

# *philosophia* JOURNAL FOR THE PHILOSOPHY OF NATURE *naturalis*

Herausgeber / Editors    Andreas Bartels  
   Bernd-Olaf Küppers  
   C. Ulises Moulines

- Monika Betzler    Macht uns die Veränderung unserer selbst  
   autonom?
- Norman Sieroka    Ist *ein* Zeithof schon genug?
- Boris Hennig    Eine Verteidigung des typologischen  
   Artbegriffs
- Vera Spillner    A Bundle Definition of Scientific  
   Understanding and its Application to  
   Quantum Physics

*philosophia*  
JOURNAL FOR THE PHILOSOPHY OF NATURE *naturalis*

46 / 2009 / 2

Herausgeber / Editors    Andreas Bartels  
                                      Bernd-Olaf Küppers  
                                      C. Ulises Moulines

Beirat / Editorial Board    Werner Diederich (Hamburg)  
                                      Michael Esfeld (Lausanne)  
                                      Don Howard (Notre Dame)  
                                      Andreas Hüttemann (Münster)  
                                      Bernulf Kanitscheider (Gießen)  
                                      Daryn Lehoux (Kingston, Ontario)  
                                      James Lennox (Pittsburgh)  
                                      Holger Lyre (Magdeburg)  
                                      Peter Mittelstaedt (Köln)  
                                      Felix Mühlhölzer (Göttingen)  
                                      Friedrich Rapp (Dortmund)  
                                      Friedrich Steinle (Berlin)  
                                      Manfred Stöckler (Bremen)  
                                      Eckart Voland (Gießen)  
                                      Gerhard Vollmer (Braunschweig)  
                                      Marcel Weber (Konstanz)  
                                      Michael Wolff (Bielefeld)

KLOSTERMANN



# Jahresinhaltsverzeichnis 2009

## *Heft 1*

Brigitte Falkenburg, Andreas Hüttemann, Manfred Stöckler	Nachruf auf Erhard Scheibe	5
Sven Walter	Wie frei sind wir eigentlich – empirisch?	8
Maria E. Kronfeldner	Meme, Meme, Meme: Darwins Erben und die Kultur	36
Matthias Rang, Olaf L. Müller	Newton in Grönland. Das umgestülpte <i>experimentum crucis</i> in der Streulichtkammer	61
Francisco Antonio Doria, Manuel Doria	On formal treatments for general relativity	115
Gregor Betz	What range of future scenarios should climate policy be based on? Modal falsificationism and its limitations	133

## *Heft 2*

Monika Betzler	Macht uns die Veränderung unserer selbst autonom? Überlegungen zur Rechtfertigung von Neuro-Enhancement der Emotionen	167
Norman Sieroka	Ist <i>ein</i> Zeithof schon genug? – Neurophänomenologische Überlegungen zum Zeitbewusstsein und zur Rolle des Auditiven	213

Boris Hennig	Eine Verteidigung des typologischen Artbegriffs	251
Vera Spillner	A Bundle Definition of Scientific Understanding and its Application to Quantum Physics	279

## Inhalt

Monika Betzler	Macht uns die Veränderung unserer selbst autonom? Überlegungen zur Rechtfertigung von Neuro-Enhancement der Emotionen	167
Norman Sieroka	Ist <i>ein</i> Zeithof schon genug? – Neurophänomenologische Überlegungen zum Zeitbewusstsein und zur Rolle des Auditiven	213
Boris Hennig	Eine Verteidigung des typologischen Artbegriffs	251
Vera Spillner	A Bundle Definition of Scientific Understanding and its Application to Quantum Physics	279
	Verzeichnis der Autoren	306
	Richtlinien zur Manuskriptgestaltung	307


The articles are indexed in *The Philosopher's Index* and *Mathematical Reviews*.

Abonnenten der Printausgabe können über Ingentaconnect auf die Online-Ausgabe der Zeitschrift zugreifen: [www.ingentaconnect.com](http://www.ingentaconnect.com)

Zurückliegende Jahrgänge sind mit einer Sperrfrist von fünf Jahren für die Abonnenten von [www.digizeitschriften.de](http://www.digizeitschriften.de) zugänglich.

© Vittorio Klostermann GmbH, Frankfurt am Main 2011

Die Zeitschrift und alle in ihr enthaltenen Beiträge und Abbildungen sind urheberrechtlich geschützt. Jede Verwertung außerhalb der Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und Einspeicherung und Verarbeitung in elektronischen Systemen.

Satz: Mirjam Loch, Frankfurt am Main / Druck: KM-Druck, Groß-Umstadt.  
Gedruckt auf alterungsbeständigem Papier  ISO 9706.

ISBN 978-3-465-04126-9

ISSN 0031-8027

Monika Betzler

## Macht uns die Veränderung unserer selbst autonom?

### Überlegungen zur Rechtfertigung von *Neuro-Enhancement* der Emotionen

#### Zusammenfassung

Die Einnahme von Psychopharmaka zur Verbesserung der emotionalen Befindlichkeit – sogenannte Massnahmen des *Neuro-Enhancement* (NE) – werden gegenwärtig mit Rekurs auf Begriffe, wie den der Autonomie oder der Authentizität, gerechtfertigt. In diesem Beitrag verfolge ich drei Ziele.

(1) Ich zeige, welcher Begriff von Autonomie überhaupt angemessen ist, um NE, zumindest prinzipiell, zu rechtfertigen. Nur ein weiter Begriff von Autonomie, der neutral gegenüber der Idee der Selbstveränderung ist, erlaubt eine nicht-zirkuläre Begründung von NE.

(2) Vor diesem Hintergrund diskutiere ich, welche Arten der emotionalen Veränderung als autonomiefördernd betrachtet werden können. Ich argumentiere dafür, dass nur die Steigerung der Rationalität emotionaler Einstellungen die Autonomie einer Person in einem weiten Sinne befördern kann.

(3) Abschliessend untersuche ich, ob NE die Rationalität emotionaler Einstellungen tatsächlich zu erhöhen vermag. Es gibt hierbei zwei Gründe zur Skepsis: Zum einen droht NE andere autonomiefördernde Bedingungen, wie die Willensstärke und das kontinuierliche Selbstverständnis der betreffenden Person, zu untergraben. Ferner bleibt es aus konzeptuellen Gründen zweifelhaft, ob Psychopharmaka zwischen rationalen und irrationalen Emotionen überhaupt unterscheiden können.

#### Abstract

In current theorizing, concepts of autonomy and authenticity are used to justify measures of neuro-enhancement (NE), i. e. the use of pharmaceuticals to enhance a person's emotions. In this article, I pursue three aims.

(1) In light of the current debate, I show which conception of autonomy is adequate to justify, in principle at least, measures of NE. I defend a wide notion of autonomy that remains neutral with regard to the idea of self-transformation, and thus makes possible a non-circular justification of self-transformation by NE.

*philosophia naturalis* 46 / 2009 / 2



(2) In light of these clarifications, I examine which kinds of emotional changes can be regarded as enhancing a person's autonomy. I suggest that only the enhancement of rational emotions can render a person autonomous (in a wide sense).

(3) Finally, I consider whether NE is really capable of enhancing the rationality of a person's emotions. There are two reasons for skepticism: On the one hand, NE threatens to undermine two further conditions of autonomy, i.e., a person's strength of will and her ongoing self-understanding. Moreover, there are conceptual grounds for doubting that pharmaceuticals can at all improve the rationality of our emotions. Pharmaceuticals are themselves incapable of distinguishing between rational and irrational emotions.

## 1. Die Fragestellung

Viele Menschen leiden an depressiven Stimmungen, die meist mit Gefühlen der Angst, Frustration, Trauer, Verzweiflung und anderen negativen Emotionen einhergehen. Diese negative emotionale Befindlichkeit ist hierbei entweder ein regelmässig wiederkehrender oder zumindest für einen längeren Zeitraum anhaltender Zustand, ohne dass die Betroffenen an einer psychischen Erkrankung leiden. Sie sind weder klinisch depressiv noch haben sie eine andere medizinisch diagnostizierbare Störung.

Sie fühlen sich jedoch aufgrund ihrer schmerzhaft erlebten, affektiven Verfassung (und möglicherweise auch wegen der Ursachen, die dazu geführt haben), unglücklich und niedergeschlagen. Aus diesem Grund scheinen sie nur unzureichend in der Lage, ein selbstbestimmtes Leben zu führen. Dies zeigt sich z.B. darin, dass sie entweder nicht das tun können, was sie selbst wirklich wertschätzen, da ihre negativen Emotionen sie daran hindern. Oder sie schätzen aufgrund ihrer affektiven Verfassung Dinge wert, die sie jedoch selbst nicht wirklich als wertvoll erachten bzw. die ihnen in anderer Hinsicht nicht gut tun. Sie sind nicht in der Lage, sich aus ganzem Herzen an Personen oder Projekte zu binden, sie erleben sich häufig als zerrissen und es ist nicht klar, ob ihnen die Orientierung fehlt, oder ob ihre Orientierung für sie falsch ist. Derartige Fälle, so allgemein sie auch noch beschrieben sind, sind relativ weit verbreitet und den meisten von uns vertraut. Sie scheinen für viele zu einem „normalen“ Leben dazuzugehören. Folgende Beispiele zeugen davon:

*Beispiel 1: Marina*

Marina war ein emotional vernachlässigtes Kind: Sie war vier, als ihre Eltern sich trennten. Ihr Vater verhielt sich seither distanziert und war wenig verfügbar. Die Mutter neigte dem Alkoholismus zu. Als ältestes Kind kümmerte sich Marina viel um ihre jüngeren Geschwister. Sie übernahm eine Fürsorge-Rolle und lernte, ihre eigenen Bedürfnisse zu verdrängen. Trotz der emotional depravierten Kindheit schaffte es Marina zu studieren. Mit höchstem Arbeitseifer machte sie Karriere als Anwältin. Ihre Beziehungen zu Männern waren jedoch von Misserfolgen geprägt. Im Berufsleben erlebte sie sich trotz allen äusseren Anscheins als unsicher, ängstlich und allzu leicht bereit, sich zu zermürben. Sie pflegte zwar einige Freundschaften und hatte Freude an Hobbies, aber sie kann nicht umhin, sich in sexuellen Phantasien zu ergehen, in denen sie sich zum blossen Sex mit weit älteren Männern trifft. Sie empfindet ihre Unsicherheit im Berufsleben als störend und verabscheut ihre Vorliebe zu Männern, die ihr nicht gut tun, sowie ihre Gefühle der Missachtung gegenüber Männern, die nett zu ihr sind und sie mit Respekt behandeln.

*Beispiel 2: Beate*

Beate ist eine überlastete, berufstätige Mutter mit einer von ihr selbst verabscheuten Neigung zum Perfektionismus, zur Ungeduld und ständigen Gereiztheit. Als Kind erlebte sie Zuneigung und Liebe meist nur aufgrund ihrer Leistungen, die sie im Laufe ihres Lebens mit beachtlicher Energie und grossem Engagement erbrachte. Sich selbst Ruhe zu gönnen, sich zu belohnen oder den Dingen einfach ihren Lauf zu lassen, war ihr kaum möglich. Sie war ständig aktiv und schien zu geradezu übermenschlichen Leistungen in Beruf und Familie fähig. Sobald sich jedoch die Verhaltensweisen anderer oder Geschehnisse ihrer Kontrolle entzogen, reagierte sie mit Tyrannei. So kommandiert sie ihre Familie herum, übt emotionalen Druck auf ihre Kinder aus, und hat kaum Toleranz gegenüber den herkömmlichen Schwächen oder einfach anderen Zielen ihrer Mitmenschen. Sie möchte selbst grosszügiger und gelassener sein, anderen mit mehr Geduld und Zuneigung begegnen und sich selbst nicht nur vom Aktivismus getrieben erleben.

Die beiden Beispiele<sup>1</sup> sind ohne Zweifel vereinfacht und skizzieren lediglich eine weit komplexere psychische Lage, in der sich die Prota-

gonistinnen befinden. Sie machen jedoch beide deutlich, dass die negativen emotionalen Befindlichkeiten der beiden Frauen sowie die eigene Ablehnung ihres Zustands – die sich u. a. in Unzufriedenheit, Schmerz und erlebter innerer Zerrissenheit äussert – nach einer Veränderung ihres Lebens verlangen.

So beschliesst Marina nach langem Grübeln über ihre Unzufriedenheit, Rast- und Orientierungslosigkeit, im Berufsleben selbstsicherer und entscheidungsfreudiger zu werden. Im Privatleben möchte sie sich weniger ausgeschlossen sowie eines guten Partners würdig fühlen. Beate fasst den Plan, mehr Geduld und Toleranz für ihre Mitmenschen, insbesondere für ihre Familie, aufzubringen und ihren übertriebenen Ehrgeiz und perfektionistischen Aktivismus zu drosseln. Da beiden eine Psychotherapie zu teuer, langwierig und nicht unbedingt erfolgsversprechend erscheint, möchten sie ein Antidepressivum verschrieben bekommen. Sie möchten auf diese Weise ihre emotionalen Befindlichkeiten verbessern.

Auf diese Weise, so scheint es, können wir dank pharmakologischer Fortschritte leicht in unsere menschliche Biologie eingreifen und zentrale Aspekte unserer selbst umgestalten. Derartige Massnahmen werden in der Regel mit dem englischen Begriff „*Neuro-Enhancement*“<sup>2</sup> bezeichnet. Darunter versteht man zunächst die Veränderung mentaler Zustände mit dem Ziel der Steigerung ihrer Funktionsweise bzw. Leistungskraft. Dies betrifft neben der Veränderung emotionaler bzw. motivationaler Eigenschaften auch die Stärkung kognitiver Fähigkeiten (wie etwa der Aufmerksamkeit, der Wachheit oder der Intelligenz und der Gedächtniskraft). Im Folgenden werde ich mich jedoch auf *Neuro-Enhancement* der Emotionen beschränken. Um den Begriff zu definieren, bedarf es noch zweier Präzisierungen:

So hat es sich eingebürgert, *Enhancement*<sup>3</sup> generell und somit auch *Neuro-Enhancement* von rein therapeutischen Massnahmen zur Wiederherstellung der Gesundheit eines Menschen oder zur Prävention von Krankheit und Behinderungen und deren Symptomen zu unterscheiden. Insofern zielt *Neuro-Enhancement* auf die (i) über das speziestypische Normalmass hinausgehende Veränderung psychischer Eigenschaften oder Zustände gesunder Menschen.<sup>4</sup>

Im Gegensatz zu einem Gespräch mit Freunden, zur Psychotherapie, zum Ausdauersport oder zum Konsum grösserer Mengen von Alkohol, die ebenfalls als Mittel zur möglichen Verbesserung emotionaler

Befindlichkeiten aufgefasst werden können, stellt *Neuro-Enhancement* (ii) eine Veränderung der physiologischen bzw. psychologischen Ausstattung mittels eines unmittelbaren, meist invasiven (z.B. genetischen, pharmazeutischen oder prosthetischen) Eingriffs in den menschlichen Körper dar.<sup>5</sup>

In diesem Sinne gelten weder Marina noch Beate als krank. Sie leiden weder an einer schweren Depression noch haben sie ein anderweitig psychiatrisch klassifizierbares Leiden. Beide wollen sich verändern, indem sie durch die Einnahme synthetischer Psychopharmaka – d. h., mittels einer direkten, nicht süchtig machenden Manipulation an ihrer neuronalen Struktur – ihre emotionale Befindlichkeit verbessern. An dieser Stelle bleibt zu betonen, dass gerade im Fall psychischer Leiden keine klare Trennlinie zwischen Krankheit und Gesundheit zu ziehen ist. Es scheint daher nicht immer hinreichend bestimmt, wie ein „kosmetischer“ (Kramer) Gebrauch von Psychopharmaka von einem nicht-kosmetischen Gebrauch zu unterscheiden ist.<sup>6</sup> Trotz dieser Schwierigkeit gehe ich im Folgenden davon aus, dass es sich um keine psychiatrischen oder neurologisch diagnostizierbaren Krankheiten handelt und die betreffenden Personen an ihren eigenen emotionalen Befindlichkeiten leiden.<sup>7</sup> Bevor ich näher darauf eingehe, wie eine Veränderung dieser Befindlichkeiten durch Psychopharmaka gerechtfertigt werden könnte und warum sie überhaupt der Rechtfertigung bedarf, möchte ich kurz die empirischen Voraussetzungen solcher Massnahmen klären. Welche Psychopharmaka bieten sich an und was bewirken sie?

Es kommen viele Medikamente in Frage, die zwar ursprünglich für die Therapie Kranker entwickelt, unterdessen aber immer häufiger ausserhalb ihres Indikationsbereichs zur Veränderung emotionaler Fähigkeiten eingesetzt werden und in den Stoffwechsel des Gehirns eingreifen. Hier finden u. a. neben dem gegen Depressionen eingesetzten selektiven Serotonin-Wiederaufnahmehemmer (wie z. B. Fluoxetin bzw. Prozac), dem Noradrenalin-Wiederaufnahmehemmer Rebosetin, dem dual wirkenden Serotonin-Noradrenalin-Wiederaufnahmehemmer (z. B. Duloxetin), auch der Betablocker Propranolol zur Abmilderung traumatischer Erinnerungen, sowie Anxiolytica (wie z. B. Dormicum) mit einer sedierenden, angstlösenden Wirkung, nicht-medizinisch indizierte Anwendung.<sup>8</sup>

Es bleibt hierbei festzuhalten, dass derzeit keine befriedigenden Forschungsergebnisse zur tatsächlichen Wirksamkeit dieser Medikamente

für *Neuro-Enhancement* vorliegen. Während manche die Glaubwürdigkeit bisheriger klinischer Studien grundsätzlich bezweifeln,<sup>9</sup> weisen andere darauf hin, dass die Wirkung von Psychopharmaka auf gesunde Menschen nicht hinreichend nachgewiesen wurde<sup>10</sup> oder nachweisbar gering ist.<sup>11</sup> Darüber hinaus wird zunehmend auf schwerwiegende Nebenwirkungen von Psychopharmaka, insbesondere von Prozac, aufmerksam gemacht.<sup>12</sup> Erstaunlich wenig wurde thematisiert, wie stabil die durch Stimmungsaufheller erzeugten Veränderungen an Gesunden sind. So ist nicht hinreichend klar, ob die dadurch erzielten Veränderungen irreversibel sind.

Diese skeptische Einschätzung mag angesichts der bereits weiten Verbreitung dieser Mittel erstaunen. Die folgenden Überlegungen sind vor diesem Hintergrund hypothetisch zu verstehen. Ich werde ein idealisiertes Psychopharmakon annehmen (ich spreche daher im Folgenden von „Prozac“ als Platzhalter): *Wenn* wir davon ausgehen, dass es synthetische Psychopharmaka gibt oder geben wird, die wirksam und ohne süchtig zu machen zur stabilen Aufhellung oder Veränderung emotionaler Eigenschaften von Gesunden eingesetzt werden können, *dann* gilt es, sie auf ihre mögliche Rechtfertigbarkeit zu prüfen.<sup>13</sup> Mir geht es darum zu fragen, inwiefern Psychopharmaka – sofern keine empirischen Gründe, die ihre Wirksamkeit bzw. Nebenwirkungen betreffen, gegen sie sprechen – als Mittel zur Verbesserung unserer emotionalen Befindlichkeit legitim sind. Und diese Frage ist besonders relevant, wenn wir davon ausgehen, dass die Wirkungen nachhaltig sind.

Es geht folglich nicht darum, das Ziel der Steigerung unserer psychischen Leistungsfähigkeit selbst zu hinterfragen. Es scheint zumindest eine weit geteilte Auffassung zu sein, dass wir nicht tadelnswürdig sind, wenn wir uns um eine Verbesserung unserer affektiven Verfassung bemühen. Schliesslich kümmerten sich Menschen seit jeher und in vielfältiger Weise um ihr eigenes Glück. Was fragwürdig scheint ist vielmehr, ob pharmakologische Mittel geeignet sind, dieses Ziel auf angemessene und damit rechtfertigbare Weise zu befördern.<sup>14</sup>

Da es aus empirischen Gründen nicht möglich ist, die genauen Wirkungen pharmakologischer Mittel zu bewerten, werde ich verschiedene mögliche Szenarien beschreiben und auf die Frage hin analysieren, welche emotionalen Veränderungen überhaupt zu einer möglichen Verbesserung (und wenn ja, welcher Art der Verbesserung) führen. Ob derartige pharmakologische Eingriffe als Mittel zur individuellen

Funktionssteigerung rechtfertigbar sind, soll nun im Folgenden genauer erörtert werden.<sup>15</sup> Dürfen bzw. sollen sich Marina und Beate auf diese Weise selbst umgestalten? Und warum sollte die Frage nach der Rechtfertigbarkeit derartiger Bemühungen überhaupt gestellt werden?

Eine Antwort auf diese Frage ist nicht nur für Marina und Beate oder für uns als Personen relevant, die ein Interesse daran haben, ein selbstbestimmtes und gelungenes Leben zu führen. Sie ist v.a. relevant für diejenigen, die derartige Massnahmen auf staatlicher oder professioneller Ebene mit verantworten und entsprechend legitimieren müssen.<sup>16</sup> Es geht folglich darum, einen Massstab zu finden, anhand dessen *Neuro-Enhancement* der Emotionen als prinzipiell erlaubt oder verboten klassifiziert werden kann. Es ist hierbei eine weit geteilte, wenn auch kaum begründete Auffassung, dass die Autonomie oder Selbstbestimmung<sup>17</sup> der Person der geeignete Massstab für eine solche Klassifizierung ist. Dieser Auffassung mag die Vorstellung zugrunde liegen, dass eine gewissermassen freiwillig und autonom gewählte Verbesserung der eigenen emotionalen Befindlichkeit, die dann aber die Autonomie der betreffenden Person letztlich untergräbt, selbstwidersprüchlich und insofern nicht rechtfertigbar ist. Dies scheint u. a. auch dann der Fall zu sein, wenn Autonomie als zentraler Wert betrachtet wird, der integraler Bestandteil eines guten Lebens ist.

Um genauer herauszufinden, wie emotionales *Neuro-Enhancement* ggf. die Autonomie einer Person erhält bzw. sogar erhöht, muss noch genauer untersucht werden, warum sich die Frage der Rechtfertigung solcher Massnahmen in brisanter Weise stellt. Warum könnte das Ansinnen von Marina und Beate zum einen problematisch, zum andern aber auch in besonderer Weise angezeigt sein?<sup>18</sup> Vor diesem Hintergrund möchte ich zum zweiten ausführen, welche Erwägungen überhaupt das Potential besitzen, derartige Massnahmen zu rechtfertigen. Um Autonomie als eine angemessene normative Quelle solcher Massnahmen zu verteidigen, muss nicht zuletzt auch im Gegensatz zur bisherigen Debatte um emotionales *Neuro-Enhancement* verdeutlicht werden, um welchen Begriff von Autonomie es geht, inwiefern es sich um einen Wert und somit um eine Quelle der Rechtfertigung handelt, und welche Art der Autonomie in Frage steht. Erst dann kann untersucht werden, welche konkreteren Veränderungen an den emotionalen Einstellungen von Personen ihre Autonomie erhöhen oder unterminieren.

## 2. Neuro-Enhancement der Emotionen als Form der Selbst-Veränderung

Die Frage nach der Rechtfertigbarkeit der Einnahme von Psychopharmaka zur sog. Aufhellung von Stimmungen erhält ihre besondere Brisanz nicht nur dadurch, dass dies – im Gegensatz zu manch anderen *Enhancement*-Massnahmen (man denke z.B. an Eingriffe in die menschliche Keimbahn) bereits eine gängige Praxis mit zunehmender Verbreitung darstellt. Sie ist insofern eine besonders realistische Option. Die Frage nach ihrer Rechtfertigbarkeit stellt sich meines Erachtens aber auch v.a. deshalb, weil *Neuro-Enhancement* der Emotionen weiterreichende Konsequenzen hat, scheint doch nach Auffassung vieler eine Veränderung der ganzen Persönlichkeit bzw. eine Art Selbst-Transformation auf dem Spiel zu stehen. Ich möchte daher in einem ersten Schritt versuchen zu klären, in welcher Weise eine Veränderung der emotionalen Befindlichkeit eine Veränderung der Persönlichkeit oder des Selbst bewirken kann. Worin bestehen genau die weiterreichenden Konsequenzen solcher Eingriffe in unsere affektive Verfassung? Warum muss überhaupt gezeigt werden, dass derartig weitreichende Veränderungen nicht verboten sind? Um mich diesen Fragen anzunähern, ist es hilfreich zu verstehen, welche Arten von Veränderung *nicht* gemeint sein können.

So handelt es sich im Fall von emotionalem *Neuro-Enhancement* nicht um eine momentane oder punktuelle Steigerung von Lebensfreude, wie dies z.B. durch eine durchtanzte Nacht oder durch ein Glas Sekt erreicht werden kann.<sup>19</sup> Ebenso wenig handelt es sich um eine Veränderung einer einzigen, episodisch vorkommenden emotionalen Einstellung. So soll Marina nicht eine spezifische Angst (wie etwa die Angst vor Spinnen) oder eine für eine bestimmte Dauer anhaltende Depression (wie etwa die Schwangerschaftsdepression) verlieren. Es geht folglich weder um eine Veränderung der emotionalen Wahrnehmung anlässlich eines bestimmten Objekts, das in der Wahrnehmung unangemessen repräsentiert wird, noch um eine Veränderung einer zeitlich begrenzten emotionalen Befindlichkeit. Diese Überlegungen könnten es nahe legen zu meinen, es ginge um eine Veränderung von Stimmungen. Schliesslich sind Stimmungen im Gegensatz zu punktuell auftretenden und zeitlich begrenzten Emotionen stabilere und länger andauernde affektive Einstellungen.<sup>20</sup>

Doch Stimmungen werden in der Regel von Emotionen dadurch unterschieden, dass ihnen kein intentionaler Gehalt zugeschrieben wird. Sie richten sich auf nichts, sondern gelten als nicht-intentionale Einstellungen, die mit spezifischen Erwartungen über zukünftige Lust und zukünftigen Schmerz einhergehen. Wer z.B. melancholisch gestimmt ist, fühlt einen nicht fokussierten negativen und damit leicht schmerzhaften Affekt, der damit einhergeht, Schlechtes zu erwarten. Die Unterscheidung zwischen Emotionen und Stimmungen ist jedoch weder trennscharf noch handelt es sich in den genannten Beispielen allein um die Aufhellung einer Stimmung, sondern ebenso um die Veränderung emotionaler Einstellungen mit intentionalem Gehalt (wie etwa die Missachtung von netten Männern im Fall von Marina, oder die Beendigung der Ungeduld gegenüber ihren Kindern im Fall von Beate).

Dies gibt zu der Vermutung Anlass, dass ein engerer Zusammenhang zwischen Stimmungen und Emotionen besteht: Wenn Beate immer wieder episodische Wut anlässlich nicht von ihr kontrollierbarer Verhaltensweisen anderer empfindet, mag dies zu ihrer anhaltenden Stimmung der Gereiztheit führen. Umgekehrt mag diese Stimmung bewirken, dass sie auf derartige Stimuli (wie eben die nicht kontrollierbaren Verhaltensweisen anderer) mit bestimmten Emotionen, wie Wut, reagiert und diese zu einer stabilen und andauernden Emotion wird.<sup>21</sup>

In den genannten Beispielen scheinen zum einen Stimmungen im Sinne bestimmter Dispositionen zu verschiedenen emotionalen Einstellungen verändert zu werden. Zum andern geht es um die Veränderung ganzer Komplexe oder Muster emotionaler Einstellungen gegenüber verschiedenen Objekten. So führt Beates gereizte Stimmung nicht nur dazu, dass sie disponiert ist, häufig und relativ schnell Wut zu empfinden. Sie scheint aufgrund dieser Stimmung auch zu einer ganzen Reihe weiterer negativ wertenden Emotionen geneigt. Sie wird sich u. a. schnell gekränkt, ängstlich und traurig fühlen. Desweiteren scheint eine solchermassen gestimmte Person disponiert, andere Emotionen, wie etwa Stolz, Freude oder Übermut *nicht* zu empfinden. Ihre Stimmung macht sie also zu bestimmten Emotionen disponiert. Dies kann zum einen eine negative Grundstimmung verstärken, zum andern aber ein Muster robuster emotionaler Einstellungen ausmachen.

Sofern es sich um stabile Dispositionen zu bestimmten Emotionen sowie um anhaltende Muster emotionaler Einstellungen handelt, wird plausibel, inwiefern die Veränderung dieser Emotionen persönlichkeits-



verändernde Wirkungen haben kann. Die Person wird verändert, da sie ihre umfassendere emotionale Wahrnehmung und damit ihre Perspektive von sich und ihrer Umwelt ändert. Dies bedeutet jedoch nicht, dass *Neuro-Enhancement* der Emotionen die Persönlichkeit in jeder Hinsicht verändert. So bleiben etwa Beates zentrale Vorlieben, viele ihrer evaluativen Überzeugungen, ihr Temperament, ihre kognitiven Eigenschaften und ihr diesbezügliches Selbstverständnis bestehen.<sup>22</sup> Die Veränderung ihrer emotionalen Befindlichkeit verändert folglich nur einen Teil ihrer Persönlichkeit. Dieser wird jedoch auch deshalb verändert, weil sich eine Person wie Beate nicht nur anders fühlen, sondern auch als andere verstehen möchte: So möchte sie nicht nur aufhören, immer gereizt und wütend zu reagieren. Sie möchte z.B. anderen Personen mehr Wert beimessen als ihr mit ihrer derzeitigen emotionalen Verfassung möglich ist. Insofern geht eine Veränderung ihrer emotionalen Befindlichkeit mit einer Veränderung ihres Selbstverständnisses bzw. ihrer normativen Identität einher: sie möchte sich als Person verstehen, die sich mit den ihr wichtigen Werten oder wertvollen Eigenschaften identifiziert und diese entsprechend (und durch ihre Emotionen reflektiert) wertzuschätzen vermag.

Wenn jedoch *Neuro-Enhancement* der Emotionen solch weitreichende Veränderungen bewirkt, dann muss die zunächst offene Frage beantwortet werden, ob eine solche Veränderung extensional äquivalent mit einer Verbesserung ist. Eine Veränderung kann schliesslich nur dann als *Enhancement* im Sinne einer „Steigerung“ betrachtet werden, wenn diese als eine bestimmte Art der Verbesserung charakterisiert werden kann. Schliesslich können einzelne emotionale Eigenschaften verändert werden und dennoch kann immer die informative Frage gestellt werden, ob eine partikulare Veränderung gut bzw. besser ist als der Ausgangszustand. Wenn wir beweisen wollen, dass sozial und selbstsicher zu sein besser ist als scheu und ängstlich zu sein, dann müssen wir erklären, wofür *Neuro-Enhancements* gut sind, bzw. was genau sie besser machen, oder inwiefern sie schlecht oder falsch sind.<sup>23</sup> Nur dann können wir mit einer begründeten Antwort auf die für Philosophinnen und Philosophen virulente Frage aufwarten, ob wir ein *Neuro-Enhancement* in einem partikularen Fall vornehmen lassen sollen, d. h., ob es erlaubt oder verboten ist. Im Fall des *Neuro-Enhancement* der Emotionen scheint eine solche Verbesserung für viele am plausibelsten als (noch näher zu charakterisierende) gesteigerte Autonomie verständlich.

Sofern *Neuro-Enhancement* der Emotionen tiefgreifende Veränderungen der normativen Identität bewirkt, stellt sich die Frage nach der Rechtfertigung solcher Massnahmen mit besonderer Brisanz. Im Folgenden wird es darum gehen, genauer zu zeigen, warum Autonomie sich als Quelle der Rechtfertigung anbietet und welcher Begriff der Autonomie hierfür geeignet ist.

### 3. Autonomie als Quelle der Rechtfertigung

Kritiker ebenso wie Befürworter von *Neuro-Enhancement* der Emotionen rekurrten bisher selbst meist auf Begriffe wie „Authentizität“ oder „Autonomie“, um derartige Massnahmen zu verteidigen oder abzulehnen. Diese Begriffe werden jedoch häufig unscharf gebraucht.<sup>24</sup>

Beide Lager scheinen hierbei anzunehmen, dass das *Enhancement* der eigenen emotionalen Eigenschaften und damit die tatsächliche Veränderung der eigenen normativen Identität dann erlaubt sein kann, wenn diese durch die Person in einer näher zu klärenden Weise selbst autorisiert ist, die Veränderungen ihrem Selbst entsprechen und ihre Autonomie dadurch erhalten oder sogar gesteigert wird.<sup>25</sup> Und in der Tat sprechen folgende Überlegungen dafür, den Beitrag von Massnahmen des *Neuro-Enhancement* zur Erlangung bzw. Steigerung von Autonomie zu prüfen und hinsichtlich ihres diesbezüglichen Beitrags zu rechtfertigen: Zum einen müssen die Gründe, die eine Entscheidung zur Veränderung der eigenen emotionalen Befindlichkeiten stützen könnten, das reflektieren, was die betreffende Person selbst als *für sich gut* erachtet und somit durch ihre eigene freie Aktivität, d.h. durch die eigene Überlegung und Wertschätzung, und nicht durch äussere Manipulation zustande gekommen ist. Nur was sie selbst (in einem näher zu definierenden Sinne) wirklich wertschätzt – als Ausdruck ihrer Wertschätzung gilt häufig, dass sie sich damit „identifiziert“ – und was ihr in diesem Sinne „eigen“ ist, kann eine begründende Auskunft darüber geben, wie sie sich verändern soll bzw. wie sie sich zu verstehen hat und wer sie sein möchte. Andernfalls würde sie ein Leben führen, das in einem emphatischen Sinne nicht *ihres* ist. Sie würde sich entfremdet und damit als passiver Zuschauer dessen fühlen, was sie tut. Nennen wir dies die „Wertungsbedingung“ von Autonomie.

Zum andern kann nur das zu ihrer Autonomie beitragen, wodurch sie sich selbst zu führen vermag. Sie muss sich an Gründen orientieren und durch sie ihr Leben, d.h. sich und ihr Tun kontrollieren können, und zwar auch dann, wenn ihre momentanen Neigungen und unmittelbaren Wertschätzungen dagegen sprechen. Andernfalls würde sie ein instabiles und nicht verständliches Leben führen.<sup>26</sup> Nennen wir dies die „Kontrollbedingung“ von Autonomie.

Ich möchte es an dieser Stelle offen lassen, ob diese beiden Bedingungen gemeinsam notwendig und hinreichend sind, um den Begriff der Autonomie zu definieren. Ebenso offen möchte ich lassen, ob der Begriff überhaupt definierbar ist. Ich gehe an dieser Stelle einfach davon aus, dass die beiden genannten Bedingungen weit geteilte Intuitionen ausdrücken, die unserem Verständnis von Autonomie zugrunde liegen.

