatzis, J./ Heider, K./ Krause, R./ LeCompte, W.A./ Pitcairn, T./ Ricci-Bitti, P.E./ Scherer, K./ Tomita, M./ Tzavaras, A. (1987): Universals and cultural differences in the judgments of facial expressions of emotion, in: *Journal of Personality and Social Psychology* 53, S. 712-717.
- Ekman, P./ Levenson, R.W./ Friesen, W.V. (1983): Autonomic nervous system activity distinguishes among emotions, in: *Science* 221, S. 1208-1210.
- Ekman, P. & Friesen, W.V. (1975): *Unmasking the face: a guide to recognizing emotions from facial clues*, Englewood Cliffs (NJ): Prentice-Hall.
- Ekman, P. & Friesen, W.V. (1978): *Investigator's guide: Facial Action Coding System*, Palo Alto (CA): Consulting Psychologist's Press.
- Ekman, P. & Friesen, W.V. (1978): *Manual: Facial Action Coding System*, Palo Alto (CA): Consulting Psychologist's Press.
- Ekman, P. & Friesen, W.V. (1982): Felt, false and miserable smiles, in: *Journal of Nonverbal Behavior* 6, S. 238-252.
- Ekman, P. & Friesen, W.V. (1986): A new pan-cultural facial expression of emotion, in: *Motivation & Emotion* 10, S. 159-168.

- Ekman, Paul (1972): Universals and cultural differences in facial expressions of emotion, in: J.K. Cole (Hg.): Nebraska Symposium on Motivation 1971, Lincoln: University of Nebraska Press, S. 207-283.
- Ekman, Paul (1977): Biological and Cultural Contributions to Body and Facial Movement in the Expression of Emotions, in: Rorty (1980), 73-102. (zuerst in: John Blacking (Hg.): The Anthropology of the Body, London: Academic Press 1977).
- Ekman, Paul (1982): Methods for measuring facial action, in: K.R. Scherer & P. Ekman (Hgg.): Handbook of methods in nonverbal behavior research, Cambridge: CUP, S. 45-90.
- Ekman, Paul (1988): Gesichtsausdruck und Gefühl: 20 Jahre Forschung von Paul Ekman, Paderborn.
- Ellis, A. (1982): Die rational-emotive Therapie, 3. Aufl., München.
- Ellsworth, P.C. (1994): William James and Emotion: Is a Century of Fame Worth a Century of Misunderstanding?, in: Psychological Review 101, 2, S. 222-229.
- Elster, Jon (1979): Ulysses and the Sirens. Studies in Rationality and Irrationality, Cambridge: CUP.
- Elster, Jon (1999): Alchemies of the Mind: Rationality and the Emotions, Cambridge: CUP.
- Elster, Jon (2000): Ulysses unbound. Studies in Rationality, Precommitment and Constraints, Cambridge: CUP.
- Engel, Pascal (1998): La logique peut-elle mouvoir l'esprit?, in: Dialogue, Canadian Philosophical Review/ Revue canadienne de Philosophie 37, no. 1, S. 35-53.
- Engstrom, Stephen & Whiting, Jennifer (Hg.) (1996): Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty, Cambridge: CUP.
- Epictetus (diatr./ enchir.): The Discourses as reported by Arrian, the Manual, and Fragments, with an English Translation by W.A. Oldfather, 2 Bde., Cambridge (Mass.): Harvard University Press 1946/ 1952.
- Epstein, Seymour (1977): Traits are alive and well, in: Magnusson, D. & Endler, N.S. (Hgg.): Personality at the Crossroads, Hillsdale (NJ): Lawrence Erlbaum Associates, S. 83-98.
- Ewert, Otto (1983): Ergebnisse und Probleme der Emotionsforschung, in: Motivation und Emotion, Bd. 1: Theorien und Formen der Motivation, hrsg. v. H. Thomae, Göttingen, S. 397-452.
- Feinstein, Howard M. (1970): William James on the Emotions, in: Journal of the History of Ideas 31, S. 133-142.
- Ferenczi, S. & Rank, O. (1924): Entwicklung der Psychoanalyse, Leipzig.

- Fink-Eitel, Hinrich (1986): Affekte. Versuch einer philosophischen Bestandsaufnahme, in: *Zeitschrift für philosophische Forschung* 40, 520-542.
- Fink-Eitel, Hinrich (1992): Die Philosophie der Stimmungen in Heideggers 'Sein und Zeit', in: *Allgemeine Zeitschrift für Philosophie* 17.3, 27-44.
- Fink-Eitel, Hinrich/ Lohmann, Georg (Hgg.) (1993): *Zur Philosophie der Gefühle*, Frankfurt a.M.
- Fischer, John Martin & Ravizza, Mark (1998): *Responsibility and Control: A Theory of Moral Responsibility*, New York: CUP.
- Fischer, John Martin (1999): Recent Work on Moral Responsibility, in: *Ethics* 110, S. 93-139.
- Flanagan, Owen & Rorty, Amélie Oksenberg (Hgg.) (1990): *Identity, Character, and Morality. Essays in Moral Psychology*, Cambridge (MA): MIT Press.
- Flanagan, Owen (1990): Identity and Strong and Weak Evaluation, in: Flanagan & Rorty (Hgg.) (1990): *Identity, Character, and Morality*, S. 37-65.
- Fletcher, Joseph (1966): *Situation Ethics*, Philadelphia.
- Fodor, J.A. (1990): A Theory of Content I & II, in: *A Theory of Content and other Essays*, Cambridge (MA): MIT Press.
- Fodor, Jerry A. (1981): Three Cheers for Propositional Attitudes, in: ders.: *Representations: Philosophical Essays on the Foundation of Cognitive Science*, Cambridge (MA): MIT Press, S. 100-123.
- Foot, Philippa (1972): Morality as a System of Hypothetical Imperatives, in: Foot (1978a), S.157-173.
- Foot, Philippa (1978): Virtues and Vices, in: Foot, Philippa (1978a): *Virtues and Vices and Other Essays in Moral Philosophy*, S. 1-18.
- Foot, Philippa (1978a): *Virtues and Vices and Other Essays in Moral Philosophy*, Oxford: Basil Blackwell
- Foot, Philippa (1983): Moral Realism and Moral Dilemma, in: *Journal of Philosophy* 80, 7, S. 379-398.
- Foot, Philippa (2001): *Natural Goodness*, Oxford: Clarendon.
- Forgas, J.P. & Bower, G.H. (1987): Mood effects on person-perception judgments, in: *Journal of Personality and Social Psychology* 53, S. 53-60.
- Fortenbaugh, William (1975): *Aristotle on Emotion*, London.
- Fortenbaugh, William (1970): Aristotle's Rhetoric on Emotions, in: *Archiv für Geschichte der Philosophie* 52, S. 40-70.

- Frank, Robert H. (1988): *Passions Within Reason. The Strategic Role of the Emotions*, New York.
- Franke, Ursula (1981): Ein Komplement der Vernunft. Zur Bestimmung des Gefühls im 18. Jahrhundert, in: Ingrid Craemer-Ruegenberg (Hg.): *Pathos, Affekt, Gefühl*, Freiburg i.Br./München 1981, 131-148.
- Frankena, William (1958): Obligation and Motivation in Recent Moral Philosophy, in: A.I. Melden (Hg.): *Essays on Moral Philosophy*, University of Washington Press.
- Frankfurt, Harry (1971): Freedom of the Will and the Concept of a Person, in: *Journal of Philosophy* 67, S. 5-20.
- Frankfurt, Harry (1987): Identification and Wholeheartedness, in: Schoeman (Hg.) (1987), S. 27-45.
- Frankfurt, Harry (1998): The Faintest Passion, in: *Necessity, Volition, and Love*, Cambridge: CUP, S. 95-107 (wiederabgedr. in Wallace (1999), S. 357-368).
- French, P./ Uehling, T. jr./ Wettstein, H. (Hgg.) (1988): *Ethical Theory: Character and Virtue*, Notre Dame (IN): University of Notre Dame Press (= *Midwest Studies in Philosophy* 13).
- French, Peter A. & Wettstein, Howard K. (Hgg.) (1998): *Philosophy of Emotions*, Notre Dame (IN): University of Notre Dame Press (= *Midwest Studies in Philosophy* 22).
- Freud, Sigmund (1917): Vorlesungen zur Einführung in die Psychoanalyse, in: *Gesammelte Werke*, Bd. XI, Frankfurt a.M.: Fischer 1999, S. 1-499.
- Fridlund, A.J. (1991): Sociality of solitary smiling: potentiation by an implicit audience, in: *Journal of Personality and Social Psychology* 60, S. 229-240.
- Fridlund, A.J. (1992): The behavioral ecology and sociality of human faces, in: M.S. Clark (Hg.): *Review of Personality and Social Psychology*, Bd. 13: *Emotion*, Newbury Park (CA): Sage, S. 90-121.
- Frijda, N. H./ Kuipers, P./ ter Schure, E. (1989): Relations among emotion, appraisal, and emotional action readiness, in: *Journal of Personality and Social Psychology* 57, S. 212-228.
- Frijda, N.H./ Mesquita, B./ Sonnemans, J./ van Goozen, S. (1991): The duration of affective phenomena or emotions, sentiments and passions, in: K.T. Strongman (Hg.): *International Review of Studies on Emotion*, Chichester: Wiley & Sons, Bd. 1, S. 187-225.
- Frijda, Nico H. (1986): *The emotions*, Cambridge: CUP.
- Früchtl, Josef (1996): *Ästhetische Erfahrung und moralisches Urteil*, Frankfurt a.M.
- Fulcher, Eamon P. (2002): Neurons with attitude. A connectionist account of human evaluative learning, in: Moore, Simon & Oaksford, Mike (Hgg.) (2002): *Emotional*

