

e-Journal Philosophie der Psychologie	<b>MORALISCHE GEFÜHLE IN  SCHWARZ-WEISS UND FARBE</b> von Ronald de Sousa (Toronto)
---	--

### Gefühle in Schwarz-Weiß

Wenn Gefühle nur als Motive herhalten sollen, dann liegt die Vorstellung nahe, dass sie lediglich ziehen und stoßen: In irgendeinem Moment bewegt uns die Gesamtheit unserer Gefühle, und dann handeln wir – oder eben nicht. Unser Erleben der Welt macht dieses Bild komplizierter, aber letzten Endes müssen unsere Gefühle doch in eine Abfolge von Entscheidungen geordnet werden, von denen uns jede einzelne zum Handeln oder Nicht-Handeln nötigt und uns das Schwarz-Weiß von 'Ja' und 'Nein' aufdrängt. Wenn wir uns hingegen auf das Erleben von Gefühlen konzentrieren, dann erweisen sie sich als so unterschiedlich, dass sie nicht einmal eine natürliche Art bilden. Jedes einzelne Gefühl trägt eine Fülle spezifischer Bedeutungen und ist um eine unüberschaubare Menge von Kontrasten angereichert: das ist das vielfarbige Bild. Die monochrome Perspektive ergibt sich folglich im Kontext von Handlung und Entscheidung; das farbige Bild von Gefühlen begreift diese stattdessen als Weisen multi-dimensionaler *axiologischer Wahrnehmung*.

Wenn wir einmal das farbige Bild aufgreifen, dann lässt sich wohl zeigen, dass wir beliebig viele verschiedene Gefühle erfahren können, gerade so, wie wir beliebig viele verschiedene Gedanken denken können.<sup>1</sup> Im Lichte unserer durch ästhetische Erfahrungen hervorgerufenen Gefühle erscheint dies plausibel.

Wenn wir z.B. eine Tanzaufführung betrachten oder ein Musikstück anhören, dann spielen Gefühle eine Rolle; allerdings droht diese Beobachtung auf zwei ungleiche und gegensätzliche Weisen ins Triviale abzugleiten. Auf der einen Seite mag man versucht sein zu vermuten, Tanz und Musik drückten die großen, bei Bedarf leicht aufzählbaren Gefühle aus: Wut, Furcht, Liebe, Ehrfurcht, Eifersucht, Traurigkeit, Verlangen. Aber warum machen wir uns dann soviel Mühe? Wenn Kunstwerke nur dazu da sind, diese Gefühle in uns wachzurufen und wenn es keine bedeutsamen Unterschiede zwischen zwei Vorkommnissen von 'Furcht' oder 'Wut' usw. gibt, dann ist doch kaum wahrscheinlich, dass Kunstwerke in all ihrer Verschiedenheit deswegen unser Interesse finden, weil sie Gefühle darstellen – und nicht aus irgendwelchen anderen Gründen.<sup>2</sup> So betrachtet ist nur schwer begreiflich, warum die meisten Kunstwerke nicht überflüssig sind. (Viele mittelmäßige Kunstwerke sind tatsächlich in genau diesem Sinne überflüssig: Wenn man einen Film aus der Sparte 'Invasion aus dem Weltall' gesehen hat, hat man alle gesehen.) Alternativ könnte man die Ansicht vertreten, jeder Moment eines Tanzes oder eines Musikstücks (oder jedes anderen Kunstwerks, von dem man annimmt, es rufe Gefühle hervor) drücke tatsächlich ein Gefühl *sui generis* – oder genauer: *sui ipsius*, also ein unwiederholbar individuelles – aus. Damit wird die Rolle der Kunst aber wieder trivial: Wenn jedes Kunstwerk notwendigerweise nicht mehr und nicht

<sup>1</sup> Für diese Ansicht hat (Campbell 1998) gute Gründe vorgebracht. Ich selbst habe sie im letzten Kapitel von (de Sousa 1987) entwickelt.

<sup>2</sup> Manfred Clynes 'sentic' bietet ein Schema zur Klassifikation und Erzeugung von Gefühlen auf der Grundlage von Berührungsrhythmen. Clynes Theorie lässt ein subtileres Repertoire an Gefühlen zu, da er glaubt, Gefühle seien irgendwie durch verschiedene physische Rhythmen erzeugt und Variationen in den Klassen der Rhythmen, die für ein Gefühl charakteristisch sind, erzeugten ein deutlich anderes Gefühl. Allerdings bleibt seine Theorie eine Art des Gefühls-Atomismus, in dem Gefühle ein begrenztes Repertoire umfassen (Clynes 1977).

weniger als das ihm entsprechende Gefühl zum Ausdruck bringt, dann wird anscheinend die Möglichkeit in Abrede gestellt, dass einige Kunstwerke Gefühle besser ausdrücken können als andere, oder dass die ausgedrückten Gefühle wertvoller, interessanter, tiefer, authentischer sind als andere. Immerhin scheinen alle diese Bewertungen vorauszusetzen, dass es etwas über den Ausdruck von Gefühl hinaus gibt, das diese Urteile ermöglicht.

Der letzte Einwand lässt sich zurückweisen, wenn wir die Annahme fallen lassen, jedes Kunstwerk oder literarische Produkt vermittele eine vorgefertigte Gefühlsbotschaft. Wenn wir stattdessen annehmen, Kunst erzeuge und verkörpere ein Gefühl ganz eigener Art, dann kann es zahllose Gefühle geben und jedes einzelne Kunstwerk erweist sich als mehr oder weniger interessant, je nachdem, wie die Qualität des von ihm vermittelten einzigartigen Gefühls ausfällt. Dann gibt es im Wortsinn *unüberschaubar viele* Gefühle.

Jetzt ergeben sich aber Probleme ganz anderer Art.

Wenn es tatsächlich unüberschaubar viele Gefühle gibt, wie lässt sich dann ihr Einfluss auf unsere Entscheidungen verstehen? Diese Frage setzt mit Aristoteles voraus, moralische Erziehung bestehe zu einem großen Teil darin zu lernen, das Richtige auf die richtige Weise zur richtigen Zeit zu fühlen (Aristoteles, NE, 1104b13). Im Zuge unserer moralischen Erziehung lernen wir irgendwie, bestimmte Gefühle zu haben; im weiteren Verlauf unseres Lebens (a) ermöglichen uns diese Gefühle, die Dinge korrekt *zu erfassen und zu verstehen*, und (b) sie *bewegen* uns vermittels dieses Verstehens dazu, richtig zu handeln. Wie aber müssen wir Gefühle verstehen, wenn sie diese vermittelnde Funktion innehaben sollen?

Eine vielversprechende Antwort konzentriert sich auf die Bedeutung von Literatur für unsere moralische Erziehung. Martha Nussbaum hat hierfür überzeugend argumentiert und dabei betont, dass uns Literatur im Gegensatz zu den dünnen Abstraktionen philosophischer Beispiele 'erschöpfend ausgemalte' Gefühle und einzelne Charaktere vor Augen führen kann (Nussbaum 1990, *passim*). Diese Idee bringt jedoch zwei Probleme mit sich. Zum einen präsentiert keine noch so detailreiche und lebendige Dichtung tatsächlich Einzelpersonen. Konkrete Einzelpersonen haben unendlich viele relationale Eigenschaften. Dagegen erben die Protagonisten der Dichtung die intrinsische Allgemeinheit der Wörter, mit denen sie vermittelt werden; und ihre bestimmenden Eigenschaften sind notwendigerweise endlich. Sie können unendlich *spezifisch* sein, aber niemals *partikulär*. Wenn wir uns zum anderen auf die einzigartigen Merkmale jeder Situation, auf die spezifische Qualität jedes Individuums, und auf die Einmaligkeit eines jeden Gefühls konzentrieren, dann ist schwer verständlich, wie sie uns überhaupt leiten können. Lernen setzt Vergleiche voraus, die wiederum Ähnlichkeiten und Unterschiede zur Grundlage haben, und diese werden durch Ausdrücke klassifiziert und auf den Begriff gebracht, die uns notwendigerweise wieder vom Einzelnen zum Allgemeinen zurückführen. Kurz gesagt: Lernen ist auf wiederholbare Muster angewiesen. Wörtlich verstanden entzieht sich so der Vorschlag, jede einzelne Situation des Lebens (und jede Episode eines Kunstwerks) erzeuge ein unwiederholbares, einzigartiges Gefühl, selbst die Grundlage. (Einzelne) Perzepte ohne (allgemeine) Begriffe können unser Wahlverhalten nicht leiten.