In der bisherigen Debatte werden die den Begriffen der Authentizität und Autonomie zugrunde liegenden Intuitionen in einer bestimmten Weise interpretiert. Dies führt dazu, dass in der Regel mehr oder weniger explizit eine bestimmte substantielle Theorie der Autonomie vertreten wird, die Selbstveränderung – verstanden als Veränderung der eigenen normativen Identität – entweder begrifflich ein- oder ausschliesst.

So rekurrieren Kritiker von *Neuro-Enhancement*-Massnahmen in der bisherigen Debatte häufig auf den Begriff der Authentizität.<sup>27</sup> Gründe, die eine Entscheidung für *Neuro-Enhancement* stützen, müssten dann gemäss der oben genannten Wertungsbedingung das reflektieren, was eine Person als für sich gut erachtet. Dies wird schliesslich so interpretiert, dass es sich hierbei um evaluative Haltungen bzw. bestimmte konative Einstellungen handeln muss, die die betreffende Person bereits hat.<sup>28</sup> Damit ist eine statische Auffassung der Person und ihrer normativen Identität verbunden, der zufolge sie ihre stabilen Einstellungen oder Eigenschaften zu akzeptieren hat. Die oben genannte Kontrollbedingung wird so aufgefasst, dass eine Person sich nur durch das stabil zu kontrollieren vermag, was sich gegenüber jedem etwaigen Veränderungsversuch als resistente Einstellung erweist. Sie übt durch ihre wesentlich unverfügbaren Eigenschaften, die es zu erkennen oder zu entdecken gilt, Kontrolle über ihr Tun aus.

Marina könnte in diesem Fall nur ihre stabilen<sup>29</sup> emotionalen Zustände registrieren. Sie wäre demzufolge sich selbst treu, wenn sie scheu und ängstlich bliebe. Sie kontrolliert ihr Tun, indem sie sich durch diese unverfügbaren Eigenschaften bestimmen lässt und diese allenfalls akzeptiert.

Diese Auffassung liegt mehr oder weniger explizit der folgenden Kritik an emotionalem *Neuro-Enhancement* zugrunde: So führt etwa George Elliott aus: „Es wäre beängstigend, wenn Prozac meine Persönlichkeit veränderte, selbst wenn es mir eine bessere Persönlichkeit bescherte – beängstigend, weil es nicht mehr *meine* Persönlichkeit wäre. Diese Art der Persönlichkeitsveränderung scheint einer Ethik der Authentizität entgegenzustehen.“<sup>30</sup> Der „President’s Council on Bioethics“ warnt davor, durch *Neuro-Enhancement* unsere Identität zu kompromittieren, die uns von Natur aus gegeben sei.<sup>31</sup> Anjan Chatterjee befürchtet eine „Erosion des Charakters“.<sup>32</sup> Und auch Peter Kramer – der generell darauf verweist, dass Stimmungsaufheller die Autonomie einer Person stärken – fragt sich: „Die Autonomie welcher Person sollen wir jetzt bewahren?“<sup>33</sup> Sofern die gewünschte Persönlichkeit eine ist, die die Werte der hochtechnisierten kapitalistischen Welt repräsentiert, wie Selbstvertrauen, Flexibilität und Schnelligkeit, kann sie Kramer zufolge Ausdruck von Selbstentfremdung sein.<sup>34</sup> Vor diesem Hintergrund wird deutlich, dass Gegner von *Neuro-Enhancement* eine Theorie der Autonomie vertreten, die auf der weiteren Intuition der Treue zu sich selbst und der Aneignung der die betreffende Person stabil kennzeichnenden Eigenschaften basiert. Diese Eigenschaften sind dann trivialerweise ihr eigen, sofern Aussagen der Art „Marina ist scheu und ängstlich“ ihre normative Identität beschreiben und, sofern sie zutreffen, wahr sind. Eine Veränderung der eigenen emotionalen Dispositionen ist vor dem Hintergrund einer solchen durch Authentizität charakterisierten Theorie der Autonomie ausgeschlossen. Schliesslich würden wir auf diese Weise unsere „Identität“ so verändern, dass wir uns selbst nicht mehr treu sein könnten.

Befürworter von *Neuro-Enhancement* vertreten dagegen eine alternative Theorie der Autonomie. Sie interpretieren die Wertungsbedingung so, dass eine Person ihre Werte wählen kann. Dies ist Ausdruck der liberalen Auffassung, der zufolge jeder sich nach selbst gewählten Werten neu gestalten kann. Die Wertungsbedingung ist dann erfüllt, wenn eine Person Werte selbst und nach ihren eigenen Vorstellungen wählt und kein äusserer Zwang vorliegt.<sup>35</sup>

Befürworter glauben daher, dass wir unser Leben selbstbestimmter gestalten können, indem wir nun selbst über das, was uns unveränderbar auszeichnet, verfügen und uns entsprechend unserer eigenen Werte verändern können.<sup>36</sup> *Neuro-Enhancement* vermag demzufolge eine Per-

son gerade autonom zu machen. Die Kontrollbedingung ist schliesslich dann erfüllt, wenn es der betreffenden Person auch gelingt, sich mit dem neu gewählten Wert oder einem wertvollen Ziel zu identifizieren.<sup>37</sup> So kommen etwa Jan Bublitz und Reinhard Merkel zu dem Ergebnis, dass selbst veranlasste pharmakologische Eingriffe die Autonomie der Person nicht untergraben: „sofern sie [die Person] sich mit ihrer Persönlichkeit identifiziert, kann sie sein, was immer sie will“.<sup>38</sup> Neil Levy bemerkt im gleichen Tenor, dass das wahre Selbst dasjenige Selbst ist, das tatsächlich existiert, und zwar „so lange die Veränderung erhalten wird und nachhaltig ist, tief genug geht – d. h., solange wir ihm [dem wahren Selbst] die neue Persönlichkeit genuin zuschreiben können – gibt es keinen Grund, sie nicht als seine authentische zu betrachten.“<sup>39</sup> Befürworter von emotionalem *Neuro-Enhancement* favorisieren eine Theorie der Autonomie, die auf der Intuition der freien Selbstgestaltung ruht.

Diese sich widerstreitenden Intuitionen der Treue zu sich selbst sowie der Selbstgestaltung oder Selbstverwirklichung,<sup>40</sup> die unterschiedlichen Theorien der Autonomie zugrunde liegen und deren jeweilige Sanktion von emotionalem *Neuro-Enhancement* dieses als Mittel zur Selbstveränderung erklären oder verbieten, werden überdies durch eine bestimmte Auffassung über die Gründe, die eine sich selbst bestimmende Person hat, gestützt.

So ist es aufschlussreich, dass eine Theorie der Autonomie, die auf einer Auffassung von Authentizität basiert, zu einer internalistischen Auffassung von Gründen neigt. Gründe sind bestimmte mentale Zustände, die eine Person zu einem bestimmten Zeitpunkt hat. Es geht hierbei darum, die richtigen konativen Einstellungen zu spezifizieren, die sich als *ihre eigenen* Gründe qualifizieren.<sup>41</sup> Eine Theorie der Autonomie, die auf der Intuition der Selbstgestaltung beruht, neigt – wenn auch weniger explizit und weniger ausschliesslich – zu einer externalistischen Auffassung von Gründen. Der zufolge reagiert eine Person auf Tatsachen in der Welt, die sie sich – sofern sie sie wertschätzt – als ihre Gründe wahrnimmt und sich entsprechend führt und ggf. umgestaltet. Beate beobachtet folglich nicht ihre mentalen Zustände, sondern denkt über Tatsachen und Sachverhalte in der Welt als ihre möglichen Gründe aus ihrer erstpersionalen Perspektive nach.<sup>42</sup> Vor dem Hintergrund ihres Selbstverständnisses, eine ungeduldige und leicht reizbare Person zu sein, nimmt Beate z. B. Gründe wahr, ihr Leben so zu gestalten, dass sie

sich souveräner, emphatischer und ausgeglichener fühlt. Sie bestimmt somit die wertvollen Eigenschaften, mit denen sie sich in Zukunft identifizieren möchte. Sofern ihr dies tatsächlich gelingt und sie diese Werte als ihre Gründe akzeptiert, hat sie sich anhand dieser selbst bestimmt.

Im Folgenden möchte ich weder eine bestimmte Auffassung über den metaphysischen Status von Gründen noch eine substantielle Theorie der Autonomie vertreten. Mir ging es bisher darum zu zeigen, dass Autonomie ein normativer Begriff ist, der wesentlich umstritten bzw. bestreitbar ist.<sup>43</sup> Der Begriff der Autonomie lässt demzufolge rivalisierende Theorien zu, die dem Begriff unterschiedliche Bedeutungen verleihen, für die eigens argumentiert werden muss. Nicht nur, dass eine solche eigene Verteidigung der jeweiligen Theorie der Autonomie in der Debatte um emotionales *Neuro-Enhancement* ausbleibt.<sup>44</sup> Die Voraussetzung einer bestimmten Theorie ist zudem mit dem Problem einer zirkulären Rechtfertigung konfrontiert:

Jede Massnahme von *Neuro-Enhancement* als Mittel zur Veränderung der eigenen normativen Identität (und in diesem Sinne seiner selbst) wäre demzufolge deshalb erlaubt, weil in der Theorie der Autonomie, die diese Massnahme rechtfertigen soll, die normative Relevanz von Selbstveränderung vorausgesetzt wird. Die jeweilige Theorie der Autonomie, die einer bestimmten Verwendung des Begriffs „Autonomie“ oder „Authentizität“ unterliegt, legt den Umfang des Begriffs so fest, dass er voraussetzt, was er erst zeigen soll: dass *Neuro-Enhancement* als Mittel zur Selbstveränderung autonom oder eben gerade nicht autonom (im Sinne von authentisch) macht, weil eine bestimmte Theorie der Autonomie Selbstveränderung impliziert oder nicht. Um als normative Quelle von *Neuro-Enhancement* zu fungieren, darf folglich nicht eine bestimmte Theorie von Autonomie und damit ein präziser Begriff von Autonomie vorausgesetzt werden, der substantielle Annahmen über Selbstveränderung enthält.<sup>45</sup> Ich gehe hierbei davon aus, dass die verschiedenen, wesentlich sich widerstreitenden Theorien der Autonomie einen allgemeineren Begriff (und damit eine allgemeinere Bedeutung) der Autonomie teilen, der in der eingangs formulierten allgemeinen Wertungs- und Kontrollbedingung besteht.

Um Massnahmen des emotionalen *Neuro-Enhancement* nicht-zirkulär zu begründen, ist nur ein Autonomiebegriff geeignet, der auf substantielle Annahmen über Selbstveränderung verzichtet. Um jedoch des Weiteren zu zeigen, welcher Begriff der Autonomie dies tatsächlich

leistet, muss dargelegt werden, warum Autonomie überhaupt ein Wert ist.<sup>46</sup> Nur dann kann Autonomie nämlich als Quelle der Rechtfertigung ausgewiesen werden. Warum genau sollten Marinas und Beates Entscheidungen zur Einnahme von Prozac mit Verweis auf ihre Autonomie legitimiert werden? Inwiefern ist Autonomie ein Wert, der das Potential hat, die Veränderung seiner selbst und damit die Veränderung der eigenen emotionalen Dispositionen zu rechtfertigen?

#### 4. Der Wert der Autonomie

Sofern es erstrebenswert ist, ein selbstgewähltes Leben zu führen und für sich selbst wertvolle Ziele zu verfolgen, und sofern es erstrebenswert ist, so zu leben, dass wir uns stabil und für uns und andere verständlich orientieren können, ist Autonomie für uns wertvoll.

Um jedoch zu zeigen, dass Autonomie wertvoll ist und insofern eine rechtfertigende Kraft besitzt, bietet sich ein Vergleich mit einem anderen, viel bemühten grundlegenden Wert – nämlich demjenigen des Wohlergehens – an. Schliesslich glauben nicht wenige, dass eine Konzeption des Wohlergehens die normative Ressource ist, die Aufklärung darüber gibt, was für eine Person bezüglich ihres eigenen Lebens gut ist zu tun.<sup>47</sup> Eine *Neuro-Enhancement*-Massnahme wäre dann rechtfertigbar, wenn sie das Leben der betreffenden Person besser macht. Dies scheint intuitiv so plausibel, dass man sich fragen kann, warum nicht Wohlergehen als der fundamentale Wert fungiert, der eine *Neuro-Enhancement*-Massnahme zu rechtfertigen erlaubt. Schliesslich könnte Autonomie auch ein für Wohlergehen konstitutiver Wert beigemessen werden, sofern die Tatsache, dass eine Person ihre eigenen wertvollen Ziele wählt, verfolgt und wertschätzt dazu beiträgt, ein gutes Leben zu haben. Dass wir selbst uns aktiv bestimmen, ist dann Voraussetzung dafür, dass bestimmte Dinge oder Werte gut für uns sind bzw. werden.<sup>48</sup> Ob Autonomie als ein Wert aufgefasst wird, der für Wohlergehen konstitutiv ist, hängt von der jeweiligen Theorie des Wohlergehens ab.<sup>49</sup> Ich kann in diesem Rahmen keine Theorie des Wohlergehens verteidigen. Es geht mir lediglich darum zu zeigen, inwiefern Autonomie (und nicht Wohlergehen) als die relevante normative Quelle für Massnahmen des *Neuro-Enhancement* betrachtet werden kann. Um dies nun näher zu klären, muss gezeigt werden, inwiefern sich der Wert der Autonomie

vom Wert des Wohlergehens unterscheidet, und inwiefern Autonomie unabhängig von seinem Beitrag zum Wohlergehen einer Person wertvoll sein kann.

Dies lässt sich verdeutlichen, wenn wir Fälle betrachten, in denen ein psychopharmakologischer Eingriff das Leben einer Person autonom, aber nicht gut, bzw. umgekehrt gut, aber nicht autonom macht. Dies zu zeigen ist nur dann möglich, wenn der Begriff des Wohlergehens nicht denjenigen der Autonomie impliziert.

Marina mag sich z.B. autonom dazu entschliessen, ein sozial ausgefüllteres Leben zu führen und sich von ihren sozialen Hemmungen zu befreien. Sie mag Eigenschaften entwickeln, die es ihr ermöglichen, offen auf andere zuzugehen und Freunde zu gewinnen. Dies mag jedoch ebenso dazu führen, dass sie andere Dinge, die ihr Leben bisher wertvoll gemacht haben, nicht mehr in derselben erfüllenden Weise verfolgen kann. So mag sie sich weniger damit identifizieren, einen guten Beruf zu haben. Möglicherweise muss sie in Kauf nehmen, als Anwältin weniger erfolgreich zu sein. Ebenso mag es sein, dass sie durch ihre Freunde grosses Leid erfährt, etwa weil sie sie verletzen, oder weil sie mit ihrem Schicksal mitfühlt. Auch wenn sie dann kein (insgesamt) besseres Leben führt, kann sie mit einiger Berechtigung für sich reklamieren, zumindest ein autonomes Leben zu führen.

Umgekehrt kann das Leben von Beate gut verlaufen, selbst wenn sie durch blossе Androhungen ihrer Familie dazu veranlasst wurde, ihre Gereiztheit und Ungeduld behandeln zu lassen und sich somit nicht autonom zu einer solchen Veränderung entschieden hätte. Selbst wenn sie durch eine solche Massnahme ihre grundsätzliche Fähigkeit verlöre, sich selbst zu bestimmen, könnte sie ein gutes Leben führen, in dem sie zufrieden ist und viele wertvolle Ziele realisiert.

Diese Möglichkeiten machen deutlich, dass Autonomie einen von Wohlergehen unabhängigen Wert besitzt, der u. a. darin besteht, sich aktiv, ohne Zwang und aus eigenen Gründen auf wertvolle Ziele festzulegen und sich entsprechend zu führen.<sup>50</sup> Massnahmen des emotionalen *Neuro-Enhancement* können gerade deshalb abzulehnen sein, weil sie zwar das Wohlergehen der Person befördern, aber ihre Autonomie unterminieren. Sie können rechtfertigbar sein, wenn sie der Autonomie der Person dienen, wenngleich sie ihr Leben möglicherweise in anderer Hinsicht schlechter bzw. nicht besser machen.<sup>51</sup> Dass wir Ziele wertschätzen und uns entsprechend führen, hat somit einen



Wert, der sowohl unabhängig von den Gründen ist, die für diese Ziele sprechen als auch von der Tatsache, dass diese unserem Wohlergehen dienen.<sup>52</sup>

Und sofern es das Wohlergehen von Beate und Marina ist, auf das als grundlegender Wert zur Rechtfertigung für etwaige Massnahmen des *Neuro-Enhancement* rekurriert wird, so scheint dies deshalb der Fall, weil die Autonomie der beiden Personen als Bestandteil ihres Wohlergehens aufgefasst wird. Vor kurzem ist jedoch eingewendet worden, dass *Neuro-Enhancement* mit Rekurs auf die Autonomie der Person gerade dann nicht rechtfertigbar sei, wenn Autonomie das Wohlergehen der betreffenden Person unterminiert. So können etwa zu viele Gestaltungsmöglichkeiten eine Person überfordern. Aus diesem Grunde, so der Einwand, sei Autonomie gerade kein Wert, der *Neuro-Enhancement* zu rechtfertigen vermöge.<sup>53</sup>

Doch diesem Einwand liegt ein bestimmter substantieller Begriff von Autonomie zugrunde. Wird Autonomie hingegen in dem hier vorgeschlagenen Sinn verstanden, dann schliesst die Wertungsbedingung gerade aus, dass eine Person mit zu vielen Optionen überfordert ist. Sie kann nur solche Optionen wählen, die sie überhaupt wertzuschätzen in der Lage ist. Einem weiteren Einwand zufolge könnten Autonomie und Wohlergehen als inkommensurable und potentiell konfligierende Werte betrachtet werden. Es sei hierbei keinesfalls entschieden, welcher dieser beiden zentralen Werte gewichtiger ist, um die Einnahme von Psychopharmaka zu rechtfertigen.<sup>54</sup> Selbst wenn Autonomie und Wohlergehen als inkommensurable, aber paritätische und somit gleichermassen fundamentale Quellen der Rechtfertigung betrachtet werden, so ging es mir lediglich darum zu zeigen, dass Massnahmen des *Neuro-Enhancement* plausibel verteidigt werden können, wenn sie die Autonomie der betreffenden Person (auch gegen ihr Wohlergehen in einem Sinn, der Autonomie nicht impliziert) erhöhen.

## 5. Minimale und maximale Autonomie

Es ist zwar eine notwendige Voraussetzung zur Rechtfertigung von emotionalem *Neuro-Enhancement*, dass die Autonomie, die dadurch befördert werden soll, einen Wert darstellt. Es muss jedoch darüber hinaus gezeigt werden, wie dieser Wert im Fall von emotionalem

*Neuro-Enhancement* realisiert werden kann. Zunächst werde ich eine weitere relevante Unterscheidung zwischen zwei Arten der Autonomie einführen und spezifizieren, welche Art befördert werden soll, und auf welche Weise dies möglich ist. In einem weiteren Schritt muss dann gezeigt werden, welche einzelnen Massnahmen emotionaler Veränderung denkbar sind, und welche dieser Massnahmen zur Realisierung der so spezifizierten Autonomie tatsächlich geeignet sind.

Beginnen wir mit der Unterscheidung zwischen zwei Arten von Autonomie: Damit wir überhaupt von *Neuro-Enhancement*, und damit von der Veränderung der emotionalen Befindlichkeiten gesunder Menschen sprechen können, müssen wir davon ausgehen, dass diese bereits über eine grundlegende Fähigkeit zur Autonomie verfügen. Andernfalls würde es sich um Personen handeln, die in einem psychiatrischen Sinne krank sind. So werden Menschen mit schweren psychischen Störungen oder geistig Behinderte häufig so definiert, dass ihnen eine grundlegende Fähigkeit zur Autonomie fehlt. Zugleich kann die autonome Entscheidung gesunder Personen für eine durch Psychopharmaka bewirkte Selbstveränderung mit Verweis auf eine Steigerung ihrer Autonomie (und damit meist auch ihres Wohlergehens, sofern Autonomie als konstitutiv für Wohlergehen betrachtet wird) begründet werden. Autonomie als Ziel oder Ergebnis einer Entscheidung muss dann als ein persönliches Ideal betrachtet werden, das sich von der Fähigkeit zur Autonomie als Bedingung einer Entscheidung unterscheidet. Es ist folglich wichtig zu beachten – und auch dies wird in der bisherigen Debatte wenig explizit gemacht – dass *Neuro-Enhancement* der Emotionen nur dann gerechtfertigt werden kann, wenn sich die prä-transformierte Person aufgrund ihrer Fähigkeiten und Eigenschaften autonom zu einer Veränderung ihrer selbst entscheiden kann. Zugleich muss die erfolgte Veränderung sowohl von der post-transformierten Person selbst als auch von der drittpersonalen Perspektive als Stärkung ihrer Autonomie bzw. als Erwerb einer neuen Autonomie anerkannt werden können.

*Neuro-Enhancement* scheint demzufolge dann eine Verbesserung darzustellen, wenn es zum einen der Ausdruck der Autonomie einer Person ist und zum andern dazu beiträgt, ihre Autonomie zu erhöhen. Massnahmen dieser Art sind folglich dann gerechtfertigt, wenn sie die Autonomie der Person mindestens erhalten bzw. idealiter steigern.

Es erweist sich in diesem Zusammenhang als hilfreich, die Unterschiede zwischen einer minimalen Konzeption der Autonomie als Bedingung

des Entscheidens von einer maximalen Konzeption der Autonomie, die als persönliches Ideal und damit als Ziel der eigenen Selbstverwirklichung fungiert, zu spezifizieren.<sup>55</sup> Diese Unterscheidung impliziert hierbei keine substantiellen Annahmen über Selbstveränderung.

Autonomie als minimale Bedingung einer Entscheidung besteht u. a. darin, dass eine Person hinreichende kognitive Fähigkeiten, hinreichende Rationalität sowie ein basales Verständnis ihrer selbst sowie ihrer Welt besitzt. Sie muss in der Lage sein, sich realistische und selbstgewählte Ziele zu setzen, die sie aufgrund ihrer evaluativen Überzeugungen oder Einstellungen ohne inneren und äusseren Zwang auszuwählen vermag.<sup>56</sup>

Autonomie als persönliches Ideal besteht darüber hinaus darin, dass es einer Person tatsächlich gelingt, ihre Ziele wertzuschätzen und sich selbst entsprechend zu bestimmen und zu führen. Als persönliches Ideal ist Autonomie folglich ein Erfolgsbegriff: es muss sich bewahren, dass die betreffende Person in der Lage ist, Ziele wertzuschätzen und sich durch diese zu kontrollieren. Ob ihr dies gelingt, hängt auch von kontingenten Faktoren ab<sup>57</sup> und kann sich erst im Verfolgen ihrer Ziele über die Zeit erweisen.

Zu einer solchen maximalen Konzeption der Autonomie gehört eine hinreichend stabile Identifikation mit dem, was eine Person wertschätzt und ein zunehmendes Verständnis dessen, was wertvoll für sie ist, sowie die Fähigkeit, ihr Tun entsprechend zu kontrollieren. Ebenso gehört dazu, auf neue, für sie evaluativ relevante Erwägungen zu reagieren und in ihre Lebensführung zu integrieren sowie eine zunehmend verbesserte Beurteilung, unter welchen Umständen sie bisherige Wertschätzungen verändern und neuen Anforderungen anpassen soll. Um sich selbst zu führen, bedarf es eines Verständnisses dessen, dass man sich aus Gründen führt. Ein solches Verständnis impliziert, dass man seine Gründe als rechtfertigbare Gründe betrachtet und insofern Rationalitätsbedingungen bewusst akzeptiert. Eine solche Konzeption lässt jedoch offen, ob eine Person sich in wesentlicher Weise verändern muss, um idealiter autonom zu sein.

Um im Folgenden prüfen zu können, ob Massnahmen des *Neuro-Enhancement* einer Person helfen, sich einem Ideal der Autonomie anzunähern, erweist sich noch eine weitere Unterscheidung als hilfreich. Sie betrifft den Unterschied zwischen Autonomie als Zustand, der einer Person zugeschrieben wird, und Autonomie als Fähigkeit einer Person.

Im Zustand minimaler Autonomie besitzt Beate lediglich die (noch nicht aktualisierte) Fähigkeit, sich autonom zu bestimmen. Sie verfügt über die bereits genannten Voraussetzungen und besitzt überdies das Potential, (i) sich festzulegen und wertend zu binden; (ii) auf Gründe zu reagieren; sowie (iii) ihr Verhalten zu kontrollieren und diese Kontrolle zu erhalten. Im Zustand maximaler Autonomie – und ich gehe hierbei davon aus, dass es sich um ein Kontinuum zwischen minimaler und maximaler Autonomie handelt – hat eine Person ihre Fähigkeit zur Autonomie aktualisiert. Sie hat sich dem Ideal der Autonomie insofern angenähert, als sie ihre Fähigkeit zur Autonomie erfolgreich umgesetzt hat: Sie hat sich dann wertend gebunden (oder „identifiziert“) und ihr Verhalten entsprechend kontrolliert. Ihr Zustand ist jedoch nur autonom, wenn ihre Fähigkeit zur Autonomie nicht nur aktualisiert ist, sondern auch weiterhin erhalten bleibt. Dies bedeutet, dass die betreffende Person weiterhin in der Lage sein muss, auf Gründe zu reagieren und ihre Wertungen entsprechend auch wieder zu hinterfragen.

Kehren wir zu einem unserer Beispiele zurück, um zu verdeutlichen, inwiefern eine Person sich einer maximalen Konzeption der Autonomie nur anzunähern vermag, wenn sie ihre Fähigkeit zur Autonomie erhält. Wenn Beate sich durch die Einnahme eines Antidepressivums emotional verändert, geduldig und empathisch wird, dann hat sie ihre Fähigkeit zur Autonomie folgendermassen aktualisiert: sie hat sich auf evaluative Eigenschaften festgelegt und schätzt sie nach der Veränderung auch wirklich wert; dies zeigt sich u. a. auch darin, dass sie ihr Verhalten entsprechend anpasst und mit ihrem Mann z. B. anders umgeht als bisher. Zugleich mag sich ihre Identifikation auch darin zeigen, dass sie Zufriedenheit in ihrer Bindung an diese wertvollen Eigenschaften erfährt. Dieser verbesserte Zustand ihrer Autonomie ist jedoch nur dann autonomieerhaltend, wenn sie weiterhin die Fähigkeit behält, auf neue Gründe zu reagieren. Wenn sie z. B. durch das Antidepressivum so determiniert werden würde, dass sie sich an jede dadurch bewirkte Veränderung wertend bindet und entsprechend kontrolliert, hätte sie ihre Autonomie genau aus diesem Grunde nicht erhalten. Nehmen wir an, Beate findet eines Tages heraus, dass ihr Mann sie schon seit einigen Jahren betrügt. Ihre relevanten Umstände haben sich dann in einer Weise geändert, dass sie keinen Grund mehr hat, geduldig und verständnisvoll mit ihrem Mann zu sein.

Diese Überlegungen sollen lediglich verdeutlichen, dass Massnahmen

von *Neuro-Enhancement* nur im Hinblick darauf rechtfertigbar sind, ob sie die Autonomie der betreffenden Person verbessern oder fördern. Ich habe hierbei versucht zu zeigen, dass dies nur dann der Fall ist, wenn sie ihre Fähigkeit zur Autonomie nicht nur besser aktualisiert, sondern auch erhält. Es geht folglich darum, ob die Veränderung einer Person sowohl ein autonomer Prozess ist als auch zu einem Ergebnis führt, dem Autonomie nicht nur punktuell, sondern für die weiteren Entscheidungen einer Person zugesprochen werden kann.

Es bleibt nun zu prüfen, welche Arten der emotionalen Veränderung überhaupt denkbar sind und inwiefern diese zu einem Ideal der Autonomie beitragen, das die soeben eingeführte Fähigkeit zur Autonomie erhält.

## 5. Welche emotionalen Veränderungen sind Verbesserungen?

Im Fall emotionaler Einstellungen scheint es nicht unmittelbar klar, inwiefern ihr *Enhancement* als Verbesserung zu verstehen ist.<sup>58</sup> Im Folgenden möchte ich daher – vor dem Hintergrund eines allgemeinen Autonomie-Begriffs und angesichts der Überlegung, dass eine Person nur dann ihre Autonomie steigert (und sich so dem Ideal der Autonomie annähert), wenn sie ihre oben beschriebene Fähigkeit zur Autonomie aktualisiert und zugleich erhält – untersuchen, inwiefern eine emotionale Veränderung einen Beitrag zur Autonomie der Person leisten kann.

Aufgrund der unzureichenden Kenntnisse über die genaue Wirkweise von Psychopharmaka, wie Prozac, werde ich zunächst verschiedene mögliche Szenarien skizzieren.<sup>59</sup> Ich werde dann der Frage nachgehen, welche Art der emotionalen Veränderung plausiblerweise autonomiefördernd ist und zu zeigen versuchen, inwiefern sie dies tut. Ich gehe hierbei davon aus, dass Emotionen mentale Zustände sind, die die folgenden drei Eigenschaften besitzen:<sup>60</sup> Sie sind (i) Gefühle, die durch ihren hedonischen Empfindungscharakter ausgezeichnet sind. Emotionen fühlen sich auf eine bestimmte Weise für die Person, die sie hat, an, und gehen mit verschiedenen physiologischen Veränderungen einher (wie etwa einer erhöhten Herzfrequenz, „weichen Knien“, „Kloss im Hals“ u. ä.).<sup>61</sup> Emotionen sind (ii) kognitive Einstellungen, die einen intentionalen Gehalt haben. Sie repräsentieren u. a. Objekte, Personen

oder Ereignisse, auf die sie gerichtet sind. Darüber hinaus sind sie mit anderen kognitiven Einstellungen, wie z.B. Überzeugungen, Urteilen und Wünschen verbunden. Emotionen haben (iii) eine spezifische epistemische bzw. evaluative Rolle. Sie vermitteln uns Zugang zu Werten bzw. sind wertende Einstellungen, die die jeweilige Perspektive der Person und die Bedeutung, die ein emotional repräsentiertes Objekt oder Ereignis für die betreffende Person hat, ausdrücken. Vor dem Hintergrund dieser relativ unkontroversen,<sup>62</sup> wenn auch noch sehr allgemeinen Annahme, dass Emotionen durch diese drei Komponenten charakterisierbar sind, werde ich nun untersuchen, welche Veränderungen durch ein Psychopharmakon überhaupt vorstellbar sind.

*Szenario 1: Dämpfende Wirkung: Eingriff in die phänomenale Komponente*

Nehmen wir in einem ersten Szenario an, das Prozac die phänomenale Wahrnehmung von Emotionen dämpft. Positive ebenso wie negative Emotionen<sup>63</sup> werden auf diese Weise weniger intensiv gefühlt, wann immer eine einzelne emotionale Episode auftritt. Es geht folglich darum, dass die Emotionen begleitende Empfindung generell geschwächt wird. Dies hat insbesondere den Vorteil, dass negative Emotionen als weniger schmerzhaft wahrgenommen werden. Die davon betroffene Person wird auf diese Weise von den damit einhergehenden bedrohlichen Gefühlen befreit. Sie fühlt sich dann von der Kraft ihrer eigenen emotionalen Zustände nicht mehr so übermannt. Sie wird unbefangener und von der schweren Last negativer Gefühle entlastet auf ihre Mitmenschen zugehen und Ereignisse in ihrer Umwelt unvoreingenommener wahrnehmen können. Marina fühlt dann einfach ihre Angst und Unsicherheit nicht mehr in derselben bedrohlichen Weise. Beate wäre von den ihre Ungeduld und Gereiztheit begleitenden Gefühlen des Getriebenseins und Irritiertseins befreit. Die Dämpfung negativer Emotionen könnte insgesamt zu mehr innerer Ruhe führen.

*Szenario 2: Stärkung positiver Emotionen: Eingriff in die Repräsentation des Gehalts*

Da sich der intentionale Gehalt – das also, worauf wir uns emotional richten – nicht selbst verändern lässt, bleibt neben der reinen Dämpfung unserer Empfindungszustände die Möglichkeit, die Art und Weise, wie wir uns auf einen bestimmten Gehalt richten, zu verändern. In einem

zweiten Szenario kann man sich entsprechend der obigen Analyse von Emotionen vorstellen, dass die Art, wie ein Objekt oder Ereignis repräsentiert wird und die jeweilige Person Zugang erhält, eine Modifikation erfährt. Um dies zu bewerkstelligen, bedarf es eines stabilen Mechanismus, der hilft, diesen Gehalt in einer gegenüber der bisherigen Wahrnehmung alternativen emotionalen Weise zu repräsentieren. Im Gegensatz zu Szenario 1 geht es daher nicht darum, generell weniger zu fühlen und somit die Intensität der phänomenalen Wahrnehmung (unabhängig vom Gehalt der Wahrnehmung) zu reduzieren, sondern ein wahrgenommenes Objekt oder Ereignis *anders* zu fühlen und damit anders wahrzunehmen.

Marina würde demzufolge dem zuvorkommenden Verhalten von Männern nicht mehr mit Missachtung begegnen, sondern mit Sympathie. Neue Herausforderungen im Beruf würde sie nicht mehr angstbesetzt wahrnehmen, sondern z. B. mit Neugier und Enthusiasmus. Beate würde den Missgeschicken ihrer Kinder nicht mehr mit Zorn, sondern mit nachsichtiger Zuneigung begegnen.

In diesem Szenario würde Prozac dazu führen, einen anderen Zugang zum Gehalt der jeweiligen Emotionen zu erwerben und das Wahrgenommene in einem anderen, d. h. positiven Wert zu deuten. Positive Emotionen zu empfinden könnte pharmakologisch so gestärkt werden, dass die bisher vorherrschenden negativen Emotionen nicht mehr (oder kaum noch) aktualisiert werden können. Ereignisse oder Objekte, auf die eine Person sich emotional richten kann, werden aufgrund der verstärkten positiven emotionalen Wahrnehmung optimistisch gedeutet. Der Vorteil einer solchen Veränderung wäre nicht nur, dass sich die Person aufgrund ihrer positiven Emotionen besser fühlt. Ihr Verhalten würde auch zu mehr sozialem Erfolg führen. Ihr Zugang zur Realität wäre möglicherweise verbessert und ihr Selbstvertrauen könnte auf diese Weise gestärkt werden.

### *Szenario 3: Elimination von Stimmungen: Eingriff in die wertende Perspektive*

Wie ich bereits in der Unterscheidung zwischen Emotionen und Stimmungen andeutete, können episodisch wiederkehrende negative Emotionen zu einer negativen Stimmung der Person führen. Dies ist häufig bei traumatischen Erfahrungen der Fall. Aufgrund einer negativen Stimmung und der daraus resultierenden evaluativen Perspektive ist

eine Person dann geneigt, jeden möglichen Gehalt als Bedrohung zu interpretieren.