- Cognition: From brain to behaviour, Amsterdam/ Philadelphia: John Benjamins (= Advances in Consciousness Research 44), S. 75-109.
- Funder, David C. (1999): *Personality Judgment: A Realistic Approach to Person Perception*, San Diego: Academic Press.
- Gamm, Gerhard (1998): Glück, Moral und ästhetische Erfahrung, in: *Philosophische Rundschau* 45, Heft 2, 129-148.
- Gaus, Gerald F. (1990): *Value and Justification: The Foundations of Liberal Theory*, Cambridge: Cambridge University Press.
- Gauthier, David (1986): *Morals by Agreement*, Oxford: Clarendon Press.
- Gazzaniga, M.S. & LeDoux, J.E. (1978): *The Integrated Mind*, New York: Plenum.
- Gazzaniga, M.S. (1972): One brain – two minds, in: *American Scientist* 60, S. 311-317.
- Geach, Peter T. (1960): Ascriptivism, in: Ders.: *Logic Matters*, 2. Aufl., Oxford 1981, S. 250-254.
- Geach, Peter T. (1965): Assertion, in: Ders.: *Logic Matters*, 2. Aufl., Oxford 1981, S. 254-269.
- Gean, William D. (1979): Emotion, Emotional Feeling and Passive Body Change, in: *Journal for the Theory of Social Behaviour* 9, S. 39-51.
- Gibbard, Alan (1990): *Wise Choices, Apt Feelings. A Theory of Normative Judgment*, Oxford.
- Gill, Frances E. (2000): Practical Identity and Practical Reason, in: *Southwest Philosophy Review* 16, 1, S. 33-39.
- Goffman, Erving (1959): *The Presentation of Self in Everyday Life*, New York: Doubleday – Anchor Books.
- Goffman, Erving (1967): *Interaction Ritual*, New York: Anchor Books.
- Goffman, Erving (1971): *Relations in Public*, New York: Harper & Row.
- Goldie, Peter (2002a): Emotions, feelings and intentionality, in: *Phenomenology and the Cognitive Sciences* 1, S. 235-254.
- Goldie, Peter (Hg.) (2002): *Understanding Emotions. Mind and Morals*, Aldershot u.a.: Ashgate.
- Goller, Hans (1992): *Emotionspsychologie und Leib-Seele-Problem*, Stuttgart u.a.: Kohlhammer (=Münchener philosophische Studien, N.F., Bd. 8).
- Gordon, Robert (1987): *The Structure of Emotions: Investigations in Cognitive Philosophy*, Cambridge UP.

- Gordon, Robert (1978): Emotion Labeling and Cognition, in: *Journal for the Theory of Social Behaviour* 8, S. 125-135.
- Gordon, Robert (1986): The Passivity of Emotions, in: *Philosophical Review* 95, S. 371-392.
- Gorevan, Patrick (1994): The Passions of the Soul: Descartes' Shadow on Theories of the Emotions, in: *American Catholic Philosophical Quarterly* 68, 4, S. 515-528.
- Goschke, Thomas/ Bolte, Annette (2002): Emotion, Kognition und Intuition: Implikationen der empirischen Forschung für das Verständnis moralischer Urteilsprozesse, in: *Die Moralität der Gefühle, Deutsche Zeitschrift für Philosophie, Sonderband 4*, hrsg. v. Sabine A. Döring u. Verena Mayer, Berlin, S. 39-57.
- Gosepath, Stefan (Hg.) (1999): *Motive, Gründe, Zwecke. Theorien praktischer Rationalität*, Frankfurt a.M.: Fischer.
- Gosepath, Stefan (2002): Practical Reason: A Review of the Current Debate and Problems, in: *Philosophical Explorations* 5, 3, S. 229-238.
- Gouinlock, James (Hg.) (1994): *The Moral Writings of John Dewey*, revised edition, Amherst: Prometheus Books.
- Green, O.H. (1992): *The Emotions. A Philosophical Theory*, Dordrecht: Kluwer Academic Publishers.
- Greenspan, Patricia S. (1993): *Emotions and Reasons. An Inquiry into emotional Justification*, New York.
- Greenspan, Patricia S. (1995): *Practical Guilt. Moral Dilemmas, Emotions, and Social Norms*, New York u.a.
- Griffin, James (1998): *Value Judgement. Improving Our Ethical Beliefs*, Oxford: OUP.
- Griffiths, P.E. (1997): *What Emotions really are. The Problem of Psychological Categories*, Chicago/ London.
- Grimm, Jakob & Wilhelm (1956): *Deutsches Wörterbuch, Band 15*, Leipzig.
- Großheim, Michael & Waschkes, Hans-Joachim (Hgg.) (1993): *Rehabilitierung des Subjektiven: Festschrift für Hermann Schmitz*, Bonn: Bouvier.
- Großheim, Michael (Hg.) (1995): *Leib und Gefühl. Beiträge zur Anthropologie*, Berlin.
- Hägerström, Axel (1953): *Inquiries into the Nature of Law and Morals*, Uppsala: Uppsala University Press.
- Halbig, Christoph (2002): Motivierende Gründe. Neue Beiträge zur Theorie praktischer Vernunft (I), in: *Deutsche Zeitschrift für Philosophie* 50, 6, S. 961-972.
- Halbig, Christoph (2003): Normative Gründe. Neue Beiträge zur Theorie praktischer Vernunft (II), in: *Deutsche Zeitschrift für Philosophie* 51, 1, S. 133-149.

- Hamann, S.B./ Stefanacci, L./ Squire, L./ Adolphs, R./ Tranel, D./ Damasio, H./ Damasio, A.R. (1996): Recognizing facial emotion, in: *Nature* 379, S. 497.
- Hammacher, Klaus (1984): La raison dans la vie affective et sociale selon Descartes et Spinoza, in: *Les Etudes Philosophiques* 1, 73-81.
- Hammacher, Klaus (2000): Spinozas Ethik als Affektenlehre: Ansatz und Entfaltung, in: K. Hammacher/ I. Reimers-Tovote/ M. Walther (Hgg.): *Zur Aktualität der Ethik Spinozas*, Würzburg: Königshausen & Neumann, S. 27-55 (= Schriftenreihe der Spinoza-Gesellschaft, Bd. 7).
- Hampton, Jean (1998): *The Authority of Reason*, Cambridge: CUP.
- Hare, Richard M. (1952): *The Language of Morals*, Oxford.
- Hare, Richard M. (1984): Supervenience, in: *Proceedings of the Aristotelian Society*, Supplementary Volume 58, S. 1-16.
- Harman, Gilbert (1984): Is There a Single True Morality?, in: David Copp & David Zimmerman (Hgg.): *Morality, Reason and Truth: New Essays on the Foundations of Ethics*, Totowa (NJ): Rowman and Allanheld, S. 112-131.
- Harman, Gilbert (1986): Moral Explanations of Natural Facts – Can Moral Claims Be Tested Against Moral Reality?, in: *Southern Journal of Philosophy* 24, Supplement, S. 57-68.
- Harman, Gilbert (1999): Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error, in: *Proceedings of the Aristotelian Society* 99, S. 315-331.
- Harré, Rom (Hg.) (1996): *The Emotions. Social, Cultural and Biological Dimensions*, London u.a.: Sage.
- Hartmann, Martin (2002): Die Repsychologisierung des Geistes. Neuere Literatur über Emotionen, in: *Philosophische Rundschau* 49, S. 195-223.
- Hartmann, Matthias (1995): Leib und Gefühl – ein Philosophem aus psychiatrischer Sicht, in: Großheim, Michael (Hg.): *Leib und Gefühl. Beiträge zur Anthropologie*, Berlin 1995, 215-228.
- Hauskeller, Michael (1995): *Atmosphären erleben. Philosophische Untersuchungen zur Sinneswahrnehmung*, Berlin.
- Hauskeller, Michael (1995a): Die Erfahrung von Anwesenheit, in: Großheim, Michael (Hg.): *Leib und Gefühl. Beiträge zur Anthropologie*, Berlin 1995, 109-118.
- Hauskeller, Michael (1999): *Auf der Suche nach dem Guten. Wege und Abwege der Ethik*, Kusterdingen.
- Hauskeller, Michael (2001): *Versuch über die Grundlagen der Moral*, München.

- Heimbrock, Hans-Günther (1981): Selbsterkenntnis als Gotteserkenntnis. Spinozas Affektenlehre im Zusammenhang mit der neueren psychoanalytischen Narzissmustheorie, in: Ingrid Craemer-Ruegenberg (Hg.): Pathos, Affekt, Gefühl, Freiburg i.Br./ München 1981, 205-230.
- Heller, Agnes (1981): Theorie der Gefühle, Hamburg.
- Helm, Bennett W. (1994): The Significance of Emotions, in: American Philosophical Quarterly 31, S. 319-331.
- Helm, Bennett W. (1996): Integration and Fragmentation of the Self, in: Southern Journal of Philosophy 34, 1, S. 43-63.
- Helm, Bennett W. (2000): Emotional Reason: How to Deliberate about Value, in: American Philosophical Quarterly 37, 1, S. 1-22.
- Helm, Bennett W. (2001): Emotions and Practical Reason: Rethinking Evaluation and Motivation, in: Noûs 35, 2, S. 190-213
- Helm, Bennett W. (2001a): Emotional Reason. Deliberation, Motivation, and the Nature of Value, Cambridge: CUP.
- Helm, Bennett W. (2002): Felt Evaluations: A Theory of Pleasure and Pain, in: American Philosophical Quarterly 39, 1, S. 13-30.
- Herman, Barbara (1985): The Practice of Moral Judgment, in: The Journal of Philosophy 82, 8, S. 414-436.
- Herman, Barbara (1993): The Practice of moral Judgment, Cambridge.
- Herman, Barbara (1996): Making Room for Character, in: Engstrom & Whiting (Hg.) (1996): Aristotle, Kant, and the Stoics, S. 36-60.
- Heuer, Ulrike (2004): Raz on Values and Reasons, in: Wallace/ Pettit/ Scheffler/ Smith (Hgg.) (2004): Reason and Value, S. 129-152.
- Höffe, Otfried (1971): Praktische Philosophie. Das Modell des Aristoteles, München/ Salzburg.
- Hohmann, G.W. (1966): Some effects of spinal cord lesions on experienced emotional feelings, in: Psychophysiology 3, S. 143-156.
- Horst, Steven (1998): Our Animal Bodies, in: French & Wettstein (Hgg.) (1998), S. 34-61.
- Hover, W. (1993): Der Begriff des Herzens bei Blaise Pascal: Gestalt, Elemente der Vorgeschichte und der Rezeption im 20. Jahrhundert, Fridingen a.D.
- Hudson, Stephen D. (1986): Human Character and Morality. Reflections from the History of Ideas, London: Routledge & Kegan Paul.