Dieses Problem stellt sich nicht nur für das Lernen eines Individuums, sondern auch für die Evolution unserer emotionalen Fähigkeiten durch natürliche Selektion. Obwohl einige Philosophen und Soziologen daran interessiert sind, den Beitrag der Biologie für unser Selbstverständnis herunterzuspielen, wird kaum jemand bestreiten, dass die natürliche Selektion beim Festlegen unseres Potenzials für Gefühle ihre Hand im Spiel hatte. Die gelegentlich als 'basal' bezeichneten Gefühle sind offenkundig für Lebensfunktionen relevant, und plausiblerweise sind sie in einigen

Spezies durch Instinkte wie die manchmal als 'vier Fs'<sup>3</sup> bezeichneten realisiert. Diese hätten nicht selektiert werden können, wenn sie nicht wiederholbare Muster mit wiederholbaren kausalen Eigenschaften erzeugt hätten.

Damit ist klar, dass die Rolle der Gefühle – nämlich: dass sie Handlungen auf der Grundlage sehr spezifischer Erfahrungen beeinflussen – auf wiederholbaren Merkmalen dieser Erfahrungen beruhen muss. Früher habe ich vorgeschlagen, dass zusätzlich zu dem halben Dutzend Mustern, die üblicherweise von Emotionentheoretikern aufgezählt werden – man denke z.B. an die von Ekman gefundenen Muster, die auf Grund ihres Ausdrucks mehr oder weniger allgemein erkennbar sind – die gemeinsamen Merkmale komplexerer Gefühle in *Schlüsselszenarien* eingebettet sind, die nicht nur Typen von Gefühlen festlegen, sondern auch ihre charakteristische Empfindungsqualität und die Bedingungen ihrer Angemessenheit. Die definierenden Wurzeln eines jeden unterscheidbaren Gefühls liegen in einer spezifischen dramatischen Situation oder einem Episodentyp, die mit einer charakteristischen Empfindung verbunden ist (im weiten Sinne, nicht im engen Sinne einer körperlichen Empfindung). Das Drama wird von Akteuren in verschiedenen *Rollen* umgesetzt, und meine ursprüngliche Hypothese war, dass diese von Person zu Person variierenden Schlüsselszenarien für uns die Bedeutungen derjenigen Gefühle festlegen, deren Prototypen sie sind. Ob ein solches Gefühl, später im Leben erfahren, der Situation angemessen ist, entscheidet sich zumindest teilweise an der Beziehung zwischen dem ursprünglichen Paradigma, welches das *formale Objekt* des Gefühls bestimmt, und der gegenwärtigen Situation. Dabei legt das formale Objekt eines Gefühls den Standard der Angemessenheit fest, den Gefühle erfüllen müssen, um Gefühle dieser Art zu sein.

Gerade so, wie sich Geschichten mehr oder weniger detailreich erzählen lassen, kann man Gefühle als unterschiedlich feinkörnig Bestimmbare und Bestimmte zusammenfassen. Für diese Ansicht stellt es kein Problem dar, dass sehr unterschiedliche Erlebnisse jeweils als Furcht oder Wut gelten können. Obwohl man die Geschichte der *Antigone* in einer Zeile wiedergeben kann (*Prinzessin gerät in einen Streit mit ihrem Onkel und geht mit ihrem Verlobten in den Tod.*), lässt sich durch diese Zusammenfassung die Wirkung des Stückes nicht erklären. Ganz ähnlich lassen sich die von uns erlebten bestimmten Gefühle nicht angemessen beschreiben, wenn man die Standardausdrücke für bestimmbare Gefühle verwendet.

Dies gibt eine Idee an die Hand, wie sich die Vielfalt der von uns erlebten Gefühle reduzieren lässt; immerhin sind die meisten unserer Gefühle entscheidend *repetitiv* und zumindest teilweise anhand eines endlichen Repertoires von Rollen in Klassen eingeteilt. Allerdings reicht dies nicht dazu hin, das von mir aufgeworfene Problem zu lösen: Wie können Gefühle unser moralisches Lernen vermitteln?

Zur Beantwortung dieser Frage schlage ich vor, auf unser Schwarz-Weiß-Bild von Gefühlen zurückgreifen: anstatt Gefühle in ihrer Vielfalt zu betrachten, sollten wir sie auf die einfache Dimension von Aversion/Attraktion abbilden. Genau wie bei der Rückführung eines Farbbildes auf eine Schwarz-Weiß-Zeichnung kann dabei der Reichtum im Detail verloren gehen, aber das Bild mag an Klarheit gewinnen. Die einfachste Schwarz-Weiß-Konzeption von Gefühlen ist der oft belächelte Psychologische Hedonismus, den ich in seiner Variante als *Unmittelbaren Hedonismus* wiederbeleben will. Wir werden später sehen, welche zusätzliche Anforderungen eine Ausweitung hin zur farbigen Perspektive der Axiologie nötig machen.

---

<sup>3</sup> Dies sind natürlich die vier Funktionen, deren Bezeichnungen mit 'F' beginnen: Entkommen, Kämpfen, Nahrungsaufnahme und Fortpflanzung.

### Der Unmittelbare Hedonismus

Der Doktrin des Psychologischen Hedonismus zufolge streben wir mit allen unseren Handlungen allein danach, Lust zu gewinnen (und Unlust zu vermeiden).<sup>4</sup> Natürlich darf man diese Doktrin nicht mit der des *Ethischen Hedonismus* verwechseln, die besagt, dass wir nach Lust streben und Unlust vermeiden sollen. Der herkömmlichen Ansicht nach sind diese beiden Konzeptionen miteinander unverträglich. Denn wenn wir nicht vermeiden können, den Neigungen unserer Lust zu folgen, macht es wenig Sinn, uns genau dies vorzuschreiben. Allerdings finden sich in jeder der Doktrinen Mehrdeutigkeiten, deren Auflösung es womöglich gestattet, die beiden Ideen miteinander zu kombinieren.

Die erste Mehrdeutigkeit ist die folgende. Der Psychologische Hedonismus besagt, letztlich treibe allein Lust unsere Handlungen an: Aber wessen Lust? Es wird vorausgesetzt, dass wir meinen: meine eigene, die des jeweiligen Akteurs. Der Ethische Hedonismus kann sich auf ein breiteres Verständnis des Wertes von Lust berufen und die utilitaristische Idee aufgreifen, dass wir die Lust der Mehrzahl maximieren sollen. Diese Idee mag tatsächlich mit dem alleinigen Streben nach meiner Lust unvereinbar sein – und so dem Akteur ein Ideal vorgeben, dem er seinen Anlagen gemäß nicht gut entsprechen kann. Aber wenn jemand seinen Anlagen gemäß einem Ideal nicht gut entsprechen kann, heißt das nicht, dass er die Vorgaben des Ideals trivialerweise erfüllt. Die zweite Mehrdeutigkeit betrifft allein den Psychologischen Hedonismus. Geht es darum, meine Lust über die Zeit hinweg zu maximieren, oder streben wir unweigerlich in jedem Augenblick den einfachst möglichen Lustgewinn an? Letzteres ist dem 'Streben nach der gegenwärtigen Lust' ähnlich, das Aristoteles dem 'Akolastos' oder 'Hemmungslosen' zuschreibt. Dies eröffnet die Möglichkeit eines Ethischen Hedonismus, der sich um andere Akteure gar nicht kümmert, sondern lediglich eine umfassendere Perspektive auf sich selbst propagiert.