Nehmen wir daher in einem dritten Szenario an, dass Prozac nicht auf einzelne emotionale Vorkommnisse, sondern auf die Disposition, Emotionen zu fühlen, wirkt. Im Gegensatz zum ersten Szenario wird nicht die Empfindung einzelner emotionaler Episoden gedämpft, und im Gegensatz zum zweiten Szenario wird auch nicht die Art, wie ein Objekt oder Ereignis im Moment seiner Wahrnehmung repräsentiert wird, verändert. Vielmehr geht es darum, eine andauernde negative Stimmung zu beseitigen (oder zumindest zu schwächen). Dies kann ggf. damit einhergehen, eine positive Stimmung zu stärken sowie die momentane emotionale Wahrnehmung (wie im Szenario 2) positiv zu wenden.

Da Marina selbst überzeugt ist, dass sie einen guten Partner verdient, verhilft ihr die Elimination ihrer widerständigen Emotion, der zufolge sie nicht liebenswert ist, und die aus ihrer in der Kindheit erworbenen negativen Stimmung resultiert, zu mehr mentaler Kohärenz und damit zu mehr Kontrolle über ihr Tun. Beate erlebt dann einzelne Ereignisse nicht mehr vor dem Hintergrund ihrer ebenfalls aufgrund vergangener Erfahrungen erworbenen Stimmung, überfordert zu sein.

Die Elimination der Disposition, negative Emotionen zu empfinden, hat den Vorteil, dass viele irrationale Emotionen, die sich einer in der Vergangenheit erworbenen Stimmung verdanken, aber nicht durch den Gehalt, auf den sie sich im einzelnen richten, angemessen gerechtfertigt sind, vermieden werden. Denn sofern negative emotionale Einstellungen die Objekte, auf die sie sich richten, nicht angemessen repräsentieren – entweder weil sie mit einer zu starken, dem Objekt nicht angemessenen Intensität einhergehen, oder weil sie auf Objekte reagieren, die überhaupt keine solche emotionale Einstellung rechtfertigen – verdanken sie sich einer irrationalen Genese. Dies hat den Vorteil, dass innere Konflikte zwischen den bewussten Überzeugungen einer Person und ihren widerständigen, häufig zwanghaften Emotionen gelöst werden können. Zugleich ist mit einer Elimination der Disposition nicht ausgeschlossen, dass eine Person im Einzelfall negative Emotionen erlebt, wenn diese einem Objekt angemessen sind.

Ich gehe im Folgenden davon aus, dass diese drei Szenarien relevante Möglichkeiten darstellen, wie ein Antidepressivum wirken kann. Sie



reflektieren hierbei die in der Philosophie der Emotionen weit vertretene Auffassung über die drei Komponenten von Emotionen.<sup>64</sup>

Vor diesem Hintergrund kann die Frage, welches dieser Szenarien die Autonomie der betreffenden Person befördert, genauer beantwortet werden. Erinnern wir uns, dass hierbei besonders zu beachten ist, ob die jeweilige, in den einzelnen Szenarien dargestellte Veränderung (i) die grundsätzliche Fähigkeit, sich festzulegen und sich entsprechend wertend zu binden, behindert; (ii) die Reaktionsfähigkeit auf Gründe verunmöglicht; oder (iii) die Kontrolle über das eigene Tun unterwandert.

Die Dämpfung der Gefühle im ersten Szenario scheint die Autonomie der Person zu untergraben. Bei einer allgemeinen Abschwächung der Empfindungskomponente von Emotionen, liegt es nahe anzunehmen, dass auch positive Emotionen dieser Dämpfung unterliegen. Eine solche Wirkweise hätte zur Folge, dass nicht nur auf das lustvoll empfundene Gefühl positiver Emotionen weitgehend verzichtet wird, sondern dass auch negative Emotionen, die angemessen sind (wie etwa die Furcht vor einem vor mir befindlichen Raubtier), kaum bzw. zu wenig der Wahrnehmung zugänglich sind.<sup>65</sup> Darüber hinaus führt eine generelle Abschwächung emotionaler Empfindung zu weniger gefühlter Zufriedenheit. Emotionen nicht bzw. nur schwach zu fühlen verhindert bzw. erschwert den Zugang zu Werten. Robert Nozick hat in seinem Gedankenexperiment der Erlebnismaschine versucht zu zeigen, dass Lust allein um der positiven Empfindung willen nicht wertvoll ist.<sup>66</sup> Analog zu Nozicks Erklärung, dass zusätzlich zu unseren Empfindungen der Zugang zur Realität einen eigenen Wert besitzt, verhindert auch Schmerzvermeidung allein um der Vermeidung von negativen Gefühlen willen einen solchen Zugang. Die mit einer solchen affektiven Dämpfung erzielte Ausgeglichenheit lässt uns den intentionalen Gehalt von Emotionen nicht mehr angemessen fühlen und die Realität wird nicht mehr angemessen emotional repräsentiert. Marina empfindet auch dann keine Scheu mehr, wenn diese ihrer Situation angemessen wäre. So wird von Carl Elliott von einer scheuen japanischen Frau berichtet, die nach der Einnahme von Prozac nur noch mit Mühe von ihren Freunden davon zurückgehalten werden kann, auf einer Feier halbnackt aufzutreten.<sup>67</sup> Auf diese Weise behindert die Dämpfung von Empfindungen, dass eine Person überhaupt etwas wichtig nehmen kann und sich entsprechend zu binden vermag. Die Person nimmt auf diese Weise sich selbst und das, was sie umgibt, in seiner Bedeutsamkeit weniger wahr

und kann ihr Tun dadurch weniger kontrollieren. So verstanden würde *Neuro-Enhancement* unserer Emotionen durch die Dämpfung unserer Empfindungen unsere normativen Maßstäbe in einer Weise verschieben, die Realität unangemessen wahrnehmen lässt.

Die Verstärkung positiver Emotionen führt im zweiten Szenario zwar zu einer verbesserten Festlegung und entschiedeneren Kontrolle des eigenen Tuns. Aber der Zugang zu Gründen scheint auf diese Weise nicht gegeben. Im Gegenteil, die bloße Verstärkung positiver Emotionen kann die Wahrnehmung bestimmter Objekte und Ereignisse in ihrem jeweiligen Wert systematisch verzerren. Eine enthusiastisch gewordene Marina mag dann keinen Zugang mehr zu Gefühlen der Angst und Unsicherheit haben, wenn diese angemessen wären. Die jeweils systematisch veränderte Repräsentation verunmöglicht einen angemessenen Zugang zur Aussenwelt.

Im dritten Szenario scheint die Kohärenz der mentalen Einstellungen ebenso erhöht wie die Reaktionsfähigkeit auf Gründe. Die Elimination oder zumindest Abschwächung der Disposition, negative Emotionen zu empfinden, erlaubt es Marina, auf Objekte ihrer Wahrnehmung unvoreingenommener und somit angemessener zu reagieren. Sie vermag sich besser selbst zu kontrollieren, da sie nicht mehr systematisch konfligierenden Einstellungen – nämlich denjenigen, die von ihrer grundsätzlichen Disposition, alles negativ zu werten und von ihren aktuellen Emotionen und Überzeugungen stammen – ausgesetzt ist. Zugleich ist die Intensität ihres Fühlens nicht beeinträchtigt und es gelingt ihr auf diese Weise und mithilfe kohärenter Einstellungen sich festzulegen und entsprechend ihrer Festlegungen zu orientieren. Sie vermag ohne die Disposition, systematisch negative Emotionen zu empfinden, einzelne emotionale Vorkommnisse angemessen wahrzunehmen.

Die Veränderung emotionaler Eigenschaften kann vor dem Hintergrund meiner Überlegungen zur maximalen Autonomie dann als Verbesserung betrachtet werden, wenn folgende autonomiefördernde Wirkungen erzielt werden: (i) die verbesserte Wahrnehmungsfähigkeit von Objekten und Umständen, die eine emotionale Reaktion verdienen, und damit die Verbesserung unserer Mechanismen, emotional Gründe aufzuspüren, (ii) die verstärkte Neigung, angemessene und kohärente Emotionen zu fühlen, und (iii) die gesteigerte Fähigkeit, seine emotionalen Einstellungen im Fall eines Irrtums zu korrigieren.

Emotionale Eigenschaften werden folglich nicht dadurch verbessert, dass wir einfach mehr Emotionen fühlen oder negative Emotionen prinzipiell überwinden, um immer im Zustand emotionaler Ausgeglichenheit und innerer Ruhe zu sein. Es hat sich vielmehr ergeben, dass letztlich die Steigerung der Rationalität unserer emotionalen Einstellungen eine Verbesserung in dem Sinne ist, dass nur sie die Autonomie einer Person befördert: Diese betrifft sowohl die Reaktion auf Gründe in der Welt mittels unserer emotionalen Wahrnehmung, als auch den kohärenten Zusammenhang zwischen unseren Emotionen und anderen Einstellungen.

Rationale Emotionen stellen auch aus folgenden Gründen eine Verbesserung dar: Eine Person, die rational verbundene Emotionen hat und deren Emotionen auf Gründe reagieren, hat mehr Zugang zu Werten und macht selbst intensivere und daher wertvollere Erfahrungen, indem sie Gefühle hat: positive Gefühle sind hierbei selbst angenehm und negative Gefühle geben ihr die Sicherheit (sofern sie auf Gründe reagieren), die Welt und sich selbst angemessen wahrzunehmen und motivieren Veränderung. Kohärente Emotionen erlauben es Personen, sich aus ganzem Herzen zu binden und ihr Handeln zu kontrollieren.<sup>68</sup> Und dies sind genau die Bedingungen, die eine Person u. a. erfüllen muss, um ihre Fähigkeit zur Autonomie zu bewahren und sich einem Ideal der Autonomie anzunähern.

## 6. Einwände und Präzisierungen

Mein Ergebnis ist folglich, dass nur rationale Emotionen die Autonomie einer Person stärken können. Ob *Neuro-Enhancement* der Emotionen erlaubt sein kann, hängt also davon ab, ob die dafür erforderlichen Massnahmen – in unserem Fall die Einnahme bestimmter Psychopharmaka – dazu dienlich sind, die Rationalität der emotionalen Einstellungen einer Person zu erhöhen. Psychopharmaka müssten folglich die Disposition zu negativen Emotionen und Stimmungen abschwächen, ohne einzelne emotionale Vorkommnisse phänomenal zu dämpfen oder deren Gehalt anders zu repräsentieren. Ob dies möglich ist, ist eine empirische Frage, die ich nicht zu beantworten vermag. Zumindest kann ich zeigen, dass Psychopharmaka, die nachweislich Emotionen nicht rationaler machen und somit die Autonomie einer Person untergraben, nicht verordnet

werden sollten. Doch selbst wenn wir hypothetisch davon ausgehen, dass es Psychopharmaka gibt, die eine solche Rationalitätsbefördernde Wirkung bzgl. unserer emotionalen Einstellungen haben, müssen die bestehenden Einwände gegen derartige Massnahmen noch genauer betrachtet werden. Wie sich zeigen wird, ist die Rationalitätssteigerung der Emotionen durch *Neuro-Enhancement* – selbst wenn diese gelingt – noch nicht hinreichend, um die Autonomie der Person maximal zu befördern.

Bisher wurde u. a. vor Antidepressiva zu Zwecken des *Neuro-Enhancement* gewarnt, weil sie eine unmittelbare Manipulation darstellen, die unsere rationalen Fähigkeiten vollständig umgehen.<sup>69</sup> Statt den Gehalt unserer negativen Emotionen als Gründe anzuerkennen, die der Veränderung bedürfen – so die Kritik – werden unsere schmerzhaften Emotionen durch *Neuro-Enhancement* einfach eliminiert.<sup>70</sup> Und statt uns selbst als verantwortliche und damit rationale Akteure zu betrachten, behandeln wir uns selbst mechanistisch wie blosse Dinge.<sup>71</sup> Diese Kritik wäre jedoch ungültig, wenn eine Person sich nicht nur autonom zur Einnahme von Psychopharmaka entschieden hat. Sie greift auch nicht, sofern Psychopharmaka tatsächlich unsere Emotionen rationaler machen können. Trotz des „mechanistischen“ Eingriffs würde uns dies dazu verhelfen, unsere Gründe emotional besser wahrzunehmen und kohärentere Einstellungen zu haben.<sup>72</sup> Und sofern diese Rationalitätssteigerung unserer Emotionen möglich ist, wäre auch die Autonomie der Person besser aktualisiert. Die Person hätte sich besser wertend gebunden und könnte sich besser, da kohärenter, kontrollieren. Darüber hinaus bliebe ihre Fähigkeit zur Autonomie erhalten, da die betreffende Person weiterhin auf veränderte Gründe zu reagieren vermag. Dieser Einwand gegen *Neuro-Enhancement* kann daher entkräftet werden, da die rationalen Fähigkeiten einer Person gar nicht umgangen werden.

Es wird zudem kritisiert, dass die mechanistische Veränderung emotionaler Einstellungen mittels eines Antidepressivums gerade verhindert, dass eine Person aktiv und selbstbestimmt die Veränderung ihrer selbst herbeiführt. Entsprechend kann sie ihre Veränderung aufgrund der blossen Einnahme einer Pille auch gar nicht als eigene Leistung schätzen.<sup>73</sup> Dies wäre nur im Fall einer Psychotherapie möglich, in der eine Person sich aktiv und in einem kontinuierlichen Prozess eigener Anstrengung umgestaltet. Die Einnahme eines Medikaments wäre demgegenüber eine illegitime Abkürzung. So könnte der psychothera-

apeutische Prozess dazu führen, dass die betreffende Person zunehmend die Gründe für eine Veränderung ihrer selbst erkennt und mehr darüber herausfindet, was sie eigentlich wertschätzt.<sup>74</sup> Zudem scheint sie ihren Veränderungsprozess selbst zu kontrollieren, da sie die Psychotherapie auch abbrechen kann, oder ihn kontinuierlich selbst begleitet.

Bei genauerer Betrachtung scheint aber die Einnahme eines Antidepressivums sowohl zu ermöglichen, dass eine Person sich wertend bindet und den Prozess ihrer Umgestaltung kontrolliert. Sie kann die Einnahme absetzen, sie kann sich dazu autonom entscheiden und ihre Veränderungen kritisch beobachten. Die Tatsache, dass dies zunächst chemisch bewirkt wird, scheint auf den ersten Blick nichts auszutragen für die Steigerung der Autonomie. Die geäußerte Kritik ist daher nur dann haltbar, wenn die eigene Anstrengung, die im Fall einer Psychotherapie notwendig ist und im Fall der Einnahme eines Antidepressivums ausbleibt, selbst autonomiesteigernd ist.

Es scheint mir nicht von der Hand weisbar, dass eine Person, die sich unter Aufbietung ihrer Willenskraft und unter möglicher Zuhilfenahme psychologischer Hilfe selbst umgestaltet, stolz auf ihre Veränderung sein kann. Sie hat dann ihren Prozess der Selbstveränderung nicht nur initiiert, sondern auch begleitet und sich somit als eine konstante Urheberin ihres Tuns erfahren. Sie gewinnt auf diese Weise Einsicht in die Ursachen ihrer emotionalen Befindlichkeit sowie in die Möglichkeiten des Umgangs. Diese Erfahrung bleibt ihr im Fall der Einnahme von Psychopharmaka verwehrt und zeigt an, dass auch wertvolle und mit Autonomie assoziierte Erfahrungen (nämlich u. a. die Erfahrung der kontinuierlichen Selbstwirksamkeit, der Anstrengung und der Auseinandersetzung mit der eigenen Befindlichkeit) dadurch verloren gehen können.

Auch wenn *Neuro-Enhancement* als Massnahme zur Steigerung der Rationalität der Emotionen gerechtfertigt erscheint, bleibt mein Ergebnis daher vorsichtig. Die chemisch bewirkte Abkürzung kann zwar einen autonomen Zustand hervorrufen und sogar die Fähigkeit der Person, sich autonom zu kontrollieren, erhalten. Dies ist wohlgemerkt nur dann der Fall, wenn die Emotionen einer Person rationaler gemacht werden. Gleichwohl gibt es zwei andere Bedingungen der Autonomie, die durch Massnahmen des *Neuro-Enhancement* nicht oder zumindest weniger erfüllt sind. Nennen wir diese die Bedingung der Willensstärke sowie die Bedingung des kontinuierlichen Selbstverständnisses. Es mag

umstritten sein, ob es sich hierbei um konstitutive oder lediglich kausale Bedingungen von Autonomie handelt. In jedem Fall kann eine Person durch die Einnahme von Psychopharmaka weder ihre Willenskraft so stärken wie im Fall einer durch eine Psychotherapie bedingten Arbeit an sich selbst, noch kann sie ihre Persönlichkeit in derselben Weise verstehen lernen. Aus diesen Gründen ist die Einnahme von Psychopharmaka nicht grundsätzlich einer Psychotherapie vorzuziehen.

Gleichwohl mag in Fällen, in denen unsere Willenskraft einfach versagt – etwa in Fällen schwerer depressiver Verstimmungen – oder in Fällen, in denen die Auseinandersetzung mit der eigenen Biographie zu traumatisch ist, die Einnahme eines solchen Mittels uns zu mehr Autonomie und mehr Rationalität verhelfen und aus diesen Gründen gerechtfertigt sein. In diesen Fällen verschwimmt allerdings die Grenze zwischen *Enhancement* und Therapie und es ist nicht mehr klar, ob es sich dann nicht um eine Behandlung einer Krankheit handelt.<sup>75</sup>

Es gibt folglich keinen Anlass zu einem uneingeschränkten Plädoyer für die pharmakologisch induzierte Aufhellung unserer Stimmungen. Schliesslich müssen wir sogar aus konzeptuellen Gründen skeptisch bleiben, ob Psychopharmaka die Rationalität unserer Emotionen zu steigern vermögen. Bisher habe ich zu zeigen versucht, in welcher Weise Emotionen verändert werden müssen, um überhaupt eine Verbesserung der Autonomie der betreffenden Person zu erzielen. Ferner habe ich ausgeführt, dass nur ein weiter Begriff der Autonomie zur Rechtfertigung solcher Massnahmen dienlich ist. Es hat sich schliesslich ergeben, dass Psychopharmaka jedoch nur dann vorzuziehen sind zur Steigerung der Autonomie, wenn eine Person in besonderem Mass unfähig ist, sich einer Psychotherapie zu unterziehen.

Die konzeptuelle Skepsis bezieht sich darauf, dass es schwer vorstellbar zu sein scheint, pharmakologisch zwischen rationalen und irrationalen Emotionen unterscheiden zu können. Zwar wird von Bioethikern darauf hingewiesen, dass nur eine monistische Konzeption unseres Geists plausibel ist, der zufolge Rationalität ein neurobiologisch realisiertes Phänomen ist. Entsprechend kann die Rationalität unserer emotionalen Einstellungen auch durch pharmakologische Interventionen verändert und moduliert werden.<sup>76</sup> Die Tatsache, dass Rationalität eine neurobiologische Grundlage hat, impliziert jedoch nicht, dass sie darauf reduzierbar ist. Ebenso wenig ist damit nahegelegt, dass durch eine chemische Veränderung eine Verbesserung unserer Rationalität bewirkt

wird. Um zu prüfen, ob eine bestimmte Emotion auf Gründe reagiert, müsste der Gehalt der entsprechenden Emotion berücksichtigt werden. Dies scheint jedoch nur aus der Perspektive der ersten Person möglich, nicht jedoch aus der drittpersonalen Perspektive durch den chemischen Eingriff in unsere mentalen Zustände.<sup>77</sup>

Dieses Ergebnis muss jedoch als vorläufig betrachtet werden. Zum einen liegt dies daran, dass die konkreten Wirkungen von synthetischen Psychopharmaka nicht hinreichend klar sind. Ich habe lediglich versucht zu zeigen, welcher Art eine pharmakologisch bewirkte Veränderung der affektiven Verfassung einer Person sein müsste, um autonomiesteigernde Wirkungen zu haben. Es bedürfte jedoch weiterer pharmakologischer Erkenntnisse, um diese Möglichkeit überhaupt empirisch nachzuweisen. Des Weiteren werden die Bedenken gegen Massnahmen des *Neuro-Enhancements* verstärkt, wenn andere fundamentale Werte mit der Autonomie der Person kollidieren. Dies kann jedoch nicht prinzipiell ausgeschlossen werden. Wenn *Neuro-Enhancement* zu gravierenden sozialen Ungleichheiten führte, oder wenn dies zu einer starken Absenkung des Wohlergehens aller führte, dann wäre es keine hinreichend rechtfertigbare Massnahme.<sup>78</sup>

Ich habe mich in diesem Beitrag darauf beschränkt zu prüfen, inwiefern die Einnahme von Psychopharmaka aus der erstpersionalen Perspektive der betreffenden Person für sie selbst rechtfertigbar sein kann (und insofern auch von Dritten für sie empfohlen werden dürfen). Dafür scheint Autonomie die relevante normative Quelle zu sein. Inwiefern die Annäherung an ein Ideal der Autonomie mit anderen fundamentalen Werten, wie Verteilungsgerechtigkeit, kollidieren kann, muss Gegenstand einer eigenen Untersuchung sein.

Allerdings glaube ich, dass mein Verständnis von der rationalen Steigerung emotionaler Befindlichkeiten wenig Anlass zu solchen Wertkonflikten gibt. Dies liegt daran, dass Autonomie kein Wert ist, der die Realisierung anderer Werte ausschliesst. Wer autonomer ist, hat nicht notwendigerweise mehr Chancen oder mehr Zugang zu knappen Ressourcen, die anderen vorenthalten werden. Vielmehr haben wir verbesserten Zugang zu Gründen und können uns mehr selbst führen, weil kohärentere Einstellungen in einer für Festlegungen und Bindungen relevanten Weise diejenigen Gründe reflektieren, die unsere eigenen sind.

Ich habe zwar zu zeigen versucht, dass die pharmakologisch induzierte Veränderung emotionaler Einstellungen die Autonomie einer Person

nicht unterminiert, sofern die Rationalität ihrer Emotionen gesteigert wird. Auf diese Weise habe ich die Bedingungen spezifiziert, unter denen *Neuro-Enhancement* der Emotionen erlaubt sein kann. Da die Rationalität emotionaler Einstellungen jedoch nur aus der Perspektive der ersten Person korrigiert werden kann, bleibt es zweifelhaft, ob chemische Mittel dies prinzipiell zu bewirken vermögen. Was sie vielleicht bewirken ist, Bedingungen zu schaffen, damit unsere Fähigkeit zur Autonomie und damit die Rationalität unserer Emotionen wiederhergestellt werden, und insofern können sie für „weniger Gesunde“ anzuraten sein.<sup>79</sup>

### Anmerkungen

- 1 Beispiele der genannten Art finden sich bei P. Kramer (1993): *Listening to Prozac. A Psychiatrist Explores Antidepressant Drugs and the Remaking of the Self*, u. a. 22–46.
- 2 Ich werde im Folgenden den englischsprachigen Begriff beibehalten. Dieser Gebrauch hat sich in der deutschsprachigen Literatur eingebürgert. Dies liegt nicht zuletzt daran, dass eine Übersetzung, wie etwa „Steigerung“ oder „Verbesserung“ v. a. nicht allen Fällen von emotionalem *Neuro-Enhancement* gerecht wird.
- 3 Neben der Veränderung psychischer Zustände im Fall von *Neuro-Enhancement* geht es in anderen Formen des *Enhancement* v. a. um die Steigerung körperlicher Fähigkeiten und Eigenschaften. Zu nennen sind hier etwa die Steigerung der sexuellen oder sportlichen Leistungsfähigkeit (wie z. B. Muskelkraft und Ausdauer), die Verlängerung des Lebens, die Erweiterung der Sinneswahrnehmungen, die Steigerung des Wachstums oder der Schönheit.
- 4 Sofern Marina und Beate auf ein besseres als ein lediglich „normales“ Leben zielen, und sofern ein „normales“ Leben unbefriedigende emotionale Befindlichkeiten der in diesen Beispielen beschriebenen Art mit sich bringt, sind ihre Versuche, sich mittels Psychopharmaka emotional zu stärken, als *Neuro-Enhancement* zu deuten.
- 5 Vgl. E. Juengst (1998): Was bedeutet *Enhancement*? 25. Siehe auch N. Bostrom und A. Sandberg (2009): Die Weisheit der Natur? 86. Auch die Grenze zwischen „invasiv“ und „nicht-invasiv“ ist nicht ganz trennscharf.
- 6 So kann zum einen die Pathologisierung von „herkömmlicher“ Traurigkeit kritisiert werden; zum andern kann davor gewarnt werden, psychische Leiden, die häufig aufgrund kultureller Prägungen erst als solche wahrgenommen werden, nicht als reale Leiden aufzufassen. Siehe etwa A. Horwitz (2007): *The Loss of Sadness*. Darüber hinaus wird vor Versuchen der pharmazeutischen Industrie gewarnt, die neue Krankheiten definiert, um mehr Medikamente verkaufen zu können. Siehe dazu M. Scheermer et



- al. (2009): *The Future of Psychopharmacological Enhancements*, 78. Für die kulturelle Bedeutung von Krankheit und Leiden siehe etwa J. Hansen (2003): *Listening to People or Listening to Prozac?* 57–62.
- 7 Dies kann zum einen daran liegen, dass es sich um negative Emotionen handelt. Zum andern könnte eine Person jedoch auch an den negativen Folgen ihrer euphorischen Zustände leiden.
  - 8 Siehe etwa D. Talbot (2009): *Pharmakologisches Enhancement*, 69 ff.
  - 9 J. P. A. Ioannidis (2008): *Effectiveness of Antidepressants*, 1–9, äussert sich v. a. auch kritisch gegenüber den Verfahren bisheriger randomisierter Untersuchungen.
  - 10 Der momentane Forschungsstand scheint im Fall der Veränderung emotionaler Eigenschaften schlichtweg nicht ausreichend zu sein, um eindeutige Aussagen über die positive Wirkung entsprechender Psychopharmaka treffen zu können. Siehe D. Repantis (2009): *Die Wirkung von Psychopharmaka bei Gesunden*, 64 ff. Siehe auch S. Weinmann (2008): *Erfolgsmythos Psychopharmaka*, v. a. Kap. 1 und 3.
  - 11 So sprechen M. Scheermer et al. (2009): *The Future of Psychopharmacological Enhancements*, 77, davon, dass Antidepressiva eine normale Stimmung nicht zu einer „glücklichen“ zu machen vermögen, auch wenn ihren Erkenntnissen nach Prozac negative Affekte, wie Traurigkeit und Angst, zu reduzieren und prosoziales Verhalten zu befördern scheint.
  - 12 Dazu zählen sexuelle Dysfunktion, Gewalttätigkeit und die Neigung zum Suizid. Siehe hierzu etwa J. Glenmullen (2001): *Prozac Backlash*; sowie D. Healy (2006): *Let Them Eat Prozac*.
  - 13 Dies gilt umso mehr, wenn die einmal erzielten Veränderungen nicht mehr ohne weiteres reversibel sind. Und sofern sie reversibel sind, bleibt die Frage, ob eine mehrfache Veränderung der normativen Identität einer Person (etwa durch die Einnahme und erfolgende Veränderung durch ein Psychopharmakon und die Absetzung des Medikaments und die erneute Veränderung) andere negative Folgen hat. Sofern die erzielten Veränderungen leicht und ohne gravierende anderweitige Folgen wieder rückgängig gemacht werden könnten, wären die von mir angestellten Überlegungen zur Rechtfertigbarkeit dieser Massnahmen weit weniger bedeutsam. Gegen diese Befürchtung spricht jedoch schon die Tatsache, dass mehrfache Veränderungen der normativen Identität selbst schwerwiegende Folgen für das Selbstverständnis einer Person haben müssten.
  - 14 Ich danke einem anonymen Gutachter, der mich angeregt hat, den Unterschied zwischen der Verbesserung unserer affektiven Verfassung als Ziel und *Neuro-Enhancement* – d.h., der Verbesserung durch bestimmte Pharmaka – als Mittel zu einem solchen Ziel zu präzisieren.
  - 15 Im Rahmen dieses Beitrags befasse ich mich aufgrund ihrer Verbreitung und einfacheren Zugänglichkeit mit synthetischen Psychopharmaka. Es gibt jedoch auch andere *Neuro-Enhancement*-Massnahmen. So werden auch magnetische oder elektronische Stimulationsverfahren, Neuro-Chips oder Schnittstellen zwischen Computer und Gehirn z.T. bereits eingesetzt oder zeichnen sich für die Zukunft ab. Synthetische Pharmaka werden von anderen chemischen Substanzen wie z. B. Kokain unterschied-

- den. Für synthetische Pharmaka scheint zu sprechen, dass sie nicht (bzw. weniger) süchtig machen, am wenigsten invasiv und leicht verfügbar sind.
- 16 Dies betrifft z.B. die weitergehenden Fragen, ob Ärzte Psychopharmaka auch ausserhalb des medizinischen Indikationsbereichs verschreiben dürfen, ob derartige Medikamente etwa durch das Internet verfügbar sein dürfen, wie pharmazeutische Konzerne kontrolliert werden, und ob Forschung an *Neuro-Enhancement* gefördert werden soll. Desweiteren kann auch gefragt werden, ob etwa künftige Arbeitgeber derartige Massnahmen gar verlangen dürfen, und mit welchen Argumenten derartige Ansinnen unterbunden werden können.
  - 17 Ich verwende die Begriffe „Autonomie“ und „Selbstbestimmung“ synonym.
  - 18 Hierbei geht es ausschliesslich um Fälle, in denen Personen derartige Massnahmen an sich selbst durchführen lassen. Eine Antwort auf die Frage, ob *Neuro-Enhancement* gesetzlich erlaubt und durch Ärzte durchgeführt werden soll, ist davon abgeleitet.
  - 19 Auch Drogen, die eine bloss punktuelle Steigerung der Lebensfreude, aber keine stabile Veränderung der Persönlichkeit bewirken, sind keine Mittel zum *Enhancement*.
  - 20 Siehe z.B. R. Rosfort und G. Stanghellini (2009): *The Person In Between Moods and Affects*, 263.
  - 21 A. Stringaris (2009): *On Emotions That Last Longer*, 279, spricht hier von „enduring emotions“, die er als „ever-recurring ways of being disposed to certain objects“ beschreibt.
  - 22 Vgl. T. Galert (2009): *Wie mag Neuro-Enhancement Personen verändern?* 161, der sich als einer der wenigen mit dieser Thematik eingehend beschäftigt hat.
  - 23 R. Chadwick (2008): *Therapy, Enhancement and Improvement*, 33 ff., ist eine der wenigen, die explizit darauf verweist, dass *Enhancement* nicht notwendigerweise als *Improvement* zu verstehen ist. Auch M. Scheermer et al. (2009): *The Future of Psychopharmacological Enhancements*, 80, erwähnen kritisch, dass *Neuro-Enhancement* vorschnell mit „beneficence“ gleichgesetzt wird.
  - 24 „Authentizität“ wird in der Regel als eine bestimmte Auffassung von Autonomie vertreten. Ich selbst gehe im Folgenden von Autonomie als dem weiteren Begriff aus.
  - 25 Zuletzt hat J.-C. Heilinger (2010): *Anthropologie und Ethik des Enhancements*, bes. Kap. IV, ausgeführt, dass eine anthropologische Konzeption die normative Grundlage zur Rechtfertigung von *Enhancements* sein sollte. Selbstbestimmung ist hierbei zentraler Bestandteil dieser Konzeption.
  - 26 Zu den dem Autonomie-Begriff in einem weiten Sinn unterliegenden Intuitionen, siehe M. Oshana (1998): *Personal Autonomy and Society*, 81 f. Vgl. M. Oshana (2003): *How Much Should We Value Autonomy*, 99 ff.
  - 27 In der Regel impliziert die Verwendung dieses Begriffs eine essentialistische und meist statische Auffassung von Autonomie.
  - 28 Repräsentativ für diese Auffassung sind v.a. die späteren Arbeiten von H. Frankfurt. Siehe etwa H. Frankfurt (1997): *On Caring*, 155–180.

- 29 Da die emotionalen Zustände stabil sind, scheinen sie ihr Wesen auszumachen.
- 30 G. Elliott (1998): Die Tyrannei des Glücklichseins, 241.
- 31 President's Council on Bioethics (2003): Beyond Therapy, 300.
- 32 A. Chatterjee (2004): Cosmetic Neurology, 971.
- 33 P. Kramer (1993): Die Botschaft in der Kapsel, 192.
- 34 P. Kramer (1993): Die Botschaft in der Kapsel, 211.
- 35 Siehe D. DeGrazia (2000): Prozac, Enhancement und Selbstgestaltung, 250ff. Vgl. D. DeGrazia (2005): *Human Identity and Bioethics*, 112.
- 36 So verweisen Verfechter der These, dass *Enhancement* intrinsisch wertvoll ist, darauf, dass es sich um eine Optimierung unserer Natur handelt. Zum einen scheint dies deshalb der Fall, weil es sich um einen Ausdruck des menschlichen Strebens nach Perfektionierung und Selbstgestaltung handelt, zum andern, weil es als an sich wertvolle Erweiterung menschlichen Bewusstseins und menschlicher Erfahrung gilt, die dazu führen könnte, die Welt neu zu erleben. Siehe z. B. A. Caplan (2003): Ist besser das Beste? 167, demzufolge es zum „Kern des Menschen“ gehört, sich selbst zu verbessern.
- 37 Siehe J. Bublitz und R. Merkel (2009): Autonomy and Authenticity of Enhanced Personality Traits, 372: „if agents who possess the minimal autonomy capacities self-initiate neuroenhancements and then identify with the results, they are autonomous. If they are manipulated or do not identify with the results, then they are not autonomous.“ Vgl. D. DeGrazia (2005): *Human Identity and Bioethics*, 102. Siehe auch N. Juth (2011): Enhancement, Autonomy, and Authenticity, 42 f.
- 38 J. Bublitz und R. Merkel (2009): Autonomy and Authenticity of Enhanced Personality Traits, 372 [meine Übers.]. Vgl. T. Runkel (2010): *Enhancement und Identität*, 265–291, demzufolge *Neuro-Enhancement* erlaubt sein sollte, wenn dadurch eine „reflexive Identifikation des Selbst“ (ebd., 291) ermöglicht werden kann. Durch eine „reflexive“ Identifikation soll verhindert werden, dass einfach soziale Wertvorstellungen blind übernommen werden.
- 39 N. Levy (2007): *Neuroethics*, 108 [meine Übers.]. Vgl. H. Schmidt-Felzmann (2009): Prozac und das wahre Selbst, 147ff., die sechs Angemessenheitskriterien für Authentizität formuliert, die u. a. in kognitiver Angemessenheit, in autobiographischer Narration und Kongruenz bestehen.
- 40 Dass sich diese beiden „Ideale“ widerstreiten, bemerkt auch G. Elliott (1998): Die Tyrannei des Glücklichseins, 245. Vgl. M. Quante (2006): Selbst-Manipulation? 112 f., der von „Authentizitätsintuition“ und „Erarbeitungsintuition“ spricht.
- 41 Besonders zentrale Vertreter sind die meisten der sog. internalistischen oder kohärentistischen Theorien der Autonomie.
- 42 Die distinkte erstpersonale Perspektive, die im Fall normativer Identität eine wesentliche Rolle spielt, hat C. Budnik (2010): Die Perspektive der ersten Person und ihre Relevanz für Theorien personaler Identität, insbes. Kap. 8 und 9, herausgearbeitet.