- Hume, David (Enquiry 1748/ 1751): Enquiry Concerning Human Understanding (1748) und Enquiry Concerning the Principles of Morals (1751), 3. Aufl., hrsg. v. L.A. Selby-Bigge mit P.H. Nidditch, Oxford 1975.
- Hume, David (Treatise): A Treatise of Human Nature, 2. Aufl., hrsg. v. L.A. Selby-Bigge mit P.H. Nidditch, Oxford 1985.
- Hursthouse, Rosalind (1997): Virtue Ethics and the Emotions, in: Statman (Hg.) (1997): Virtue Ethics, S. 99-117.
- Hursthouse, Rosalind (1999): On Virtue Ethics, Oxford: OUP.
- Iambi et Elegi Graeci ante Alexandrum cantati (West), hrsg. v. M.L. West, 2 Bde., 2. Aufl., Oxford 1989-92.
- Isen, A.M. & Means, B. (1983): The influence of positive affect on decision making strategy, in: Social Cognition 2, S. 18-31.
- Isen, A.M. & Patrick, R. (1983): The effect of positive feelings on risktaking: when the chips are down, in: Organizational Behavior and Human Performance 31, S. 194-202.
- Isen, A.M. (1993): Positive affect and decision making, in: M. Lewis & A.J.M. Haviland (Hgg.): Handbook of emotions, New York: Guilford Press, S. 261-277.
- Isen, A.M./ Daubmann, K.A./ Nowicki, G.P. (1987): Positive affect facilitates creative problem solving, in: Journal of Personality and Social Psychology 52, S. 1122-1131.
- Isen, A.M./ Shalke, T.E./ Clark, M.S./ Karp, L. (1978): Affect, accessibility of material in memory, and behavior: a cognitive loop?, in: Journal of Personality and Social Psychology 36, S. 1-12.
- Izard, Carrol (1990): Facial Expressions and the regulation of emotions, in: Journal of Personality and Social Psychology 58, 3, S. 487-498.
- Jackson, F. & Pettit, P. (1995): Moral functionalism and moral motivation, in: Philosophical Quarterly 45, S. 451-472.
- Jackson, Frank (1986): What Mary didn't know, in: Journal of Philosophy 82, S. 291-295.
- Jackson, Frank (1992): Critical Notice of Susan Hurley's „Natural Reasons: Personality and Polity“, in: Australasian Journal of Philosophy 70, S. 475-488.
- Jackson, Frank (1996): Mental Causation, in: Mind 105, S. 377-413.
- Jackson, Frank/ Pettit, Philip (1995): Moral Functionalism and Moral Motivation, in: The Philosophical Quarterly 45, S. 21-40.
- Jacobi, Klaus (1981): Aristoteles über den rechten Umgang mit Gefühlen, in: Ingrid Craemer-Ruegenberg (Hg.): Pathos, Affekt, Gefühl, Freiburg i.Br./ München 1981, 21-52.

- Jacobs, Jonathan A. (2001): *Choosing Character: Responsibility for Virtue and Vice*, Ithaca/London: Cornell University Press.
- Jacques, S. (1979): Brain stimulation and reward: „pleasure centers“ after twenty-five years, in: *Neurosurgery* 5, 2, S. 277-283.
- James, William (1884): What is an emotion?, in: *Mind* 9, S. 188-205.
- James, William (1890), *The Principles of Psychology*, New York: Holt.
- James, William (1894): The physical basis of emotion, in: *Psychological Review* 1, S. 516-529.
- James, William (1920): *Psychologie*, 2. Aufl., Leipzig.
- James, William (1950): *The Principles of Psychology*, 2 Bde., New York: Dover. (-1890).
- James, William (1984): *Psychology: Briefer Course*, *The Works of William James*, Bd. 13, Cambridge (MA)/ London: Harvard University Press (zuerst 1892).
- Janssen, Jan-Peters (1995): Leib und Gefühl bei sportlicher Tätigkeit. Eine sportpsychologische Perspektive, in: Großheim, Michael (Hg.): *Leib und Gefühl. Beiträge zur Anthropologie*, Berlin 1995, 283-294.
- Jenkins, J./ Oatley, K./ Stein, N. (Hgg.) (1998): *Human Emotion: A Reader*, London: Blackwell.
- Johnson, Robert (1999): Internal Reasons and the Conditional Fallacy, in: *The Philosophical Quarterly* 49, S. 53-71.
- Johnson-Laird, P.N. (1988): *Der Computer im Kopf: Formen und Verfahren der Erkenntnis*, München: dtv 1996.
- Johnston, M. (2001): The Authority of Affect, in: *Philosophy and Phenomenological Research* 63, S. 181-214.
- Jones, Karen (2003): Emotion, Weakness of Will, and the Normative Conception of Agency, in: *Philosophy and the Emotions*, hrsg. v. Anthony Hatzimoysis (=Royal Institute of Philosophy supplement 52), S. 181-200.
- Kaegi, Dominic (1999): Ein gutes Gefühl. Aristoteles über den Zusammenhang von Affekt und Tugend, in: Hübsch & Kaegi (Hgg.): *Affekte. Philosophische Beiträge zur Theorie der Emotionen*, Heidelberg 1999.
- Kambouchner, Denis (1995): *L'homme des passions. Commentaire sur Descartes*, Paris (Bd. 1: „Analytique“, Bd. 2: „Canonique“).
- Kant, Immanuel (Grundlegung): *Grundlegung zur Metaphysik der Sitten*, hrsg. v. Bernd Kraft u. Dieter Schönecker, Hamburg: Meiner 1999.

- Kant, Immanuel (KpV): Kritik der praktischen Vernunft, hrsg. v. Karl Vorländer, 10. Aufl., Hamburg: Meiner 1990.
- Kapp, B.S./ Pascoe, J.P./ Bixler, M.A. (1984): The amygdala: A neuroanatomical systems approach to its contributions to aversive conditioning, in: N. Buttlers & L.R. Squire (Hgg.): Neuropsychology of memory, New York: Guilford, S. 473-488.
- Katz, David & Katz, Rosa (Hgg.) (1960): Handbuch der Psychologie, 2., erw. Aufl., Basel/ Stuttgart.
- Kaufmann, Peter (1992): Gemüt und Gefühl als Komplement der Vernunft. Eine Auseinandersetzung mit der Tradition und der phänomenologischen Ethik, besonders Max Schelers, Frankfurt a.M.
- Kay, S.R. (1991): Positive and negative syndromes in Schizophrenia, Assessment and Research, Clinical and Experimental Psychiatry, New York: Brunner & Mazel.
- Keinan, G. (1987): Decision making under stress: scanning of alternatives under controllable and uncontrollable threats, in: Journal of Personality and Social Psychology 52, S. 639-644.
- Keltner, D./ Ellsworth, P.C./ Edwards, K. (1993): Beyond simple pessimism: effects of sadness and anger on social perception, in: Journal of Personality and Social Psychology 64, S. 740-752.
- Kenny, Anthony (1963): Action, Emotion, and Will, London.
- Kenny, Anthony (1989): The Metaphysics of Mind, Oxford.
- Kentridge, R.W. & Heywood, C.A. (1999) The Status of Blindsight. Near-threshold vision, islands of cortex and the Riddoch phenomenon, in: Journal of Consciousness Studies 6, 5, S. 3-11.
- Kihlstrom, John F. (1987): The cognitive unconscious, in: Science 237, S. 1445-1452.
- Kleinginna, Paul R. & Kleinginna, Anne M. (1981): A Categorized List of Emotion Definitions, with Suggestions ofr a Consensual Definition, in: Motivation and Emotion, 5, 4, S. 345-379.
- Kluwe, Rainer H. (2001): Kognition, in: Lexikon der Psychologie in fünf Bänden, Heidelberg/ Berlin: Spektrum, Bd. 2, S. 352-355.
- Koch, Philip J. (1987): Bodily Feeling in Emotion, in: Dialogue 26, 1, S. 59-75.
- Kochinka, Alexander (2004): Emotionstheorien. Begriffliche Arbeit am Gefühl, Bielefeld: transcript.
- Kolnai, Aurel (1977): Deliberation is of Ends, in: Ethics, Value and Reality, Oxford 1977, S. 44-62 (wiederabgedr. in: Millgram (2001), S. 259-278).

- Korsgaard, Christine (1986): Skepticism about Practical Reason, in: *Journal of Philosophy*, S. 5-25.
- Korsgaard, Christine (1996): *The Sources of Normativity*, hrsg. v. Onora O'Neill, Cambridge.
- Korsgaard, Christine (1996a): *Creating the Kingdom of Ends*, Cambridge.
- Korsgaard, Christine (1997): The Normativity of Instrumental Reason, in: Cullity & Gaut (1997), S. 215-254.
- Korsgaard, Christine (1996): From Duty and for the Sake of the Noble: Kant and Aristotle on Morally Good Action, in: Engstrom & Whiting (Hg.) (1996): *Aristotle, Kant, and the Stoics*, S. 203-236.
- Kosman, L.A. (1980): Being properly affected. Virtues and Feelings in Aristotle's Ethics, in: Amélie Oksenberg Rorty (Hg.): *Essays on Aristotle's Ethics*, Berkeley/ Los Angeles/ London: University of California Press, S. 103-116.
- Kraut, R.E. & Johnston, R.E. (1979): Social and emotional messages of smiling: an ethological approach, in: *Journal of Personality and Social Psychology* 37, S. 1539-1553.
- Kristjánsson, Kristján (2001): Some Remaining Problems in Cognitive Theories of Emotion, in: *International Philosophical Quarterly* 41, 4, S. 393-410.
- Kuhn, Thomas (1962): *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.
- Kunst-Wilson, W.R. & Zajonc, R.B. (1980): Affective Discrimination of Stimuli that cannot be recognized, in: *Science* 207, S. 557-558.
- Kupperman, Joel (1991): *Character*, New York/ Oxford: OUP.
- Kupperman, Joel J. (2001): The Indispensability of Character, in: *Philosophy* 76, S. 239-250.
- Kupperman, Joel J. (1988): Character and Ethical Theory, in: *Midwest Studies in Philosophy* 13, S. 115-125.
- Kupperman, Joel J. (1999): *Value... and what follows*, Oxford/ New York: OUP.
- LaBar, K.S./ LeDoux, J.E./ Spencer, D.D./ Phelps, E.A. (1995): Impaired fear conditioning following unilateral temporal lobectomy in humans, in: *Journal of Neuroscience* 15, S. 6846-6855.
- Laird, J.D. & Bresler, C. (1992): The process of emotional experience: a self-perception theory, in: M.S. Clark (Hg.): *Review of Personality and Social Psychology*, Bd. 13: *Emotion*, Newbury Park (CA): Sage, S. 213-234.