Einer der Einwände gegen den Psychologischen Hedonismus stützt sich auf das Gedankenexperiment der *Erlebnismaschine* (Nozick 1974, 42-45). Man bittet die Versuchspersonen, sich eine Wahlsituation vorzustellen: Sie können entweder ihr bisheriges Leben mit allen seinen Unwägbarkeiten und Risiken weiterführen, oder sie können sich an eine Maschine anschließen lassen, die ihnen für den Rest ihres Lebens uneingeschränkt Lust vermittelt oder ihnen, wenn das zu langweilig klingt, genau die richtige Mischung garantiert aus Antizipation, Frustration und Erfolg, genau die perfekte Mischung aus Momenten der Hoffnung, Erwartung, und letztendlicher Erfüllung. Wie sich herausstellt, lehnen die meisten Menschen die Maschine ab. Wenn wir einmal die langweilige Hypothese außer Acht lassen, unsere Versuchspersonen könnten sich

---

<sup>4</sup> Der Dualismus in dieser Zielvorstellung erzeugt zuweilen eine unerwartete Schwierigkeit, insbesondere wenn man bedenkt, dass gerade die Einfachheit des Erklärungsschemas für ihn spricht. In der Erkenntnistheorie ist die Situation ganz ähnlich: Überzeugung zielt sowohl auf Wahrheit als auch auf das Vermeiden von Falschheit; die beiden Ziele, die keineswegs zwei Aspekte eines einzigen Ziels sind, können aber in der Praxis in Konflikt geraten. 'Glaube alles' wäre eine Politik, die das erste Prinzip auf Kosten des zweiten erfüllt; 'Glaube nichts' würde den Erfolg des zweiten Prinzips unter Aufgabe des ersten garantieren. Ganz analog hat (Nussbaum 1994) gezeigt, dass das stoische Ideal der *ataraxia* auf der plausiblen Annahme beruht, es sei psychologisch unmöglich, nur freundliche Gefühle zu behalten und die üblen aufzugeben; aber das rechtfertigt für sich keineswegs, das vollständige Fehlen von Gefühlen einer Gefühlslage vorzuziehen, in der Lust die Unlust überwiegt. Wenn jede Zuneigung und jedes Verlangen das Risiko quälender und destruktiver negativer Gefühle in sich birgt – so wie Liebe Eifersucht mit sich bringt, oder wie die Liebe zur Gerechtigkeit die Neigung zu destruktivem Zorn nährt – mag sie es doch wert sein: und wir könnten Senecas Ideal sehr wohl Yeats' Motiv entgegenhalten: "Was macht es, es gibt also nichts als Feuer – in dir, in mir."

schlicht nicht dazu bringen, das geschilderte Versprechen zu glauben, stellt dies ein Problem für den Psychologischen Hedonismus dar. Wie Sober und Wilson hervorgehoben haben, ist dieses Argument jedoch gegen einen Psychologischen Hedonismus wirkungslos, demzufolge wir unsere Handlungen stets am unmittelbaren Lustgewinn anstatt an Berechnungen über den maximalen zukünftigen Lustgewinn ausrichten (Sober und Wilson 1998, 275-291). Ich werde diese Variante des Psychologischen Hedonismus den *Unmittelbaren Hedonismus* nennen.

Der Unmittelbare Hedonismus schreibt der unmittelbaren Lust – und d.h. insbesondere: dem Erlebnis des Wählens im Gegensatz zu den künftigen Konsequenzen der Wahl – die entscheidende kausale Kraft zu. Er ist leicht mit dem Ergebnis des Gedankenexperiments vereinbar. Dazu muss er lediglich voraussetzen, die Entscheidung gegen die Maschine sei als Handlung lustvoller als die Entscheidung für sie, um wieviel lustvoller die Konsequenzen der alternativen Entscheidung auch erscheinen mögen.

Ein Vorteil dieser Hypothese ist, dass sie Wahlentscheidungen zulässt, deren Konsequenzen weniger Lust erzeugen als die Konsequenzen der alternativen Wahl. Immerhin folgt aus dem folgenden Prinzip kein Widerspruch:

(1) *A* wird *B* vorgezogen, aber *B* zu wählen wird *A* zu wählen vorgezogen.

Auf diese Weise kann der Unmittelbare Hedonismus dem überzeugendsten Argument gegen die üblichen Varianten des Psychologischen Hedonismus entgehen.

Das heißt nicht, dass sich keine Einwände gegen diese Doktrin finden lassen. Man kann ihr immer noch vorwerfen, trivial zu sein. Bevor wir allerdings der Versuchung erliegen, den Unmittelbaren Hedonismus aus diesen Gründen zurückzuweisen, beachte man, dass (1) zwar konsistent sein mag, aber trotzdem *irrational* ist.

### Das Philebus Prinzip

Das Prinzip (1) ist irrational, weil es ein plausibles Rationalitätsprinzip verletzt, das man das *Philebus Prinzip* nennen könnte.

In seinem *Philebus* unternimmt Platon eine umfassende Untersuchung der Lust, wobei er die Möglichkeit erwägt, Lust als eine Art der Wahrnehmung, und nicht bloß als eine Weise der Empfindung aufzufassen. Dies ist eine wichtige Unterscheidung, da Wahrnehmung im Gegensatz zu Empfindung vorgibt, Informationen über etwas außerhalb unserer selbst zu vermitteln. Insbesondere kann die Lust, die an die Erwartung einer zukünftigen Lust geknüpft ist, Informationen über das Ausmaß der zukünftigen Lust liefern. Dieser Umstand rechtfertigt wohl nicht, die an die Erwartung geknüpfte Lust mit Platon als im Wortsinne wahr oder falsch anzusehen. Aber er stützt das folgende Rationalitätsprinzip:

(PP) Die Intensität der an eine Erwartung geknüpften Lust sollte zu derjenigen erwarteten Lust proportional sein, die sie abbildet.

Biologisch betrachtet ist dies ein vernünftiges Prinzip. Dem Unmittelbaren Hedonismus zufolge wird die Intensität der an die Erwartung geknüpften Lust darüber bestimmen, wie beharrlich meine gegenwärtige Motivation ist, die erwartete Lust zu verfolgen. Wenn die Intensität meiner gegenwärtigen Lust jedoch nicht dem Maß zukünftiger Lust entspricht, die sie mich anzustreben motiviert, dann steht meine gegenwärtige Motivation in keinem fixen Verhältnis zu meiner zukünftigen Befriedigung. Unser Leben wäre immer wieder wie das von jemandem, der für einen bestimmten Ort und eine bestimmte Zeit ein Picknick plant und sorgsam auf Wettervorhersagen für beliebige Orte und Zeiten achtet.

## Gefühle und Lernen

Wir können nun einige der Arten und Weisen unterscheiden, wie Gefühle, projiziert in die 'einfarbige' Dimension von Lust und Unlust, unseren Fortschritt in verschiedenen Bereichen motivieren können. *De facto* tun sie dies auf drei verschiedene Weisen.

Aristoteles Beobachtung, jeder Mensch strebe von Natur aus nach Wissen (Met 980a21), ist offenkundig eine korrekte empirische Beobachtung unseres Lebens von Kindesbeinen an. Die Ausübung aller möglichen Fähigkeiten geht mit Lust einher; Gleiches gilt für die Erweiterung unserer Fähigkeiten. Alle Kinder zeigen diese Art von Freude am Lernen – zumindest bis die Schule ihre Neugier beschneidet und ihre Freude in Ablehnung umschlagen lässt. Evolutionär betrachtet ist die Lust am Lernen nicht weniger erstaunlich als die Lust am Sex.