- 43 Ich beziehe mich hier auf W. B. Gallie (1956): *Essentially Contested Concepts*, 167ff.
- 44 N. Janssen (2010): Der Authentizitätsbegriff in der Enhancementdebatte, 135 ff., hat fünf verschiedene Verständnisweisen des Authentizitätsbegriffs herausgearbeitet. Diese sind so unterschiedlich, dass sie die Tauglichkeit dieses Ideals zur Rechtfertigung von *Neuro-Enhancement* in Frage stellt.
- 45 Ich gehe davon aus, dass sowohl Kritiker als auch Befürworter eine für manche oder sogar viele Fälle verteidigbare Theorie der Autonomie vertreten. Im werde daher nicht für eine bestimmte Theorie zu Lasten einer anderen Theorie argumentieren, sondern den Begriff der Autonomie selbst so weit wie möglich fassen. Für eine Art Komplementarität verschiedener Theorie scheint auch J. Bublitz (2008): *The Value of Authenticity in the Age of Pharmaceutical Neuroenhancement*, 2, zu plädieren. E. Parens (2005): *Authenticity and Ambivalence*, 37ff. bietet eine Art Metaanalyse der verschiedenen Theorien, die im Fall der Debatte um *Neuro-Enhancement* vertreten wurden und kommt zu dem allerdings sehr vagen Ergebnis, dass die Ambivalenz zwischen beiden Theorien ein Zeichen der Offenheit ist (ebd., 38).
- 46 Nur N. Juth (2011): *Enhancement, Autonomy, and Authenticity*, 36, macht überhaupt darauf aufmerksam, dass nur Autonomie als Wert Massnahmen des *Neuro-Enhancement* zu rechtfertigen vermag. Eine nähere Analyse dieses Werts bleibt er jedoch schuldig.
- 47 Viele Befürworter gehen davon aus, dass *Enhancement* generell das individuelle Wohlergehen erhöht. Siehe etwa R. Ranisch und J. Savulescu (2007): *Ethik und Enhancement*, 49. Vgl. N. Bostrom (2008): *Why I Want to be a Posthuman when I Grow Up*, 120. Siehe auch B. Gesang (2007): *Perfektionierung des Menschen*, 93ff. Siehe T. Tännsjö (2009): *Ought We To Enhance Our Cognitive Capacities?* 421–432, der z.B. dagegen argumentiert, dass kognitives *Enhancement* Wohlergehen erhöht, jedoch auch von Wohlergehen als grundlegendem Wert ausgeht. Andere bemühen sich zu zeigen, dass *Enhancement* das kollektive Wohlergehen bzw. das Wohlergehen unseres Genpools erhöht. Vgl. A. Buchanan et al. (2000): *Why Not the Best?* 156–203. Vgl. auch N. Bostrom und A. Sandberg (2009): *Die Weisheit der Natur?* 85.
- 48 Diesen Aspekt betont C. Rosati (2006): *Personal Good*, 113.
- 49 Auch an dieser Stelle gehe ich daher von einem möglichst weiten Begriffsumfang (und nicht von einer bestimmten Theorie des Wohlergehens) aus. Die beiden Begriffe erweisen sich dann als überlappend, wenn auch nicht als koextensiv. Je nach dem, welcher Theorie des Wohlergehens man zugeneigt ist, könnte eine Person (zumindest im Rahmen einer subjektivistischen Theorie des Wohlergehens – und subjektivistische Theorien gehören zu den meist vertretenen) ihr Leben für gut befinden, auch wenn es ein schlechtes, weil z.B. nicht-autonomes Leben ist. Dies führt zu einer Reformulierung subjektivistischer Theorien, denen zufolge Autonomie als authentische Einstellungen Bestandteil des guten Lebens ist (siehe L. W. Sumner (1996): *Welfare, Happiness & Ethics*, 156–171). Andere führt dies zu einer Kritik an subjektivistischen Theorien des Wohlerge-

- hens. Siehe hierzu J. Hawkins (2008): Well-Being, Autonomy, and the Horizon Problem, inbes. 156ff. Sofern Autonomie als konstitutiv für Wohlergehen betrachtet wird, scheint man auf ein nicht-subjektivistisches Verständnis von Wohlergehen festgelegt. Siehe etwa C. Rosati (2006): Personal Good, 124f., die Wohlergehen begrifflich an Autonomie bindet: ein Leben ist dann gut, wenn es uns gelingt, unser Leben durch das, was wir selbst wirklich wertschätzen und selbst als wertvolle Ziele gewählt haben, zu führen. Die autonome Tätigkeit ist hierbei ein wesentlicher Teil dessen, was unser Leben gut macht.
- 50 M. Oshana (2003): How Much Should We Value Autonomy, 101, bestimmt Autonomie folgendermassen: „Being autonomous is not simply a matter of having values and preferences that mirror those a person holds under conditions in which control is absent. Rather, being autonomous is a matter of directing one’s life according to such values and preferences.“
- 51 Dies gilt, wie gesagt, wenn man einen Begriff des Wohlergehens verteidigt, der Autonomie nicht einschliesst.
- 52 S. Darwall (2006): The Value of Autonomy and Autonomy of the Will, 270ff., zufolge hat z.B. Autonomie einen persönlichen Wert für uns, sobald wir der Tatsache, dass Personen Ziele verfolgen, Wert zuschreiben. Seiner Meinung nach involviert der Wert der Autonomie sogar den Anspruch, Respekt von anderen für die eigene Autonomie einzufordern.
- 53 Siehe S. Nagel (2010): Too Much of a Good Thing? 111 ff.
- 54 Diesen Einwand verdanke ich Jon Leefmann und Thomas Schramme.
- 55 Bereits J. Feinberg (1986): Autonomy, 28, unterscheidet u. a. zwischen der Fähigkeit, sich selbst zu bestimmen und einem Charakterideal. Im Gegensatz zu Feinberg gehe ich davon aus, dass die Fähigkeit zur Autonomie mehr enthalten muss als rationale Entscheidungen zu treffen. Vgl. auch R. Merkel u. a. (2007): *Intervening in the Brain*, v.a. S. 337ff.
- 56 J. Bublitz und R. Merkel (2009): Autonomy and Authenticity of Enhanced Personality Traits, 361, sprechen von minimalen Fähigkeiten der Autonomie, machen jedoch den Unterschied zu einer maximaleren Konzeption nicht deutlich. So rubrizieren sie auch die Tatsache, dass sich eine Person mit den neu erhaltenen Eigenschaften identifiziert, als Bestandteil dieser minimalen Fähigkeiten. Da Identifikation mit einer erfolgten Veränderung jedoch nicht zu den Voraussetzungen für eine autonome Entscheidung zu einer solchen Veränderung gehört, fasse ich Identifikation als Bestandteil einer maximaleren Autonomie-Konzeption (und somit als Ziel von *Neuro-Enhancement*). L. Bolt (2007): True to Oneself? 296, plädiert dafür, dass Authentizität als „ideal of human flourishing“ eine relevante Rolle in der Debatte um *Neuro-Enhancement* spielen sollte, bleibt jedoch auch eine genauere Definition schuldig.
- 57 So wird in diesem Zusammenhang immer wieder auf ein hinreichendes Selbstwertgefühl, auf das Bestehen hinreichend vieler Optionen, auf funktionierende soziale Beziehungen und hinreichende ökonomische Bedingungen hingewiesen. Ob diese Bedingungen notwendig sind, um ein persönliches Ideal der Autonomie zu erlangen, kann ich hier nicht diskutieren.
- 58 N. Bostrom (2009): Why I Want to Be a Posthuman When I Grow up,

- 119 ff., gesteht z. B. zu, dass es schwieriger ist, emotionales *Enhancement* zu spezifizieren. Auch F. Krämer (2009): *Neuro-Enhancement von Emotionen*, 213, bemerkt, dass Emotionen nicht „steiger- oder verbesserbar“ sind.
- 59 Allerdings gehe ich davon aus, dass sie keine oder geringe Nebenwirkungen besitzen, und schliesse sozialpolitisch motivierte Vorbehalte etwa gegen die damit verbundene Unterstützung einer anrücklich gewordenen Pharmaindustrie aus.
- 60 Die Literatur zur Philosophie der Emotionen ist mittlerweile fast unüberschaubar geworden. Im Folgenden gehe ich von mittlerweile weit geteilten Einsichten über die Natur von Emotionen aus, ohne mich mit der Frage zu beschäftigen, wie die verschiedenen Charakteristika zusammenhängen.
- 61 Es handelt sich hierbei um vier verschiedene physiologische Merkmale: Gesichtsausdruck, muskuläre Veränderungen, Veränderung der Stimme sowie des Nervensystems. Siehe J. Deonna und F. Teroni (2008): *Qu'est-ce qu'une émotion?* 58.
- 62 Es gibt eine ausgedehnte Debatte innerhalb der Philosophie der Emotionen, was genau unter Emotionen zu verstehen ist. Ich kann mich in diesem Rahmen dieser Debatte nicht stellen. Im Folgenden gehe ich davon aus, dass die von mir angeführten Komponenten die plausibelsten Kandidaten für eine Definition emotionaler Einstellungen sind. Nicht alle Emotionstheorien teilen diese Auffassung.
- 63 Positive Emotionen sind solche, die eine billigende Wertung ausdrücken (z. B. Freude). Negative Emotionen sind solche, die eine ablehnende Wertung ausdrücken (z. B. Neid). Positive Emotionen werden häufig als angenehm, bisweilen als aufwühlend empfunden. Negative Emotionen gehen mit verschiedenen Arten unangenehmer bis hin zu schmerzhaften Gefühlen einher.
- 64 Ich möchte nicht ausschließen, dass es noch andere mögliche Szenarien gibt. Die drei genannten scheinen mir angesichts der drei Komponenten von Emotionen jedoch besonders naheliegend und plausibel. Für die Einschränkung auf diese drei Szenarien spricht überdies, dass eine feinkörnigere, auf neurowissenschaftlichen Erkenntnissen basierende Mikronalyse der Manipulationsmöglichkeiten von Emotionen nicht erlaubt zu verstehen, wie diese zu einer Veränderung der normativ zu verstehenden Makroebene – nämlich derjenigen der Erhöhung der Autonomie einer Person – beizutragen vermögen. Hierzu bedarf es bereits einer normativ anbindungsfähigen Analyse von Emotionen, die grobkörniger ist als diejenige der Neurowissenschaften.
- 65 Eine Dämpfung wäre lediglich dann autonomiefördernd, wenn die Emotionen der betreffenden Person prinzipiell von zu starken Empfindungen begleitet werden. Ich danke Christian Budnik, der mich auf diese Möglichkeit aufmerksam gemacht hat.
- 66 Siehe R. Nozick (1974): *Anarchy, State, and Utopia*, 42–45.
- 67 Siehe C. Elliott (2003): *Better Than Well*, 74.
- 68 Die genauen Rationalitätsbedingungen von Emotionen sowie der Zusammenhang zwischen diesen Bedingungen sind nicht leicht auszubuchstabieren. Es scheint mir jedoch plausibel anzunehmen, dass Emotionen zum

- einen auf ihre sozial oder objektiv verstandene Angemessenheit (objektive Bedingung) als auch auf ihre Kohärenz (subjektive Bedingung) hin rational kritisiert werden können. Vgl. F. Kraemer (2010): *Authenticity Anyone?*
- 69 Diesen Einwand führt N. Levy (2007): *Neuroethics*, 75 ff., an.
- 70 Siehe B. Manninen (2006): *Medicating the Mind*, 100–105.
- 71 Diese Kritik führt C. Freedman (1998): *Aspirin for the Mind?* 140, aus. So lässt sich auch M. Sandel (2006): *The Case Against Perfection*, 26f. sowie 85 ff., verstehen, der v.a. den Drang zur Beherrschung kritisiert und dafür plädiert, die menschlichen Eigenschaften als uns gegebenes Geschenk zu betrachten.
- 72 In diesem Sinne warnt M. Rego (2005): *What Are (And What Are Not) the Existential Implications of Antidepressant Use*, 124, ganz richtig vor einer falschen Dichotomie zwischen Autonomie und biologischer Dysfunktion, wenngleich er dies nicht systematisch ausführt.
- 73 Siehe etwa E. Juengst (1998): *Was bedeutet Enhancement?* 37, der sich allerdings gegen diese Position wendet.
- 74 F. Svenaues (2009): *The Ethics of Self-Change: Becoming Oneself by Way of Antidepressants or Psychotherapy?* 176ff. scheint den Unterschied letztlich auch nicht für wesentlich zu halten, auch wenn er betont, dass Psychotherapie in stärkerem Masse „das ganze Selbst“ involviert. Auch P. Biegler (2010): *Autonomy and Ethical Treatment in Depression*, 179–189, vertritt die Auffassung, dass Psychotherapie im Gegensatz zu Psychopharmaka Selbsterkenntnis fördert, die ihrerseits Autonomie befördert.
- 75 M. Olsen (2006): *Depression, SSRIs, and the Supposed Obligation to Suffer Mentally*, 292, betont, dass es auf die Ursache der depressiven Verstimmung ankommt: „Depression, has a momentum that is separate from its initial cause.“ In diesem Fall macht es Sinn, Medikamente zu benutzen und die damit behandelten Emotionen nicht als Reaktion auf Gründe zu betrachten.
- 76 Siehe B. Schöne-Seifert (2006): *Pillen-Glück statt Psycho-Arbeit*, 289.
- 77 Wenn eine Person auf die Frage, warum sie ihre emotionale Befindlichkeit verändert hat, Gründe angeben kann, die ihre veränderten Emotionen als rational auszeichnen, scheint dies (zum Teil) diese Veränderung zu rechtfertigen. Vgl. K. Deligiorgi: *Happy Pills and Magic Potions*, 16, die auf die Bedeutung dieser Frage verweist.
- 78 Ich habe bisher Fragen bezüglich der Reichweite und Machbarkeit anderer *Enhancement*-Massnahmen sowie die Bewertung sozialer Folgen ebenso ausgeklammert wie Fragen, die die möglichen Objekte derartiger Massnahmen betreffen, wie etwa die Auswirkungen von *Enhancement* an unseren Nachkommen und gar am Genpool zukünftiger Generationen durch Keimbahntherapie.
- 79 Für wertvolle schriftliche Kommentare zu früheren Versionen dieses Beitrags danke ich herzlich Susanne Boshammer, Christian Budnik, Bernhard Gesang, Felicitas Kraemer, Jon Leefmann und Thomas Schramme sowie zwei anonymen Gutachtern. Sie haben mich vor vielen Irrtümern bewahrt. Ebenso danke ich den TeilnehmerInnen des Kolloquiums für

praktische Philosophie an der Universität Zürich sowie den TeilnehmerInnen der Interdisziplinären Woche der Fachschaft der theologischen Fakultät an der Universität Fribourg zum Thema „Kontroverse zwischen Theologie und Bioethik“ für wichtige Anregungen zu vorgetragenen Versionen dieses Beitrags im Oktober 2010 und Mai 2011. Für fruchtbare Diskussionen der hier verhandelten Thematik bin ich zudem André Chapuis dankbar.

### Literatur

- Betzler, Monika, 2009: Authenticity and Self-Governance. In: V. Mayer und M. Salmela (Hgg.): *Emotions, Ethics, and Authenticity*. Amsterdam: John Benjamins Publishing Company, S. 51–68.
- Biegler, Paul, 2010: Autonomy and Ethical Treatment in Depression. In: *Bioethics* 24, S. 179–189.
- Bolt, L. L. E., 2007: True to Oneself? Broad and Narrow Ideas on Authenticity in the Enhancement Debate. In: *Theoretical Medicine and Bioethics* 28, S. 285–300.
- Bostrom, Nick, 2008: Why I Want to be a Posthuman when I Grow Up. In: B. Gordijn und R. Chadwick (Hgg.), S. 107–136.
- Bostrom, Nick; Sandberg, Anders, 2009: Die Weisheit der Natur? Eine Evolutionäre Heuristik für Enhancement am Menschen. In: N. Knoepffler und J. Savulescu (Hgg.), S. 83–126.
- Bublitz, Jan, 2008: The Value of Authenticity in the Age of Pharmaceutical Neuroenhancement. In: *Newsletter Europäische Akademie zur Erforschung wissenschaftlich-technischer Entwicklungen Bad Neuenahr-Ahrweiler* 80, S. 1–3.
- Bublitz, Jan; Merkel, Reinhard, 2009: Autonomy and Authenticity of Enhanced Personality Traits. In: *Bioethics* 23, S. 360–374.
- Buchanan, Allen et al., 2000: Why Not the Best? In: dies. (Hgg.): *From Chance to Choice*. Genetics and Justice. Cambridge: Cambridge University Press, S. 156–203.
- Budnik, Christian, 2010: *Die Perspektive der ersten Person und ihre Relevanz für Theorien personaler Identität*. Diss. Universität Bern (unveröffentlichtes Manuskript).
- Caplan, Arthur L., 2003: Ist besser das Beste? Ein renommierter Ethiker plädiert für Enhancement des Gehirns. In: S. Schöne-Seifert und D. Talbot (Hgg.), 2009, S. 165–168.



- Chadwick, Ruth, 2008: Therapy, Enhancement and Improvement. In: B. Gordijn und dies. (Hgg.), S. 25–37.
- Chatterjee, Anjan, 2004: Cosmetic Neurology. The Controversy Over Enhancing Movement, Mentation, and Mood. In: *Neurology* 63, S. 968–974.
- Darwall, Stephen, 2006: The Value of Autonomy and Autonomy of the Will. In: *Ethics* 116, S. 263–284.
- DeGrazia, David, 2000: Prozac, Enhancement und Selbstgestaltung. In: B. Schöne-Seifert und D. Talbot (Hgg.), 2009, S. 249–263.
- DeGrazia, David, 2005: *Human Identity and Bioethics*. Cambridge: Cambridge University Press.
- Deligiorgi, Katerina: Happy Pills and Magic Potions: Cosmetic Neurology and the Explanation of Actions [unveröffentlichtes Manuskript].
- Deonna, Julien A; Teroni, Fabrice, 2008: *Qu'est-ce qu'une émotion?* Paris: Vrin.
- Elliott, Carl, 1998: Die Tyrannei des Glücklichseins: Ethik und kosmetische Psychopharmakologie. In: B. Schöne-Seifert und D. Talbot (Hgg.), 2009, S. 235–247.
- Feinberg, Joel, 1986: Autonomy. In: J. Christman (Hg.), 1989: *The Inner Citadel*. Essays on Individual Autonomy. New York: Oxford University Press, S. 27–51.
- Frankfurt, Harry G., 1997: On Caring. In: ders., 1999: *Necessity, Volition, and Love*. Cambridge. Cambridge University Press, S. 155–180.
- Freedman, Carol, 1998: Aspirin for the Mind? Some Ethical Worries About Psychopharmacology. In: E. Parens (Hg.): *Enhancing Human Traits. Ethical and Social Implications*. Washington, D.C.: Georgetown University Press, S. 135–150.
- Galert, Thorsten, 2009: Wie mag Neuro-Enhancement Personen verändern? In: B. Schöne-Seifert, D. Talbot, U. Opolka und J. Ach (Hgg.), S. 159–187.
- Galert, Thorsten; Bublitz, Christoph; Heuser, Isabella; Merkel, Reinhard; Repantis, Dimitris; Schöne-Seifert, Bettina; Talbot, Davinia, 2009: Memorandum Neuro-Enhancement. Das optimierte Gehirn. In: *Gehirn & Geist*, S. 40–52.
- Gallie, Walter G., 1956: Essentially Contested Concepts. In: *Proceedings of the Aristotelian Society* 56, S. 167–198.

- Gesang, Bernward, 2007: *Perfektionierung des Menschen*. Berlin: de Gruyter.
- Glenmullen, Joseph, 2001: *Prozac Backlash: Overcoming the Dangers of Prozac, Zoloft, Paxil, and Other Antidepressants with Safe, Effective Alternatives*. New York: Simon & Schuster.
- Gordijn, Bert; Chadwick, Ruth (Hgg.), 2008: *Medical Enhancement and Posthumanity*. Heidelberg: Springer.
- Hansen, Jennifer, 2003: Listening to People or Listening to Prozac? Another Consideration of Causal Classifications. In: *Philosophy, Psychiatry, and Psychology* 10, S. 57–62.
- Hawkins, Jennifer S., 2008: Well-Being, Autonomy, and the Horizon Problem. In: *Utilitas* 20, S. 143–168.
- Healy, David, 2006: *Let Them Eat Prozac: The Unhealthy Relationship Between the Pharmaceutical Industry and Depression*. New York: New York University Press.
- Heilinger, Jan-Christoph, 2010: *Anthropologie und Ethik des Enhancements*. Berlin/New York: deGruyter.
- Horwitz, Allan V., 2007: *The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder*. Oxford: Oxford University Press.
- Ioannidis, John P.A., 2008: Effectiveness of Antidepressants. An Evidence Myth Constructed from Thousand Randomized Trials? In: *Philosophy, Ethics, and Humanities in Medicine* 3, S. 1–9.
- Janssen, Nathalie, 2010: *Der Authentizitätsbegriff in der Enhancementdebatte*. Münster: LIT Verlag.
- Juengst, Eric T., 1998: Was bedeutet *Enhancement*? In: B. Schöne-Seifert und D. Talbot (Hgg.), 2009, S. 25–45.
- Juth, Niklas, 2011: Enhancement, Autonomy, and Authenticity. In: J. Savulescu, R. ter Meulen und G. Kahane (Hgg.): *Enhancing Human Capacities*. Oxford: Blackwell, S. 34–48.
- Knoepffler, Nikolaus; Savulescu, Julian (Hgg.), 2009: *Der neue Mensch? Enhancement und Genetik*. Freiburg und München: Alber Verlag.
- Kramer, Peter D., 1993: *Listening to Prozac. A Psychiatrist Explores Antidepressant Drugs and the Remaking of the Self*. New York und London: Penguin.
- Kramer, Peter D., 1993: Die Botschaft in der Kapsel. In: B. Schöne-Seifert und D. Talbot, 2009, S. 183–212.

- Kraemer, Felicitas, 2009: Neuro-Enhancement von Emotionen. Zum Begriff emotionaler Authentizität. In: B. Schöne-Seifert, D. Talbot, U. Opolka und J. Ach (Hgg.), S. 189–217.
- Kraemer, Felicitas, 2010: Authenticity Anyone? The Enhancement of Emotions via Neuro-Psychopharmacology. In: *Neuroethics* [im Erscheinen].
- Levy, Neil, 2007: *Neuroethics. Challenges for the 21st Century*. Cambridge: Cambridge University Press.
- Levy, Neil, 2011: Enhancing Authenticity. In: *Journal for Applied Philosophy* 28, S. 308–318.
- Manninen, Bertha A., 2006: Medicating the Mind: A Kantian Analysis of Overprescribing Psychoactive Drugs. In: *Journal of Medical Ethics* 32, S. 100–105.
- Merkel, Reinhard et al., 2007: *Intervening in the Brain. Changing Psyche and Society*. Berlin u. a.: Springer.
- Nagel, Saskia K., 2010: Too Much of a Good Thing? Enhancement and the Burden of Self-Determination. In: *Neuroethics* 3, S. 109–119.
- Nozick, Robert, 1974: *Anarchy, State, and Utopia*. New York: Basic Books.
- Olsen, Mark J., 2006: Depression, SSRIs, and the Supposed Obligation to Suffer Mentally. In: *Kennedy Institute of Ethics Journal* 16, S. 283–303.
- Oshana, Marina, 1998: Personal Autonomy and Society. In: *Journal of Social Philosophy* 29, S. 81–102.
- Oshana, Marina, 2003: How Much Should We Value Autonomy? In: *Social Philosophy & Policy*, S. 99–126.
- Parens, Erik, 2005: Authenticity and Ambivalence: Toward Understanding the Enhancement Debate. In: *Hastings Center Report* 35, S. 34–41.
- Quante, Michael, 2006: Selbst-Manipulation? Neuro-Enhancement und personale Autonomie. In: C. Kaminsky und O. Hallich (Hgg.): *Verantwortung für die Zukunft. Zum 60. Geburtstag von Dieter Birnbacher*. Münster: LIT Verlag 2006, S. 103–118.
- Ranisch, Robert; Savulescu, Julian, 2007: Ethik und Enhancement. In: N. Knoepffler und J. Savulescu, S. 21–53.
- Rego, Mark D., 2005: What Are (And What Are Not) the Existential Implications of Antidepressant Use. In: *Philosophy, Psychiatry, and Psychology* 12, S. 119–128.

- Repantis, Dimitris, 2009: Die Wirkung von Psychopharmaka bei Gesunden. In: A. Wienke et al. (Hgg.), S. 63–68.
- Rosati, Connie S., 2006: Personal Good. In: T. Horgan und M. Timmons (Hgg.): *Metaethics After Moore*. Oxford: Clarendon, S. 107–131.
- Rosfort, René; Stanghellini, Giovanni, 2009: The Person In Between Moods and Affects. In: *Philosophy, Psychiatry, and Psychology* 16, S. 251–266.
- Runkel, Thomas, 2010: *Enhancement und Identität. Die Idee einer biomedizinischen Verbesserung des Menschen als normative Herausforderung*. Tübingen: Mohr Siebeck.
- Savulescu, Julian; Bostrom, Nick (Hgg.), 2009: *Human Enhancement*. Oxford: Oxford University Press.
- Sandel, Michael, 2007: *The Case Against Perfection. Ethics in the Age of Genetic Engineering*. Cambridge, Mass.: Harvard University Press Belknap.
- Schermer, Maartje; Bolt, Ineke; de Jongh, Reinoud; Olivier, Berend, 2009: The Future of Psychopharmacological Enhancements: Expectations and Policies. In: *Neuroethics* 2, S. 75–87.
- Schmidt-Felzmann, Heike, 2009: Prozac und das wahre Selbst: Authentizität bei psychopharmakologischem Enhancement. In: B. Schöne-Seifert, D. Talbot, U. Opolka und J. Ach (Hgg.), S. 143–158.
- Schöne-Seifert, Bettina, 2006: Pillen-Glück statt Psycho-Arbeit. Was wäre dagegen einzuwenden? In: J. Ach und A. Pollmann (Hgg.): *no body is perfect. Baumassnahmen am menschlichen Körper*. Bielefeld: transcript Verlag, S. 279–291.
- Schöne-Seifert, Bettina; Talbot Davinia (Hgg.), 2009: *Enhancement. Die ethische Debatte*. Paderborn: mentis.
- Schöne-Seifert, Bettina; Talbot, Davinia; Opolka, Uwe; Ach, Johann S. (Hgg.), 2009: *Neuro-Enhancement. Ethik vor neuen Herausforderungen*. Paderborn: mentis.
- Stringaris, Argyris, 2009: On Emotions That Last Longer. In: *Philosophy, Psychiatry, and Psychology* 16, S. 277–281.
- Sumner, L. Wayne, 1996: *Welfare, Happiness & Ethics*. Oxford: Oxford University Press.
- Svenaesus, Fredrik, 2009: The Ethics of Self-Change: Becoming Oneself by Way of Antidepressants or Psychotherapy? In: *Medicine, Health Care and Philosophy* 12, S. 169–178.
- Talbot, Davinia, 2009: Pharmakologisches Enhancement – Eine Ein-

- führung in nichtmedizinische Anwendungen von Arzneimitteln zu Verbesserungszwecken. In: A. Wienke et al. (Hgg.), S. 69–78.
- Tännsjö, Torbjörn, 2009: Ought We To Enhance Our Cognitive Capacities? In: *Bioethics* 23, S. 421–432.
- The President's Council on Bioethics (Hg.), 2003: *Beyond Therapy: Biotechnology and the Pursuit of Perfection*. Washington: Dana Press.
- Weidmann, Stefan, 2008: *Erfolgsmythos Psychopharmaka: Warum wir Medikamente in der Psychiatrie neu bewerten müssen*. Bonn: Psychiatrie-Verlag.
- Wienke, Albrecht; Eberbach, Wolfram; Kramer, Hans-Jürgen; Janke, Kathrin (Hgg.), 2009: *Die Verbesserung des Menschen. Tatsächliche und rechtliche Aspekte der wunscherfüllenden Medizin*. Heidelberg u. a.: Springer.

Norman Sieroka

## Ist *ein* Zeithof schon genug? – Neurophänomenologische Überlegungen zum Zeitbewusstsein und zur Rolle des Auditiven

### Zusammenfassung

Dieser Aufsatz zeigt und betont die Bedeutung phänomenologischer und neurowissenschaftlicher Untersuchungen des Auditiven für ein adäquates Verständnis des Verhältnisses zwischen subjektiver und objektiver Zeit. Anhand von Beispielen aus dem Bereich des Hörens werden strukturelle Analogien herausgearbeitet zwischen Konzepten einer ausgedehnten Gegenwart, wie sie in der Phänomenologie, Psychologie, Musiktheorie auftreten, und Mechanismen und Modellen zeitlicher Integration, wie sie in der Neurophysiologie und Neuroinformatik angenommen werden. Es wird dabei nicht nur für die Möglichkeit einer fruchtbaren heuristischen Wechselwirkung zwischen Phänomenologie und Neurowissenschaft argumentiert, sondern auch für eine breitere systematische Relevanz dieser Analogien für eine Theorie des Zeitbewusstseins. Gerade die spezifischen Prozesse und Mechanismen, die beim Hören involviert sind, haben Psychologen und Neurowissenschaftler immer wieder veranlasst, hier von der „(Sinnes-)Modalität der Zeit“ zu sprechen. Dementsprechend ist dieser Aufsatz zugleich eine implizite Kritik an der Fokussierung eines Großteils der gegenwärtigen phänomenologischen Literatur auf das Visuelle. Der Aufsatz schließt mit einigen – eher programmatischen – Überlegungen zum genaueren Verhältnis zwischen phänomenologischen und neurowissenschaftlichen Aspekten der Wirklichkeit. Hier wird gegen eine reduktionistische Sichtweise argumentiert, der zufolge diese subjektiven und objektiven (modal- und lagezeitlichen) Aspekte strikt miteinander identifizierbar sind. Stattdessen wird angedeutet, inwiefern diese Aspekte lediglich partial isomorph sind und die Zeit selbst als ein „Knoten“ (diesen Begriff übernehme ich von Peter Rohs) aufgefasst werden kann, in dem sich die grundlegenden Dimensionen des Geistigen und des Physischen treffen.

### Abstract

This paper aims to show the importance of both phenomenological and neuroscientific investigations into hearing with the intention of gaining an adequate understanding of the relation between subjective and objective time. On

*philosophia naturalis* 46 / 2009 / 2

the basis of examples taken from the auditory domain, I will present structural analogies between the concepts of an extended present, as argued for in phenomenology, psychology, and music theory, and of the mechanisms and models of temporal integration, as known from neurophysiology and neuroinformatics. Using these, I will demonstrate not only the possibility of a fruitful heuristic interaction between phenomenology and neuroscience, but also the wider systematic relevance of these structural analogies for providing a satisfying account of time consciousness. Indeed the specific processes and mechanisms involved in hearing have led psychologists and neuroscientists to speak about hearing as being the “modality of time”. Hence, the present paper is also an implicit critique of the focus on vision in (most of) contemporary phenomenological literature. The paper closes by providing some – rather programmatic – considerations about the exact relationship between the phenomenological and neurophysiological aspects of reality. Here I will oppose a reductionist view, according to which these subjective and objective (tensed and tenseless) aspects are strictly identifiable with one another. Instead, I will suggest that those aspects are only partially isomorphic and that time itself should be viewed as a kind of “knot” (I borrow this term from Peter Rohs) in which the fundamental dimensions of the mental and of the physical meet.

## 1. Einleitung

Ein mitreißendes Konzert und eine langweilige Vorlesung werden in der Regel als deutlich unterschiedlich lang empfunden. Aber der Blick auf eine Uhr mag belegen, dass das eine wie das andere neunzig Minuten gedauert hat. Die äußere (oft auch als „objektiv“ bezeichnete) Zeit, wie sie die Uhr aufgrund sich wiederholender physikalischer Abläufe anzeigt, hat offenbar nicht immer die gleichen Maßverhältnisse wie die empfundene, die subjektive Zeit.

Stellt man die Frage, wieso das so sein kann, sind Reduktionsstrategien schnell bei der Hand. Eine von beiden, entweder die subjektive oder die objektive Zeit, wird schon die richtige sein, und man wird die Diskrepanz zur anderen schon „wegerklären“ können. Eine gängige Reduktionsstrategie für das gerade beschriebene Phänomen im Rahmen der Psychologie lautet beispielsweise wie folgt: Der Mensch hat eine bestimmte Menge, ein bestimmtes Reservoir an Aufmerksamkeit zur Verfügung, das er unterschiedlich nach „innen und außen“ aufteilen kann.<sup>1</sup> Im Fall des Konzerts sei die Aufmerksamkeit fast vollständig nach außen auf die Musik gerichtet gewesen, während es im Fall der Vorlesung Mühe bereite, die Aufmerksamkeit im gleichen Maß nach

außen auf das Gesprochene zu fokussieren. Ein erheblicher Teil des Vorrats an Aufmerksamkeit habe sich hier vielmehr nach innen gerichtet und damit – so die Hypothese – auch auf das Verstreichen der inneren Zeit selbst. Entsprechend war die Vorlesung subjektiv länger als das Konzert.

An dieser Stelle muss es nicht weiter interessieren, wie plausibel diese Erklärung ist bzw. welche Feinheiten noch zu ergänzen wären. Entscheidend ist die Strategie: Am Ende steht letztlich die Behauptung, dass es gar keinen *wirklichen* Zeitunterschied zwischen den beiden Ereignissen gab. Man habe sie *nur unterschiedlich empfunden*, weil diese und jene Mechanismen abgelaufen seien. Vielleicht kann dieses psychologische Erklärungsmuster dann noch auf ein physiologisches reduziert werden, und vermeintlich stellt sich heraus, dass die Ereignisse nur deshalb als unterschiedlich lang empfunden wurden, weil diese und jene Neuronen gefeuert haben. Nimmt man schließlich noch an, die Physiologie sei vollständig auf die Physik reduzierbar, so ergibt sich das in sich geschlossene Weltbild eines Naturalismus (oder genauer eines Physikalismus) – einer Welt, in der es ausschließlich physikalische Prozesse oder Zustände gibt.