- Laird, J.D. (1974): Self-attribution of emotion: the effects of expressive behavior on the quality of emotional experience, in: *Journal of Personality and Social Psychology* 29, S. 475-486.
- Laird, J.D. (1984): The real role of facial response in the experience of emotion: a reply to Tourangeau and Ellsworth, and others, in: *Journal of Personality and Social Psychology* 47, S. 909-917.
- Lamb, Roger (1987): Objectless Emotions, in: *Philosophy and Phenomenological Research* 48, S. 107-117.
- Lane, R. & Nadel, L. (1999): *The Cognitive Neuroscience of Emotion*, Oxford.
- Lang, Hermann (1988): Zum Begriff des Unbewussten, in: *Die Philosophen und Freud*, hg. v. H. Vetter & L. Nagel, Wien/ München: Oldenbourg, S. XXX (= Wiener Reihe, Bd. III).
- Lang, Hermann (1993): Zur Phänomenologie der Affektivität in der Psychotherapie, in: Fink-Eitel & Lohmann (1993), 293-306.
- Lang, P.J./ Bradley, M.M./ Cuthbert, B.N. (1990): Emotion, attention and the startle reflex, in: *Psychological Review* 97, S. 377-395.
- Lange, Carl & James, William (1922): *The Emotions*, hrsg. v. Knight Dunlap, Baltimore: Williams & Wilkins.
- Lange, Carl (1885): *Die Gemütsbewegungen, ihr Wesen und ihr Einfluss auf körperliche, besonders auf krankhafte Lebenserscheinungen*, 2. Aufl., Würzburg 1910 (dänisches Orig. 1885).
- Lanzetta, J.T. & Kleck, R.E. (1970): Encoding of nonverbal affect in humans, in: *Journal of Personality and Social Psychology* 16, S. 12-19.
- Laux, L. & Weber, H. (1990): Bewältigung von Emotionen, in: K.R. Scherer (Hg.): *Enzyklopädie der Psychologie*, Bd. C IV 3: *Psychologie der Emotionen*, Göttingen: Hogrefe, S. 560-629.
- Lazarus, R.S./ Averill, J./ Opton, E. (1970): *Towards a Cognitive Theory of Emotion*, in: *Feelings and Emotions*, hrsg. v. Magda B. Arnold, Academic Press 1970.
- Lazarus, Richard S. (1966): *Psychological stress and the coping process*, New York: McGraw-Hill.
- Lazarus, Richard S. (1982): Thoughts on the relations between emotion and cognition, in: *American Psychologist* 37, S. 1019-1024.
- Lazarus, Richard S. (1984): On the primacy of cognition, in: *American Psychologist* 39, S. 124-129.

- Lazarus, Richard S. (1990): Constructs of the mind in adaptation, in: N.L. Stein/ B. Leventhal/ T. Trabasso (Hgg.): Psychological and biological approaches to emotion, Hillsdale (NJ): Lawrence Erlbaum Associates, S. 3-19.
- Lazarus, Richard S. (1991): Emotion and Adaptation, New York: OUP.
- Lazarus, Richard S. (1991a): Cognition and motivation in emotion, in: American Psychologist 46, S. 352-367.
- Lazarus, Richard S. (1991b): Progress on a cognitive-motivational-relational theory of emotion, in: American Psychologist 46, S. 819-834.
- LeDoux, Joseph (1993): Emotional networks in the brain, in: M. Lewis & J.M. Haviland (Hgg.): Handbook of emotions, New York: Guilford Press, S. 109-117.
- LeDoux, Joseph (1996): Das Netz der Gefühle. Wie Emotionen entstehen, 2. Aufl., München: dtv 2003 (eng. Orig.: The Emotional Brain. The Mysterious Underpinnings of Emotional Life).
- Leighton, Stephen R. (1982): Aristotle and the Emotions, in: Phronesis 27, S. 144-174.
- Leighton, Stephen R. (1984): Feelings and Emotion, in: Review of Metaphysics 38, S. 303-320.
- Levenson, R.W. (1992): Autonomic nervous system differences among emotions, in: Psychological Science 3, S. 23-27.
- Levi, A. (1964): French Moralists. The Theory of the Passions 1585 to 1649, Oxford.
- Lewis, David (1989): Dispositional Theories of Value, in: Proceedings of the Aristotelian Society, Supplementary Volume LXIII, S. 113-137.
- Lewis, David (1988): Desire as belief, in: Mind 97, S. 323-332.
- Lewis, David (1996): Desire as belief II, in: Mind 105, S. 303-313.
- Lloyd, A.C. (1978): Emotion and Decision in Stoic Psychology, in: The Stoics, hrsg. v. John M. Rist, Berkeley u.a.: University of California Press, S. 233-246.
- Long, A.A./ Sedley, D.N. (LS): Die hellenistischen Philosophen: Texte und Kommentare, übers. v. Karlheinz Hülsler, Stuttgart/ Weimar 2000.
- Louden, Robert B. (1986): Kant's Virtue Ethics, in: Philosophy 61, S. 473-489.
- Lyons, William (1974): Physiological Changes and Emotions, in: Canadian Journal of Philosophy 3, S. 603-617.
- Lyons, William (1980): Emotions, Cambridge.
- Macann, Christopher (1995): Ein-fühlung: Schlüsselbegriff der persönlichen Beziehung, in: Großheim, Michael (Hg.): Leib und Gefühl. Beiträge zur Anthropologie, Berlin 1995, 97-108.

- Mackie, J.L. (1977): *Ethics. Inventing Right and Wrong*, Harmondsworth.
- Madell, Geoffrey & Ridley, Aaron (1997): *Emotion and Feeling*, in: *Proceedings of the Aristotelian Society Supplementary Volume 71*, S. 147-176.
- Magnusson, D. & Endler, N. (1977): *Interactional Psychology: Present Status and Future Prospects*, in: *Ibid.* (Hgg.): *Personality at the Crossroads*, Hillsdale (NJ): Lawrence Erlbaum Associates, S. 3-31.
- Manstead, A.S.R. (1991): *Expressiveness as an individual difference*, in: R.S. Feldman & B. Rimé (Hgg.): *Fundamentals of nonverbal behavior*, Cambridge: CUP, S. 285-328.
- Marañon, Gregorio (1924): *Contribution à l'étude de l'action emotive de l'adrénaline*, in: *Revue Française d'Endocrinologie 2*, S. 301-325.
- Margolis, Diane R. (1998): *The Fabric of Self. A Theory of Ethics and Emotions*, New Haven.
- Marks, Joel (1982): *A Theory of Emotion*, in: *Philosophical Studies 42*, S. 227-242.
- Marshall, G.D. & Zimbardo, P.G. (1979): *Affective consequences of inadequately explained physiological arousal*, in: *Journal of Personality and Social Psychology 37*, S. 970-988.
- Marshall, John C. & Halligan, Peter W. (1988): *Blindsight and insight in visuo-spatial neglect*, in: *Nature 336*, S. 766f.
- Martin, L.L./ Ward, D.W./ Achee, J.W./ Wyer, R.S., Jr. (1993): *Mood as input: people have to interpret the motivational implications of their moods*, in: *Journal of Personality and Social Psychology 64*, S. 317-326.
- Maslach, C. (1979): *Negative emotional biasing of unexplained arousal*, in: *Journal of Personality and Social Psychology 37*, S. 953-969.
- Mauro, R. (1992): *Affective dynamics: opponent processes and excitation transfer*, in: M.S. Clark (Hg.): *Review of Personality and Social Psychology*, Bd. 13: *Emotion*, Newbury Park (CA): Sage, S. 150-173.
- Mayer, Verena (2002): *Tugend und Gefühl*, in: *Die Moralität der Gefühle*, Deutsche Zeitschrift für Philosophie, Sonderband 4, hrsg. v. Sabine A. Döring u. Verena Mayer, Berlin, S. 125-149.
- Mazis, Glen (1993): *Emotion and Embodiment: Fragile Ontology*, New York: Peter Lang.
- McDowell, John (1978): *Are Moral Requirements Hypothetical Imperatives?*, in: *Proceedings of the Aristotelian Society, Supplementary Vol. 52*, S. 13-29.
- McDowell, John (1979): *Virtue and Reason*, in: Crisp, Roger & Slote, Michael (Hgg.) (1997): *Virtue Ethics*, S. 141-162 (zuerst in: *The Monist 62*, S. 331-350).

- McDowell, John (1981): Non-Cognitivism and Rule-Following, in: Steven Holtzman & Christopher Leich (Hgg.): Wittgenstein: To Follow a Rule, Routledge, S. 141-162.
- McDowell, John (1985): Values and Secondary Qualities, in: Ders.: Mind, Value, and Reality, Cambridge (Mass.) 1998, S. 131-150.
- McDowell, John (1996): Deliberation and Moral Development in Aristotle's Ethics, in: Engstrom & Whiting (Hg.) (1996): Aristotle, Kant, and the Stoics, S. 19-35.
- McKinnon, Christine (1999): Character, Virtue Theories, and the Vices, Ontario: Broadview Press.
- McNaughton, David (1988): Moral Vision, Oxford: Basil Blackwell.
- Mees, Ulrich (1985): Was meinen wir, wenn wir von Gefühlen reden? Zur psychologischen Textur von Emotionswörtern, in: Sprache und Kognition 1, S. 2-20.
- Mees, Ulrich (1991): Die Struktur der Emotionen, Göttingen/ Toronto/ Zürich: Hogrefe.
- Meier-Seethaler, Carola (1997): Gefühl und Urteilskraft. Ein Plädoyer für die emotionale Vernunft, München.
- Mele, Alfred R. (1989): Motivational Internalism: The Powers and Limits of Practical Reasoning, in: *Philosophia* 19, S. 417-436 (wiederabgedr. in: Wallace (1999): Reason, Emotion and Will).
- Mele, Alfred R. (1996). Internalist moral cognitivism and listlessness, in: *Ethics* 106, S. 727-753.
- Merten, Jörg (2003): Einführung in die Emotionspsychologie, Stuttgart: Kohlhammer.
- Metzger, Wolfgang (1975): Psychologie, 5. Aufl., Darmstadt.
- Meyer, Wolf-Uwe/ Schützwohl, Achim/ Reizenzein, Rainer (2001): Einführung in die Emotionspsychologie, Bd. 1, 2., vollständig überarbeitete Auflage, Bern u.a.: Huber.
- Meyer, Wolf-Uwe/ Schützwohl, Achim/ Reizenzein, Rainer (2003): Einführung in die Emotionspsychologie, Bd. 2, 3. korrigierte Auflage, Bern u.a.: Huber.
- Meyer, Wolf-Uwe/ Schützwohl, Achim/ Reizenzein, Rainer (2002): Einführung in die Emotionspsychologie, Bd. 3, Bern u.a.: Huber.
- Michon, H. (1996): L'ordre du cœur: philosophie, théologie et mystique dans les pensées de Pascal, Paris.
- Midgley, Mary (1981): Heart and Mind. The Varieties of Moral Experience, London.
- Milgram, Stanley (1963): Behavioral Study of Obedience, in: *Journal of Abnormal and Social Psychology* 67, 4, S. 371-378.
- Milgram, Stanley (1974): Das Milgram-Experiment, Reinbek: Rowohlt.
- Milgram, Stanley (1974): Obedience to Authority, London: Tavistock.