Dieses Muster ermöglicht uns, zweitens, das Phänomen der *Akrasia* auf den Begriff zu bringen: *Akrasia* lässt sich begreifen als die Wahl von *A* anstelle von *B* zu einem Moment, an dem man einräumt, dass *B* der Alternative *A* vorzuziehen ist. Wie ich erwähnt habe, kann dies nach dem oben skizzierten Muster begriffen werden: Im Moment der Wahl sehe ich ein, dass die Aussicht auf *B* attraktiver ist als die auf *A* – in unserem beschränkten, zweidimensionalen Schwarz-Weiß-Schema gesprochen: *B* bereitet mehr Lust als *A*. Aber aufgrund eines, wie Platon es nennen könnte, falschen Lustgefühls der Erwartung bereitet mir die Handlung, *A* zu wählen, mehr Lust als die Handlung der Wahl von *B*.

Wenn meine bisherigen Ausführungen richtig sind, dann findet sich hier ein Muster, das zugleich sowohl möglich als auch irrational ist. Aber Irrationalität weckt Hoffnung auf Rationalität. Positiv an *Akrasia* ist, dass sie Zeugnis ablegt für die Möglichkeit von moralischem und emotionalem Fortschritt. Fortschritt setzt Konflikt voraus: Er setzt voraus, dass es sowohl in Fragen moralischen Handelns als auch in Fragen erlebter Gefühle möglich ist, Fehler zu machen und diese zu bedauern. Soweit sich diese Fehler in Zukunft vermeiden lassen, künden sie von Verbesserung. (Völlig zu Recht bemerkt Aristoteles, dass der 'Akolastos', der 'nach der gegenwärtigen Lust strebt' und das Falsche tut ohne innere Konflikte zu erleben, nichts bedauert und folglich 'unheilbar' ist (EN 1150b29). Er ist sozusagen ein glücklicher moralischer Trottel ohne Aussicht auf Besserung.)

In gewissem Maße erklärt unser beschränktes Schwarz-Weiß-Schema des Unmittelbaren Hedonismus, wie Gefühle bestimmte Weisen des Lernens erleichtern. Daneben gibt es einen direkteren, rein physiologischen Weg, der sich zu den verschiedenen Theorien von Motivation neutral verhält, weil seine Wirkung rein kausal und weniger motivational ist. Die rein physiologische Stimulation des limbischen Systems reicht aus, um gegenwärtige Ereignisse im Gedächtnis festzuhalten (Rosenfield 1992, 165). So weiß jeder Amerikaner, wo sie oder er die Nachricht von der Ermordung Kennedys erhalten hat. (Mordanschläge sind inzwischen eine Alltäglichkeit, sie sind nicht länger unvergesslich.)

Das reduzierte Schwarz-Weiß-Bild von Motivation als Lust lässt zu, dass aufgrund von instrumentellen Überlegungen eine ganze Reihe von Gefühlen unser Verhalten beeinflussen. Auf diese Weise können Nachahmung, Bewunderung und Stolz auf unser Verhalten wirken, Gleiches gilt für eine Vielzahl von Gefühlen, die wir ansonsten ungern zugeben und die man wohl gut und gerne für verzichtbar halten mag: Gier, Neid, Eitelkeit, Eifersucht, Erniedrigung und Furcht können uns zu Anstrengungen antreiben, die den Umfang unserer Fähigkeiten zu erweitern im Stande sind. Auf diese Weise bewegen sie uns, Dinge zu lernen, bei denen der Prozess des Lernens selbst nicht die reine Freude ist. Von diesen gibt es leider viele, angefangen mit der Mühsal des Lesenlernens, das jedem Kind und jedem Elternteil hin und wieder als eine Art der Folter erschienen sein muss, wobei gerade die am meisten um das Wohlergehen ihres Kindes besorgten Eltern am wenigsten

bereit sind, es ihrem Kind zu ersparen. Dass diese Gefühle solche wertvollen Ziele befördern, ist allerdings zufällig; nichts prädestiniert sie dazu, gerade dem kognitiven, emotionalen oder moralischen Fortschritt zu dienen.

### Moralische Gefühle

Von einigen Gefühlen nimmt man jedoch traditioneller Weise genau dies an, weswegen man versucht sein mag, sie als intrinsisch moralisch zu bezeichnen. Die wichtigsten dieser sind Güte (*benevolence*) und Mitleid (*compassion*), aber ebenso gehören Schuld, Scham und Entrüstung auf die Liste. Insbesondere Letzteres ist von Francois Itard als das entscheidende moralische Gefühl angesehen worden. Itard, ein Arzt des 18. Jahrhunderts, berichtete über das in freier Natur aufgewachsene Kind, das er im Wald von Aveyron gefunden hatte. Er entwarf ein Experiment, mit dem er herausfinden wollte, ob das Kind einen Gerechtigkeitsinn erworben hatte. Das Experiment bestand darin, das Kind für korrekte Handlungen zu 'bestrafen', und zu Itards Befriedigung reagierte das Kind mit Zorn und Entrüstung, während es Bestrafungen üblicherweise einfach hinnahm (Itard 1964, Abschnitt 54).

Der interessanteste Aspekt des Experiments ist, dass Itards Interpretation des Verhaltens voraussetzt, er könne den Ausdruck von Entrüstung oder Zorn klar von dem der Scham, Traurigkeit, oder Frustration unterscheiden – immerhin hat das Naturkind nie zu Sprechen gelernt. Die Interpretation setzt mit anderen Worten voraus, dass diese Gefühle und ihr Ausdruck natürlich und vorsprachlich sind. Diese Annahme Itards war vermutlich korrekt, wenn man sich an den späteren Beobachtungen Darwins oder den neueren Forschungen Paul Ekman orientiert, denen zufolge es für ein gutes halbes Dutzend Gefühle, zu denen auch Zorn und Angst gehören, klar über Kulturschranken hinweg identifizierbare Gesichtsausdrücke gibt. Es lässt sich allerdings schwer sagen, was sich von der Einstellung eines Kindes zu Bestrafung auf andere Aspekte von Gerechtigkeit übertragen lässt. Das Fehlen von Sprache macht es in diesem Fall besonders schwer zu ermitteln, ob das Kind auch angesichts von Ungerechtigkeiten gegen andere mit Entrüstung reagiert hätte, oder ob diese sich gegen es selbst gerichtet hätte, wenn das Kind bei einer gerechten Verteilung etwas zu verlieren gehabt hätte. Es erscheint allerdings wahrscheinlich, dass sich, so Sprache zur Verfügung steht, aus solchen 'natürlichen' Gefühlsregungen ein allgemeinerer Gerechtigkeitsinn entwickeln kann, der den Test des Altruismus besteht.

Ursprünglich ist 'altruistisch' ein Charakteristikum von Motivation und es hat einen fruchtlosen Streit darüber gegeben, ob Menschen tatsächlich jemals altruistisch handeln oder ob sie 'letztendlich' anderen nur helfen, um daraus Befriedigung zu ziehen. Der Streit war fruchtlos, weil ich natürlich auch dann aus dem Erlangen eines Gutes durch eine andere Person Lust schöpfen kann, wenn ich aufrichtig nur ihr Bestes will. Der Unmittelbare Hedonismus, der bislang den Rahmen meiner Überlegungen abgegeben hat, ist offenkundig darauf festgelegt, dass die Lust, die der Akteur bei seiner Handlung verspürt, die unmittelbar motivierende Ursache seiner wohlwollenden Handlung ist, aber dies ist vollständig damit vereinbar, dass der Handelnde aufrichtig um sein Gegenüber besorgt ist. Aus diesem Grund erscheint die Debatte wenig interessant.