Doch auch andere reduktionistische Erklärungsstrategien scheinen möglich. So könnte beispielsweise auch ein Phänomenologe (im Sinne von Husserls Analyse des intentionalen Bewusstseins) die Spannung zwischen den Maßverhältnissen der subjektiven und der physikalischen Zeit im obigen Beispiel „wegerklären“.<sup>2</sup> Er könnte sich dabei sogar zum Teil der gleichen strukturellen Elemente bedienen wie ein Psychologe. „Aufmerksamkeit“, so könnte der Phänomenologe argumentieren, sei ein sehr diffuser Begriff in der Psychologie. Womit man eigentlich anheben sollte und was laut dem Phänomenologen ein Psychologe „eigentlich meint“ (oder eigentlich meinen sollte), wenn er davon spricht, man könne seine Aufmerksamkeit auf etwas richten, sei eine Frage der Intentionalität. Und nun hätte der Phänomenologe keine Probleme, den Unterschied in der empfundenen Zeit über komplexe Vernetzungen verschiedener Arten von Intentionalität zu erklären. Die zeitlichen Muster der neuronalen Aktivität spielen hierbei keine Rolle. Denn diese Muster gibt es nur in der physikalischen Zeit und von der ist beim Phänomenologen zunächst gar nicht die Rede bzw. müsste er deren Existenz (oder zumindest ihre Möglichkeit) erst über die Begriffe der Intentionalität und des *Zeitbewusstseins* ableiten.



Ich werde in diesem Aufsatz diskutieren, wie sich ein bestimmtes naturalistisches (physikalistisches) und ein bestimmtes, letztlich transzendentalphilosophisches Zeitverständnis zueinander verhalten. Dazu werde ich das vorherrschende Verständnis in der Neurophysiologie mit dem der Husserlschen Phänomenologie vergleichen und beide daraufhin untersuchen, ob es möglich und sinnvoll ist, das eine Zeitverständnis auf das andere zu reduzieren. (Im Zuge dieser Darstellung werde ich auch auf „neurophänomenologische“ Ansätze von Varela, Thompson und van Gelder eingehen.)

Allerdings wird es im Folgenden, was die Größenordnung der betrachteten Zeitintervalle betrifft, weniger um ganze Vorlesungen und Konzerte als vielmehr um einzelne Verbal- und Melodiephrasen gehen – also auf physikalischer Ebene um den Zeitbereich weniger Sekunden. Die Frage der unterschiedlichen Maßverhältnisse in der physikalischen und der subjektiven Zeit stellt sich auch hier; und sie wird sogar verschärft durch die Existenz einer ausgedehnten erlebten Gegenwart. Denn während offensichtlich weder eine Melodie- noch eine Verbalphrase als Ganze zu einem einzelnen Zeitpunkt physikalisch an meiner Ohrmuschel als Schalldruck gegenwärtig ist, ist sie *mir* (als einem Subjekt oder einem Bewusstsein) eben doch als Ganze im Moment präsent. Auch hier stellt sich also die Frage, ob man diese Unterschiede mittels eines Reduktionismus sinnvoll und überzeugend auflösen kann und sollte.

Zwar scheint weder eine Naturalisierung der Phänomenologie noch so etwas wie eine transzendente Reduktion der Neurowissenschaften auf die Phänomenologie ein von vornherein inkonsistentes Unternehmen zu sein. Doch das allein ist kein genügendes Argument, einen dieser beiden Reduktionismen betreiben zu wollen. Auch die strukturellen Analogien, die im Folgenden zwischen Phänomenologie und Neurophysiologie aufgezeigt werden, sind keine ausreichende Begründung für die Reduktion in eine bestimmte Richtung. Vielmehr werde ich dafür argumentieren, keinen dieser Reduktionismen zu betreiben und stattdessen den Zeitbegriff selbst als einen gemeinsamen Bezugspunkt zu betrachten. Ich werde mich deshalb gegen Ende dieses Aufsatzes auf Peter Rohs beziehen, der in seiner „feldtheoretischen Transzendentalphilosophie“ aufgezeigt hat, wie die Zeit als ein „ontologischer Knoten“ betrachtet werden kann, von dem aus sich zum einen eine Theorie der Subjektivität entwickeln kann und zum anderen ein Rahmen für die naturwissenschaftlichen Einzeldisziplinen unter einem Primat des phy-

sikalischen Feldbegriffs. Streng bewusstseinstheoretische oder naturalistische Reduktionen werden dabei nicht betrieben, und die Welt wird in ihren natürlichen und geistigen Facetten ernst genommen und ausgehend von der Zeit geordnet. Damit ist schon die Begründung geliefert, warum in diesem Aufsatz die Analogien zwischen Phänomenologie und Neurophysiologie anhand der *Zeit*wahrnehmung behandelt werden und nicht etwa anhand von Farbwahrnehmungen oder Schmerzempfindungen. (Zudem kommt auch innerhalb der Husserlschen Phänomenologie dem Zeitbewusstsein eine fundamentale Rolle zu, werden doch alle geistigen Akte und insbesondere alle Bedeutungen über die zeit-spezifischen intentionalen Momente der Retention, Urimpression und Protention konstituiert – siehe unten.)

Zum Abschluss dieser Einleitung muss noch einem möglichen Einwand begegnet werden. Man könnte der Meinung sein, dass für die folgenden Überlegungen bereits das Beispiel mit dem Konzert- und Vorlesungsbesuch ausreiche. Alles konzeptionell Relevante müsse sich bereits daran aufzeigen lassen. Das mag von einem theoretischen Standpunkt aus betrachtet „im Prinzip“ richtig sein. Allerdings werden, ohne einen kurzen Blick auf die Details zu werfen, die Attraktivität und die Motivation für einen Reduktionismus nicht genügend deutlich. Denn wirklich verlockend, so werde ich unten darzustellen versuchen, wird er gerade in dem Moment, in dem sich die strukturellen Parallelen zwischen Neurophysiologie und Phänomenologie im Einzelnen zeigen und diverse, sehr spezielle Hirnantworten mit spezifischen Formen der Intentionalität korrelieren. Es gibt eine Reihe neuronaler Antworten des menschlichen Gehirns, deren unterschiedliche zeitliche Verlaufsmuster den Unterschieden in den Arten und Qualitäten geistiger Akte zu „entsprechen“ scheinen: So werden im Folgenden einige Ergebnisse aus der Elektrophysiologie der Zeitwahrnehmung berichtet und anschließend mögliche „Kandidaten physiologischer Korrelate“ zu Größen des inneren Zeitbewusstseins (Retention, Urimpression, Protention) diskutiert. Gerade wenn es darum geht, einem Reduktionismus auf möglichst überzeugende Weise zu widerstehen, muss man ihn zunächst mit all seinen intuitiven Stärken ernst nehmen.

Umgekehrt hat das Vernachlässigen solch detaillierter struktureller Parallelen zu oberflächlichen und naiven Argumenten in der Philosophie des Geistes geführt. So hat etwa Peter Lanz an mehreren Beispielen zu verdeutlichen versucht, wie sich subjektiv empfundene Qualitä-

ten von physikalischen Kenngrößen unterscheiden:<sup>3</sup> Der Mensch höre doch, so Lanz, Töne und nicht etwa Schalldrücke, was man schon an der unterschiedlichen Struktur ihrer Eigenschaften erkenne. Noch eindrücklicher sei dies beim Farbsehen: Man komme doch beim Farbkreis, wenn man von blau über violett weitergehe, wieder zu rot; etwas Entsprechendes passiere aber nicht, wenn man die Frequenz einer elektromagnetischen Welle, die man zunächst als blau und dann als violett wahrnimmt, weiter erhöht. Damit wäre nicht nur der qualitative Unterschied zwischen sekundären und primären Eigenschaften nachgewiesen, sondern insbesondere auch die Unlösbarkeit des Leib-Seele-Problems.

Doch diese Argumentation ist fragwürdig. Denn zunächst einmal sind der Farbkreis und das elektromagnetische Frequenzspektrum unterschiedliche *Formen der Darstellung*, nämlich eine zirkuläre und eine lineare. Solche verschiedenen Darstellungsformen kann man auf eine vermeintliche Außenwelt, auf eine Verarbeitung im Gehirn und auf die Wahrnehmung anwenden. Aber das allein belegt keine irgendwie gearteten intrinsischen Unterschiede zwischen physikalischen, physiologischen und psychischen Zuständen. Den Eindruck eines Bruches in der Form der Repräsentation zwischen Leib und Seele kann Lanz nur dadurch so stark machen, dass er ein wichtiges Zwischenglied unterschlägt; nämlich eine Repräsentation der physiologischen Verarbeitung. Bereits im Auge selbst gibt es bekanntermaßen drei Arten von Farbrezeptoren, deren Differenzierung nach kurz-, mittel-, und langwellig bzw. blau, grün und rot man offensichtlich und je nach Interessenlage durchaus mit Berechtigung mit dem (linearen) Frequenzspektrum aber eben auch mit dem Farbkreis in Beziehung setzen kann. Außerdem kommt es auf kortikaler Ebene (zusammen mit einer retinotopen) zu einer chromotopen Form der Repräsentation;<sup>4</sup> d. h. die Farbdimensionen korrespondieren hier Aktivitäten an spezifischen Orten des (mehrdimensionalen) visuellen Kortex, womit erneut eine strikt eindimensional-lineare Form der Frequenzdarstellung durchbrochen wird.

Diese Hinweise allein genügen selbstverständlich nicht, um ein vermeintliches Leib-Seele-Problem zu lösen (bzw. aufzulösen) oder um den Zusammenhang zwischen sekundären und primären Eigenschaften adäquat zu fassen. Ein Vertreter der Lanzschen Position könnte weiterhin entgegnen: So viel oder so wenig, wie der Mensch Schalldrücke hört, hört er die neuronale Aktivität in seinem Großhirn. Doch an die-

ser Stelle ging und geht es gar nicht um ein mögliches Leib-Seele-Problem, sondern darum aufzuzeigen, wie ein oberflächlicher Blick auf die Naturwissenschaften zu einem zweifelhaften philosophischen Argument umgearbeitet wird.

Um nicht denselben Fehler zu begehen, diskutiere ich zunächst einige Befunde aus der Neurophysiologie und Psychophysik der Zeitwahrnehmung, für die gerade das Hören, die auditorische Wahrnehmung, eine herausragende Rolle spielt. – Damit stellt die vorliegende Arbeit zugleich eine wichtige Ergänzung dar zu Arbeiten im Bereich der Neuropänomenologie wie auch allgemeiner der Philosophie des Geistes, in denen im Allgemeinen das Hören gegenüber dem Sehen stark vernachlässigt wird.<sup>5</sup> Gerade die Besonderheiten der auditorischen Wahrnehmung im Vergleich zur visuellen sind aber höchst relevant für Diskussionen nicht nur zum Zeitbewusstsein, sondern auch zu vorbegrifflichen Zuständen, zu nicht-propositionaler Intentionalität u. dgl.<sup>6</sup> Auch deshalb ist also die folgende, etwas ausführlichere Darstellung von Relevanz.

## 2. Neurophysiologie und Psychophysik der Zeitwahrnehmung

Aus dem mittlerweile fast unüberschaubar gewordenen Zoo von Hirnantworten sollen hier nur drei (*mismatch negativity*, *sustained field* und *contingent negative variation*) diskutiert werden, da sie m. E. genügen, um einen Einblick zu geben in die Spezifität unterschiedlicher neuronaler Prozesse und auch in die herausragende Rolle des auditorischen Systems für die Verarbeitung von Zeitinformationen.

Neurowissenschaftliche Experimente zur Zeitwahrnehmung des Menschen untersuchen oft die Fähigkeit, Töne unterschiedlicher Dauer voneinander zu unterscheiden. Hier kann man auf psychophysischer Seite Daten über Diskriminationsfähigkeiten der Probanden gewinnen, und man kann sie gegebenenfalls mit physiologischen Daten vergleichen, die man etwa über elektrophysiologische oder andere bildgebende Verfahren gewinnt.<sup>7</sup>

Zumeist wird dabei das auditorische System untersucht, weil es so etwas wie die „Sinnesmodalität der Zeit“ ist. Zum einen sind Menschen es eher gewohnt und auch besser dazu in der Lage, akustische Reize ihrer Auftretensdauer nach zu beurteilen als etwa Gerüche oder Bil-

der. Zum anderen korrespondieren diverse, sehr unterschiedliche Qualitäten, die der Mensch beim Hören erfährt, gerade mit den zeitlichen Regularitäten physikalischer Reize. Perzeptiv so unterschiedliche Phänomene wie die Lokalisation von Geräuschen im Raum, die Wahrnehmung von Tonhöhen wie auch diejenige eines Rhythmus unterscheiden sich, wenn man sie als physikalische Phänomene betrachtet, allein durch das Zeitfenster oder die Skala ihrer zeitlichen Regelmäßigkeit. Ohne die Unterschiede in der Laufzeit, die der Schall braucht, um vom Mund eines anderen in mein linkes bzw. in mein rechtes Ohr zu gelangen, wüsste ich (bei geschlossenen Augen) nicht, ob ich gerade von vorne links oder von hinten rechts angesprochen werde. Ein hohes C unterscheidet sich physikalisch von einem tiefen G (angenommen selbstverständlich, dass es beide Male vom gleichem Instrument oder Gerät mit gleichem Klangspektrum produziert wird) nur in der Dauer der Aufeinanderfolge zweier Wellenberge des Schalldrucks. Und ein Rhythmus ist nichts anderes als ein bestimmtes zeitliches Muster der Abfolge von Schlägen und Pausen.

Dabei verarbeitet das auditorische System diese zeitlichen Regularitäten als ebensolche Regularitäten – soll heißen: Hier gibt es keine vorgeschalteten Transformationen wie beim visuellen System, bei dem die Stäbchen und Zäpfchen bereits in der Netzhaut die Original-Frequenzen des Lichtes aufheben. Und diese Besonderheit des Hörens beruht physiologisch darauf, dass die relevanten Zeitskalen akustischer Reize im Bereich von Millisekunden und Sekunden (statt wie beim Licht von Femtosekunden) liegen und somit identisch sind mit der Zeitskala neuronaler Prozesse. Damit hängt zusammen, dass sich auf der für diese Arbeit besonders relevanten Zeitskala von etwa ein bis zehn Sekunden das auditorische System gegenüber allen anderen sensorischen Systemen des Menschen durch eine von der Aufmerksamkeit unbeeinflusste „Erinnerungsspur“ (*auditory memory trace*) auszeichnet. Diese Erinnerungsspur ist wichtig für die Sprach- und Musikwahrnehmung. Und sie ist unabhängig von Leistungen des Langzeitgedächtnisses, bei denen sozusagen zuvor abgelegte Informationen wieder abgerufen werden (und die im Allgemeinen auch nicht auditorisch spezifisch sind).

Das physikalische Korrelat dieser Erinnerungsspur ist die *mismatch negativity* (MMN), die mit Methoden der Elektrophysiologie gemessen werden kann.<sup>8</sup> Die MMN ist eine spezifische Differenz in der neuronalen Aktivität, die durch einen aus einer Reihe gleichartiger Töne heraus-

fallenden Ton (*deviant* oder „Abweichler“) hervorgerufen wird – durch einen Ton also, der beispielsweise kürzer ist oder höher oder eine andere Klangfarbe besitzt als seine Vorgänger. Das Auftreten einer MMN kann man sich entsprechend über einen Adaptationsprozess plausibel machen: Der MMN geht so etwas wie das reduzierte Antwortverhalten (Adaptation) bestimmter neuronaler Verbände auf regelmäßig wiederkehrende Reize gleicher Art voraus. Tritt nun ein anderer Reiz auf (ändert der Ton etwa seine Frequenz) kommt es zu einer Hirnantwort der Größe, wie sie für nicht-adaptierte Reize normal ist. Doch in ihrer Umgebung sticht diese Antwort damit hervor, und der Vergleich zu den Hirnantworten auf den adaptierten Standardreiz liefert einen spezifischen Unterschied: die MMN. Wie bereits erwähnt, gibt es eine MMN nur für das auditorische System. Zeigt man einem Probanden beispielsweise immer wieder ein rotes Dreieck und zwischendurch auch einmal ein blaues, so gibt es auf diesen „Abweichler“ keine spezifische Hirnantwort in der Art einer MMN.

Weiterhin tritt die MMN unabhängig davon auf, ob sich der Proband auf die Tonreihe konzentriert oder nicht. Eine MMN wird also auch dann evoziert, wenn der Proband im Vorfeld gar nicht weiß, dass in der Präsentation neben bestimmten Standardtönen noch „Abweichler“ auftreten, oder wenn der Proband schläft.<sup>9</sup> In der Literatur wird die MMN deshalb auch als „prä-attentive“ Hirnantwort bezeichnet; manchmal ist auch von „primitiver (vorbewusster) Intelligenz“ die Rede, da MMNs nicht nur bei Abweichungen von direkten Klangeigenschaften wie Tonhöhe oder Timbre auftreten, sondern auch bei der Verletzung abstrakter Regeln wie etwa „je höher der Ton, desto lauter ist er“ oder „der zweite Ton eines Paares ist jeweils höher als der erste Ton“.

Insbesondere, um nun wieder genauer auf die Zeitwahrnehmung zurückzukommen, wird eine MMN durch Töne abweichender Zeitdauer evoziert. Zudem ist das Auftreten einer solchen MMN mit den diskriminativen Fähigkeiten der Probanden korreliert. Liegt der Unterschied zwischen der Dauer des Standardtons und des „Abweichlers“ unterhalb der (in einem separaten Experiment gemessenen) Diskriminationsschwelle eines Probanden, so findet man bei diesem Proband in der Regel für diesen Fall auch keine MMN. Außerdem wird eine MMN nur dann evoziert, wenn die Pause zwischen den Tönen nicht mehr als einige Sekunden beträgt. Spätestens nach zehn Sekunden bricht also die kortikale auditorische Erinnerungspur zusammen.

Vor diesem Hintergrund wird nachvollziehbar, dass und warum man in der neurophysiologischen Literatur immer wieder Beschreibungen findet, die mit intentionalen Begriffen aufgeladen sind. Da wird nämlich aus der MMN (also aus einem sensorischen Antwortverhalten, das Adaptation involviert) eine prä-attentive „Erwartung“, der zufolge der nächste Ton die Eigenschaften seines Vorgängers haben „sollte“; und wird diese Erwartung „frustriert“, dann „antworte“ das Gehirn mit einer MMN. Denn so irreführend solche Beschreibungen auch sind, da eine MMN zunächst ein physiologisches und kein geistiges Ereignis ist: es sind eben genau die psychophysischen Korrelationen und die strukturell-funktionalen Analogien zwischen Physiologie und Psychologie (und Phänomenologie), die dazu verführen, bestimmte Charakteristika des Physiologischen „auch mal im anderen Jargon zu illustrieren“.<sup>10</sup>

Was die Ausbildung der kortikalen auditorischen Erinnerungsspur betrifft, so wird die MMN sozusagen funktional ergänzt durch das *sustained field* (SF), eine sensorische Hirnantwort, die einen einkommenden Reiz über seine gesamte Dauer hinweg begleitet. Wie die MMN ist das SF aufmerksamkeits*unabhängig*; und auch die Zeitskala, bei der es auftritt (bis etwa zehn Sekunden), ist gleich. Aber anders als die MMN ist das SF keine transiente Hirnantwort – also keine, die nach einer bestimmten Zeit relativ zum Stimulusende einmalig auftritt. Das SF ist vielmehr direkt an den „Echtzeit-Verlauf“ des Reizes gebunden.

Ebenfalls auf der gleichen Zeitskala und ebenfalls in diesem Sinne eine „Echtzeit-Antwort“, ist die *contingent negative variation* (CNV). Dabei handelt es sich um ein in seiner Amplitude linear zunehmendes negatives Potential, das aber – und dies ist nun der wichtige Unterschied zu MMN und SF – aufmerksamkeits*abhängig* ist. Die CNV baut sich in der Folge eines Anstoßreizes auf und endet mit dem Auftreten eines zu beachtenden Folge- oder Signalreizes.<sup>11</sup> Dabei kann es sich (anders als bei der MMN und dem auditorischen SF) sowohl um auditorische als auch um visuelle Reize handeln.

Inbesondere tritt eine CNV dann auf, wenn die aufmerksam verfolgte Eigenschaft der Reize deren zeitliche Ausdehnung ist.<sup>12</sup> Ihr linearer Verlauf, der kurz nach Ende des tatsächlich vorgegebenen Reizes zusammenbricht, kann hier funktional als ein neuronales Begleiten (*monitoring*) des Stimulus interpretiert werden. Die CNV ist somit insbesondere nicht so etwas wie ein „Vorlagenabgleich“, bei dem die Dauer des Standardreizes in irgendeiner Weise neuronal implementiert wäre

und die Hirnantwort also unabhängig vom einkommenden Reiz immer nach Länge des Standardintervalls zusammenbrechen müsste.<sup>13</sup>

Die Größe oder Stärke auditorisch evozierter CNVs ist außerdem mit der Diskriminationsfähigkeit im Experiment korreliert: Wer Töne besser ihrer Dauer nach unterscheiden kann als jemand anderes, wird in der Regel auch eine größere CNV haben, und umgekehrt. Eine solche Korrelation gibt es bei visuell evozierten CNVs (also wenn Probanden Bildpräsentationen und nicht Töne ihrer Dauer nach unterscheiden müssen) nicht.<sup>14</sup> Dies zusammen mit der bereits erwähnten Tatsache, dass die Diskriminationsfähigkeit für Zeitintervalle im Auditorischen signifikant besser ist als bei den anderen Sinnesmodalitäten, kann als Hinweis auf einen gemeinsamen Kernmechanismus gedeutet werden. Im Falle der Töne, so die Spekulation, würde dann ein zusätzlicher Prozess (eine zusätzliche, „bessere“ Repräsentation) relevant werden. Auch diese Untersuchungen im Kontext der CNV belegen also die ausgezeichnete Rolle des Auditorischen für die neuronale Repräsentation zeitlicher Information.

Außerdem korreliert die in diesen Experimenten bestimmte Diskriminationsfähigkeit der Probanden bei den verschiedenen Aufgaben signifikant mit ihrer rhythmischen Begabung (wie diese gemessen wurde und was der konzeptionelle Hintergrund dieses Tests ist, wird im nächsten Abschnitt zu diskutieren sein). Das legt die Interpretation nahe, die rhythmische Begabung stehe in einem engeren Zusammenhang zu dem, was gerade als „gemeinsamer Kernmechanismus“ zur Unterscheidung von Zeitintervallen bezeichnet wurde. Und dass es überhaupt eine so enge Verbindung von Musikalität und zeitlicher Diskriminationsfähigkeit gibt, verwundert vermutlich wenig, da, wie erwähnt, die Wahrnehmung von Rhythmus und auch Tonhöhen physikalisch bzw. physiologisch eine Wahrnehmung zeitlicher Regularitäten ist.<sup>15</sup>

Spekulationen dieser Art können und sollten genauer zu den oben diskutierten Erkenntnissen über die MMN in Beziehung gesetzt werden. Denn auch dort gibt es Zusammenhänge zur Musikalität, wie u. a. Studien zur prä-attentiven Verarbeitung von Akkordfolgen nachweisen.<sup>16</sup> Weitere Untersuchungen und Überlegungen zum Zusammenhang von empirisch bestimmter Musikalität mit anderen psychometrischen Daten und mit neuronalen Aktivitäten sind also ein Desiderat, um das allgemeine Verständnis von Musikalität wie auch die konkrete Entwicklung beispielsweise musikpädagogischer Konzepte voranzubringen.



Anstatt dem hier allerdings weiter nachzugehen, möchte ich eine kurze Überlegung zu den psychophysischen Korrelationen und Analogien anschließen. Denn gerade diese Korrelationen führen, wie im Fall der MMN oben bereits illustriert, immer wieder zu Übertragungen intentionaler Redeweisen auf das Physiologische. Doch lässt sich eine solche, zunächst einmal rhetorische, Übertragung vielleicht auch systematisch verstehen? Hier könnte man vermuten, die Zusammenhänge zwischen einzelnen Hirnantworten und Wahrnehmungen oder Diskriminationsfähigkeiten deuteten philosophisch auf eine *Type-Identity-Theory*, die einen (im mathematischen Sinne) funktionalen Zusammenhang von physischen und psychischen Zuständen bestimmten Typs annehmen. Doch diese Annahme wird durch die statistische Natur der Zusammenhänge unterlaufen.<sup>17</sup> In der Neuropsychologie hat man es nicht mit funktionalen sondern mit statistischen Beziehungen zu tun. Konditionale in der strikten Form etwa von „Wahrnehmungsunterschiede in der Tonlänge gibt es dann und nur dann, wenn die Amplitude der MMN den Wert  $x$  übersteigt“, die für eine *Type-Identity-Theory* nötig wären, lassen sich eben nicht aufstellen.

Mit statistischen Korrelationen, so scheint es, lässt sich allenfalls eine *Token-Identity-Theory* im Sinne Davidsons motivieren:<sup>18</sup> Der vom Probanden  $P$  wahrgenommene Tonlängenunterschied im Experiment vom 14.07.2010 um 14.07 Uhr *ist identisch* mit dem konkreten Hirnzustand (*token brain state*) des Probanden zu diesem Zeitpunkt. Und dieser Hirnzustand muss keineswegs der gleiche für verschiedene Probanden sein, noch auch immer der gleiche für den Probanden  $P$ . – Doch auch diese Form der Identitätstheorie ist nicht überzeugend. Denn nun verlieren die statistischen Korrelationen zwischen Physiologie und Phänomenologie ihre systematische Relevanz. Wenn es im strengen Sinne allein um *token identity* ginge, so verlöre man letztlich die Grundlage für die Untersuchung nicht nur funktionaler, sondern auch statistischer psychophysischer Zusammenhänge. Doch das scheint nach den zahlreichen Beispielen, die oben aus dem Bereich des Hörens diskutiert wurden, nicht plausibel. Außerdem basiert Davidsons Verneinung der Möglichkeit psychophysischer Gesetze gerade darauf, dass er das Geistige (im Gegensatz zum Physikalischen) nicht für ein „comprehensive closed system“ hält. Doch genau dieser These muss an dieser Stelle aus phänomenologischer Perspektive widersprochen werden.<sup>19</sup> – Ich werde auf diese Thematik im letzten Teil dieses Aufsatzes zurückkommen,

wenn ich dafür argumentiere, die statistischen Korrelationen (oder genauer: die zeitlichen Dimensionen, in denen sich physikalische und geistige Zustände aufeinander beziehen lassen) als „Knoten“ von Natur und Subjektivität zu verstehen.

### 3. Der Zeithof als multidisziplinäres Konzept

In diesem Abschnitt sollen kurz die Begriffe vorgestellt werden, die aus unterschiedlichen wissenschaftlichen Teildisziplinen stammen und die für jeweils spezifische Aspekte der Zeitwahrnehmung von besonderer Relevanz sind. Insbesondere gehe ich ein auf die Musikpädagogik und den Begriff der musikalischen Begabung bei Edwin Gordon, auf eine Arbeit von Jerrold Levinson aus der analytischen Philosophie der Musik, auf Edmund Husserls Phänomenologie des inneren Zeitbewusstseins, sowie auf Dan Lloyds Neuroinformatik temporaler Informationsspeicherung und die oben präsentierten Ergebnisse aus der Elektrophysiologie.<sup>20</sup> Die Motivationen für die jeweiligen Untersuchungen sind dabei sehr unterschiedlich, die Struktur der Ergebnisse hingegen analog. (Wenn in der Überschrift dieses Kapitels der Husserlsche Terminus „Zeithof“ auftritt, so möchte ich damit also keinesfalls behaupten, dass alle Autoren, die hier behandelt werden, Phänomenologie im engeren Sinne betreiben.)

Der Musikpädagoge und Jazzmusiker Edwin Gordon hat einen Musikalitätstest entwickelt, dessen Ergebnisse oben im Zusammenhang zeitlicher Diskriminationsfähigkeit erwähnt wurden.<sup>21</sup> Gordon beschreibt diesen Test als *Advanced Measure of Musical Audiation* (oder kurz: AMMA-Test). Es geht in diesem Test darum, einen numerischen Wert für das abzuleiten, was Gordon „Audiation“ nennt (ein Kunstwort, das nicht mit dem englischen Ausdruck „audition“ zu verwechseln ist). Etwas vage umschreibt Gordon das Verhältnis von Audiation und Musik in Analogie zum Verhältnis von Sprache und Denken bzw. Bedeutung: „sound becomes music only through audiation, when, as with language, you translate the sounds in your mind to give them meaning“.<sup>22</sup> Gordon nimmt hier eine Hierarchie von Prozessen an, die je nach Musikalität bei verschiedenen Menschen unterschiedlich weit reicht. Insgesamt nennt er sechs Stufen der Audiation: das Behalten einer kurzen Serie von Tonhöhen und Dauern, das Organisieren des Gehörten anhand einer mitt-

leren Tonhöhe und eines unterliegenden Rhythmus, die Etablierung von Tonalität und Meter, eine Restrukturierung oder Überprüfung des Gehörten anhand dieser etablierten Größen, ein In-Beziehung-Setzen des Gehörten zu anderen Musikstücken und schließlich die Vorhersage oder Erwartung einer bestimmten Fortsetzung des Gehörten. Deshalb und aufgrund der Zeitskala, die laut Gordon für die Audiation relevant ist und die er mit etwa drei bis acht Sekunden beziffert, ergeben sich eine Reihe musikpädagogischer Konsequenzen (Einsatz von „pattern vocabulary“, Solfege etc.). Die Details sind an dieser Stelle nicht weiter relevant, und es mag der allgemeine Hinweis genügen, dass diese Konsequenzen oftmals den Standardmethoden der etablierten Musikpädagogik und -wissenschaft zuwiderlaufen.

Auch das Buch *Music in the Moment* von Levinson aus dem Jahre 1997, das eine breite Diskussion im englischsprachigen Raum innerhalb der analytischen Ästhetik auslöste, kann als Polemik gegen bestimmte (Vor-)Urteile in der Musikwissenschaft gelesen werden. Levinson argumentiert detailliert anhand diverser Beispiele gegen die Meinung, musikalisches Verstehen sei nur dann möglich, wenn der Hörer die übergeordnete architektonische Struktur eines Musikstücks bewusst verfolge. Ein Großteil musikalischen Verstehens ist laut Levinson bereits durch die vorbewusste Verknüpfung der gegenwärtigen mit gerade vergangenen sowie unmittelbar bevorstehenden Passagen möglich; und solche Verknüpfungen werden immer dann automatisch gebildet, so Levinson, wenn man seine Aufmerksamkeit auf den gegenwärtigen Moment der Musik richtet.<sup>23</sup> Die zeitliche Umgebung wird „gleichsam mitgehört“ – Levinson spricht von *quasi-hearing*. Das „Gleichsam-Hören“ besteht damit aus insgesamt drei Komponenten: dem, was zu einem bestimmten Zeitpunkt gerade akustisch präsent ist; dem, was gerade nicht mehr physikalisch aktual ist, aber noch lebhaft erinnert wird (*vivid remembering, vivid memory*); und dem, was noch nicht präsent ist, aber schon lebhaft antizipiert wird (*vivid anticipation*).<sup>24</sup>

Man nimmt also laut Levinson in jedem Moment – „Moment“ hier verstanden als die subjektive und nicht als physikalische Grundeinheit der Zeit – eine zeitlich ausgedehnte musikalische Gestalt wahr. Die zeitliche Ausdehnung selbst liege im Bereich der Dauer einer einzelnen Melodie – also typischerweise im Bereich von Sekunden.<sup>25</sup>

Die Wahrnehmung einer Melodie sollte aber nicht nur innerhalb der Musikwissenschaft und der Philosophie der Musik von besonderem

Interesse sein. Und zumindest zu Lebzeiten Husserls war die auditorische Wahrnehmung tatsächlich das Standardbeispiel in Untersuchungen zur Zeit bzw. zum Zeitbewusstsein.<sup>26</sup> Insbesondere schreibt Husserl, auf den die Diskussion an dieser Stelle beschränkt werden soll, dass „das Wahrnehmen einer Melodie ein zeitlich ausgebreiteter, sich allmählich und stetig entfaltender Akt“ ist, bei dem zu einem jeweiligen und immer neuen Jetzt, in dem ein Ton gehört wird, „zugleich ein Soeben-vergangen, und wieder ein Noch-weiter-vergangen in einigen Gliedern gegenständlich ist; und vielleicht auch ein oder das andere gegenständlich als ‚künftig‘.“<sup>27</sup>

Für Husserl – wie für Levinson – ist die Wahrnehmung einer Melodie immer ein sich zeitlich erstreckender Prozess. Man nimmt nicht nur den gegenwärtig *physikalisch* präsenten Ton wahr, sondern auch die gerade vergangenen und bevorstehenden Töne sind *dem Hörer präsent*. Statt *vivid memory* und *vivid anticipation* benutzt Husserl hier die Begriffe von Retention und Protention; und den mit der vermeintlich punktförmigen physikalischen Gegenwart zusammenhängenden Teil der Wahrnehmung, den Levinson mit *actual hearing* bezeichnet, nennt Husserl Urimpression. Retention, Urimpression und Protention zusammen – also das, was nach Levinson „gleichsam-gehört“ wird – bilden bei Husserl den „Zeithof“.<sup>28</sup>

Allerdings sind diese Phänomene für Husserl von viel allgemeinerer Bedeutung als für Levinson. Levinson geht es allein um eine Analyse musikalischen Verstehens, während Husserl eine philosophische Phänomenologie vom reinen Bewusstsein überhaupt zu etablieren versucht, an deren Basis die gerade genannten Zeitbegriffe stehen. Retention, Urimpression und Protention sind die wesentlichen Bestandteile aller geistigen (intentionalen) Akte – Bestandteile dessen, wie einem Bewusstsein die Welt erscheint. Der Zeithof ist laut Husserl nicht nur für die Wahrnehmung von Musik wesentlich, sondern beispielsweise auch dafür, dass man Sprache versteht oder dass Gegenstände uns als kontinuierlich existierende erscheinen. Da der Zeithof phänomenologisch fundamental ist, dürfen die entsprechenden intentionalen Momente (Retention, Urimpression und Protention) insbesondere nicht mit anderen, „höherstufigen“ wie etwa einer bewussten Wiedererinnerung und Erwartung verwechselt werden. In der Retention bleibt etwas, das zuvor ins Bewusstsein eingegangen ist, in modifizierter Form weiterhin präsent (in Form einer „Abschattung“, wie Husserl

es nennt). Während laut Husserl eine Wiedererinnerung eine *Re*-Präsentation im strengen Wortsinne ist, bei der etwas zuvor Erfahrenes gleichsam wieder hervorgeholt wird, ist eine Retention in Husserls Redeweise „präsentativ“; d.h. sie trägt konstitutiv zum aktuell Gegebenen bei.<sup>29</sup>

Obwohl ihm die Melodiewahrnehmung als Standardbeispiel dient, macht Husserl keine Angaben über die Ausdehnung des Zeithofs in physikalischen Maßeinheiten. Denn immerhin ist das, was in Retention, Urimpression und Protention gegeben ist, etwas, das es mit Aufmerksamkeit, mit Intentionalität zu tun hat. Von daher ist es nicht identisch mit physikalischen Gegebenheiten und auch nicht ohne Weiteres in deren Maßeinheiten angebbar. Vielmehr soll ja die Phänomenologie eine allgemeine und grundlegende Wissenschaft des Bewusstseins sein, aus der die Möglichkeit physikalischer Betrachtungen überhaupt erst abzuleiten ist. Das gilt sowohl allgemein für die Physik als Einzelwissenschaft wie auch insbesondere für den physikalischen Zeitbegriff, der bei Husserl erst über den Zeithofbegriff konstituiert wird.<sup>30</sup>

Da solche grundlegenden Überlegungen für die thematisch deutlich stärker eingeschränkten Untersuchungen von Gordon und Levinson nicht relevant sind, erfährt man bei ihnen auch entsprechend wenig darüber, wie fundamental Audiation bzw. Gleichsam-Hören im Vergleich zu anderen geistigen Prozessen außerhalb der Melodiewahrnehmung sind. Wichtig ist allerdings für beide, Levinson und Gordon, dass es sich bei Audiation bzw. Gleichsam-Hören um einen Prozess handelt, der zumindest *fundamentaler* ist als das bewusste Verfolgen der architektonischen Struktur etwa einer Symphonie. Denn sonst liefe ihre Kritik an den althergebrachten Meinungen in der Musikpädagogik und -wissenschaft ins Leere.