- Millgram, Elijah (1993): Pleasure in practical reasoning, in: *The Monist* 76, S. 394-415 (wiederabgedr. in: Millgram (2001), S. 331-354).
- Millgram, Elijah (1996): Williams' Argument Against External Reasons, in: *Noûs* 30, S. 197-220 (wiederabgedr. in: Wallace (1999): *Reason, Emotion and Will*).
- Millgram, Elijah (1997): *Practical Induction*, Cambridge (MA): Harvard UP.
- Millgram, Elijah (Hg.) (2001): *Varieties of Practical Reasoning*, Cambridge (MA): Harvard UP.
- Montmarquet, James (2003): Moral Character and Social Science Research, in: *Philosophy* 78, 355-368.
- Moore, G.E. (1903): *Principia Ethica*, Cambridge: CUP.
- Moore, Simon & Oaksford, Mike (2002): Emotional Cognition. An introduction, in: Moore, Simon & Oaksford, Mike (Hgg.) (2002): *Emotional Cognition: From brain to behaviour*, Amsterdam/ Philadelphia: John Benjamins (= *Advances in Consciousness Research* 44), S. 1-8.
- Moore, Simon & Oaksford, Mike (Hgg.) (2002): *Emotional Cognition: From brain to behaviour*, Amsterdam/ Philadelphia: John Benjamins (= *Advances in Consciousness Research* 44).
- Morreal, John (1993): Fear without Belief, in: *Journal of Philosophy* 90, 359-366.
- Morsbach, H. & Tyler, W.J. (1986): A Japanese emotion: Amai, in: Rom Harré (Hg.): *The social construction of emotions*, New York: Blackwell, S. 289-307.
- Morton, Adam (2002): Beware stories: Emotions and Virtues, in: Goldie, Peter (Hg.): *Understanding Emotions. Mind and Morals*, Aldershot u.a.: Ashgate, S. 55-64.
- Müller, Anselm Winfried (2003): Handeln, in: *Zeitschrift für philosophische Forschung* 57, 3, S. 327-349.
- Mulligan, Kevin (1998): From Appropriate Emotions to Values, in: *The Monist* 81, S. 161-188.
- Murphy, Jeffrie G. (2003): *Getting even. Forgiveness and its Limits*, New York/ Oxford: OUP.
- Murphy, Jeffrie G./ Hampton, Jean (1988): *Forgiveness and Mercy*, Cambridge: CUP.
- Murphy, S. & Zajonc, R. (1993): Affect, cognition, and awareness: Affective priming with suboptimal and optimal stimuli, in: *Journal of Personality and Social Psychology* 64, S. 723-739.
- Nagel, Thomas (1970): *The Possibility of Altruism*, Princeton University Press.

- Nagel, Thomas (1976): Moral Luck, in: ders., *Mortal Questions*, Cambridge: CUP 1979, S. 24-38 (zuerst in: *Proceedings of the Aristotelian Society, Supplementary Volume 50*, S. 137-152).
- Nagel, Thomas (1986): *The View from Nowhere*, Oxford: OUP.
- Nash, R.A. (1989): Cognitive Theories of Emotion, in: *Nous* 23, S. 481-504.
- Neill, D.B./ Fenton, H./ Justice, J.B. Jr. (2002): Increase in accumbal dopaminergic transmission correlates with response cost not reward of hypothalamic stimulation, in: *Behavioural Brain Research* 137, 1-2, S. 129-138.
- Neu, Jerome (1977): *Emotion, Thought, and Therapy*, London: Routledge.
- Nickl, Peter (2001): *Ordnung der Gefühle. Studien zum Begriff des habitus*, Hamburg.
- Nida-Rümelin, Julian (2001): *Strukturelle Rationalität. Ein philosophischer Essay über praktische Vernunft*, Stuttgart: Reclam.
- Nieuwenburg, Paul (2002): Emotion and Perception in Aristotle's *Rhetoric*, in: *Australasian Journal of Philosophy* 80, 1, S. 86-100.
- Nisbett, R. & Ross, L. (1980): *Human Inference: Strategies and Shortcomings of Social Judgment*, Englewood Cliffs (NJ): Prentice Hall.
- Nozick, Robert (1974): *Anarchy, State, and Utopia*, New York.
- Nussbaum, Martha C. (1985): Finely Aware and Richly Responsible, in: *The Journal of Philosophy* 82, 516-529.
- Nussbaum, Martha C. (1990): *Love's Knowledge*, New York: OUP.
- Nussbaum, Martha C. (1994): *The Therapy of Desire. Theory and Practice in Hellenistic Ethics*, Princeton.
- Nussbaum, Martha C. (1996): *Upheavals of Thought. A Theory of the Emotions*, Gifford Lectures 1993, Cambridge.
- Nussbaum, Martha C. (1996a): Compassion: The Basic Social Emotion, in: *Social Philosophy and Policy* 13, S. 27-58.
- O'Neill, Onora (1989): *Constructions of Reason. Explorations of Kant's Practical Philosophy*, Cambridge: CUP.
- Oakley, Justin (1992): *Morality and the Emotions*, New York.
- Oakley, Justin (1996): Varieties of Virtue Ethics, in: *Ratio* 9, S. 128-152.
- Oatley, Keith & Johnson-Laird, P.N. (1987): Towards a Cognitive Theory of Emotions, in: *Emotions and Cognition* 1, S. 29-50.
- Oatley, Keith & Johnson-Laird, P.N. (2002): Emotion and reasoning to consistency. The case of abductive inference, in: Moore, Simon & Oaksford, Mike (Hgg.) (2002): *Emotional*

- Cognition: From brain to behaviour, Amsterdam/ Philadelphia: John Benjamins (= Advances in Consciousness Research 44), S. 157-182.
- Oatley, Keith (1992): *Best Laid Schemes: The Psychology of Emotions*, Cambridge: CUP.
- Öhman, Arne (1987): *The Psychophysiology of Emotion. An evolutionary-cognitive Perspective*, in: *Advances in Psychophysiology* 2, S. 79-127.
- Ortony, A./ Clore, G.L./ Collins, A. (1988): *The Cognitive Structure of Emotions*, Cambridge UP.
- Otto, J.H. & Schmitz, B.B. (1993): *Veränderung positiver Gefühlszustände durch analytische und kreative Informationsverarbeitung*, in: *Zeitschrift für Experimentelle und Angewandte Psychologie* 40, S. 235-266.
- Otto, Jürgen H./ Euler, Harald A./ Mandl, Heinz (Hgg.) (2000): *Emotionspsychologie. Ein Handbuch*, Stuttgart: KNV.
- Palmer, Frank (1992): *Literature and Moral Understanding*, Oxford.
- Parfit, Derek (1984): *Reasons and Persons*, Oxford.
- Parfit, Derek (1997): *Reasons and Motivation*, in: *Proceedings of the Aristotelian Society Supplementary Volume 77*, S. 99-130 (wiederabgedr. in: Wallace (1999): *Reason, Emotion and Will*).
- Pascal, Blaise (*Pensées*): *Œuvres Complètes*, hrsg. v. J. Chevalier, Paris: Pléiade 1941.
- Pauer-Studer, Herlinde (Hg.) (2003): *Constructions of Practical Reason. Interviews on Moral and Political Philosophy*, Stanford: Stanford University Press.
- Perler, Dominik (2000): *Der kognitive Gehalt von Emotionen. Erklärungsansätze des klassischen Rationalismus*, in: *Studia Philosophica* 59, S. 83-107.
- Petrus, Klaus (2000): *Die Autonomie der Emotionen*, in: *Studia Philosophica* 59, S. 129-154.
- Pettit, Philip & Smith, Michael (1990): *Backgrounding Desire*, in: *Philosophical Review* 99, S. 565-592.
- Pettit, Philip & Smith, Michael (1993): *Practical Unreason*, in: *Mind* 102, S. 53-79.
- Pettit, Philip (1993): *The Common Mind*, Oxford: OUP.
- Pfaffman, Car (1960): *The Pleasure of Sensation*, in: *Psychological Review* 67, S. 253-268.
- Pitcher, George (1965): *Emotion*, in: *Mind* 74, S. 326-346.
- Platts, Mark (1979): *Ways of Meaning*, Oxford.
- Platts, Mark (1981): *Moral Reality and the End of Desire*, in: ders. (Hg.): *Reference, Truth and Reality*, London: Routledge & Kegan Paul, S. 69-82.
- Pocai, R. (1996): *Heideggers Theorie der Befindlichkeit. Sein Denken zwischen 1927 und 1933*, Freiburg/ München.