Altruismus ist jüngst in der Biologie präzise als eine Art von Verhalten definiert worden, das einen Netto-Verlust an Fitness für den Handelnden in Kombination mit einem Fitness-Gewinn für andere erzeugt. So verstanden bringt es ein Paradox für den Darwinismus mit sich und wirft gewichtige Fragen auf. Seinem eigenen Anspruch nach (kurz gesagt: jede erbliche Tendenz, seine eigene Fitness zu verringern, muss ihr eigenes Verschwinden garantieren) muss der Darwinismus

Mechanismen finden, die dazu geeignet sind, die altruistischen Handlungen von Organismen hervorzubringen. Damit wird das Spektrum von *kin selection*, *reciprocal altruism* und *group selection* verständlich, *reciprocal altruism* und *group selection* verständlich, um die in der Literatur der letzten Jahrzehnte ein so erbitterter Streit ausgefochten wurde. Mithilfe dieses Begriffs können wir uns fragen, ob die Evolution in uns bestimmte Gefühlsdispositionen angelegt hat, deren Funktion es ist, nicht nur einen Gemeinschaftssinn, sondern auch altruistisches Verhalten zu befördern (Sober und Wilson 1998).

Vom evolutionären Standpunkt aus, so bemerken Sober und Wilson, "müssen wir den menschlichen Geist als einen proximalen Mechanismus betrachten, der dazu dient, in Organismen adaptives Verhalten zu erzeugen" (Sober und Wilson 1998, 200). Insbesondere scheinen intrinsische oder grundlegende Wünsche, ob sie sich nun als unweigerlich auf das eigene Wohl konzentriert erweisen oder nicht, vom konkreten Mittel ihrer Erfüllung psychologisch weit entfernt zu sein; trotzdem sind sie die proximalen Mittel, mit deren Hilfe die Evolution ihre 'Ziele' erreicht. In diesem Lichte besehen könnten die Gefühle auf meiner Liste plausibler 'moralischer Gefühle' sehr wohl teilweise für diesen Zweck selektiert worden sein.

Martha Nussbaum (1996) hat dafür argumentiert, dass Mitgefühl zu den wenigen benannten Gefühlen gehört, die einen speziellen Wert als spezifisch moralische Gefühle haben: Gefühle, die mit einiger Sicherheit von sich aus moralisches Verhalten befördern. Nun betrachtet es die soziobiologische Perspektive, die ich in den Vordergrund gestellt habe, keineswegs als methodologisch verfehlt, den Netto-Gesamtertrag einer Art von Verhalten als Schlüssel zu seiner Funktion zu verstehen. Wirklich von Nussbaum lernen können wir aber, dass wir die Art und Weise betrachten müssen, wie Mitgefühl in einem größeren Zusammenhang funktioniert. Wir müssen es in voller Farbe betrachten, d.h. in seinem realen multi-dimensionalen Kontext. Im Folgenden wird in der Hauptsache darzustellen sein, warum dies so ist.

### Umschalten auf Farbe

Nussbaum vermutet "der Schmerz des Anderen geht mich nur an, wenn ich eine Gemeinschaft zwischen mir und ihm zugebe, wenn ich verstehe, wie es für mich wäre, diesem Schmerz ausgesetzt zu sein" (Nussbaum 1996, 35). Und dazu: "Die Annahme ähnlicher Möglichkeiten ist Teil des Konstrukts, das die Kluft überwindet zwischen umsichtiger Sorge und Altruismus"(36). Ich frage mich allerdings, ob dies nicht gerade verkehrt herum ist. Das heißt, ich frage mich, ob nicht eine bestimmte natürliche Fähigkeit zum Erleben des Schmerzes des Anderen – mit der Folge, dass wir tatsächlich dazu bewegt werden, ihn zu berücksichtigen – zumindest Teil der Ursache unserer Fähigkeit ist, die anscheinend so viel komplexere mentale Handlung des 'Anerkennens von Gemeinschaft' auszuführen. Ich verstehe nicht, warum wir die Berichte, denen zufolge die uns verwandten Säugetiere wie z.B. Wale, Tümmler und Elefanten die Fähigkeit zur Empathie mit den Mitgliedern ihrer eigenen oder denen anderer Säugetierspezies haben, nicht für bare Münze nehmen sollten (Masson 1995, 79-82). Meiner Ansicht nach müssen wir dazu keineswegs den Walen so etwas Aufwendiges wie die Fähigkeit zuschreiben, eine Gemeinschaft anzuerkennen, die sie und uns umfasst.

Dieser Vorschlag greift die Idee auf, die unter anderem von Robert Gordon stammt, wonach die Simulation des mentalen Zustands meines Gegenübers der praktischste Weg ist, seinen mentalen Zustand zu ermitteln (Gordon 1986, vgl. Davies und Stone 1995). Diese spekulative Idee ist nicht gänzlich ohne empirische Basis. Empirischen Befunden zufolge ist eine solche Simulation durch bestimmte spezialisierte Neurone vermittelt, die – als Ergebnis von Lernen oder Heranwachsen –



gerade sowohl durch die Beobachtung bestimmter zielgerichteter motorischer Handlungen als auch bei der Ausführung eben dieser Handlungen aktiviert werden. Die Entdeckung dieser 'Spiegel-Neurone' hat zu der Vermutung geführt, dass wir die Fähigkeit des 'Zuschreibens intentionaler Zustände' mit den nicht-menschlichen Primaten teilen.

Obwohl eine solche Zuschreibung hinter einem vollständigen Erfassen der mentalen Zustände des Anderen zurückbleibt, handelt es sich bei ihr um "einen notwendigen phylogenetischen Schritt auf dem evolutionären Weg, der zu der vollständigen Fähigkeit des Gedankenlesens bei Menschen führt" (Gallese und Goldman 1998, 500). Diese Forschungen betreffen nicht speziell Gefühle; trotzdem scheint es nicht abwegig zu sein, unsere Spekulation auf das Erfassen der Gefühle des Anderen zu übertragen. Um ihr Plausibilität zu verleihen, mag der Verweis darauf genügen, dass der Ausdruck von Gefühlen motorische Aktivität einschließt, wobei die faszinierenden Forschungen der Gruppe um Paul Ekman über Gefühlsausdrücke gezeigt haben, dass ein durch erlernte direkte Kontrolle über Gesichtsmuskeln erzeugter Gesichtsausdruck, der dem Ausdruck eines typischen Gefühls präzise entspricht, mit großer Wahrscheinlichkeit dazu führt, dass das entsprechende Gefühl verspürt wird (Ekman und Friesen 1975) (Ekman 1999).

Wenn wir vor diesem Hintergrund Mitleid als eine eher primitive Fähigkeit zu emotionaler Kognition ansehen, können wir vielleicht sogar Nussbaum zustimmen, wenn sie sagt, Mitleid sei 'das basale soziale Gefühl', und dabei der Debatte über die 'Evolution des Altruismus' ein Stück weit ausweichen. Immerhin kann die Fähigkeit zum Mitleid ein Nebeneffekt der klarerweise adaptiven Fähigkeit des 'Gedankenlesens' sein – ein 'Bogenzwickel' im Sinne von S.J. Gould (Gould 1979)(Gould und Vbra 1982).<sup>5</sup> Es ist offenkundig ein entscheidender Vorteil, wenn man sich ein genaues Bild des mentalen Zustands von jemandem machen kann, der das eigene Wohlergehen beeinflussen könnte. Allerdings mag sich dies ohne ein gewisses Maß an Empathie nicht auf effiziente Weise bewerkstelligen lassen. Um dieser Ansicht zu sein, muss man keineswegs eine endgültige Antwort auf die Frage parat haben, worin dies bestehen könnte: "Glaubt man tatsächlich ... man selbst leide? Stellt man sich die eigenen Reaktionen als auf mysteriöse Weise mit denen des Leidenden verbunden vor?" (Nussbaum 1996, 34). Es ist lediglich erforderlich, dass das Erfassen des mentalen Zustands des Anderen präzise ist, und diese Präzision muss als Nebeneffekt den Gedankenleser zu Handlungen bewegen können. Dieser Nebeneffekt könnte kostspielig sein, wenn er zu altruistischem Verhalten führt, aber diese Kosten können nicht schwerer wiegen als die generellen Vorteile, die er mit sich bringt.