Die Ansätze von Gordon, Levinson und Husserl sind also zumindest in den folgenden Aspekten miteinander vergleichbar: Audiation, Gleichsam-Hören und der Zeithof sind allesamt Vorbedingungen für musikalisches Verstehen. In allen drei Fällen handelt es sich um einen geistigen Prozess, der in jedem Moment existiert, dessen Inhalt sich aber über eine gewisse physikalische Zeitspanne in die Vergangenheit (und Zukunft) erstreckt. Zugleich betonen alle drei Autoren (Levinson und Gordon mit Bezug auf die Musik, Husserl allgemeiner) den Unterschied zu einem lang-andauernden, bewussten Verfolgen zeitlicher Strukturen und zu Akten expliziter Wiedererinnerung.

Dieser thematische Bogen, den ich gerade von Gordon über Levinson zu Husserl geschlagen habe, lässt sich sogar zu einem Kreis schließen. Denn die Arbeiten von Husserl und Gordon werden durch eine Reihe musikphänomenologischer Ansätze miteinander verbunden, die sich um eine Vereinigung von musikalischer Erfahrung und Musiktheorie auf der Basis der Phänomenologie bemühen.<sup>31</sup> Und auch hier münden die Ergebnisse zumeist in einer starken Kritik an musiktheoretischen Schulmeinungen und im Vorschlag konkreter Lernkonzepte, wie sie auch von Gordon vertreten werden.<sup>32</sup>

Aber kann man den Bogen von den Autoren, deren jeweiliges Zeithof-Konzept in diesem Abschnitt behandelt wurde, auch wieder zurück-schlagen zu den im vorigen Abschnitt behandelten Neurowissenschaften? Zumindest einmal gibt es eine starke äußere Ähnlichkeit zwischen der Gestalt der elektro-physiologisch gemessenen CNV-Kurve und den von Husserl selbst gegebenen schematischen Darstellungen des Zeithofs. Der zeitliche Verlauf einer CNV gleicht dem, wie Husserl den Abschattungsmechanismus (also den Übergang von Urimpression zu Retention und dann zu Retention der Retention etc.) darstellt. Und auch der Befund, bei der CNV handle es sich um einen begleitenden Prozess (*monitoring*) und nicht um das Wieder-Aufrufen einer Intervallvorlage, findet sein Analogon in Husserls Betonung, Retentionen seien präsentativ und nicht re-präsentativ.

Solche strukturellen Analogien deuten auf die Möglichkeit hin, dass sich Phänomenologie und Neurowissenschaft (zumindest) *heuristisch* befruchten könnten. Um dies genauer illustrieren zu können, sollen zunächst noch weitere strukturelle Analogien kurz benannt und skizziert werden: Man kann beispielsweise mögliche Unterfunktionen der kortikalen auditorischen Erinnerungsspur (MMN) zu unterscheiden versuchen, indem man Diskussionen zu Unteraspekten der Retention aufgreift.<sup>33</sup> Ergänzend dazu könnte die Untersuchung des Verhältnisses zum *sustained field* als einer „Echtzeit“-Antwort fruchtbar sein, da man hier im Vergleich zur transienten MMN von einem zusätzlichen oder anderen „retentionalen Charakter“ sprechen könnte – zumindest wenn man bereit ist, das „zeitliche Mitlaufen“ eines *sustained field* (heuristisch) in Beziehung zu einem Prozess der zeitlichen Abschattung zu setzen. Weiterhin mag auch die phänomenologische Diskussion um die Rolle der Protentionen, deren Existenz schon unter den direkten Brentano- und Husserl-Schülern für Kontroversen sorgte

und die immer noch stark diskutiert wird, heuristische Impulse durch einen Vergleich mit der Neurophysiologie erfahren.<sup>34</sup> Und schließlich sind viele funktionelle Unterscheidungen, die von Neurophysiologen eingeführt werden und mit denen sie eine allgemeine Ordnung in den Zoo der neuronalen Prozesse zu bringen hoffen, selbst durch einen im gewissen Sinne phänomenologischen Hintergrund geprägt.<sup>35</sup> Dass es für diesen oder jenen geistigen Prozess (beispielsweise des Erinnerns) diese und jene Subfunktionen geben müsste, ist eine quasi phänomenologische Vorentscheidung, um dann im Experiment nach zeitlich und räumlich separierten neuronalen Aktivitäten Ausschau zu halten. Denn unterschiedliche Funktion meint in einer naturalistisch motivierten Neurophysiologie nichts anderes als Instanzierung an einem anderen Raum-Zeit-Punkt.

Neben den Hirnantworten, wie sie tatsächlich empirisch an der Schädeldecke als Ergebnis der Überlagerung der elektrischen Aktivität von Millionen von Neuronen abgeleitet werden, gibt es mittlerweile auch Computersimulationen, die auf rekurrenten, neuronalen Netzwerken basieren. Insbesondere hat hier Dan Lloyd unlängst die CNV zumindest ihrer allgemeinen Verlaufsform nach erfolgreich simuliert.<sup>36</sup> Das ist an dieser Stelle deshalb bemerkenswert, weil sich damit nicht nur eine strukturelle Analogie zwischen Neuroinformatik und -physiologie zeigt, sondern weil Lloyd selbst davon spricht, dass er „Neurophänomenologie“ betreibe und den fundamentalen Charakter der (von ihm offensichtlich naturalistisch verstandenen) Retention für sämtliche Hirnprozesse aufzeigen will.<sup>37</sup> – Wie sinnvoll die Behauptung ist, Retentionen seien im Gehirn zu finden, wird mich bei der Diskussion des Naturalismus im nächsten Abschnitt beschäftigen. Aber auch unabhängig davon gibt es problematische Aspekte in Lloyds Bezugnahme auf die Phänomenologie:

Selbst wenn eine Naturalisierung der Phänomenologie möglich ist, so wird doch eine CNV nur dann evoziert, wenn der Proband bestimmte Stimuli oder Stimuluseigenschaften aufmerksam verfolgt. Der Zeithof hingegen ist in der Husserlschen Systematik notwendig ein universelles Phänomen und unabhängig von zusätzlicher mentaler Fokussierung. Wollte man insbesondere diesen Aspekt in eine vermeintliche Neurophänomenologie retten, so müsste man die Mechanismen der prä-attentiven MMN und des *sustained field* und nicht die der CNV als grundlegende Prozesse betrachten.

Allerdings kann man Lloyds „Neurophänomenologie“ sicherlich insofern ein Husserlsches Erbe eingestehen, als dass sie eben versucht, die (naturalistisch verstandenen) Retentionen als *die* fundamentalen Grundbausteine geistiger bzw. kortikaler Prozesse zu verstehen. In dieser Beziehung steht Lloyd dem Husserlschen Projekt also sicherlich näher als es die bescheideneren, musikbezogenen Projekte von Levinson und Gordon tun.

Eine weitere wichtige Frage betrifft das „Zeitfenster“ der verschiedenen diskutierten Phänomene. Eine MMN wird nur bei auditorischen Stimulussträngen ausgelöst, die keine Unterbrechungen von mehr als zehn Sekunden haben, und auch die stetig ansteigende Aktivität der CNV bricht nach etwa dieser Zeit zusammen. Weiterhin hatten auch Gordon und Levinson explizit diese Zeitskala als Größenordnung der Audiation bzw. des Gleichsam-Hörens genannt. Husserl macht hier, wie erwähnt, keine konkreten Angaben. Allerdings gleichen seine Beispiele (kurze Melodien, Sätze) denen von Levinson und Gordon.

Weiterhin ist dieses Zeitfenster von mehreren Sekunden auch in der Psychologie von entscheidender Bedeutung, in der es als *psychological present* (auch *perceptual* oder *specious present*) diskutiert wird.<sup>38</sup> Für die dortigen Diskussionen ist ebenfalls das Hören als Modalität der Zeit zentral (im Gegensatz etwa zum Sehen als Modalität des Raumes): Die Details dieser „psychologischen Gegenwart“ werden einmal mehr durch Melodie- und Satzwahrnehmungen expliziert, und sie selbst wird wiederum akribisch von Prozessen der Wiedererinnerung unterschieden.<sup>39</sup>

Die Charakteristika und Ausführungen zur psychologischen Gegenwart decken sich also in diversen Aspekten mit den Gegenwarts Konzepten bei Husserl, Gordon und Levinson. Aber die Frage nach der Naturalisierung – ob oder inwiefern man also Zeithöfe mit der Stoppuhr messen kann – wurde damit noch nicht berührt. Bisher wurden lediglich gewisse strukturelle Gemeinsamkeiten von Zeitkonzepten diskutiert, wie sie in der Neuroinformatik, Neurophysiologie, Musikpädagogik und in verschiedenen Bereichen der Philosophie auftreten (wobei diese Liste keinen Anspruch auf Vollständigkeit erhebt) und wie sie einander womöglich heuristisch befruchten können. Es stellt sich aber nun die Frage, ob solche analogischen Übertragungen *ausschließlich* heuristische Relevanz haben. Oder hat es zudem eine systematische Bedeutung, wenn bestimmte physikalische Charakteristika einer



MMN untereinander so verknüpft sind wie bestimmte phänomenologische Charakteristika einer Retention oder wenn an Husserl anknüpfende Musikwissenschaftler, ein analytischer Philosoph sowie ein Jazz-Musiker und Musikpädagoge mit gleichen Argumenten gegen dieselben musiktheoretischen Vorstellungen und Lernkonzepte angehen?

#### 4. Naturalistische und transzendente Reduktionen

An dieser Stelle sind vielleicht Reduktionismen das Nächstliegende: Wieso sollte das Zeithofkonzept so multidisziplinär sein, wenn nicht aus dem Grund, dass es eine fundamentale Disziplin gibt, für die es gilt und durch die es sich auf alle anderen Disziplinen überträgt? Wieso separat einen neurophysiologischen, einen musiktheoretischen und einen phänomenologischen Zeithof annehmen und somit zusätzlich über mögliche Zusammenhänge reflektieren müssen?

Die beiden Reduktionismen, die ich behandeln möchte, sind die vermeintlich typischen des Neurophysiologen und des Phänomenologen, die im Folgenden kurz als „Naturalismus“ und als „transzendentaler Reduktionismus“ bezeichnet werden. Naturalismus verstehe ich hier als die Annahme, das Geistige oder Mentale sei eine Erweiterung der Natur, wobei Natur die Menge aller physikalischen und damit insbesondere auch aller neurophysiologischen Abläufe meint.<sup>40</sup> (Auf den Begriff des transzendentalen Reduktionismus gehe ich später näher ein.)

Es soll hier kein strenges Argument gegen diese beiden Reduktionismen vorgebracht werden. Allerdings werde ich in meinen folgenden, zum Teil eher programmatischen Überlegungen zu zeigen versuchen, dass und warum beide nicht attraktiv sind bzw. ihr eigentliches Ziel verfehlen und einen jeweils erheblichen Teil der menschlichen Erfahrungswelt diskreditieren.

Naturalistische Ambitionen im Kontext der Neurophysiologie werden von einigen ihrer Vertreter (siehe etwa den oben erwähnten Lloyd) durchaus als Versuche verstanden, die Husserlsche Philosophie zu stärken und zu einer Neurophänomenologie umzubauen. Ihre Argumentation verläuft dabei grob wie folgt: Die Neurophysiologie sei eine weit hin beachtete und erfolgreiche Disziplin, und wenn es also gelänge, die sozusagen etwas angestaubte Phänomenologie durch die Neurophysiologie (oder zumindest mit ihr) auf eine „solide experimentelle Grund-

lage“ zu stellen, um so besser. Genau vor diesem Hintergrund erklärt sich also die enorme Anzahl an Veröffentlichungen, die sich gerade im englischsprachigen Raum in den letzten Jahren der Naturalisierung der Phänomenologie widmen.<sup>41</sup> Dabei ist den Autoren in den meisten Fällen durchaus bewusst, dass der transzendentalphilosophische Hintergrund der Phänomenologie zunächst allen Naturalisierungsversuchen zuwiderlaufen muss. Tim van Gelder beispielsweise nennt einen seiner Aufsätze über Phänomenologie und Kognitionswissenschaften deshalb auch „Wooden Iron?“.

An einer Stelle in diesem Aufsatz handelt van Gelder von einer gegenseitigen Bereicherung zwischen Phänomenologie und Neurowissenschaft und von einander ergänzenden wissenschaftlichen Beiträgen.<sup>42</sup> Das klingt zunächst wie das, was oben als Möglichkeit der gegenseitigen heuristischen Befruchtung bezeichnet wurde. Doch einige Seiten weiter behauptet van Gelder, seine zentrale These sei „the bold one that cognitive science can tell us what retention and protention *actually are* and in that sense deepen our understanding of them and of time consciousness in general.“<sup>43</sup> Mit dieser Erkenntnis würde sich dann auch unser Verständnis der Phänomenologie vertiefen und verbessern. Hier vertritt van Gelder also eindeutig einen neurophysiologischen Reduktionismus. Es ist die Neurophysiologie, die uns sagen kann, was denn Retentionen „nun wirklich sind“ und worum es in der Phänomenologie „eigentlich“ geht.

Ein solches Verständnis von gegenseitiger Bereicherung, die in einen Reduktionismus mündet, ist allerdings nur dann möglich, wenn man eine Unterscheidung trifft zwischen der tatsächlichen momentanen Situation in der Neurowissenschaft und ihrem (idealen) Endzustand. Gemäß einer solchen Unterscheidung gäbe es eine vollständige Reduktion im idealen Endzustand, von dem die Neurophysiologie in ihrem gegenwärtigen Stand aber als noch weit entfernt angenommen werden müsste. Eine vollständige physiologische Reduktion etwa von Retentionen, so die These, ist *noch nicht* möglich, weshalb die Phänomenologie (noch) ein sinnvolles Ergänzungsprojekt ist. Anders formuliert: Während sich der Neurophysiologe bei der Beschreibung des Bewusstseins quasi von unten hocharbeitet, arbeitet der Phänomenologe vorläufig und von oben nach unten. Auf Dauer würden mehr und mehr vermeintlich genuin phänomenologische Begriffe auf neurophysiologische zurückgeführt werden können, und einzig relevant bliebe der Erklä-

rungsverlauf von unten nach oben. – Solch eine argumentative Zuflucht auf einen idealen Endzustand einer Einzelwissenschaft ist allerdings gerade im Zusammenhang der Diskussion von Naturalismus bzw. Physikalismus überzeugend kritisiert worden.<sup>44</sup>

Einen etwas anderen Ansatz für eine Neurophänomenologie vertreten hier Varela und Thompson. Stärker als van Gelder betonen sie die besondere Rolle der Theorie dynamischer Systems in der Analyse neurophysiologischer Daten. Es sei aufgrund dieser mathematischen Theorie möglich, sich allein auf die geometrischen und topologischen Aspekte neuronaler Aktivität bzw. von Aktivität ganz allgemein zu beziehen, wodurch diese Analysemethode in gewisser Weise ontologisch neutral und also auch auf das Phänomenale anwendbar sei.<sup>45</sup> (Im nächsten Abschnitt wird darauf kurz zurückzukommen sein.)

Doch letztlich bleiben die Behauptungen, dass die Neurowissenschaften klären könnten, was Retentionen „eigentlich sind“, bzw. dass alles letztlich auf die Theorie dynamischer Systeme zurückführbar sei, unzureichend begründet. So ist unklar, inwiefern allein die Theorie dynamischer Systeme dazu geeignet sein sollte, die Phänomenologie zu formalisieren. Und selbst innerhalb der Neurowissenschaften ist diese explizite Vorrangstellung problematisch, da es hier durchaus andere Auswertemethoden für neurophysiologische Daten gibt.

Auch ließe sich allgemeiner einwenden, dass ich mit Sicherheit nicht *eigentlich* Hirnpotentiale meine, wenn ich über musikalisches Verstehen spreche. Und auch wenn ich mit einer EEG-Haube auf dem Kopf in der nächsten langweiligen Vorlesung und im nächsten mitreißenden Konzert sitze: Die Vorlesung werde ich erneut als länger empfinden als den Konzertbesuch. (Eine gewisse Ausnahme bildet hier allerdings das so genannte Neurofeedback, bei dem zu therapeutischen Zwecken Epilepsie- und Schlaganfall-Patienten mithilfe der Betrachtung einer Echtzeit-Analyse ihrer eigenen Gehirnströme instrumentell konditioniert werden; siehe unten.)

Die Frage, warum mir die beiden Ereignisse unterschiedlich lang vor kamen, wird damit nicht befriedigend beantwortet. „Im einen Fall war diese und im anderen Fall jene Aktivität in ihrem Parietallappen“ ist schlichtweg keine vernünftige Antwort auf die Frage „Warum war das Konzert mitreißender als die Vorlesung?“ Im einen Fall geht es um wirkursächliche Zusammenhänge, im anderen Fall um Zusammenhänge im Wahrgenommenen, in subjektiven Zuständen, die durchdrungen sind

von Bewertungen und Gründen. Doch diese beiden Arten, Zusammenhänge zu erzeugen, sind nicht als wechselseitige Bedingungen miteinander gekoppelt. Deshalb ist es auch umgekehrt nicht sinnvoll, jemandem auf seine Frage „Wie lange benötigt das Licht, um die Strecke zwischen Sonne und Erde zurückzulegen?“ zu antworten: „So lange, wie ihnen gestern das Anstehen an der Supermarktkasse vorkam.“

Und selbst wenn bestimmte neurophysiologische Prozesse *notwendig* sind, um sich an etwas zu erinnern, um eine Tonhöhe als solche wahrzunehmen oder um Töne unterschiedlicher Dauer voneinander unterscheiden zu können: Das Aufzeigen notwendiger Bedingungen ist nicht gleichbedeutend mit dem Aufzeigen eines „Wesens“ von etwas, damit, dass man zeigt, was etwas „eigentlich“ oder „in Wirklichkeit“ ist. Es mag wohl notwendig sein, Beine zu haben, um ein guter Fußballspieler zu sein. Aber „das Wesen“ des Fußballspiels von Mesut Özil ist deshalb nicht in seinen Beinen angesiedelt; und es sind auch nicht seine Beine, die „eigentlich“ Fußball spielen.<sup>46</sup>

Übertragen auf die Zeit bedeutet das: Diejenigen Hirnaktivitäten, die mit der Zeitwahrnehmung korreliert sind, sind nicht dasjenige, was Zeit bzw. Zeitwahrnehmung „wirklich“ oder „wesentlich“ ist. Auch zeigen sie nicht, dass es in „Wirklichkeit“ nur eine physikalische und keine subjektive Zeit gibt. Denn auch der Blick auf die Uhr, der in beiden Fällen die gleiche physikalische Dauer anzeigt, lässt das erwähnte Konzert und die erwähnte Vorlesung nicht subjektiv gleich lang werden – zumindest in der Regel (auf mögliche Ausnahmefälle komme ich unten kurz zurück).

Auch um das reduktionistische Projekt in umgekehrter Richtung – also transzendental-phänomenologisch statt naturalistisch – ist es nicht viel besser bestellt. Nur zu Erinnerung: Phänomenologie wird von Husserl verstanden als eine anschauliche Methode der Analyse und Beschreibung von bewussten mentalen Repräsentationen, von Urteilen, Wahrnehmungen usw. auf allgemeiner Ebene.<sup>47</sup> Phänomenologie hat es nicht mit besonderen geistigen Zuständen eines einzelnen Menschen zu tun. Als formale Ontologie ist sie ein viel allgemeineres Unternehmen als etwa die empirischen Wissenschaften. Phänomenologisch betrachtet befassen sich Naturwissenschaften mit einem spezifischen, sachhaltigen Teilaspekt davon, wie dem Menschen die Welt erscheint – und insofern könnte man eine Disziplin wie die Neurowissenschaft als eine *regionale* Ontologie im Sinne Husserls begreifen.<sup>48</sup> Ihren jeweiligen

Teilaspekt kann eine Einzelwissenschaft nämlich nur deshalb so genau beleuchten, weil sie im Vergleich zur philosophischen Fundamentaldisziplin der Phänomenologie, die ihr systematisch vorangeht, zusätzliche empirische Annahmen machen muss. Umgekehrt reduzieren sich also sämtliche Einzelwissenschaften, sobald man von ihren spezifischen Zusatzannahmen absieht, auf Aspekte der Phänomenologie – und das ist es, was oben mit dem Begriff eines transzendentalen Reduktionismus bezeichnet werden sollte.

Dieser Reduktionismus tritt bei Husserl vor allem dann deutlich hervor, wenn er gegen eine Gleichsetzung von Phänomenologie und Psychologie argumentiert. Dabei führt er als Analogie meistens das Verhältnis von Mathematik und Naturwissenschaft an. Für Husserl ist die Phänomenologie das „Fundament“, die Grundlage der Psychologie, so wie die Mathematik die Grundlage aller Naturwissenschaften sei.<sup>49</sup> Die Möglichkeit eines gleichberechtigten Wechselverhältnisses zwischen den beiden Disziplinen wird von ihm damit abgelehnt. Gerade das ist aber vor dem Hintergrund der obigen Beispiele aus dem auditorischen Bereich und auch aufgrund anderer jüngerer neurophänomenologischer Arbeiten nicht überzeugend.<sup>50</sup> Was sich in diesen Arbeiten andeutet, ist weniger die Hoffnung auf eine phänomenologische Fundierung der Neurophysiologie und Psychophysik als eben eine fruchtbare und gleichberechtigte Wechselbeziehung zwischen diesen Disziplinen und der Phänomenologie. – Ganz ähnlich verhält es sich übrigens auch für den von Husserl bemühten Vergleich zum Verhältnis von Mathematik und Physik. Hier hat nämlich gerade in den letzten Jahrzehnten eine ausgeprägte *gegenseitige* konzeptionelle Befruchtung etwa zwischen Teilchenphysik und Topologie stattgefunden.

Außerdem stellt sich die Frage, was es überhaupt für einen Sinn haben sollte, die Möglichkeit aufzuzeigen, dass es sich bei der Neurophysiologie um eine regionale Ontologie im Sinne Husserls handelt. Ebenso wie sich der Phänomenologe weiterhin dagegen wehren wird, wenn man ihm unterstellt, er meine doch *eigentlich* bestimmte Hirnaktivitäten, wenn er von „Retentionen“ spricht, so wird sich auch umgekehrt der Neurophysiologe gegen die Behauptung wehren, dass er, wenn er über „MMN“ und „CNV“ spricht, doch *eigentlich* bestimmte Intentionalitäten meine. Sicherlich gilt: In den Fällen, in denen man sich heuristisch befruchten kann, soll man es auch tun. Aber ein Neurophysiologe, der jede seiner Arbeiten streng phänomenologisch ableiten möchte, wird

aller Voraussicht nach nicht lange ein guter Neurophysiologe bleiben – und das Entsprechende gilt für einen Phänomenologen, der seine Untersuchungen allein und verbindlich auf die Neurophysiologie aufbaut.<sup>51</sup>

Naturwissenschaften und Philosophie sind wichtige kulturelle Unternehmen, die ganz unterschiedliche Funktionen erfüllen. Beide helfen auf ihre Weise, dass Menschen sich in der Welt handelnd zurechtfinden. Und sie können sich heuristisch befruchten, wie hoffentlich die vielen konzeptionellen Analogien, die oben diskutiert wurden, gezeigt haben (und wie es von neurophänomenologischer Seite etwa Varela und Thompson systematisch weitergedacht haben). Vielleicht fördert ein ernsthaftes, zunächst heuristisches Interesse an einer anderen Disziplin auch den gegenseitigen Respekt und baut naturalistischen bzw. transzendentalphilosophischen „Übernahmeversuchen“ der jeweils anderen Disziplin vor.<sup>52</sup> Im letzten Abschnitt möchte ich deshalb einen begrifflichen Rahmen skizzieren, der es möglich erscheinen lässt, das jeweils andere kulturelle Unternehmen als ein Eigenständiges ernst zu nehmen.

## 5. Pluralismus mit Knoten

Statt nach einer monistischen Welt Ausschau zu halten, in der alle Erscheinungen letztlich von gleicher Art sind, sollte man vielmehr nach einem gemeinsamen Bezugspunkt suchen, von dem aus die verschiedenen Perspektiven als jeweils eigenständig denkbar sind. Es sollte nicht darum gehen, die Kluft, die zwischen physikalischen und bewusstseinstheoretischen Beschreibungen besteht, quasi durch Gleichmacherei zuzuschütten, sondern man sollte diese Kluft ernst nehmen und eher etwas zwischen ihnen Vermittelndes suchen, das Aspekte sowohl der einen wie der anderen Seite in sich trägt, selbst aber kein Bestandteil einer dieser Seiten ist.

Nun ist das Gehirn sicherlich kein guter Kandidat für eine gemeinsame, gleichberechtigte Beschreibung des Physikalischen und des Geistigen. Die geistige Wirklichkeit würde immer ins Hintertreffen geraten, da mit dem Gehirn der Ausgangspunkt selbst ja ein physikalischer wäre. Und Entsprechendes gilt für das Bewusstsein als Ausgangspunkt für die physikalische Welt. Will man eine Stellung einnehmen, von der aus man auf beides, auf Transzendentalphilosophie wie auch auf die Naturwis-

senschaften als gleichberechtigt blicken kann, dann ist das Bewusstsein oder das Ich ein ähnlich schwieriger Ausgangspunkt wie das Gehirn.

Eine aussichtsreiche Alternative bietet stattdessen der zentrale Gegenstand dieses Aufsatzes: die Zeit. Sie erlaubt eine nicht-reduktionistische Verbindung zwischen der physikalischen und der subjektiven Welt. Die Analogien und Korrespondenzen zwischen den unterschiedlichen philosophischen und physikalischen Disziplinen, die oben diskutiert wurden, wären demzufolge nicht direkt aufeinander reduzierbar, liefen allerdings in der Dimension der Zeit, die sowohl physikalische als auch bewusstseinstheoretische Charakteristika besitzt, als einem „Knoten“ zusammen.

Doch bevor ich diesen Begriff nun weiter ausführe, möchte ich einer Kritik vorbeugen; nämlich der, dass gerade eine solch metaphorische Redeweise wie die von einem „Knoten“ niemals etwas Anderes sein könne als der Ausdruck einer Hilflosigkeit gegenüber dem Problem selbst. Diese Kritik wäre nur dann berechtigt, wenn sich die beiden erfolgreichen Unternehmen Naturwissenschaft und Transzendentalphilosophie (bzw. Phänomenologie) tatsächlich in einem strengen Sinne aufeinander reduzieren ließen. Doch genau das erscheint nicht plausibel. Ihr Verhältnis zueinander lässt sich allerdings beschreiben; und das schließt die Verwendung metaphorischer Ausdrücke ein (wie auch die Bezugnahme auf statistische Korrelationen und funktionale Zusammenhänge).

Die nicht-reduktionistische Redeweise von der Zeit als einem Knoten ist ebenso wenig Ausdruck einer Hilflosigkeit, wie diejenige von den beiden (miteinander korrelierten) Seiten einer Münze. Statt also schlicht zu behaupten, eine weitere Reduktion müsse möglich sein, wäre an dieser Stelle beispielsweise eine historisch systematische Rekonstruktion des Leib-Seele-Zusammenhangs im Sinne einer Blumenbergschen Metaphorologie philosophisch aufschlussreicher.<sup>53</sup> Doch das kann an dieser Stelle nicht geleistet werden, so dass vielleicht der kurze Verweis auf ein paar Autoren genügen mag, die ein nicht-reduktionistisches Verhältnis zwischen Geist und Materie mit Hilfe einer prägnanten Metapher beschrieben haben: So spricht Leibniz von einer „prästabilierten Harmonie“, die sich bis in die Details seiner Pneumatologie einerseits und seiner Dynamik (Physik) andererseits auswirkt.<sup>54</sup> Schelling (in den Jahren um 1800) ordnet die Natur- und Transzendentalphilosophie als gleichberechtigte, aber gegenläufige Projekte ein.<sup>55</sup> Und bei Viktor von Weizsäcker wird in der ersten Hälfte des vergangenen Jahrhunderts

die Psychosomatik über die metaphorische Rede von einer Drehtür motiviert, die zwar eine Innen- und eine Außenseite habe, aber dennoch immer die gleiche Tür sei. Oder man denke an Carl Friedrich von Weizsäckers Theorie der Ure, die sogar eine gewisse Ähnlichkeit mit der Neurophänomenologie von Varela aufweist – und zwar in der Art und Weise, wie auf die Möglichkeit und Bedeutung einer Mathematisierung des philosophischen Ansatzes Bezug genommen wird. In der Theorie der Ure geht es zwar nicht um neurophysiologische Prozesse im Besonderen (und dementsprechend auch nicht um die Überhöhung bestimmter neurophysiologischer Auswertemethoden; siehe oben), aber auch von Weizsäcker nimmt in gewisser Weise eine „ontologische Neutralität“ der Mathematik an und führt seine Theorie der Natur wie der Subjektivität auf ein und dieselbe mathematische Gruppenstruktur zurück.<sup>56</sup>

Diese kurze Aneinanderreihung einiger Autoren bedeutet selbstverständlich nicht die Gleichheit ihrer Beschreibungsinteressen. Hier gibt es offensichtlich spezifische Unterschiede; und Ähnliches gilt auch für den gerade benutzten Begriff des Knotens, bei dessen Verwendung ich mich teilweise an diejenige bei Peter Rohs anschließe und die ich noch kurz skizzieren möchte.

In seiner durch Fichte und Spinoza beeinflussten „feldtheoretischen Transzendentalphilosophie“ schlägt Rohs vor, die Zeit als „ontologischen Knoten“ zu betrachten, von dem aus sich sowohl die physikalische wie auch die geistige Wirklichkeit verstehen lasse.<sup>57</sup> Es geht hier nicht darum, die subjektive auf die physikalische Zeit zu reduzieren oder umgekehrt, was ohnehin nicht sehr erfolgversprechend wäre.<sup>58</sup> Es ist aber auch gar nicht nötig, da sowohl die bewusst erlebte Wirklichkeit wie auch diejenige physikalischer Prozesse sich zeitlich ordnen. Die Zeit, einmal als subjektive und einmal als physikalische, ist Bezugsordinate dieser Aspekte der Wirklichkeit und bildet auch deren Verknüpfung.

Die Perspektive der subjektiven Zeit ist diejenige des Bewusstseins, des Ich. In ihrer Wirklichkeit gibt es ein Werden, gibt es eine gegenüber Vergangenheit und Zukunft ausgezeichnete Gegenwart. In der subjektiven Zeit wird gehandelt und gibt es begründbare Handlungen. Demgegenüber spannt die physikalische Zeit laut Rohs das physikalische Feld auf, das kontinuierlich ist und in dem es nur Sein und kein Werden gibt.<sup>59</sup> Hier wird nicht gehandelt oder erlebt, hier gibt es nur kausale Zusammenhänge, nur lagezeitliche früher-später-Relationen,



aber keine genuin modalzeitliche Ordnung von vergangen, gegenwärtig und zukünftig.

Soweit Rohs. Doch wie lässt sich das nun genauer auf die neurowissenschaftlichen und phänomenologischen Überlegungen und Befunde dieses Aufsatzes beziehen? – Meines Erachtens müssen hier die oben behaupteten strukturellen Analogien zwischen Phänomenologie und Neurophysiologie zurückgeführt werden auf eine strukturelle Analogie zwischen den Ordnungen der subjektiven und der physikalischen Zeit selbst; und zwar im Sinne einer (partiellen) Isomorphie zwischen den Relationen einer modalzeitlichen (subjektiven) und einer lagezeitlichen (physikalischen) Ordnung. Die zugrundeliegende Annahme wäre also die, dass es eine (teilweise) struktur-erhaltende Abbildung von den lagezeitlichen Relationen zwischen physikalischen Zuständen oder Ereignissen (der Form:  $x$  ist früher oder später als  $y$ ) auf die modalzeitliche Ordnung bewusster Zustände (der Form:  $x$  ist mehr oder weniger vergangen/zukünftig als  $y$ ) gibt. Plausibilisiert bzw. motiviert wird diese Betonung der modalzeitlichen Ordnung des Geistigen dabei gerade durch den Bezug auf die Phänomenologie. Denn laut ihr zeichnen sich bewusste Zustände gegenüber physikalischen durch ihre Intentionalität aus; also durch ihre Gerichtetheit, die sich auf grundlegender Ebene einzelner Akte genau durch die zeitlich intentionalen Momente der Retention, Urimpression und Protention manifestiert.

Die Frage, warum die Vorlesung langweilig war, mit dem Hinweis auf eine Hirnaktivität zu beantworten, oder die Ankunftszeit des Busses mit dem Hinweis auf ein Erlebnis der Dauer zu beantworten, wäre demzufolge deshalb unsinnig, weil die Ausprägung der Rahmenkoordinate Zeit nicht die gleiche ist. Intentionalität – und damit auch Gründe für Handlungen – existieren nicht in der physikalischen oder objektiven Zeit, Kausalursachen (strikt betrachtet) nicht in der subjektiven. Doch man kann diese beiden Ausprägungen von Zeit beispielsweise über Untersuchungen in der Psychophysik miteinander in Beziehung setzen. Dabei müssen diese Untersuchungen nicht wie bei Varela und Thompson direkt und exklusiv auf einen bestimmten mathematischen Formalismus wie etwa die Theorie dynamischer Systeme beschränkt werden. Man kann hier allgemeiner auf statistische Betrachtungen zurückgreifen, die als solche zunächst einmal weder physikalisch-wirkursächlich sind noch intentional-begründungstheoretisch (finalursächlich).

Somit lässt sich auch verstehen, was das Besondere ist an den oben

erwähnten Prozessen des Neurofeedback oder des permanenten Blicks auf die Uhr in der Vorlesung oder dem Konzert. In diesen Fällen ist man nämlich – anders als sonst im Alltag, in dem man nur selten solcher Art von introspektiven und reflexiven Betrachtungen nachgeht – mit der (partiellen) Isomorphie zwischen subjektiver und objektiver Zeit explizit konfrontiert. Es scheinen sich die körperlich-physiologische und die intentional-geistige Ebene auf besondere Weise anzunähern – und zwar genau deshalb, weil man selbst sie auf der Basis des Erlebens einer Korrelation zeitlicher Ordnungen miteinander in Beziehung setzt.