- Prinz, Jesse J. (2003): Emotion, Psychosemantics, and Embodied Appraisals, in: *Philosophy and the Emotions*, hrsg. v. Anthony Hatzimoysis (=Royal Institute of Philosophy supplement 52), S. 69-86.
- Prinz, Jesse J. (2002): Consciousness, computation, and emotion, in: Moore, Simon & Oaksford, Mike (Hgg.) (2002): *Emotional Cognition: From brain to behaviour*, Amsterdam/ Philadelphia: John Benjamins (= Advances in Consciousness Research 44), S. 137-155.
- Pylyshyn, Zenon (1984): *Computation and Cognition: Toward a Foundation for Cognitive Science*, Cambridge (MA): MIT Press.
- Quinn, Warren (1993): Putting Rationality in its Place, in: ders., *Morality and Action*, hg. v. Philippa Foot, Cambridge: CUP, S. 228-255.
- Quinton, Anthony (1982): *Thoughts and Thinkers*, New York: Holmes & Meier.
- Railton, Peter (1986): Moral Realism, in: *Philosophical Review* 95, S. 163-207.
- Rangell, L. (1954): Similarities and Differences between Psychoanalysis and Dynamic Psychotherapy, in: *Journal of the American Psychoanalytical Association* 2, 734-747.
- Rapp, Christof (2002): Vorbemerkung zu Kap. II 2-11: Die ‚Emotionen‘, in: Aritstoteles, *Rhetorik*, übers. u. erläutert v. Ch. Rapp, Berlin: Akademie-Verlag, Bd. 4.2, S.543-583.
- Ratcliffe, Matthew (2002): Heidegger's attunement and the neuropsychology of emotion, in: *Phenomenology and the Cognitive Sciences* 1, S. 287-312.
- Raz, Joseph (1990): *Practical Reason and Norms*, Princeton: Princeton University Press.
- Raz, Joseph (1999): *Engaging Reason. On the Theory of Value and Action*, Oxford: OUP.
- Raz, Joseph (1999a): Incommensurability and Agency, in: Raz (1999), S. 46-66.
- Raz, Joseph (1999b): Explaining Normativity: On Rationality and the Justification of Reason, in: Raz (1999), S. 67-89.
- Raz, Joseph (1999c): Explaining Normativity: Reason and the Will, in: Raz (1999), S. 90-117.
- Reber, Arthur S. (1995): *The Penguin Dictionary of Psychology*, 2. Aufl., London u.a.: Penguin.
- Reisenzein, Rainer (2000): Wundt's Three-Dimensional Theory of Emotion, in: Balzer, Wolfgang (Hg.) (2000): *Structuralist Knowledge Representation: Paradigmatic Examples*, Atlanta: Rodopi.
- Rey, Georges (1980): Functionalism and the Emotions, in: Rorty (1980), 163-196.
- Richardson, Henry (1994): *Practical Reasoning about Final Ends*, Cambridge: CUP.
- Roberts, Robert C. (1996): Propositions and Animal Emotion, in: *Philosophy* 71, S. 147-156.

- Roberts, Robert C. (1988): What an Emotion is: A Sketch, in: *The Philosophical Review* 97, S. 183-209.
- Roberts, Robert C. (1995): Feeling One's Emotions and Knowing Oneself, in: *Philosophical Studies* 77, 2-3, S. 319-338.
- Roberts, Robert C. (1999): Emotions as Judgments, in: *Philosophy and Phenomenological Research* 59, 3, S. 793-798.
- Robinson, Jenefer (1983): Emotion, Judgment, and Desire, in: *The Journal of Philosophy* 80, S. 731-741.
- Rorty, Amélie Oksenberg & Wong, David (1990): Aspects of Identity and Agency, in: Flanagan & Rorty (Hgg.) (1990): *Identity, Character and Morality*, S. 19-36.
- Rorty, Amélie Oksenberg (1980a): Introduction, in: Rorty (1980), S. 1-7.
- Rorty, Amélie Oksenberg (1988): Explaining Emotions, in: dies.: *Mind in Action. Essays in the Philosophy of Mind*, Boston.
- Rorty, Amélie Oksenberg (Hg.) (1976): *The Identities of Persons*, Berkeley/ Los Angeles: University of California Press.
- Rorty, Amélie Oksenberg (Hg.) (1980): *Explaining Emotions*, Berkeley u.a.
- Rorty, Richard (1991): *Kontingenz, Ironie und Solidarität*, Frankfurt a.M.
- Rorty, Richard (1993): Menschenrechte, Rationalität und Gefühl, in: Stephen Shute & Susan Hurley (Hgg.): *Die Idee der Menschenrechte*, Frankfurt a.M. 1996, 144-170.
- Roseman, I.J. (1984): Cognitive Determinants of Emotion: A Structural Theory, in: *Review of Personality and Social Psychology* 5, S. 11-36.
- Roskies, Adina (2003): Are ethical judgments intrinsically motivational? Lessons from „acquired sociopathy“, in: *Philosophical Psychology* 16, 1, S. 51-66.
- Ross, L. & Nisbett, R. (1991): *The Person and the Situation: Perspectives of Social Psychology*, New York: McGraw Hall.
- Ryle, Gilbert (1949): *The Concept of Mind*, London: Hutchinson's University Library.
- Sabini, John & Silver, Maury (1982): *Moralities of Everyday Life*, New York: OUP.
- Sabini, John & Silver, Maury (1987): Emotions, Responsibility, and Character, in: Schoeman (Hg.) (1987), S. 165-175.
- Salovey, P. & Birnbaum, D. (1989): Influence of mood on health-relevant cognitions, in: *Journal of Personality and Social Psychology* 57, S. 539-551.
- Sartre, Jean-Paul (1964): Entwurf einer Theorie der Emotionen, in: *Die Transzendenz des Ego. Drei Essays*, Hamburg.

- Saver, J.L. & Damasio, A.R. (1991): Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage, in: *Neuropsychologia* 29, S. 1241-1249.
- Sayre-McCord, Geoffrey (1997): The Metaethical Problem, in: *Ethics* 108, S. 55-83.
- Scanlon, Thomas (1982): Contractualism and Utilitarianism, in: Amartya Sen & Bernard Williams (Hgg.): *Utilitarianism and Beyond*, Cambridge: CUP, S. 103-128.
- Scanlon, Thomas (1999): *What we Owe To Each Other*, Cambridge (MA): Harvard UP.
- Scarano, Nico: *Moralische Überzeugungen. Grundlinien einer antirealistischen Theorie der Moral*, Paderborn: mentis 2001.
- Schachter, S./ Singer, J.E. (1962): Cognitive, social, and physiological determinants of emotional states, in: *Psychological Review* 69, S. 379-399.
- Schaub, Harald (1997): Denken, in: Jürgen Straub/ Wilhelm Kempf/ Hans Werbik (Hgg.): *Psychologie. Eine Einführung. Grundlagen, Methoden, Perspektiven*, München: dtv, S. 374-400.
- Scheler, Max (1980): *Der Formalismus in der Ethik und die materielle Wertethik*, 6., durchges. Aufl., Bern/ München.
- Scherer, Klaus R. (1984): On the Nature and Function of Emotion: A Component Process Approach, in: Scherer, K.R./ Ekman, P. (Hgg.): *Approaches to Emotion*, Hillsdale (NJ): Erlbaum, S. 293-318.
- Scherer, Klaus R. (1988): Criteria for emotion-antecedent appraisal: a review, in: V. Hamilton/ G.H. Bower/ H. Frijda (Hgg.): *Cognitive perspectives on emotion and motivation*, Dordrecht: Kluwer Academic Publishers, S. 89-126.
- Scherer, Klaus R. (1990): Theorien und aktuelle Probleme der Emotionspsychologie, in: ders. (Hg.): *Psychologie der Emotion. Enzyklopädie der Psychologie, Themenbereich C: Theorie und Forschung, Serie IV: Motivation und Emotion, Bd. 3: Psychologie der Emotion*, Göttingen u.a.: Hogrefe, S. 1-38.
- Schlossberger, Eugene (1986): Why we are responsible for our emotions, in: *Mind* 95, S. 37-56.
- Schmidt-Atzert, Lothar (1996): *Lehrbuch der Emotionspsychologie*, Stuttgart u.a.: Kohlhammer.
- Schmitz, Hermann (1972): Über leibliche Kommunikation, in: Schmitz (1992), 175-218. (zuerst in: *Zeitschrift für klinische Psychologie und Psychotherapie* 20, 4-32.)
- Schmitz, Hermann (1974): Das leibliche Befinden und die Gefühle, in: Schmitz (1992), 107-124. (zuerst in: *Zeitschrift für philosophische Forschung* 28, 325-338.)

- Schmitz, Hermann (1978): Leib und Seele in der abendländischen Philosophie, in: Schmitz (1992), 289-316. (zuerst in: Philosophisches Jahrbuch 85, 221-241).
- Schmitz, Hermann (1985): Phänomenologie der Leiblichkeit, in: Hilarion Petzold (Hg.): Leiblichkeit, Paderborn 1985, 71-106.
- Schmitz, Hermann (1992): Leib und Gefühl. Materialien zu einer philosophischen Therapeutik, hrsg. v. H. Gausebeck & G. Risch, Paderborn.
- Schmitz, Hermann (1993): Gefühle als Atmosphären und das affektive Betroffensein von ihnen, in: Fink-Eitel & Lohmann (1993), 33-56.
- Schmitz, Hermann (1998): Der Leib, der Raum und die Gefühle, Ostfildern (Stuttgart).
- Schmoll, Dirk (1995): Leib und Psychose, in: Großheim, Michael (Hg.): Leib und Gefühl. Beiträge zur Anthropologie, Berlin 1995, 229-240.
- Schneider, M. (1993): Das mechanistische Denken in der Kontroverse: Descartes' Beitrag zum Geist-Maschine-Problem, Stuttgart.
- Schoeman, Ferdinand David (Hg.) (1987): Responsibility, Character, and the Emotions: New Essays in Moral Psychology, Cambridge/ New York: Cambridge University Press.
- Schönecker, Dieter/ Wood, Allen W. (2002): Kants ‚Grundlegung zur Metaphysik der Sitten‘. Ein einführender Kommentar, Paderbon.
- Schroeter, François (1995): Tugend und Moraltheorie, in: Zeitschrift für philosophische Forschung 49, 1, S. 104-123.
- Schueler, G.F. (1995): Desire: Its Role in Practical Reason and the Explanation of Action, Cambridge (Mass.).
- Schwartz, Alan (2002): Expected feelings about risky options, in: Moore, Simon & Oaksford, Mike (Hgg.) (2002): Emotional Cognition: From brain to behaviour, Amsterdam/ Philadelphia: John Benjamins (= Advances in Consciousness Research 44), S. 183-195.
- Scoville, William B. & Milner, Brenda (1957): Loss of recent memory after bilateral hippocampal lesions, in: Journal of Neurology and Psychiatry 20, S. 11-21.
- Searle, John R. (1979): Expression and Meaning. Studies in the Theory of Speech Acts, Cambridge.
- Searle, John R. (1983): Intentionality. An Essay in the Philosophy of Mind, Cambridge: CUP.
- Sedikides, C. (1992): Mood as a determinant of attentional focus, in: Cognition and Emotion 6, S. 129-148.
- Seewald, Jürgen (1989): Leiblichkeit und symbolische Entwicklung. Implizite Sinnprozesse in systematischer und genetischer Betrachtung, Marburg.