Der Prozess, mit Hilfe dessen sich die Fähigkeit herausgebildet hat, die Gedanken und Gefühle einer anderen Person zu 'lesen', könnte nach dem faszinierenden Mechanismus des 'Baldwin Effekt' funktionieren (Baldwin 1896, Hall 2001, Deacon 1997, 322ff). Der Baldwin Effekt wurde ursprünglich erdacht, um lamarcksche Phänomene mit strikt darwinschen Theoremen zu erklären. In seiner allgemeinsten Form besteht er aus einer Rückkopplungs-Schleife, die mit nicht-instinktivem Verhalten beginnt, das eine Veränderung in der Umwelt erzeugt, die ihrerseits zu neuem positiven Selektionsdruck für eine genetische Disposition führt, welche wiederum bestimmtes Verhalten begünstigt. Wenn das Verhalten am Ende der Schleife vom selben Typ ist wie das am Anfang, dann entsteht der Eindruck einer Episode lamarckscher Evolution, da die Entscheidung eines Organismus (oder einer Gruppe von Organismen) für ein bestimmtes Verhalten letztlich dazu führt, dass eine Neigung zu diesem Verhalten genetisch kodiert wird. Das geschieht

---

<sup>5</sup> Einen guten Überblick über die Debatte über Adaptionen, Nebeneffekte, Ex-Adaptionen etc. gibt (Griffith 1996).

allerdings nur in besonderen Fällen; die indirekte Weise des Einflusses garantiert üblicherweise die völlige Unvorhersagbarkeit der Resultate, so dass der von Baldwin identifizierte Effekt letztlich beschreibt, wie die Evolution von verhaltensbefähigten Organismen in völlig neue Richtungen führen kann.

Auf unseren Fall übertragen möchte ich vorschlagen, den Gemeinschaftssinn nicht als eine Vorbedingung, sondern vielmehr als die unvorhersagbare entfernte Wirkung einer Veränderung der Lebensweise unserer Vorfahren aufzufassen, die ihre gegenseitige Abhängigkeit vergrößerte. So könnte es gewesen sein: Stellen wir uns vor, in einer Population, die gerade erst an der Schwelle zu einer Gemeinschaft steht, tritt eine Verhaltensänderung auf. In diesem Fall wird den 'Gedankenlesern' (Humphrey 1984) in der Population ein klarer Vorteil zuteil werden. Unter diesen Umständen kann es sehr gut sein, dass die effektivste Methode des Gedankenlesens diejenige ist, die sich Gefühle zu Nutze macht. Im gegenteiligen Fall haben wir es versäumt, wie Nussbaum zu Recht bemerkt, "zwischen dem Akzeptieren einer Proposition und dem bloßen Aussprechen der Wörter zu unterscheiden"; weswegen Nussbaum mit Bezug auf Rousseaus Diktum "Wer es sieht, ohne es zu fühlen, kennt es nicht" bemerkt, dass "dies ... eine klare Bedeutung hat: das Leiden anderer ist kein Teil von Emiles kognitivem Repertoire geworden ..." (Nussbaum 1996, 38). Dadurch erreichen wir eine höhere Stufe von Sozialität, eine Stufe auf der, wie Nussbaum mit Blick auf die Zuschauer der Tragödie bemerkt, man durch das Mitleid, das der Tragiker oder Schriftsteller hervorruft, dazu gebracht werde, "Ungleichheiten in Verletzlichkeit zu sehen: deshalb ... hat man Gründe, die unterste Ebene an Sicherheit für alle zu erhöhen"(45). Auf diese Weise werden die angeborenen Fähigkeiten zum empathischen Mitleid dazu führen, uns dem stoischen Ideal der moralischen Entwicklung näher zu bringen, wobei "die [konzentrischen] Kreise [der Belange, die fortschreitend weiter von der eigenen, rein egoistischen entfernt sind,] nach und nach näher an das Zentrum rücken, so dass die eigenen Eltern einem Selbst gleichkommen, Verwandte den Eltern, Fremde den Verwandten, und so fort" (Nussbaum 1996, 48).

So wie ich sie ausgemalt habe, präsentiert die Geschichte den notwendigen Übergang von der Schwarz-Weiß-Projektion unserer Gefühle zum farbigen Bild. Die Geschichte beginnt mit der allein instrumentell interessanten Fähigkeit des Gedankenlesens; letztlich bin ich jedoch in meiner Erzählung nur insoweit von Nussbaums eigener Formulierung abgewichen, als ich die Möglichkeit des einfachen Ursprungs von Mitleid betont habe, so dass wir anstatt zu fragen: "Wie kann ein so komplexes Gefühl erlernt sein?" (Nussbaum 1996, 39) die Frage erwägen müssen: "Wie hat sich dieses einfache und robuste Gefühl herausbilden können?". Um aber das volle durch ein solches Verständnis implizierte Bild zu erhalten, müssen wir über die eindimensionale Notwendigkeit der Wahl hinausgehen: Selbstverständlich ist mein Wissen um die mentalen Zustände anderer für meine praktischen Interessen wichtig, aber wenn Nussbaums Beschreibung (und mein Begriff des Schlüsselszenarios) unser Gefühlsleben auch nur im Ansatz richtig charakterisiert, dann lässt sich das Erfassen des mentalen Zustands einer anderen Person nicht angemessen in Schwarz-Weiß wiedergeben, d.h. mithilfe des zweidimensionalen Vokabulars der Wahl zwischen Handeln und nicht-Handeln. Sie begründen reiche Erfahrungen von inkommensurablen Wertdimensionen, und sie sind das, was Mitleid oder Empathie letzten Endes zu erfassen suchen.

### **Das Problem niederträchtiger Emotionen**

Nussbaum bringt Argumente zur Verteidigung der moralischen Relevanz des Mitleids vor, die dem Einwand begegnen sollen, wenn man Mitleid zulasse, müsse man auch weniger akzeptable Leidenschaften anerkennen. Jemand, der genuin Bemitleidenswertes ausmacht und der die darin

implizit enthaltenen Werte anerkennt, "akzeptiert, dass er einen Scheck auf die Zukunft ausgestellt hat. Und wer einen Scheck auf die Zukunft ausstellt, muss sich nicht nur auf Mitleid einrichten, sondern auch auf Angst und Beklemmung und Trauer – und ... auf Zorn und die Neigung zu Vergelten" (Nussbaum 1996, 43). Nussbaum selbst gibt hierauf zu bedenken, der Anhänger der Idee, Gefühle seien moralisch bedeutsam, sei "nicht darauf festgelegt, dass jeder Schaden, jeder Affront oder jede Beleidigung ein angemessener Anlass für wütende Vergeltung ist. Die mit Abstand meisten sozialen Übel, die durch Vergeltung verursacht werden, betreffen Schäden an Vermögen, Status, Macht und Ehre – und diesen misst der Anhänger des Mitleids wenig Wert zu" (Nussbaum 1996, 49). Diese Antwort erscheint mir hinreichend plausibel; allerdings möchte ich sie in eine größere Perspektive einordnen.

Die Anhänger der Bedeutsamkeit von Gefühlen für das vernünftige und moralische Leben werden sich erwartungsgemäß mit dem Vorwurf konfrontiert sehen, sie ignorierten die Tatsache, dass einige Emotionen *niederträchtig* sind. Ein Gefühl kann typischerweise als niederträchtig verstanden werden, wenn es jede der folgenden Bedingungen erfüllt:

- (i) das Erleben des Gefühls wird strikt vermieden
- (ii) es wird stark missbilligt; und
- (iii) es hat immer oder meistens schlechte Konsequenzen.