In diesem Zusammenhang sollte abschließend auch nochmals die zentrale Rolle der Hofstruktur der Zeit betont werden: Intentionale Zustände (Begründungen und Bewertungen) basieren nicht auf punktförmigen Ereignissen in der subjektiven Zeit, und Ursachen sind keine punktförmigen Ereignisse in der physikalischen Zeit. Auf Ersteres weist u. a. die Phänomenologie hin, letzteres wird neben der oben diskutierten Neurophysiologie auch durch die neuere Physik unterstützt.<sup>60</sup> Zeit als „ontologischer Knoten“ meint also keinen „Kristallisationspunkt“ für Physik und Bewusstseinstheorie, sondern eine „Umgebung“ – einen (physikalisch wie subjektiv prä-disjunktiven) Zeithof.

### Anmerkungen

- 1 Ähnlich argumentieren auch Horgan und Kriegel 2007, S. 133, 137, in ihrem Ansatz einer analytischen Theorie des Selbstbewusstseins.
- 2 Um hier einem möglichen Missverständnis vorzubeugen: Wenn im Folgenden auf phänomenologischer Seite von „Reduktionismus“ oder „Fundierung“ die Rede ist, so ist dies immer in einem transzendentalphilosophischen Rahmen (und eben in Bezug auf Husserls Analyse des intentionalen Bewusstseins) gemeint. Es geht also nicht um die Frage einer Fundierung jeglicher Wissenschaft in der Lebenswelt (dazu habe ich mich an anderer Stelle ausführlich geäußert; siehe Sieroka, 2010b, S. 324–332).
- 3 Siehe Lanz, 1996.
- 4 Siehe Conway und Tsao, 2009.
- 5 Siehe etwa Varela, 1999, Gallagher und Zahavi, 2008, Thompson, 2007. Vgl. aber andererseits auch Schmicking, 2003.
- 6 Siehe z. B. Crane, 2003, und Coates, 2007.
- 7 Elektrophysiologische Messverfahren sind die Elektro- und Magneto-Enzephalographie (EEG, MEG), weitere bildgebende Verfahren bspw. die funktionelle Kernspinresonanz- und die Positronenemissions-Tomographie (fMRI, PET). Eine einführende Übersicht zu diesen Verfahren bietet etwa Jäncke, 2005.

- 8 Siehe zu Folgendem insbesondere Näätänen et al., 2001.
- 9 Umgekehrt kommen Koma-Patienten, sobald bei ihnen eine MMN gemessen wird, in der Regel in den darauf folgenden Tagen wieder zu Bewusstsein.
- 10 Weiterhin haben wohl auch die gesteigerten räumlichen und zeitlichen Auflösungsvermögen der neuen Messmethoden erheblich zur Verwendung dieser Rhetorik beigetragen – einer Rhetorik, die offensichtlich mit einem allgemeinen Interesse zusammenläuft, geistige Zustände wie Erwartungen oder gar den freien Willen in einem Bild, einer Abbildung dingfest zu machen. Siehe hierzu die Kritik von Hampe, 2006.
- 11 Zur CNV und weiteren aufmerksamkeitsabhängigen Hirnantworten siehe Näätänen, 1992.
- 12 Siehe Sieroka et al., 2003, und Sieroka, 2004.
- 13 Als weitere transiente Hirnantwort sei noch die N<sub>400</sub> erwähnt, da sie für die oben erwähnte Diskussion um vorbegriffliche bzw. nicht-propositionale Zustände relevant ist. Die N<sub>400</sub> wird in Antwort auf „semantische Abweichter“ evoziert – also wenn man beispielsweise statt „Der Laptop steht auf dem Tisch“ den Satz hört „Der Laptop steht auf dem Geburtstag“. Allerdings geht diese Hirnantwort deutlich über das rein sensorische Niveau hinaus. Im Vergleich zu den oben behandelten (spezifisch auditorischen) Hirnantworten tritt die N<sub>400</sub> relativ spät auf und ihr Ursprung liegt größtenteils außerhalb des auditorischen Kortex bzw. gibt es hier Anzeichen für „Rückflüsse“ neuronaler Aktivität von höheren zu den primären sensorischen Arealen. Von Seiten einiger Neurowissenschaftler werden solche Rückflüsse (auch als *top-down processing* oder *re-entry* bezeichnet) für die wichtigste – oder sogar „wesentliche“ – Bedingung erachtet für die Entstehung bewusster Zustände (siehe etwa Edelman und Tononi, 2002). Eine Zurückweisung dieser letzten These ist implizit in der Naturalismuskritik im übernächsten Abschnitt enthalten.
- 14 Siehe Sieroka, 2004.
- 15 Mittlerweile gibt es auch zahlreiche vergleichende Untersuchungen zu Hirnaktivitäten bei Musikern und Nicht-Musikern in Antwort auf Rhythmik, Klangfarben etc. (siehe bspw. Perrot et al., 1999, und Pantev et al., 2001).
- 16 Siehe Koelsch et al., 1999; vgl. auch Rüsseler et al. 2001.
- 17 Siehe Crane und Mellor, 1990.
- 18 Siehe Davidson, 1980.
- 19 Auch gilt umgekehrt, dass physikalische Beschreibungen, insofern sie oft gleichzeitig auf unterschiedliche Theorien zurückgreifen müssen (bspw. Newtonsche Mechanik und Elektrodynamik oder Elektrodynamik und Quantenmechanik etc.), keineswegs so „comprehensively closed“ sind, wie Davidson es voraussetzt; ganz zu schweigen davon, dass die Quantenphysik selbst genuin probabilistisch ist. – Zu diesen wie auch den obigen Kritikpunkten an einer Token-Identity-Theory siehe ausführlicher Crane und Mellor, 1990.
- 20 Für einen ausführlichen „musikphilosophischen“ Vergleich von Gordon,

- Levinson und Husserl (ohne Bezug auf Neuroinformatik, Neurophysiologie und Naturalismusdebatte) siehe Sieroka, 2005.
- 21 Der Test besteht aus einer Reihe von Melodiepaaren und der Testkandidat ist jeweils aufgefordert zu bestimmen, ob die erste Melodie eines Paares mit der zweiten identisch ist oder ob sie sich tonal bzw. rhythmisch von dieser unterscheidet. Der Test wird mittlerweile auch von Konservatorien verwendet, und sowohl seine Reliabilität als auch seine Validität sind ausgiebig überprüft worden. Insbesondere ist sichergestellt worden, dass der AMMA-Test nicht etwa die individuelle Erinnerungsfähigkeiten anstelle der musikalischen Begabung misst. Ob man also beispielsweise ein gutes oder schlechtes Zahlengedächtnis hat, ist (statistisch) davon unabhängig, wie gut man sich die Melodien in diesem Test einprägen und miteinander vergleichen kann. (Hierfür wie auch für die folgenden Details siehe Gordon, 1997, und Gordon, 1998.)
  - 22 Gordon, 1997, S. 5.
  - 23 Siehe Levinson, 1997, S. ix, 18.
  - 24 Ebd., S. 16.
  - 25 Ebd., S. 17 (als maximal mögliche Ausdehnung nennt er wenige Minuten).
  - 26 Eine Hoch-Zeit erlebte die Diskussion der Melodiewahrnehmung um etwa 1900 bei Autoren wie Brentano, Wundt, Stumpf, Husserl und James.
  - 27 Husserl, 1985, S. 33–34. (Zu Husserls Phänomenologie des Zeitbewusstseins vgl. auch Miller, 1984, sowie Kortooms, 2002.)
  - 28 Siehe Husserl, 2000, S. 396. Husserl verwendet diesen Begriff in Anlehnung an die Alltags-Redeweise von einem Mondhof. So wie dieser Begriff das Fehlen einer scharfen räumlichen Begrenzung beim Anblick des Mondes ausdrückt, soll „Zeithof“ die aufgeweichte Erstreckung in der Wahrnehmung alles Zeitlichen aufzeigen.
  - 29 Siehe ebd., S. 397–401, sowie Husserl, 1985, S. 30–31.
  - 30 Siehe Husserl, 2000, S. 420–426.
  - 31 Siehe etwa Clifton, 1983, S. 296 („contribute to the effort of reuniting music theory with musical experience“), Lewin, 1986, S. 327 („recent years have seen an increasing influence on the music theory of perceptual investigations that can be called phenomenological in the sense of Husserl, either explicitly or implicitly“), sowie Smith, 1979, und Orlik, 1994.
  - 32 Siehe bspw. Bartholomew, 1985, und Bartholomew, 1991.
  - 33 Siehe etwa die von McInerney, 1988, S. 614, vorgeschlagene Untergliederung der Retention in „interestless maintaining“ und „present involvement“.
  - 34 Unter den Husserl-Schülern hat z.B. Anton Marty die Existenz von Pro-  
tentionen verneint (vgl. Simons, 1990, S. 165). Zur aktuellen inner-phänomenologischen Debatte siehe u.a. McInerney, 1988, S. 615–616, und Rinofner-Kreidl, 2000, S. 269–271; auf neurophänomenologischer Seite (im Anschluss v.a. an Merleau-Ponty) wurden Pro-  
tentionen als „Tönungen“ zukünftiger Wahrnehmung begriffen, wie sie durch neurophysiologische Prozesse entstehen, die mit Emotionen korreliert sind (siehe Varela, 1999).
  - 35 Vgl. hierzu Marbach, 2007, der aus ähnlichen Gründen eine phänomeno-

- logische Reflexion in den Neurowissenschaften für wichtig erachtet und dies ebenfalls durch zahlreiche konkrete Beispiele belegt.
- 36 Siehe Lloyd, 2004.
- 37 Lloyd will zeigen, „that retention is a pervasive aspect of the brain“ – ebd., S. 326.
- 38 Siehe Fraisse, 1985, S. 84–85.
- 39 Vgl. Clarke, 1999, S. 476, und Block, 1990, S. 5.
- 40 Dass die hierin implizite Reduktion der Biologie bzw. Neurophysiologie auf die Physik nicht trivial ist, muss an dieser Stelle nicht weiter interessieren; bzw. würde dies die Argumentation gegen einen Reduktionismus allenfalls zusätzlich unterstützen.
- 41 Siehe bspw. die Aufsatzsammlung *Naturalizing Phenomenology* (Petitot et al., 1999), Thompson, 2007, sowie diverse Aufsätze in der vor wenigen Jahren gegründeten Zeitschrift *Phenomenology and the Cognitive Sciences*.
- 42 „[...] phenomenology and cognitive science should be regarded not only as compatible, but as mutually constraining and enriching approaches to the study of mind. [...] both phenomenology and cognitive science can shed light on the phenomenon of time consciousness, and [...] their respective contributions can inform each other“ – van Gelder, 1999, S. 246.
- 43 Ebd., S. 259–260.
- 44 Siehe Crane und Mellor, 1990.
- 45 Siehe Varela, 1999, S. 283, und Thompson, 2007, S. 333, 357.
- 46 Dieses Beispiel ist lediglich eine Variante des berühmten Violinisten-Beispiels aus Ryle, 1954.
- 47 Siehe z. B. Husserl, 1993b, Vol. II/1, S. 18.
- 48 Vgl. Ryckman, 2005, S. 111, der die (Gravitations-)Physik in diesem Sinne als regionale Ontologie versteht.
- 49 Siehe Husserl, 1993a, S. 2; vgl. auch Husserl, 1993b, Vol. II/1, S. 18.
- 50 Einen Überblick zu diversen neurophänomenologisch inspirierten Arbeiten bieten Gallagher und Zahavi, 2008. Siehe außerdem die Aufsatzsammlung von Gallagher und Schmicking, 2010.
- 51 Historisch eindruckliche Belege für solch eine übertriebene Bezugnahme sind einerseits die „romantische Physik“ von Johann Wilhelm Ritter und im direkten Gegenzug Schellings verquere Nobilitierung von Magnetismus und Galvanismus innerhalb seiner Naturphilosophie. Man vergleiche auch die neueren Versuche, physikalisch-chemische Selbstorganisations-theorien zum naturphilosophischen Paradigma zu stilisieren (siehe Heuser-Keßler, 1986, und, vornehmlich kritisch, Sieroka, 2010b, S. 219–220, 253–256).
- 52 Siehe hierzu erneut auch Marbach, 2007, sowie Marbach, 1993, der sich aus phänomenologischer Perspektive um das Wechselverhältnis von Neurowissenschaft und Phänomenologie bemüht, ohne deshalb aber einen transzendentalen Reduktionismus zu vertreten.
- 53 Vgl. Blumenberg, 1998.
- 54 Siehe Leibniz, 1996, und Leibniz, 1982; vgl. Sieroka und Dosch, 2008.
- 55 Siehe etwa „Vorrede“ und „Einleitung“ zu Schelling, 2000.
- 56 Siehe von Weizsäcker, 1988. Dabei diskutiert er auch – letztlich ganz ähnl-

- lich wie Schelling – Übergänge zwischen diesen Theorien (von Weizsäcker spricht hier von einem „Kreisgang“). – Man beachte außerdem, dass somit der neurophänomenologische Ansatz von Varela und Thompson etwas ganz anderes ist als etwa der Versuch einer Fundierung jeglicher Wissenschaft in der Lebenswelt.
- 57 Siehe Rohs, 1996, S. 12.
- 58 Beispielsweise muss selbst Mellor, 1998, der für das ontologische Primat eines physikalisch-realistischen Zeitbegriffs (einer lagezeitlich geordneten Datenreihe) argumentiert, zur Beschreibung des Erlebens auf die subjektiven Zeitformen verfallen, gegenwärtig und zukünftig zurückgreifen.
- 59 Rohs' Feldbegriff ist dabei derjenige der klassischen (nicht der Quanten-) Physik und müsste also sicherlich noch genauer überarbeitet werden; vgl. hierzu Sieroka, 2010a.
- 60 Auf der Grundlage von Quantenmechanik und Relativitätstheorie wurde ein Zeithofkonzept etwa von Weyl, 1968, vertreten. Eine zusätzliche Motivation hierfür stammte aus seiner mathematischen Analyse des Kontinuums, in der Punktförmigkeit als etwas aus einem primären Umgebungsbegriff Abgeleitetes begriffen wurde (vgl. Sieroka, 2010b, S. 70–72, 228, 237, 240). – Ganz Ähnliches gilt übrigens auch für Whitehead, 1955.

### Literatur

- Bartholomew, D., 1985: Preamble to a Phenomenology of Music. In: *Journal of Musicological Research* 5, S. 319–360.
- Bartholomew, D., 1991: Whole/Part Relations in Music: An Exploratory Study. In: *Journal of Aesthetic Education* 25, S. 175–191.
- Block, R. A., 1990: Models of Psychological Time. In: Block, R. A. (Hg.): *Cognitive Models of Psychological Time*. Hillsdale (NJ): Lawrence Erlbaum, S. 1–35.
- Blumenberg, H., 1998: *Paradigmen zu einer Metaphorologie*. Frankfurt a. M.: Suhrkamp.
- Clarke, E. F., 1999: Rhythm and Timing in Music. In: Deutsch, D. (Hg.): *Psychology of Music*. London: Academic Press, S. 473–500.
- Clifton, T., 1983: *Music as Heard*. New Haven: Yale University Press.
- Coates, P., 2007: *The Metaphysics of Perception*. London/New York: Routledge.
- Conway, B. R.; D. Y. Tsao, D. Y., 2009: Color-tuned Neurons Are Spatially Clustered According to Color Preference within Alert Macaque Posterior Inferior Temporal Cortex. In: *Proceedings of the National Academy of Sciences* 106, S. 18034–18039.
- Crane, T., 2003: The Intentional Structure of Consciousness. In: Jokic,

- A. und Smith, Q. (Hg.): *New Philosophical Perspectives*. Oxford: Oxford University Press, S. 33–56.
- Crane, T.; Mellor, D. H., 1990: There is no Question of Physicalism. In: *Mind* 99, S. 185–206.
- Davidson, D., 1980: *Essays on Actions and Events*. Oxford: Clarendon Press.
- Edelman, G. M.; Tononi, G. 2002: *Gehirn und Geist – Wie aus Materie Bewußtsein entsteht*. München: C. H. Beck.
- Fraisse, P., 1985: *Psychologie der Zeit*. München: Reinhard.
- Gallagher, S.; Schmicking, D. (Hg.), 2010: *Handbook of Phenomenology and Cognitive Science*. Dordrecht: Springer.
- Gallagher, S.; Zahavi, D., 2008: *The Phenomenological Mind*. London/ New York: Routledge.
- Gordon, E. E., 1997: *Learning Sequences in Music*. Chicago: GIA Publications.
- Gordon, E. E., 1998: *Introduction to Research and the Psychology of Music*. Chicago: GIA Publications.
- Hampe, M., 2006: Sichtbare Wesen, deutbare Zeichen, Mittel der Konstruktion: zur Relevanz der Bilder in der Wissenschaft. In: *Angewandte Chemie* 118, S. 1044–1048.
- Heuser-Kessler, M.-L., 1986: *Die Produktivität der Natur: Schellings Naturphilosophie und das neue Paradigma der Selbstorganisation in den Naturwissenschaften*. Berlin: Duncker & Humblot.
- Horgan, T.; Kriegel, U., 2007: Phenomenal Epistemology: What is Consciousness that We May Know It so Well?. In: *Philosophical Issues* 17, S. 123–144.
- Husserl, E., 1985: *Texte zur Phänomenologie des inneren Zeitbewußtseins (1893–1917)*. Hamburg: Meiner.
- Husserl, E., 1993a: *Ideen zu einer reinen Phänomenologie und phänomenologische Philosophie*, Bd. I. Tübingen: Niemeyer.
- Husserl, E., 1993b: *Logische Untersuchungen*. Tübingen: Niemeyer.
- Husserl, E., 2000: *Vorlesungen zur Phänomenologie des inneren Zeitbewußtseins*. Tübingen: Niemeyer.
- Jäncke, L., 2005: *Bildgebende Verfahren in den kognitiven Neurowissenschaften und der Psychologie*. Stuttgart: Kohlhammer.
- Koelsch, S.; Schröger, E.; Tervaniemi, M., 1999: Superior Pre-attentive Auditory Processing in Musicians. In: *NeuroReport* 10, S. 1309–1313.

- Kortooms, T., 2002: *Phenomenology of Time. Edmund Husserl's Analysis of Time-Consciousness*. Dordrecht: Kluwer.
- Lanz, P., 1996: *Das phänomenale Bewußtsein: Eine Verteidigung*. Frankfurt a. M.: Klostermann.
- Leibniz, G. W., 1982: *Specimen Dynamicum*. Hamburg: Meiner.
- Leibniz, G. W., 1996: *Neue Abhandlungen über den menschlichen Verstand*. Hamburg: Meiner.
- Levinson, J., 1997: *Music in the Moment*. Ithaca: Cornell University Press.
- Lewin, D., 1986: Music Theory, Phenomenology, and Modes of Perception. In: *Music Perception* 3, S. 327–392.
- Lloyd, D., 2004: *Radiant Cool. A Novel Theory of Time Consciousness*. Cambridge (Mass.): MIT Press.
- Marbach, E., 1993: *Mental Representation and Consciousness. Towards a Phenomenological Theory of Representation and Reference*. Dordrecht: Kluwer.
- Marbach, E., 2007: Towards Integrating Husserlian Phenomenology with Cognitive Neuroscience of Consciousness. In: *Synthesis Philosophica* 44, S. 385–400.
- McInerney, P. K., 1988: What is Still Valuable in Husserl's Analyses of Inner Time-Consciousness. In: *Journal of Philosophy* 85, S. 605–616.
- Mellor, D. H., 1998: *Real Time II*. London: Routledge.
- Miller, I., 1984: *Husserl, Perception, and Temporal Awareness*. Cambridge (Mass.): MIT Press.
- Näätänen, R., 1992: *Attention and Brain Function*. Hillsdale (NJ): Lawrence Erlbaum.
- Näätänen, R.; Tervaniemi, M.; Sussman, E.; Paavilainen, P.; Winkler, I., 2001: "Primitive Intelligence" in the Auditory Cortex. In: *Trends in Neurosciences* 24, S. 283–288.
- Orlik, F., 1994: „Inneres Zeitbewußtsein“ und „attentionale Modifikation“ – Ein Beitrag zur Klärung des Verhältnisses von Zeit und musikalischer Gestalt im Anschluß an Husserl. In: *Archiv für Musikwissenschaft* 51, S. 253–273.
- Pantev, C.; Roberts, L. E.; Schulz, M.; Engelien, A.; Ross, B., 2001: Timbre-specific Enhancement of Auditory Cortical Representations in Musicians. In: *NeuroReport* 12, S. 169–174.
- Perrot, X.; Micheyl, C.; Khalfa, S.; Collet, L., 1999: Stronger Bilateral Efferent Influences on Cochlear Biomechanical Activity in Musicians.



- cians than in Non-musicians. In: *Neuroscience Letters* 262, S. 167–170.
- Petitot, J.; Varela, F. J.; Pachoud, B.; Roy, J.-M. (Hg.), 1999: *Naturalizing Phenomenology*. Stanford: Stanford University Press.
- Rinofner-Kreidl, D., 2000: *Edmund Husserl. Zeitlichkeit und Intentionalität*. Freiburg: Alber.
- Rohs, P., 1996: *Feld – Zeit – Ich: Entwurf einer feldtheoretischen Transzendentalphilosophie*. Frankfurt a.M.: Klostermann.
- Rüsseler, J.; Altenmüller, E.; Nager, W.; Kohlmetz, C.; Münte, T. F., 2001: Event-related Brain Potentials to Sound Omissions Differ in Musicians and Non-musicians. In: *Neuroscience Letters* 308, S. 33–36.
- Ryckman, T., 2005: *The Reign of Relativity: Philosophy and Physics 1915–1925*. Oxford: Oxford University Press.
- Ryle, G., 1954: The World of Science and the Everyday World. In: Ryle, G. (Hg.): *Dilemmas*. Cambridge: Cambridge University Press, S. 68–81.
- Schelling, F. W. J., 2000: *System des transzendentalen Idealismus*. Hamburg: Meiner.
- Schmicking, D., 2003: *Hören und Klang: Empirisch phänomenologische Untersuchungen*. Würzburg: Königshausen & Neumann.
- Sieroka, N., 2004: *Neurophysiological Aspects of Time Perception*. Dissertation, Fakultät für Physik, Universität Heidelberg.
- Sieroka, N., 2005: Quasi-Hearing in Husserl, Levinson, and Gordon. In: *Journal of the British Society for Phenomenology* 36, S. 4–22.
- Sieroka, N., 2010a: Feldmetaphysik und physikalisches Materieverständnis. In: *Allgemeine Zeitschrift für Philosophie* 35, S. 105–122.
- Sieroka, N., 2010b: *Umgebungen. Symbolischer Konstruktivismus im Anschluss an Hermann Weyl und Fritz Medicus*. Zürich: Chronos.
- Sieroka, N.; Dosch, H. G., 2008: Leibniz’s “Perceptions Insensibles” and Modern Neurophysiology. In: *Studia Leibnitiana* 40, S. 14–28.
- Sieroka, N.; Dosch, H. G.; Specht, H.-J.; Rupp, A., 2003: Additional Neuromagnetic Source Activity Outside Auditory Cortex in Duration Discrimination Correlates with Behavioural Ability. In: *NeuroImage* 20, S. 1697–1703.
- Simons, P., 1990: Marty on Time. In: Mulligan, K. (Hg.): *Mind, Meaning and Metaphysics: The Philosophy and Theory of Language of Anton Marty*. Dordrecht: Kluwer, S. 157–170.

- Smith, F. J., 1979: *The Experiencing of Musical Sound. Prelude to a Phenomenology of Music*. New York: Gordon & Breach.
- Thompson, E., 2007: *Mind in Life: Biology, Phenomenology, and the Sciences of the Mind*. Cambridge (Mass.): Harvard University Press.
- van Gelder, T., 1999: Wooden Iron? Husserlian Phenomenology Meets Cognitive Science. In: Petitot et al. 1999, S. 245–265.
- Varela, G. (1999). The Specious Present: A Neurophenomenology of Time Consciousness. In: Petitot et al. 1999, S. 266–314.
- von Weizsäcker, C. F., 1988: *Aufbau der Physik*. München: dtv.
- Weyl, H., 1968: Zeitverhältnisse im Kosmos, Eigenzeit, gelebte Zeit und metaphysische Zeit. In: Brightman, E. S. (Hg.): *Proceedings of the Sixth International Congress of Philosophy (Harvard 1926)*. Nendeln: Kraus, S. 54–58.
- Whitehead, A. N., 1955: *Enquiry into the Principles of Natural Knowledge*. Cambridge: Cambridge University Press.



Boris Hennig

## Eine Verteidigung des typologischen Artbegriffs

### Zusammenfassung

Es wird gezeigt, dass der biologische Artbegriff, den Ernst Mayr gegen den typologischen Artbegriff ausspielt, tatsächlich eine Variante des typologischen Artbegriffes voraussetzt. Denn man kann nur dann beurteilen, ob zwei Lebewesen miteinander Nachkommen erzeugen können, wenn man bereits weiß, was als Nachkomme zählen würde. Also muss man wissen, welche nicht-relationalen Merkmale typische Nachkommen von Lebewesen einer bestimmten Art haben, um den biologischen Artbegriff anzuwenden. Der typologische Artbegriff, der hier zu Grunde liegt, ist der Aristotelische.

### Abstract

The paper demonstrates that the biological species concept that Mayr contrasts with the typological one in fact presupposes a version of the typological species concept. For one cannot assess whether two living beings are capable of producing offspring without already knowing what would count as offspring. Therefore, one must know non-relational features of typical offspring of a kind of living beings in order to be able to apply the biological species concept. The typological species concept that is at stake here is the Aristotelian one.

### 1. Einleitung

Von den vielen Artbegriffen, die in der Theorie der Biologie diskutiert werden, möchte ich mich hier auf drei grundlegende Varianten beschränken: den typologischen, den morphologischen, und den so genannten biologischen Artbegriff. Über das Verhältnis des typologischen Artbegriffes zum biologischen schreibt Ernst Mayr folgendes:

Der biologische Artbegriff hat sich in der zweiten Hälfte des 19. Jahrhunderts entwickelt. Bis zu dieser Zeit, von Platon und Aristoteles bis zu Linné und Autoren des frühen 19. Jahrhunderts, hat man einfach „species“, eide (Platon),

oder Arten (Mill) anerkannt. Da weder die Taxonomen noch die Philosophen eine strenge Unterscheidung zwischen unbelebten Dingen und biologischen Arten machten, waren die Artdefinitionen, die sie gegeben haben, eher unterschiedlich und nicht sehr spezifisch. Das Wort „Art“ vermittelte die Vorstellung einer Klasse von Dingen, deren Mitglieder gewisse definierende Eigenschaften miteinander teilen. Ihre Definition unterschied eine Art von allen anderen. Eine solche Klasse ist gleichbleibend, sie ändert sich nicht im Verlauf der Zeit, und alle Abweichungen von der Definition der Klasse sind bloße „Akzidentien“, also unvollkommene Erscheinungsformen des Wesens (eidos). (Mayr, 1996, 268)

Dem Vernehmen nach spricht Mayr hier auch über Aristoteles,<sup>1</sup> und er sagt drei Dinge über ihn. Aristoteles, der ja zugleich Taxonom und Philosoph war, habe erstens nicht klar zwischen belebten und unbelebten Dingen unterschieden. Dies habe dazu geführt, dass er zweitens nicht sehr genau definiert habe, was eine biologische Art ist. Vielmehr habe er sich unter einer Art so etwas wie eine Klasse von Dingen vorgestellt, die alle gewisse Merkmale aufweisen, durch deren Angabe die Klasse definiert werden kann. Eine solche Definition sehe drittens nicht vor, dass sich die Mitglieder einer Art wesentlich ändern. Ein Ding gehöre zur Art, wenn es die definierenden Merkmale aufweist, und nicht, wenn es das nicht tut. Daran ändere sich nichts im Verlauf der Zeit, und daher könne eine Evolution nicht stattfinden, oder jedenfalls nur so, dass sich die definierenden Eigenschaften nicht ändern. Von den drei Aussagen, die Mayr hier beiläufig Aristoteles zuschreibt, werde ich mich im Folgenden auf die ersten beiden konzentrieren. Wie es Aristoteles mit der Evolution hält, möchte ich nicht abschließend beurteilen.

## 2. Selbstbewegung

Zuerst will ich mich dem zuwenden, was Mayr zu Beginn des zitierten Abschnittes sagt: dass die Philosophen und Taxonomen, Aristoteles offenbar eingeschlossen, nicht klar zwischen belebten und unbelebten Dingen unterschieden hätten. Was Aristoteles angeht, ist das eine gewagte Unterstellung. Ziemlich offensichtlich ist jedenfalls, dass Aristoteles klar und deutlich zwischen natürlichen und anderen Dingen unterscheidet. Er schreibt zu Beginn des zweiten Buches der *Physik*:

Vom Seienden ist einiges von Natur, anderes dagegen aus anderen Ursachen. Von Natur sind die Lebewesen und ihre Teile, die Pflanzen, sowie die einfachen unter den Körpern, wie Erde, Feuer, Luft und Wasser. (*Physik* II 1, 192b8-11)<sup>2</sup>

Ob die Unterscheidung zwischen natürlichen und anderen Dingen klar ist, hängt davon ab, wie Aristoteles sie genau zieht, und wie gut man sie verwenden kann. Aristoteles schreibt, dass ein Ding dann natürlich ist, wenn es einen Ursprung von Bewegung und Ruhe in sich hat. Diese grobe Bestimmung erläutert er in zwei Hinsichten. Erstens fasst er auch Wachstum, Schrumpfen, und Eigenschaftsveränderungen unter den Begriff der Bewegung. Ich werde es ihm im Folgenden gleich tun. Zweitens betont Aristoteles, dass der Ursprung der Bewegung eines natürlichen Dinges in diesem Ding als solchem liegen muss, und nicht etwa allein in seinen Bestandteilen. Er schreibt:

Die Dinge, die von Natur sind, scheinen in sich selbst einen Ursprung der Bewegung und Ruhe zu haben, teils bezogen auf den Ort, teils auf das Wachsen und Schrumpfen, teils auf das Anderswerden. Ein Bett oder Kleid dagegen, und was es dergleichen Gattungen sonst noch gibt, insofern es so bezeichnet wird und durch Kunst hergestellt ist, hat gar keinen ihm eingeborenen Anstoß zu einer Veränderung in sich. Nur nebenher, indem es steinern oder irden oder gemischt ist, hat es einen von daher, und nur darauf bezogen. Denn die Natur ist ein gewisser Ursprung und eine Ursache des Bewegtseins und Ruhens in dem, dem sie in erster Linie und als solchem zukommt, und nicht etwa nur nebenher. (*Physik* II 1, 192b13-23)

Aristoteles sagt also, dass ein Ding dann natürlich ist, wenn es den Ursprung seiner Bewegung oder Ruhe in sich hat. Nahe liegende Gegenbeispiele hierzu liefern Maschinen und Automaten, die sich von selbst bewegen (Waterlow, 1982, 51). Denn in einem gewissen Sinn hat auch ein Auto den Anstoß seiner Ortsbewegung in sich. Sicher, es bewegt sich nur dann, wenn jemand es fährt, aber das kann man leicht ändern. Man braucht nur ein Auto zu konstruieren, das von selbst fährt, und anhand von Sensoren und Navigationssystemen sogar selbstständig Richtung und Tempo wechselt. Natürlich braucht das Auto dann noch Sprit und einen festen Untergrund, um zu fahren. Aber ein Tier braucht auch Nahrung, und ohne äußerlichen Widerstand kann es sich auch nicht bewegen.

Eine zu schnelle Ausflucht wäre, darauf hinzuweisen, dass Aristoteles von solchen Automaten nichts wissen konnte. In *De Motu Animalium*

7 erwähnt er nämlich Puppen, die sich von Fäden gezogen so ähnlich bewegen wie Tiere (701b1-7). Nun ist klar, dass diese Puppen keine Automaten sind, weil sie sich immer nur dann bewegen, wenn sie von außen gezogen werden. Von Tieren gilt aber, Aristoteles zufolge, letztlich dasselbe. Ihre Beine bewegen sich, wenn sie von Sehnen gezogen werden, und die Sehnen bewegen sich, wenn die Tiere etwas anstreben oder vermeiden. Daher schreibt Aristoteles:

Der Ursprung der Bewegung ist, wie gesagt, das Gesuchte oder Gemiedene im Bereich der Handlungen. (*De Motu Animalium* 8, 701b33-34)<sup>3</sup>

Was die Tiere suchen oder meiden, das verursacht eine kleine Veränderung in diesen Tieren. Diese führt zu der Bewegung ihrer Sehnen und Muskeln, durch die sie sich auf das zu bewegen, was sie suchen, und von dem fortbewegen, was sie meiden. Man kann die Bewegung eines Tieres also komplett mechanisch beschreiben. Eigentlich werden sie von dem angezogen, was sie suchen, und von dem abgestoßen, was sie meiden, so ähnlich wie eine Marionette von dem gezogen wird, der sie führt. Wenn das so ist, dann scheinen sie den Ursprung ihrer Bewegung gar nicht in sich selbst zu haben; denn das, was sie suchen und meiden, ist ja meist außerhalb ihrer selbst.<sup>4</sup>

Es könnte also sein, dass die Unterscheidung, die Aristoteles zwischen natürlichen und anderen Dingen macht, gar nicht klar ist; und dann hätte Mayr letztlich recht. Da die Unklarheit sich aber bereits direkt durch das ergibt, was Aristoteles selbst schreibt, handelt es sich wohl eher um eine Unklarheit unserer Interpretation, und nicht wirklich um eine Unklarheit dessen, was Aristoteles eigentlich sagen will. Vielleicht lesen wir ihn falsch, wenn wir meinen, er spreche von einem inneren Anstoß zur Bewegung.

Aristoteles hebt hervor, dass natürliche Dinge den Ursprung ihrer Bewegungen nicht deswegen in sich haben, weil sie aus einem Material wären, das einen solchen Ursprung in sich trägt. Er erwähnt Antiphon, der sich vorstellen kann, dass ein Bett Wurzeln schlägt, wenn es in die Erde vergraben wird. Dann wächst etwas von selbst, aber es ist eben nicht das Bett als solches, das wächst. Was da wächst, ist bestenfalls ein Baum oder Strauch (*Physik* II 1, 193a12-14). Das Bett, das Wurzeln schlägt, hat nun in einem gewissen Sinne den Ursprung des Wachsens in sich. Denn es hat ja Holz in sich, und das Holz hat den Ursprung seines Wachsens in sich. Das Bett hat diesen Ursprung aber eben nicht *als Bett*

in sich. Aus diesem Grund formuliert Aristoteles sehr vorsichtig: Ein Bett hat den Ursprung seiner Bewegung nicht in sich „insofern es als solches bezeichnet wird und durch Kunst hergestellt ist“ (192b17-18). Der aus dem Bett wachsende Baum dagegen ist ein Naturding, und er hat den Ursprung seines Wachsens in sich insofern er als Baum bezeichnet wird und nicht durch Kunst hergestellt ist.