- Seligman, Martin E. P. (1971): Phobias and Preparedness, in: *Behavior Therapy* 2, S. 307-320.
- Sellars, Wilfried (1956): *Empiricism and the Philosophy of Mind*, Cambridge (MA): Harvard University Press 2000.
- Sherman, Nancy (1989): *The Fabric of Character. Aristotle's Theory of Virtue*, Oxford: Clarendon Press.
- Sherman, Nancy (1990): The Place of Emotions in Kantian Morality, in: Flanagan & Rorty (Hgg.) (1990): *Identity, Character, and Morality*, S. 158-170.
- Sherman, Nancy (1993): The Role of Emotions in Aristotelian Virtue, in: *Proceedings of the Boston Area Colloquium in Ancient Philosophy* 9, S. 1-33.
- Sherman, Nancy (1994): The Role of Emotions in Aristotelian Virtue, in: *Proceedings of the Boston Area Colloquium in Ancient Philosophy* 9, S. 1-33.
- Sherman, Nancy (1997): *Making a Necessity of Virtue. Aristotle and Kant on Virtue*, Cambridge.
- Sherman, Nancy (1998): Empathy and Imagination, in: French & Wettstein (Hgg.) (1998), S. 82-119.
- Sherrington, C.S. (1900): Experiments on the value of vascular and visceral factors for the genesis of emotion, in: *Proceedings of the Royal Society of London* 66, S. 390-403.
- Sihvola, Juha & Engberg-Pedersen, Troels (Hgg.) (1998): *The Emotions in Hellenistic Philosophy*, Dordrecht u.a.
- Sihvola, Juha (1996): Emotional Animals: Do Aristotelian Emotions Require Beliefs?, in: *Apeiron* 29, 2, S. 105-144.
- Singer, J.A. & Salovey, P. (1988): Mood and memory: evaluating the network theory of affect, in: *Clinical Psychology Review* 8, S. 211-251.
- Slote, Michael (1990): Some Advantages of Virtue Ethics, in: Flanagan & Rorty (Hgg.) (1990): *Identity, Character and Morality*, S. 429-448.
- Slote, Michael (1992): *From Morality to Virtue*, New York/ Oxford: OUP.
- Slote, Michael (1997): Agent-Based Virtue Ethics, in: Crisp & Slote (Hgg.) (1997): *Virtue Ethics*, S. 239-262.
- Slote, Michael (2001): *Morals from Motives*, Oxford/ New York: OUP.
- Smith, Adam (1982): *The Theory of Moral Sentiments*, Indianapolis.
- Smith, C.A. & Lazarus, R.S. (1993): Appraisal Components, Core Relational Themes, and the Emotions, in: *Cognition and Emotion* 7, S. 233-269.

- Smith, Michael (1987): The Humean Theory of Motivation, in: *Mind* 96, S. 36-61 (wiederabgedr. in: Wallace (1999): Reason, Emotion and Will).
- Smith, Michael (1989): Dispositional Theories of Value, in: *Proceedings of the Aristotelian Society*, Supplementary Volume LXIII, S. 89-111.
- Smith, Michael (1992): Valuing: Desiring or Believing?, in: *Reduction, Explanation, and Realism*, hrsg. v. David Charles u. Kathleen Lennon, Oxford, S. 323-360.
- Smith, Michael (1994): *The moral problem*, Oxford: Blackwell.
- Smith, Michael (1994a): Why Expressivists about Value Should Love Minimalism about Truth, in: *Analysis* 54, S. 1-12.
- Smith, Michael (1995): Internal Reasons, in: *Philosophy and Phenomenological Research* 55, 1, S.109-131 (wiederabgedr. in: Wallace (1999): Reason, Emotion and Will).
- Smith, Michael (1995a): Internalism's wheel, in: *Ratio* 8, S. 277-302.
- Solomon, R.L. (1980): The opponent-process theory of acquired motivation: the cost of pleasure and the benefits of pain, in: *American Psychologist* 35, S. 691-712.
- Solomon, Robert C. (1973): Emotions and Choice, in: Solomon, Robert C. (2003a): *Not Passion's Slave*, S. 3-24 (zuerst in: *Review of Metaphysics* 28, 1; Nachwort zuerst in: Rorty, Amélie Oksenberg (Hg.) (1980): *Explaining Emotions*, Berkeley u.a.).
- Solomon, Robert C. (1976): *The Passions*, New York.
- Solomon, Robert C. (1984): Emotions' Mysterious Objects, in: Solomon, Robert C. (2003a): *Not Passion's Slave*, S. 57-75 (zuerst in: Myers, G.E & Irani, K.D (Hgg.): *Emotion: Philosophical Studies*, New York: Haven Books 1984).
- Solomon, Robert C. (1988): On Emotions as Judgments, in: Solomon, Robert C. (2003a): *Not Passion's Slave*, S. 92-114 (zuerst in: *American Philosophical Quarterly* 25, 2).
- Solomon, Robert C. (2003): Emotions, Thoughts and Feelings: What is a 'Cognitive Theory' of the Emotions and Does it Neglect Affectivity?, in: *Philosophy and the Emotions*, hrsg. v. Anthony Hatzimoysis (=Royal Institute of Philosophy supplement 52), S. 1-18.
- Solomon, Robert C. (2003a): *Not Passion's Slave. Emotions and Choice*, Oxford: OUP.
- Sorabji, Richard (2000): *Emotion and Peace of Mind. From Stoic Agitation to Christian Temptation: The Gifford Lectures*, Oxford: OUP.
- Specht, R. (1966): *Commercium mentis et corporis. Über Kausalvorstellungen im Cartesianismus*, Stuttgart/ Bad Cannstatt.
- Squire, L. (1987): *Memory and the brain*, New York: OUP.
- Sreenivasan, Gopal (2002): Errors about Errors: Virtue Theory and Trait Attribution, in: *Mind* 111, S. 47-68.

- Stampe, Dennis W. (1987): The Authority of Desire, in: *Philosophical Review* 96, S. 335-381 (wiederabgedr. in: Wallace (1999), S. 309-355).
- Stark, Susan (2001): Virtue and Emotion, in: *Noûs* 35, 3, S. 440-455.
- Statman, Daniel (1997): Introduction to Virtue Ethics, in: Statman, Daniel (Hg.) (1997): *Virtue Ethics*, S. 1-41.
- Statman, Daniel (Hg.) (1997): *Virtue Ethics*, Edinburgh: Edinburgh University Press.
- Steinfath, Holmer (2001): Orientierung am Guten, Frankfurt a.M.
- Steinfath, Holmer (2001a): Gefühle und Werte, in: *Zeitschrift für philosophische Forschung* 55, S. 196-220.
- Steinfath, Holmer (2002): Emotionen, Werte und Moral, in: *Die Moralität der Gefühle, Deutsche Zeitschrift für Philosophie, Sonderband 4*, hrsg. v. Sabine A. Döring u. Verena Mayer, Berlin, S. 105-122.
- Stephan, Achim & Walter, Henrik (Hgg.) (2003): *Natur und Theorie der Emotion*, Paderborn: mentis.
- Stevenson, Charles L. (1937): The Emotive Meaning of Ethical Terms, in: *Mind* 46, S. 14-31.
- Stevenson, Charles L. (1944): *Ethics and Language*, New Haven.
- Stocker, Michael (1976): The Schizophrenia of Modern Ethical Theories, in: *Journal of Philosophy* 14, S. 453-466.
- Stocker, Michael (1979): Desiring the Bad: An Essay in Moral Psychology, in: *Journal of Philosophy* 76, S. 738-753 (wiederabgedr. in: Wallace (1999), S. 269-284).
- Stocker, Michael (1987): Emotional Thoughts, in: *American Philosophical Quarterly* 24, S. 59-69.
- Stocker, Michael (1997): Emotional Identification, Closeness and Size: Some Contributions to Virtue Ethics, in: Statman (Hg.) (1997): *Virtue Ethics*, S. 118-127.
- Stocker, Michael (2002): Some Ways to Value Emotions, in: Goldie, Peter (Hg.): *Understanding Emotions. Mind and Morals*, Aldershot u.a.: Ashgate, S. 65-80.
- Stocker, Michael with E. Hegeman (1996): *Valuing Emotions*, Cambridge.
- Stoicorum Veterum Fragmenta (SVF), collegit I. ab Arnim, 4. Bde., Leipzig 1903-24, Nachdr. Stuttgart 1964.
- Stout, G.F. (1929): *A Manual of Psychology*, 4. Aufl., London: University Tutorial Press.
- Strack, F./ Martin, L.L./ Stepper, S. (1988): Inhibiting and facilitating conditions of facial expressions. A nonobstrusive test of the Facial Feedback Hypothesis, in: *Journal of Personality and Social Psychology* 54, S. 768-777.

- Strasser, Stephan (1981): Rezension zu Lyons (1980), in: *Philosophische Rundschau* 28, 247-251.
- Stroud, Barry (1977): *Hume*, London: Routledge.
- Stroud, Barry (1979): Inference, Belief, and Understanding, in: *Mind* 88, S. 179-196.
- Sturgeon, Nicholas (1985): Moral Explanations, in: David Copp & David Zimmerman (Hgg.): *Morality, Reason and Truth*, Rowman and Allanheld, S. 49-78.
- Sundararajan, Louise (2002): Shifting Paradigms in the Energy Theory of Emotions: Toward a Synthesis, in: *Ultimate Reality and Meaning* 25, 4, S. 295-306.
- Svavarsdóttir, S. (1999): Moral Cognitivism and Motivation, in: *Philosophical Review* 108, S. 161-219.
- Tappolet, Christine (1995): Are Emotions Mental or Physical?, in: *Revue de Théologie et de Philosophie* 127, 3, S. 251-259.
- Taylor, Charles (1976): Responsibility for Self, in: Rorty (Hg.) (1976), S. 281-299.
- Taylor, Charles (1989): *Sources of the Self: The Making of the Modern Identity*, Cambridge.
- Taylor, Gabriele (1985): *Pride, Shame and Guilt: Emotions of Self-Assessment*, Oxford: Clarendon Press.
- Tellenbach, Hubert (1968): *Geschmack und Atmosphäre*, Salzburg.
- Thalberg, Irving (1964): Emotion and Thought, in: *American Philosophical Quarterly* 1, S. 45-55.
- Thalberg, Irving (1973): Constituents and Causes of Emotion and Action, in: *Philosophical Quarterly* 23, S. 1-13.
- Thomas, Laurence (1988): Moral Motivation: Kantians versus Humeans (and Evolution), in: *Midwest Studies in Philosophy* 13, S. 367-383.
- Tilley, John J. (1997): Motivation and Practical Reason, in: *Erkenntnis* 47, S. 105-127.
- Tomkins, Silvan S. (1962): *Affect, imagery, consciousness*, Bd. 1: *The Positive Affects*, New York: Springer.
- Tomkins, Silvan S. (1963): *Affect, imagery, consciousness*, Bd. 2: *The Negative Affects*, New York: Springer.
- Tomkins, Silvan S. (1984): *Affect Theory*, in: K.R. Scherer & P. Ekman (Hg.), *Approaches to emotion*, Hillsdale (NJ): Laurence Erlbaum Associates, S. 163-195.
- Tomkins, Silvan S. (1991): *Affect, imagery, consciousness*, Bd. 3: *The Negative Affects. Anger and Fear*, New York: Springer.
- Tomkins, Silvan S. (1992): *Affect, imagery, consciousness*, Bd. 4: *Cognition*, New York: Springer.