Für den Anhänger der Idee, Gefühle seien Wertwahrnehmungen, bedeutet die Existenz niederträchtiger Gefühle eine doppelte Herausforderung. Zum einen, wie lässt sich die axiologische Perspektive mit der Existenz niederträchtiger Gefühle vereinbaren? Und zum anderen, könnte die Existenz niederträchtiger Gefühle dazu geeignet sein, die Idee zu untergraben, Gefühle seien Wahrnehmungen von axiologischen Werten?

Aus den Unterschieden zwischen dem Schwarz-Weiß-Schema und der farbigen Konzeption von Gefühlen lässt sich hierauf eine Antwort gewinnen. In einer rein schwarz-weißen Welt, in der nur die unmittelbar praktische Dimension Beachtung findet (eine Perspektive bei der, wie betont, die simple Konzeption des Unmittelbaren Hedonismus ein erstaunliches Maß an Einsicht bringt), würden niederträchtige Gefühle schlicht deswegen verurteilt werden, weil ihre Existenz – ob direkt oder zusammen mit anderen Konsequenzen – einen negativen Einfluss auf Nutzen hat. Ob 'Schäden an Vermögen, Status, Macht und Ehre' als wichtig angesehen werden oder nicht, es handelt sich hier um Schäden, und sie sind die üblichen Folgen von Gefühlen, die sich auf Vergeltung richten. Der axiologischen Konzeption zufolge ist die entscheidende Frage aber gar nicht, ob diese Gefühle schlechte Konsequenzen haben, genau so wenig wie die Erkenntnistheorie uns aufträgt, zu überlegen, ob eine bestimmte Überzeugung schlechte Konsequenzen hat. Wahre Überzeugungen können entmutigend sein oder gar zu Depressionen führen (außerdem gibt es immer wieder interessante Belege für die Idee, dass seine Erkrankung den Depressiven wahre Überzeugungen bevorzugen lässt (Pacini et. al 1998)); aber das ist erkenntnistheoretisch irrelevant: die Konsequenzen des Habens einer Überzeugung sind für ihre Korrektheit unwesentlich. Ganz ähnlich mag das Haben eines 'korrekten' Wunsches, d.h.: das Haben eines Wunsches mit Bezug auf etwas 'Gutes' oder tatsächlich Wünschenswertes, höchst unerwünschte Konsequenzen haben: Aber die Konsequenzen des Habens des Wunsches sind für die Frage irrelevant, ob das Objekt des Wunsches tatsächlich so war, wie der Wunsch es repräsentiert hat, also wünschenswert und gut. Ob ein Zustand seiner eigenen Maßgabe nach 'korrekt' ist, hängt davon ab, ob sein Ziel mit dem formalen Objekt des Zustand-Typs übereinstimmt. Im Fall von Überzeugung ist die analoge Dimension die der Wahrheit; im Fall des Wunsches ist es Güte oder

Wünschbarkeit. Im Fall des Gefühls wird dies durch das definierende Schlüsselszenario bestimmt: das formale Objekt spannt die axiologische Dimension auf. Allerdings müssen wir der Versuchung widerstehen, anzunehmen, es gäbe eine einzige Dimension von so etwas wie emotionaler Angemessenheit, denn in Wahrheit hat jedes Gefühl seinen eigenen Angemessenheits-Standard.<sup>6</sup> Ein weiteres Problem, das man von jeder speziellen epistemologischen Frage mit Bezug auf eine bestimmte Überzeugung unterscheiden kann, ist, ob diese Überzeugung in einem bestimmten Kontext wahrscheinlich zu einem guten Leben gehört oder nicht. Ganz ähnlich stellen sich für jedes Gefühl zwei verschiedene Fragen, von denen die eine das Ausmaß betrifft, in dem das Gefühl seinem Objekt entspricht, wobei dies an der Beziehung des Objekts zum formalen Objekt des Gefühls festgemacht wird, und die andere den Beitrag betrifft, den ein Gefühl dieser Art im gegenwärtigen Kontext zu einem guten Leben leisten kann.

Wir benötigen hier eine Konzeption, die ich an anderer Stelle den Axiologischen Holismus genannt habe (de Sousa 2000). In unserer vielfarbigen Welt mit unbegrenzt vielen Gefühlsdimensionen können wir ein bestimmtes Gefühl nur in dem umfassenden Kontext beurteilen, in dem es erlebt wird. Jedes Gefühl ist auf eigene Weise entstanden, und hat sein eigenes spezielles formales Objekt geformt; aber weder die natürliche Selektion noch die jeweils individuelle Biografie führt zu einer intrinsischen Teleologie, die vorschreibt, dass alle Gefühle nützlich, angenehm, altruistisch oder sonst wie lobenswert zu sein haben. Trotzdem trägt allein der Umstand der Existenz eines Gefühls den Anspruch an uns heran, dass es Teil eines Lebens bilden kann, das jemand leben möchte. Dieser Anspruch kann durch andere Ansprüche aufgewogen werden; und dies mag auch dann der Fall sein, wenn das Gefühl im Allgemeinen gut in einen vernünftigen Entwurf des guten Lebens passt. Sozial nützliche Gefühle wie Mitleid oder Güte können uns gelegentlich in die Irre führen, gerade so, wie die offenkundige Nützlichkeit unseres Wahrnehmungsapparates optische Täuschungen nicht ausschließt. Dass Gefühle im Allgemeinen Wahrnehmungen axiologischer Wirklichkeit sind, impliziert nicht, dass alle Gefühle gleich vertrauenswürdig sind. Dies gilt auf beiden der gerade unterschiedenen Bewertungsebenen: Einige Gefühle mögen in ihrem Kontext in die Irre führen, obwohl sie in sich stimmig sind; andere mögen systematisch irreführend oder niederträchtig sein. In beiden Fällen könnte man von axiologischen Fehlern sprechen. Das zeigt aber weder, dass nicht jedes Gefühl etwas richtig hinbekommt. Noch zeigt dies, dass ein Gefühl, das etwas richtig hinbekommt, in jedem Fall eine gute Sache ist.

Einer weiteren Komplikation müssen wir uns noch annehmen. Obwohl die meisten Philosophen Gefühle eher als natürlich entstanden denn als erlernt begreifen, würden sie einräumen, dass die für die moralische Entwicklung zentralen Gefühle gelernt werden. Zum Teil mag dies auch für die allgemeinen, 'basalen' Gefühle zutreffend sein; aber für alle anderen gilt es ganz sicher.

Das bedeutet nicht notwendigerweise, dass sie in jeder Kulturgemeinschaft einheitlich ausfallen; aber es bedeutet, dass sie entweder in Konformität mit dem Gruppendruck oder in Reaktion auf ihn entstehen: In beiden Fällen kann sich das Gefühlsrepertoire eines Kindes wohl kaum den kulturellen Einflüssen entziehen, die es geformt haben, auch wenn sich die Vorzeichen einiger

---

<sup>6</sup> Ich habe in (de Sousa 1974) dafür argumentiert, dass sich einige der zentralen Unterschiede in den Rationalitätsstandards für Überzeugungen und für Wünsche auf die Unterschiede in ihren formalen Objekten zurückführen lassen; und ich habe zu erklären versucht, wie dies sich auf die unbegrenzte Dimensionalität von Gefühlen übertragen lässt (de Sousa 1987, insbes. Kapitel 6). (D.Arms und Jacobson 2001) führen auf treffende und durchsichtige Weise aus, wie wichtig es ist, zwischen der Beziehung des Übereinstimmens eines Gefühls mit seinem Ziel und seinem moralischen Wert oder seiner Angemessenheit zu unterscheiden. Für das Versäumnis, dies zu unterscheiden, prägen sie den treffenden Ausdruck 'Moralistischer Fehlschluss'.

Vektoren umkehren mögen. Die eiserne Regel der meisten Gesellschaften ist Konformität. Das wird in Sober und Wilsons Überblick über die Kulturen der Welt deutlich: Gesellschaften wie die unsrige, in denen das Individuum nicht durch die Tyrannei der Gruppe gezwungen ist, den sozialen Erwartungen zu entsprechen, sind außerordentlich selten, insbesondere in allen Fragen, die Fortpflanzung, Verwandtschaft und Sexualität auch nur entfernt berühren (Sober und Wilson, Kapitel 5). Aktuelle Beispiele lassen sich leicht beibringen. Vor einiger Zeit habe ich ein Interview mit einigen Männern gesehen, die eine Schwester oder Tochter wegen einer 'Ehrensache' getötet hatten – d.h., weil das Opfer vergewaltigt worden war. Das Erstaunlichste an der Einstellung dieser Männer war, dass sie ob ihrer durch ihr 'Ehrgefühl' motivierten Handlung auch nicht von den leisesten Anfällen von Zweifel oder gar Schuld geplagt wurden. Hier sieht man die Kraft sozialer Konditionierung.

Der Axiologische Holismus, so die Moral, kann sich nicht darin erschöpfen, dass wir über die Gefühle des Einzelnen hinaus die von der Gesellschaft gebilligten Gefühle in Betracht ziehen müssen; denn nach unseren Maßstäben sind viele der von den meisten Gesellschaften akzeptierten Gefühle unsagbar abscheulich. Wir können nur dadurch herausfinden, ob ein Gefühl richtig oder falsch ist, indem wir auf die gleiche Weise verfahren wie beim Test von Wahrnehmungen auf Wahrhaftigkeit: Wir müssen dasselbe Objekt von verschiedenen Standpunkten aus betrachten, was letztlich bedeutet, dass wir uns auf andere Wahrnehmungen oder, *mutatis mutandis*, Gefühle berufen. So etwas wie die Methode des Überlegungsgleichgewichts ist Gemeingut in Wissenschaft wie Ethik; allerdings wird selten bemerkt, was insbesondere Nussbaums Arbeit deutlich gemacht hat, dass es nämlich in der Ethik typischerweise nicht darum geht, ein Gleichgewicht zwischen empirischen Tatsachen zu erzeugen, sondern darum, Gefühlsreaktionen in ein Gleichgewicht zu bringen. Diese Gefühlsreaktionen lassen sich am Besten als axiologische Wahrnehmungen verstehen; wenn man ihnen vollständig gerecht werden will, reicht es nicht aus, ihre zweidimensionalen Projektionen in den Raum von Lust und Schmerz oder das halbe Dutzend Dimensionen zu beachten, die durch die 'Standard'-Gefühle aufgespannt werden. Die weniger-dimensionale Projektion einer vieldimensionalen Welt kann nicht nur irreführend, sondern im strikten Sinne inkohärent sein, gerade so, wie die zweidimensionalen Bilder von Escher etwas darstellen, das im Dreidimensionalen, das sie abzubilden vorgeben, unmöglich ist. Darum ist in einer vielfarbigem Welt echter Gefühle der hohe Genauigkeitsgrad 'erschöpfend ausgemalter' Kunst und Literatur nicht nur hilfreich, sondern wesentlich für die Aufgabe des moralischen und emotionalen Lernens.

### Literatur

- Baldwin, J. M. 1896. Consciousness and Evolution. *Psychological Review* 3:300-08.
- Campbell, S. 1998. *Interpreting the Personal: Expression and the Formation of Feeling*. Ithaca: Cornell University.
- Clynes, M. 1977. *Sentics: The Touch of Emotions*. Garden City: Anchor Press.
- D.Arms, J., and D. Jacobson 2001. The Moralistic Fallacy. *Philosophy and Phenomenological Research* 61, 65-90.
- Darwin, Charles. 1998 <1872,1889>. Ed. with Introduction, Afterword and Cmmentaries by Paul Ekman. *The Expression of the Emotions in Man and Animals*. London: Harper Collins.
- Davies, M., and T. Stone, eds. 1995. *Mental Simulation: Evaluations and Applications*. Readings in Mind and Language. Oxford: Blackwell.
- de Sousa, R. 2000. Learning to be Natural. In *Being Humans*, ed. Neil Roughley. Berlin and New York: De Gruyter. 287-307.

- 1974. The Good and the True. *Mind* 83:534-51.
- 1987. *The Rationality of Emotion*. Cambridge (Mass.): MIT Press.
- Deacon, T. W. 1997. *The Symbolic Species: The Coevolution of Language and the Brain*. New York: Norton.
- Ekman, P. 1999. Emotional and Conversational Nonverbal Signals. Conference Presentation. San Sebastian.
- Ekman, P., and W. Friesen. 1975. *Unmasking the Face: A Guide to Recognizing Emotions from Facial Expressions*. Englewood Cliffs, NJ: Prentice-Hall.
- Gallese, V., and A. Goldman. 1998. Mirror Neurons and the Simulation Theory of Mind- Reading. *Trends in Cognitive Sciences* 2(12):493-501.
- Gallop, D. 1960. True and False Pleasures. *Philosophical Quarterly* 10:331-342.
- Gordon, R. M. 1986. Folk Psychology and Simulation. *Mind and Language* 1:158-71.
- Gould, S. J., and E. Vrba. 1982. Exaptation: A Missing Term in the Science of Form. *Paleobiology* 8:4-15
- Gould, S. J. 1979. The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society of London* 205(581-598).
- Griffiths, P. E. 1996. The Historical Turn in the Study of Adaptation. *British Journal for the Philosophy of Science* 47:511-32. 26
- Hall, B. K. 2001. Organic Selection: Proximate Environmental Effects on the Evolution of Morphology and Behaviour. *Biology and Philosophy* 16: 215-237.
- Humphrey, N. 1984. Nature's Psychologists. In *Consciousness Regained: Chapters in the Development of Mind*, 29-41. Oxford: Oxford University Press.
- Itard, J. 1964. Mémoire et Rapport sur Victor de l'Aveyron. L. Malson. In *Les Enfants Sauvages: Mythe et Réalité*. Paris: Union Générale d'éditions.
- Masson, J. M. 1995. *When Elephants Weep: The Emotional Lives of Animals*. New York: Delacorte Press.
- Nozick, R. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.
- Nussbaum, M. 1990. *Love's Knowledge*. Oxford, New York, Toronto: Oxford University Press.
- 1994. *The Therapy of Desire: Theory and Practice in Hellenistic Ethics*. Princeton: Princeton University Press.
- 1996. Compassion: The Basic Social Emotion. *Social Philosophy and Policy* 13:27-58.
- 1997. *Cultivating Humanity : a Classical Defense of Reform in Liberal Education*. Cambridge, MA: Harvard University Press.
- Pacini, R.; Muir, F.; Epstein, S. 1998. Depressive realism from the perspective of cognitive-experiential self-theory. *Journal of Personality and Social Psychology* 74: 1056-1068
- Penner, T. 1970. False Anticipatory Pleasures in the Philebus 3613-41a6. *Phronesis*:166-78.
- Rosenfield, I. 1992. *The Strange, Familiar and Forgotten*. New York: Alfred A. Knopf.
- Sober, E., and D. S. Wilson. 1998. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.

**Übersetzung: Christian Nitz**

\*

**Erstpublikation in: Achim Stephan/Henrik Walter (Hrsg.): Moralität, Rationalität und die Emotionen (Bausteine zur Philosophie, Band 22). Interdisziplinäre Schriftenreihe Humboldt-Studienzentrum: Universität Ulm 2004.**

**Nachdruck mit freundlicher Genehmigung der Herausgeber.**