- Trapp, Robert/ Petta, Paolo/ Payr, Sabine (Hgg.) (2002): *Emotions in Humans and Artifacts*, Cambridge (MA): MIT Press.
- Trepel, Martin (1999): *Neuroanatomie. Struktur und Funktion*, 2., überarbeitete Aufl., München u.a.: Urban & Fischer.
- Trianosky, Gregory V. (1990): What is Virtue Ethics all about?, in: Statman (Hg.) (1997): *Virtue Ethics*, S. 42-55 (zuerst in: *American Philosophical Quarterly* 27, S. 335-344).
- Trianosky, Gregory V. (1990a): *Natural Affections and Responsibility for Character: A Critique of Kantian views of the Virtues*, in: Flanagan & Rorty (Hgg.) (1990): *Identity, Character, and Morality*, S. 93-110.
- Tugendhat, Ernst (1993): *Vorlesungen über Ethik*, Frankfurt a.M.: Suhrkamp.
- Tzschentke, T.M. (2000): The medial prefrontal cortex as a part of the brain's reward system, in: *Amino Acids* 19, 1, S. 211-219.
- Ulich, D. (1982): *Das Gefühl. Einführung in die Emotionspsychologie*, München.
- Valins, Stuart (1966): Cognitive effects of false heart-rate feedback, in: *Journal of Personality and Social Psychology* 4, S. 400-408.
- Velleman, David J. (1989): *Practical Reflection*, Princeton: Princeton University Press.
- Velleman, David J. (1992): The Guise of the Good, in: *Noûs* 26, S. 3-26 (wiederabgedr. in: Wallace (1999): *Reason, Emotion and Will*).
- Velleman, David J. (1996): The Possibility of Practical Reason, in: *Ethics* 106, S. 694-726 (wiederabgedr. in: Wallace (1999): *Reason, Emotion and Will*).
- Velleman, David J. (1999): *The Possibility of Practical Reason*, Oxford: OUP.
- Walker, A.D.M. (1989): Virtue and Character, in: *Philosophy* 64, S. 349-362.
- Wallace, James D. (1988): Ethics and the Craft Analogy, in: *Midwest Studies in Philosophy* 13, S. 222-232.
- Wallace, James D. (1988): Ethics and the Craft Analogy, in: *Midwest Studies in Philosophy* 13, S. 222-232.
- Wallace, R. Jay (James) (1990): How to Argue about Practical Reason, in: *Mind* 99, S. 355-380 (wiederabgedr. in: Wallace (1999), S. 105-135).
- Wallace, R. Jay (James) (1994): *Responsibility and the Moral Sentiments*, Cambridge (MA): Harvard University Press.
- Wallace, R. Jay (James) (1999a): Three Conceptions of Rational Agency, in: *Ethical Theory and Moral Practice* 2, S. 217-242.
- Wallace, R. Jay (James) (2000): An Anti-Philosophy of the Emotions?, in: *Philosophy and Phenomenological Research* 60, 2, S. 469-477.

- Wallace, R. Jay (James) (Hg.) (1999): Reason, Emotion and Will, Aldershot u.a.: Ashgate–Dartmouth (= The International Research Library of Philosophy 23).
- Wallace, R. Jay/ Pettit, Philip/ Scheffler, Samuel/ Smith, Michael (Hgg.) (2004): Reason and Value. Themes from the Moral Philosophy of Joseph Raz, Oxford: Clarendon Press.
- Walton, Kendall (1993): Mimesis as Make-Believe: On the Foundations of the Representational Arts, Harvard UP.
- Watson, Gary (1975): Free Agency, in: ders. (Hg.): Free Will, Oxford: OUP 1982, S. 96-110.
- Watson, Gary (1984): Virtues in Excess, in: Philosophical Studies 46, S. 57-74.
- Watson, Gary (1990): On the Primacy of Character, in: Flanagan & Rorty (Hgg.) (1990): Identity, Character and Morality, S. 449-483.
- Watson, J.B. & Rayner, R. (1920): Conditioned emotional reactions, in: Journal of Experimental Psychology 3, S. 1-14.
- Wedgwood, Ralph (2002): Practical Reason and Desire, in: Australasian Journal of Philosophy 80, 3, S. 345-358.
- Weiskrantz, L. (1986): Blindsight: A case study and implications, Oxford: OUP.
- Wiggins, David (1987): Needs, Values, Truth. Essays in the Philosophy of Value, Oxford.
- Wiggins, David (1987a): Deliberation and Practical Reason, in: Wiggins (1987) (wiederabgedr. in: Millgram (2001), S. 279-300).
- Wiggins, David (1998): A Sensible Subjectivism?, in: Ders.: Needs, Values, Truth, 3. Aufl., Oxford, S. 185-214.
- Wiland, Eric (2002): Theories of Practical Reason, in: Metaphilosophy 33, 4, S. 450-467.
- Wiland, Eric (2003): Psychologism, Practical Reason and the Possibility of Error, in: The Philosophical Quarterly 53, 210, S. 68-78.
- Williams, Bernard (1973): Problems of the Self, Cambridge.
- Williams, Bernard (1976): Persons, Character, and Morality, in: Rorty (Hg.) (1976), S. 197-216.
- Williams, Bernard (1976a): Moral Luck, in: Williams (1981), S. 20-39 (zuerst in: Proceedings of the Aristotelian Society, Supplementary Volume 50, S. 115-136).
- Williams, Bernard (1980): Internal and External Reasons, in: Ders.: Moral Luck, Cambridge 1981, S. 101-113.
- Williams, Bernard (1981): Moral Luck, Cambridge.
- Williams, Bernard (1981a): Philosophy, in: Finley, M. (Hg.) (1981): The Legacy of Greece, Oxford: OUP, S. 202-255.
- Williams, Bernard (1985): Ethics and the Limits of Philosophy, Cambridge (Mass.).

- Williams, Bernard (1993): *Shame and Necessity*, Berkeley.
- Williams, Bernard (1995): *Acting as the Virtuous Person Acts*, in: Heinamen, R. (Hg.) (1995): *Aristotle and Moral Realism*, London: UCL Press, S. 13-23.
- Williams, Bernard (1995a): *Internal Reasons and the Obscurity of Blame*, in: *Making Sense of Humanity*, Cambridge: CUP, S. 5-45.
- Wolf, Ursula (1984): *Das Problem des moralischen Sollens*, Berlin/ New York: de Gruyter.
- Wolf, Ursula (1988): *Haben wir moralische Verpflichtungen gegen Tiere?*, in: Angelika Krebs (Hg.): *Naturethik*, Frankfurt a.M. 1997, 47-75.
- Wolf, Ursula (1993): *Gefühle im Leben und in der Philosophie*, in: Fink-Eitel & Lohmann (1993), 112-135.
- Wollheim, Richard (1999): *On the Emotions – The Ernst Cassirer Lectures 1991*, New Haven/ New York.
- Wollheim, Richard (2003): *The Emotions and their Philosophy of Mind*, in: *Philosophy and the Emotions*, hrsg. v. Anthony Hatzimoysis (=Royal Institute of Philosophy supplement 52), S. 19-38.
- Woods, D.J. (1977): *The repression-sensitization variable and self-reported emotional arousal: effects of stress and instructional set*, in: *Journal of Consulting and Clinical Psychology* 45, S. 173-183.
- Woods, Michael (1972): *Reasons for Action and Desire*, in: *Proceedings of the Aristotelian Society Supplementary Volume* 46, S. 189-201.
- Worcester, W.L. (1883): *Observations on some points in James's Psychology. II. Emotion*, in: *The Monist* 3, S. 285-298.
- Wörner, Markus H. (1981): *'Pathos' als Überzeugungsmittel in der Rhetorik des Aristoteles*, in: Ingrid Craemer-Ruegenberg (Hg.): *Pathos, Affekt, Gefühl*, Freiburg i.Br./ München 1981, 53-78.
- Wright, Crispin (1988): *Moral Values, Projection and Secondary Qualities*, in: *Proceedings of the Aristotelian Society, Supplementary Volume* 42, S. 1-26.
- Wright, G.H. von (1963): *The Varieties of Goodness*, London.
- Wundt, Wilhelm (1909): *Grundriß der Psychologie*, 9., verbesserte Aufl., Leipzig (zuerst 1896).
- Zajonc, Robert B. (1968): *Attitudinal effects of mere exposure*, in: *Journal of Personality and Social Psychology Monograph Supplement* 9, S. 1-27.
- Zajonc, Robert B. (1980): *Feeling and thinking: Preferences need no inferences*, in: *American Psychologist* 35, S. 151-175.

- Zajonc, Robert B. (1984): On the Primacy of Affect, in: *American Psychologist* 39, S. 117-123.
- Zajonc, Robert B./ Murphy, S.T./ Inglehart, M. (1989): Feeling and Facial Reference: Implications of the Vascular Theory of Emotion, in: *Psychological Review* 96, S. 395-416.
- Zemach, Eddy M. (2001): What is Emotion?, in: *American Philosophical Quarterly* 38, 2, S. 197-207.
- Zillmann, D. (1978): Attribution and misattribution of excitatory reactions, in: J.H. Harvey/ W. Ickes/ R.F. Kidd (Hgg.): *New directions in attribution theory*, Hillsdale (NJ): Lawrence Erlbaum Associates, Bd. 2, S. 335-368.
- Zimbardo, Philip G. (1973): On the Ethics of Intervention in Human Psychological Research. With Special Reference to the Stanford Prison Experiment, in: *Cognition* 2, S. 243-256.
- Zimmerman, David (2001): Thinking with Your Hypothalamus: Reflections on a Cognitive Role for the Reactive Emotions, in: *Philosophy and Phenomenological Research*, 63, 3, S. 521-541.