Es kommt also darauf an, *als was* da etwas wächst. Das Bett wächst nicht als Bett, sondern bestenfalls als Baum oder Strauch. Etwas ist nicht einfach dadurch natürlich, dass es sich von selbst bewegt, sondern erst dadurch, dass es sich *als es selbst* bewegt. Ein Ding bewegt sich als es selbst, wenn der Grund zu seiner Bewegung allein darin liegt, dass es ist, was es ist (also in seiner Form).<sup>5</sup> Der Grund für das Wachsen des Bettes liegt zum Beispiel nicht darin, dass es ein Bett ist. Anders steht es mit einem Tier, das mechanisch von dem angezogen wird, was es erstrebt. Zwar ist der Anlass seiner Bewegung von außen gegeben. Dass das Tier sich aus diesem Anlass bewegt, liegt dennoch in erster Linie daran, dass es ist, was es ist. Es liegt nicht in erster Linie daran, dass das Tier aus bestimmten Stoffen besteht, oder gar daran, dass es mit einer bestimmten Absicht hergestellt worden wäre.<sup>6</sup>

Damit sind wir dem Gegenbeispiel noch nicht ganz ausgewichen, denn auch der Grund dafür, dass sich ein selbstfahrendes Auto bewegt, scheint allein darin zu liegen, dass es ein selbstfahrendes Auto ist. Man scheint seine Bewegung verstehen und erklären zu können, wenn man nichts weiter weiß, als dass es sich um ein solches Auto handelt. Das scheint aber nur so zu sein. Denn auch selbstfahrende Autos müssen irgendwie konstruiert und programmiert werden. Ihre Form kommt gewissermaßen von außen. Wenn sie auch noch von selbst wachsen und aufwachsen würden, spräche nichts dagegen, sie zu den Naturdingen zu zählen. Selbstfahrende Autos geben aber nur dann ein Gegenbeispiel zu Aristoteles Bestimmung natürlicher Dinge ab, wenn man das nicht tut.

Der Punkt ist also, dass man nicht wirklich verstehen kann, warum selbstfahrende Autos sich bewegen, ohne die Lebewesen zu verstehen, die sie bauen, programmieren und benutzen. Die Geschichte, die man erzählen kann, um die Bewegung selbstfahrender Autos zu erklären, muss auch und sogar in erster Linie von den Wesen handeln, die sie konstruieren und verwenden, also von Menschen. Nicht alle Geschichten dieser Art müssen vom Menschen handeln, denn auch andere Lebewesen können Dinge herstellen und benutzen. Wir verstehen Spinnennetze



nur dadurch, dass wir sie auf Spinnen beziehen, und Schwalbennester nur dadurch, dass wir sie auf Schwalben beziehen.

Genau das ist anders bei Naturdingen. Ihre Bewegungen kann man anhand von Geschichten erklären, die allein oder in erster Linie von ihnen selbst handeln. Zwar kommen in den Geschichten, anhand derer wir uns die Bewegungen von Naturdingen erklären, oft auch andere Naturdinge vor. Zum Beispiel verstehen wir die Lebensweise bestimmter Ameisen besser, wenn wir die Lebensweise von Blattläusen kennen. Darauf kommt es aber nicht an. Die Geschichte, die wir über Ameisen erzählen, handelt bestimmt nicht in erster Linie von Blattläusen. Was wir an einem selbstfahrenden Auto nicht verstehen, wenn wir den Menschen außer Betracht lassen, ist auch nicht, wie und durch welche Mechanismen es sich bewegt. Das können wir gut physikalisch erklären, so wie ja auch Aristoteles meint, man könne die Bewegung eines Tieres mechanisch erklären. Was wir nicht ohne Verweis auf den Menschen tun können ist etwas anderes. Wir können nicht beurteilen, ob das Auto sich so bewegt, wie es sich bewegen soll. Der Standard, anhand dessen wir das beurteilen können, liegt nicht in der Natur des Autos, sondern in der Natur des Menschen. Wenn Aristoteles sagt, dass natürliche Dinge den Ursprung ihrer Bewegung in sich haben, dann meint er einen solchen Standard. Natürliche Dinge sind derart, dass sich die Standards, denen ihre Bewegungen unterliegen, allein aus der Betrachtung ihrer Art ergeben – und nicht aus der Betrachtung ihrer Materie oder aus der Betrachtung von anderen Wesen, die sie herstellen und benutzen. Deshalb nennen wir nur Lebewesen gesund oder krank, denn gesund ist etwas dann, wenn es den Standards entspricht, die sich aus seiner eigenen Natur ergeben. Bei leblosen Dingen und Teilen lebender Dinge sprechen wir dagegen davon, dass sie funktionieren oder kaputt sind, denn etwas funktioniert dann, wenn es Standards gerecht wird, die außerhalb seiner eigenen Natur liegen.

Das ist also der Sinn, in dem natürliche Dinge den Ursprung ihrer Bewegungen in sich haben: Man kann ihre Bewegungen dadurch hinsichtlich ihres Erfolges beurteilen, dass man sie in eine Geschichte einbettet, die in erster Linie von diesen Dingen handelt (und von dem, was sie herstellen und benötigen). So verstanden ist die Unterscheidung, die Aristoteles trifft, klar. Es ist aber noch nicht ganz die von Mayr gewünschte Unterscheidung, denn Aristoteles unterscheidet nicht zwischen lebenden und leblosen Dingen, sondern zwischen natürlichen und

anderen; und zu den natürlichen Dingen zählt er in der oben zitierten Passage auch die Elemente. Diese sind offenbar nur deshalb natürlich, weil sie einfach sind; denn von zusammengesetzten Dingen erwähnt Aristoteles nur die, die auch belebt sind. Laut Aristoteles ist etwas dann ein Element, wenn alle seine Teile von derselben Art sind wie es selbst (*De Caelo* III 3, 302a15-18). Dass sie einfach sind, erklärt, warum Aristoteles die Elemente nicht zu den unnatürlichen Dingen zählen will. Wenn sie nicht natürlich wären, dann könnte man ihre Bewegungen nur dadurch beurteilen, dass man auf etwas verweist, das nicht von ihrer Art ist. Nun kann man die Bewegungen derjenigen nichtnatürlichen Dinge, die keine Kunstprodukte sind, normalerweise durch Verweis auf die Natur ihrer Teile oder ihrer Materie erklären. Da die Teile eines Elements aber von derselben Art sind wie das Element selbst, geht das in ihrem Fall nicht. Da sie aber auch keine Kunstprodukte sind, kann man ihre Bewegungen auch nicht dadurch erklären, dass man sie in eine Geschichte über Dinge anderer Art einbettet. Wenn man also der Meinung ist, dass die Bewegungen der Elemente überhaupt irgendwelchen Standards unterliegen, anhand derer sie als erfolgreich beurteilt werden können, muss man sie wohl zu den natürlichen Dingen rechnen. Da die Elemente aber einen klar umgrenzten Sonderfall darstellen, können wir sie hier ganz gut beiseite lassen.<sup>7</sup> Wenn wir das tun, dann entspricht die Unterscheidung, die Aristoteles trifft, genau der zwischen belebten und unbelebten Dingen. Belebte Dinge sind solche, aus deren Form sich die Standards ableiten lassen, anhand derer ihre Bewegungen als natürlich und erfolgreich oder erfolglos beurteilt werden können. Unbelebte Dinge sind solche, bei denen diese Standards in erster Linie aus etwas anderem als ihrer eigenen Form hervorgehen, nämlich entweder aus der Form der Dinge, die sie herstellen oder benutzen, oder aber aus der Form dessen, woraus sie bestehen.

### 3. Zwei Aristotelische Artbegriffe

Wenn Aristoteles klar zwischen belebten und unbelebten Dingen unterscheidet, kann das kein Grund für ihn sein, den Artbegriff zu vernachlässigen. Er könnte es natürlich aus anderen Gründen getan haben. Deshalb sollten wir jetzt fragen, wie er es mit den Arten von Lebewesen hält. Wir wissen bereits, dass der Begriff der Art für Aristoteles im

Wortsinne grundlegend ist. Denn der Ursprung der Selbstbewegung eines belebten Dinges liegt ja in seiner Art. Es bewegt sich im relevanten Sinne von selbst, wenn es sich als Ding dieser Art bewegt, wenn man also die Bewegung hinsichtlich ihres Erfolges beurteilen kann, indem man in erster Linie diese Art betrachtet.

Aristoteles verwendet vor allem zwei Begriffe, die man als Artbegriff bezeichnen könnte. Im zweiten Kapitel der *Kategorienschrift* unterscheidet er zwei Sachen, die man von etwas aussagen kann (1a20-b1). Von Sokrates kann man zum Beispiel sagen, dass er ein Mensch ist, und dass er blass ist. Blässe ist etwas, das Sokrates zukommt, sie ist *in ihm* als eine seiner Eigenschaften. Ein Mensch ist er nicht so, dass es ihm als eine seiner Eigenschaften zukommt. Ein Mensch ist einfach das, was er ist. Wäre er nicht Mensch, dann wäre da nichts, von dem man sagen könnte, dass ihm die Eigenschaft, Mensch zu sein, fehle.<sup>8</sup>

Sachen, die man von einem Ding aussagen kann, ohne dass sie in diesem Ding sind, nennt Aristoteles in der *Kategorienschrift* zweite Substanzen. Er nennt sie *zweite* Substanzen, um sie von den Einzeldingen zu unterscheiden, die weder in etwas sind noch von etwas gesagt werden können. Er nennt sie ebenso wie diese Einzeldinge *Substanzen*, weil sie in keiner weiteren Sache sind. Denn Substanzen sind Dinge, in denen Eigenschaften sind, und die nicht in anderen Dingen in derselben Weise sind (*Kategorien* 5, 2a11-13 und 34-35).<sup>9</sup> Sowohl das Einzelding Sokrates als auch die zweite Substanz „Mensch“ haben Eigenschaften (2b37-3a6). Sokrates ist rational, zweibeinig, und stupsnasig, die zweite Substanz „Mensch“ ist ebenfalls rational und zweibeinig, aber nicht stupsnasig. Sie ist das, was man beschreibt, wenn man beschreibt, was Menschen sind. Die zweite Substanz ist die biologische Art Mensch.

Aristoteles hat noch ein zweites Wort für biologische Arten. In der *Physik* schreibt er, dass man bei der Beschreibung natürlicher Dinge vier Faktoren beachten müsse. In der Passage, die ich gleich zitieren werde, spricht Aristoteles von vier *αἴτια* (Sg. *αἴτιον*).<sup>10</sup> Die Griechen bezeichnen dann etwas als *αἴτιον*, wenn es für eine Sache in einem sehr allgemeinen Sinne verantwortlich ist.<sup>11</sup> Wenn jemand für eine Sache verantwortlich ist, dann bedeutet das unter anderem, dass er auf bestimmte Fragen antworten kann (oder soll). Man kann man ihn meist fragen, wie diese Sache zustande gekommen ist, um was für eine Sache es sich handelt, oder wozu sie gut ist. Etwas ähnliches gilt für die vier Faktoren, die Aristoteles auflistet. Es sind Faktoren, die man gewissermaßen

befragen kann, um zu erfahren, wie es sich mit einer Sache verhält.<sup>12</sup> Einer dieser Faktoren ist später als „Formalursache“ bezeichnet worden. Über die Formalursache schreibt Aristoteles folgendes:

Ein anderes [αἴτιον] ist die Art (εἶδος) oder das Vorbild (παράδειγμα), also die Erklärung des „was es war zu sein“ (ὁ λόγος ὁ τοῦ τί ἦν εἶναι), sowie die Gattungen davon. (*Physik* II 3, 194b26-27)

Der Zusatz „sowie die Gattungen davon“ ist nicht so gemeint, dass die Formalursache eines Dinges dasselbe sei wie seine Gattung. Was Aristoteles sagen will, ist nur, dass auch die zur Art, dem Vorbild, und der Erklärung des „was es war zu sein“ gehörigen Gattungen als Formalursache zählen.

Mit εἶδος meint Aristoteles keine biologischen Arten, jedenfalls verwendet er dieses Wort nicht terminologisch in seinen biologischen Schriften (Balme, 1962, 84–85). Vielleicht sollte man es der Vorsicht halber erst einmal als „Form“ übersetzen.<sup>13</sup> Aber auch wenn Aristoteles mit εἶδη keine biologischen Arten im Gegensatz zu biologischen Gattungen, Familien oder Klassen meint, kann man seine Formen doch mitunter als Arten von Lebewesen auffassen (vgl. Cho, 2010, 301). Wenn man das tut, fällt auf, dass Aristoteles diese Arten mit dem gleichsetzt, was etwas wird, wenn es sich seiner Natur gemäß entwickelt. In Buch Λ 3 der *Metaphysik* schreibt er:

Denn bei jeder Veränderung ändert sich etwas durch etwas zu etwas. Das, wodurch [es sich ändert], ist das erste Bewegende; das, was [sich ändert], ist die Materie; das, wozu [es wird], ist das εἶδος. (*Metaphysik* Λ 3, 1069b36-1070a2)

Wenn etwas zu etwas anderem wird, dann ist also das εἶδος das, was es wird. Wenn Sokrates blass wird, dann wird er nicht Blässe, sondern er wird ein blasser Sokrates. Das, was er wird, ein blasser Sokrates, ist ein εἶδος. Wenn ein natürliches Ding sich seiner Natur entsprechend entwickelt, dann ist das εἶδος ebenfalls das, was es wird. Das, was ein Welpen wird, ist ein Hund; ein Hund ist also das εἶδος des Welpen. In diesem Sinne bezeichnet Aristoteles die Formalursache auch als Vorbild (παράδειγμα).<sup>14</sup> Denn in Kapitel 7 des zweiten Buchs der *Physik* schreibt er:

Die drei letztgenannten [εἶδος, Ursprung der Bewegung und Worumwillen] fallen oft zusammen. Denn was etwas ist und worumwillen es ist, ist eines;

und das, woher die erste Bewegung kommt, ist der Form nach dasselbe: Es ist ein Mensch, der einen Menschen zeugt. (*Physik* II 7, 198a24-27)

Wenn ein Mensch sich seiner Natur entsprechend entwickelt, ist seine Form (sein εἶδος) ein und dasselbe wie das Ziel seiner Entwicklung, und das, wozu er sich entwickelt, ist der Form nach dasselbe wie das Lebewesen, das den Menschen gezeugt hat. Die Form des Erzeugers ist hier buchstäblich ein Vorbild, jedenfalls insoweit der Erzeuger sich selbst so entwickelt hat, wie sich Menschen ihrer Natur entsprechend entwickeln sollten.

Außerdem setzt Aristoteles die Formalursache eines Dinges noch mit der Erklärung des „was es war zu sein“ gleich, dem λόγος seines τί ἦν εἶναι. „Was es war zu sein“ ist eine der Wendungen, die Aristoteles prägt, weil ihm Worte fehlen. Da nicht ganz einfach auszumachen ist, was er damit sagen will,<sup>15</sup> behelfe ich mir hier mit zwei Stellen: einer aus der *Metaphysik* Z 4 und einer aus der *Zweiten Analytik* II 11. In der *Metaphysik* schreibt Aristoteles, dass nur diejenigen Dinge ein „was es war zu sein“ besitzen, deren Formel (λόγος) eine Definition ist (1030a6-7). In der *Zweiten Analytik* schreibt er, dass das „was es war zu sein“ das ist, was durch die Formel (λόγος) eines Dinges bezeichnet wird (II 11, 94a34-5). Also ist das „was es war zu sein“ eines Dinges wohl das, was seine Definition bezeichnet. Wir wissen „was es war Mensch zu sein,“ wenn wir die Definition des Menschen kennen. Wieder ergibt sich, dass das „was es war zu sein“ eines Dinges das ist, wozu sich das Ding entwickelt, wenn es sich normal entwickelt. Denn die Definition einer Sache steht für diese Sache selbst, und nicht für eine seiner Eigenschaften (vgl. *Kategorien* 5, 2a19-34). Wenn man zum Beispiel annimmt, dass die Definition von Mensch „rationales Lebewesen“ ist, dann kann man jeden Verweis auf Menschen durch einen Verweis auf rationale Lebewesen ersetzen, und also kann man von rationalen Lebewesen alles das sagen, was man von typischen Menschen sagen kann. Das wäre nicht möglich, wenn das „was es war zu sein“ bloß eine Eigenschaft des definierten Dinges wäre. Denn von der komplexen Eigenschaft, rational und lebendig zu sein, kann man nicht alles das sagen, was man von typischen Menschen sagen kann. Zum Beispiel ist sie nicht zweibeinig.

Wir können auf jeden Fall festhalten, dass die folgenden drei Fragen für Aristoteles auf dasselbe hinauslaufen:

- (1) Was ist die Formalursache dieses Dinges?
- (2) Zu was entwickelt sich dieses Ding, wenn es sich seiner Natur entsprechend entwickelt?
- (3) Was wird durch die Definition dieses Dinges bezeichnet?

Nun betont Aristoteles in der *Metaphysik* Z 15, dass sich Einzeldinge nicht (als einzelne) definieren lassen (1040a7). Man kann die Definition von Sokrates zwar an Stelle seines Namens verwenden, um Sokrates zu bezeichnen, aber es ist immer auch die Definition von Kallias. Es ist die Definition dessen, was beide sind: die Definition der zweiten Substanz Mensch. Sokrates ist ein Mensch, und Kallias ist auch ein Mensch. Also ist das, was durch die Definition eines Dinges bezeichnet wird, zwar keine Eigenschaft, aber doch etwas Allgemeines.<sup>16</sup> Es ist die Art, zu der das Ding gehört. Die Beschreibung dieser Art ist, wie wir eben gesehen haben, zugleich eine Beschreibung dessen, was die Mitglieder der Art werden, wenn sie sich ihrer Art entsprechend entwickeln.

#### 4. Probleme mit Mayrs typologischen Artbegriff

Man wird nach all dem nicht sagen können, Aristoteles habe den Artbegriff vernachlässigt. Ich werde mich nun den beiden Artbegriffen zuwenden, die Mayr gegeneinander ausspielt, und am Rande auch den morphologischen Artbegriff diskutieren. Mayr schreibt Aristoteles einen typologischen Artbegriff zu und meint, dass der typologische und der biologische Artbegriff einander vollständig entgegengesetzt seien (1969b, 200). Ich möchte zeigen, dass der Aristotelische Artbegriff nicht in der Weise typologisch ist, wie Mayr es beschreibt, und dass der richtig verstandene typologische Artbegriff dem biologischen zu Grunde liegt. Um das zu tun, werde ich nun eine Weile lang nicht über Aristoteles reden.

Die wichtigsten Probleme, die man mit Mayrs Version des typologischen Artbegriffs haben kann, sind schnell dargelegt.<sup>17</sup> Viele Arten belebter Dinge kann man nicht einfach dadurch definieren, dass man Merkmale angibt, die alle und nur die Instanzen dieser Art teilen. Genauer muss man sagen: Es gibt oft kein *nicht-relationales* Merkmal dieser Art. Ein relationales Merkmal ist eines, das allein darin besteht, dass das Ding, das es hat, in einer bestimmten Beziehung zu etwas anderem steht. Anhand solcher Merkmale kann man biologische Arten

relativ leicht definieren. Zum Beispiel ist die Art „Mensch“ die Klasse der Lebewesen, die von den ersten Menschen abstammen. Diese Definition mag operational schwer zu handhaben sein, aber sie trifft die Sache. Wenn man den Stammbaum aller Menschen kennen würde, könnte man wenigstens theoretisch genau alle Menschen in eine Klasse sortieren, und zwar anhand eines relationalen Merkmals, das allen und nur den Menschen zukommt.

Mayr unterstellt, dass typologische Arten stets anhand nicht-relationaler Merkmale definiert werden (1969a, 312). Sein Problem ist daher, dass sich keine definierenden Merkmale von Menschen angeben lassen, die man ihnen mehr oder weniger direkt ansehen könnte. Die Merkmale, die mit Sicherheit allen Menschen zukommen, sind immer auch Merkmale, die anderen Lebewesen ebenso zukommen. Die Merkmale, die allein einem Menschen zukommen können, kommen nicht mit Sicherheit allen Menschen zu (vgl. Mayr, 1969a, 313). Zum Beispiel sind nicht nur Menschen lebendig, und nicht alle Menschen sind rational. Wenn man Menschen als rationale Lebewesen definiert, und damit die Klasse aller Wesen meint, die leben und rational sind, dann hat man falsch definiert.

Aber auch wenn man solche Merkmale angeben könnte, die alle und nur die Instanzen einer Art aufweisen, dann würde man sich doch damit auf zwei Dinge festlegen. Erstens wäre dann ausgeschlossen, dass sich die Art evolutionär so entwickelt, dass eines dieser Merkmale verschwindet. Zweitens würde gelten, dass alle Dinge zur Art gehören, sofern sie nur diese Merkmale aufweisen. Das wollen Biologen aber nicht. Wenn eine Art ausstirbt und sich später eine merkmalsgleiche Art wieder entwickelt, dann soll es sich um zwei verschiedene Arten handeln (Hull, 1976, 184).

## 5. Ein morphologischer Artbegriff

Nun könnte man sich fragen, was eigentlich der Unterschied zwischen Mayrs typologischem Artbegriff und dem morphologischen Artbegriff sein soll.<sup>18</sup> Morphologische Merkmale sind nämlich nicht-relationale: solche, die man dem Ding, das sie hat, mehr oder weniger direkt ansehen kann. Aber nicht jeder morphologische Artbegriff basiert auf der Annahme, dass alle und nur die Mitglieder einer Art bestimmte nicht-

relationale Merkmale aufweisen. Der morphologische Artbegriff, den Douglas Gasking vorschlägt, funktioniert zum Beispiel anders. Um ihn zu verstehen, müssen wir den Begriff einer Referenzklasse definieren.

Eine Referenzklasse ist eine Menge von Dingen, die definiert werden kann durch Angabe eines bestimmten einzelnen Elementes dieser Menge, sowie einer Relation, in der alle anderen Elemente der Menge zu diesem Element stehen müssen (Gasking, 1996, 146). Das ausgezeichnete Element soll hier als Fokus bezeichnet werden. Im Fall einer morphologischen Art ist die relevante Relation die der seriellen hinreichenden Ähnlichkeit (1996, 150). Wann zwei Dinge einander hinreichend ähnlich sind, kann zunächst offen bleiben. *Seriell* hinreichend ähnlich sind zwei Dinge, wenn es eine Kette von Dingen gibt, von denen eines dem jeweils nächsten hinreichend ähnlich ist, welche die beiden zuerst genannten Dinge miteinander verbindet. Eine morphologische Art ist also eine Menge von Dingen, für welche es jeweils eine Kette von einander hinreichend ähnlichen Dinge gibt, die sie mit einem ausgezeichnetem Element der Menge, dem Fokus, verbindet. Es muss also nicht so sein, dass alle und nur die Mitglieder einer morphologischen Art bestimmte Merkmale miteinander teilen. Sie müssen sich noch nicht einmal paarweise hinreichend ähneln, und es müssen noch nicht einmal alle direkt dem Fokus hinreichend ähneln.

Nun ist es offenbar so, dass ein solcher Artbegriff leicht aus dem Ruder läuft, wenn man nicht genau genug festlegt, wann zwei Dinge einander hinreichend ähneln. Wenn man zum Beispiel annimmt, dass sich zwei Dinge dann hinreichend ähneln, wenn sie mehr Gemeinsamkeiten als Unterschiede aufweisen, landet man schnell bei der Klasse aller Dinge. Zu einem beliebigen Ding lässt sich immer eine Kette von Dingen angeben, die einander in dieser Weise hinreichend ähneln, welche dieses Ding mit einem beliebigen Fokus verbindet. Es gibt zwei nahe liegende Methoden, dieses Ausufern des morphologischen Artbegriffs zu verhindern. Erstens könnte man bestimmte Hinsichten festlegen, in denen die betreffenden Dinge einander ähneln sollen. Wenn man das tut, unterstellt man aber bereits, dass die Dinge in bestimmten anderen Hinsichten nicht nur ähnlich sind, sondern gleich. Wenn man zum Beispiel vorgibt, dass sich die Instanzen einer morphologischen Art hinsichtlich ihres Knochenbaus ähneln sollen, unterstellt man, dass sie alle Knochen haben. Zweitens könnte man festlegen, wie lang die Ketten sein dürfen. Es ist gut möglich, dass man zum Beispiel alle Muscheln



einer Art zu fassen kriegt, indem man die Klasse der Dinge definiert, die einem besonders typischen Exemplar dieser Art zu höchstens  $x$ -tem Grad seriell hinreichend ähneln. So ähnlich könnten Biologen vorgehen, wenn sie es mit Lebewesen zu tun habe, die sich nicht geschlechtlich vermehren.<sup>19</sup> Ein solcher morphologischer Artbegriff vermeidet eines der wichtigeren Probleme, die Mayr mit dem typologischen hat: Es muss nicht gelten, dass alle und nur die Exemplare einer Art bestimmte nicht-relationale Merkmale aufweisen. Er bleibt aber eine Notlösung, der die meisten Biologen sofort den biologischen Artbegriff vorziehen würden, wenn er anwendbar wäre. Diesem Artbegriff sollten wir uns nun also zuwenden.

## 6. Ein biologischer Artbegriff

Der biologische Artbegriff heißt biologisch, weil er auf Kriterien basiert, die nur auf belebte, nicht aber auf unbelebte Dinge Anwendung finden können (Mayr, 1969a, 314). Eine biologische Art kann man gut als Cluster definieren. Gasking zufolge ist ein Cluster so etwas Ähnliches wie eine Referenzklasse, nur dass gilt, dass jedes beliebige Element der Klasse als Fokus verwendet werden kann (1996, 155). Man kann also ein beliebiges Element eines Clusters herausgreifen, und alle anderen Elemente dieses Clusters stehen dann in der einschlägigen Relation zu dem herausgegriffenen. Die Relation, die für biologische Arten relevant ist, ist die der seriellen möglichen Paarungsfähigkeit. Diese Relation kann nur zwischen Lebewesen bestehen. Zwei Lebewesen sind paarungsfähig, wenn sie durch Paarung Nachkommen zeugen können, und sie sind möglicherweise paarungsfähig, wenn sie ohne weiteres durch Paarung Nachkommen zeugen könnten, falls sie sich begegnen würden. Zwei weibliche oder zwei männliche Lebewesen derselben Art sind untereinander nicht einmal möglicherweise paarungsfähig, aber dieses Problem lässt sich durch Serialisierung beheben. Serielle mögliche Paarungsfähigkeit liegt vor, wenn man zu zwei Lebewesen eine Kette von Lebewesen angeben kann, welche beide verbindet, so dass jedes Glied der Kette ohne weiteres mit dem jeweils nächsten Nachkommen zeugen könnte, wenn sich diese beiden begegnen würden. Innerhalb von Populationen und Arten kann man solche Ketten angeben, über die Grenzen von Arten hinweg kann man keine solche Ketten angeben.

So ungefähr gehen Biologen meist vor, wenn sie Arten beschreiben und bestimmen (vgl. Mayr et al., 1953, Kap. 12; Winston, 1995, Kap. 9). Es ist schon länger üblich und seit einigen Jahren Pflicht, jeder Beschreibung einer neuen Art einen Verweis auf den Verbleib des Exemplars beizufügen, das der Beschreibung zu Grunde lag. Dieses Exemplar wird (u. a.) Holotyp genannt, und die Regeln zur biologischen Nomenklatur sehen vor, dass der offizielle Name der Art all das bezeichnet, was mit diesem Exemplar in der richtigen Weise verwandt ist. Dieses Vorgehen hat nichts mit dem zu tun, was Mayr als „typologisches Denken“ brandmarkt (vgl. Mayr et al., 1953, 15). Es geht vielmehr einfach darum, die Bezeichnung einer Art vollständig von ihrer Beschreibung zu trennen. Wenn es darauf ankommt, zu entscheiden, ob ein Lebewesen einer Art angehört oder nicht, ist die letzte Instanz der Holotyp selbst, und nicht etwa dessen veröffentlichte Beschreibung. Also zählt letztlich auch keines der Merkmale, die diese Beschreibung den Instanzen einer Art zuschreibt. Auch der Holotyp wird tunlichst nicht anhand allgemeiner Merkmale identifiziert. Die einzigen Merkmale, die zu seiner Identifikation relevant sein sollen, sind eben die Merkmale, die er mit keinem anderen Lebewesen teilt: sein Standort zu einer gegebenen Zeit und die Nummer, unter der er im Katalog einer Sammlung aufgeführt ist. Der Holotyp zur Art *Tyrannosaurus Rex* trägt zum Beispiel die Nummer CM 9380. (Er steht, schwer zu übersehen, im *Carnegie Museum of Natural History* in Pittsburgh.)

Der Holotyp einer Art ist also ein Fokus, so dass man diese Art als die Klasse aller Lebewesen bestimmen kann, die in der richtigen Weise mit ihm verwandt sind. Die Regeln der Nomenklatur etablieren ein System, das es im Prinzip möglich machen soll, den Fokus jeder biologischen Art zu lokalisieren und bei Bedarf direkt zu untersuchen. Sie legen dagegen nicht fest, wann zwei Lebewesen in der richtigen Weise miteinander verwandt sind. Naheliegend ist es, eben hier die Relation der seriellen möglichen Paarungsfähigkeit in Anschlag zu bringen.

## 7. Probleme mit dem biologischen Artbegriff

Die Relation der seriellen möglichen Paarungsfähigkeit hat noch einen kleinen Mangel: Sie geht an den unfruchtbaren Mitgliedern einer Art vorbei. Diese können mit keinem anderen Lebewesen Nachkommen

zeugen, also sind sie auch in keiner Weise seriell möglicherweise paarungsfähig. Man kann sie vielleicht ins Boot holen, indem man den Möglichkeitsbegriff vorsichtig erweitert. Zum Beispiel könnte man sagen, dass zwei Lebewesen dann möglicherweise paarungsfähig sind, wenn sie Nachkommen zeugen könnten, falls sie sich begegnen würden und beide gesund wären. Wenn man das tut, muss man allerdings dazu sagen, was es für ein Lebewesen heißt, gesund zu sein, und das könnte in der definitiven Situation, in der wir uns befinden, leicht einen Kurzschluss bewirken. Denn Gesundheit ist keine Eigenschaft, die man einfach so an einem Lebewesen feststellen könnte, ohne bereits zu wissen, zu welcher Art es gehört. Man muss dafür die relevanten Standards kennen, denen das betreffende Lebewesen unterliegt, und die leiten sich bekanntlich aus seiner Natur ab. Offenbar bedeutet es für eine Spinne etwas anderes als für einen Menschen, gesund zu sein. Also kann man nur dann entscheiden, ob etwas gesund ist, wenn man bereits weiß, ob es eine Spinne oder ein Mensch ist (vgl. Thompson, 2004). Das aber stand eigentlich noch in Frage; denn ob es eine Spinne ist, wollten wir ja davon abhängig machen, ob es mit einer anderen Spinne Nachkommen zeugen könnte, wenn es gesund wäre. Es dürfte also besser sein, Verweise auf Gesundheit ganz zu unterlassen und die unfruchtbaren Lebewesen einer Art einfach zu ignorieren. Sie verursachen ohnehin nur Probleme, denn es kommt ja auch vor, dass Pferde und Esel oder Tiger und Löwen miteinander Nachkommen zeugen. Da diese meist unfruchtbar sind, wäre es ganz gut, wenn sie gar nicht als Nachwuchs zählten.<sup>20</sup>

Leider stellt sich ein sehr ähnliches Problem ein, wenn es um die Frage geht, was es für zwei Lebewesen überhaupt heißt, gemeinsam Nachkommen zu zeugen. Diese Frage muss man beantworten, bevor man sagen kann, unter welchen Umständen zwei Lebewesen seriell möglicherweise paarungsfähig sind. Nun gilt für das Zeugen von Nachkommen dasselbe wie für die Gesundheit: Verschiedene Lebewesen tun es auf verschiedene Art und Weise. Um zu wissen, ob sich zwei Lebewesen paaren, muss man also wissen, um welche Art Lebewesen es sich handelt. Wichtiger noch ist, dass man nur dann von der wirklichen Zeugung von Nachkommen sprechen kann, wenn auch Nachkommen entstehen, und also auch nur dann von der möglichen Zeugung von Nachkommen, wenn dies möglich ist. Dazu muss man wiederum wissen, wann etwas als Nachkomme von Lebewesen einer gegebenen Art zählen würde.

Das Problem ist, dass wir Paarung von anderen Interaktionen zwi-

schen Lebewesen unterscheiden können müssen, um den biologischen Artbegriff anzuwenden. Um das zu tun, müssen wir eine Vorstellung davon haben, was geschieht, wenn Fortpflanzung erfolgreich ist. Wenn wir aber beschreiben, was dann geschieht, müssen wir auch das beschreiben, was dabei herauskommt. Es reicht außerdem nicht, dass bei der Fortpflanzung einfach nur etwas herauskommt, das lebt. Denn erstens ist nicht jede Interaktion zwischen zwei Lebewesen, bei der ein weiteres Lebewesen entsteht, ein Fall von Fortpflanzung. Es liegt auf der Hand, dass bei der Fortpflanzung nicht einfach nur etwas gepflanzt wird, sondern eine Art *fort*gepflanzt wird, und das bedeutet: Die Art besteht dadurch fort, dass Nachkommen entstehen, die zu ihr gehören. Zweitens verhält es sich mit dem Leben selbst wieder genau so wie mit der Gesundheit und der Paarung: Verschiedene Lebewesen tun es auf verschiedene Weise. Es gibt wahrscheinlich kein allgemeines Merkmal, das alle und nur die Wesen aufweisen, die leben (vgl. Thompson, 2008, Kap. 2–3). Um festzustellen, ob eine Spinne lebt, muss man etwas über Spinnen wissen, und um festzustellen, ob ein Löwe lebt, muss man etwas über Löwen wissen.

Das, so könnte man hier einwenden, stimmt aber so nicht ganz. Denn man kann bereits feststellen, ob ein Löwe lebt, wenn man ganz allgemeine Dinge über Säugetiere weiß. Dazu muss man nichts über die Art wissen, der die betreffenden Lebewesen angehören, sondern vielleicht nur über ihre Gattung, Familie oder Klasse. Das oben gegebene Argument funktioniert nur, wenn man den Artbegriff allgemein fasst, also als Sorte von Ding. Wenn man dagegen genauer formuliert, scheint heraus zu kommen, dass man auch wissen kann, was es für zwei Lebewesen bedeutet, sich erfolgreich fortzupflanzen, ohne genau zu wissen, welcher *biologischen* Art diese Lebewesen angehören. Es reicht, wenn der Nachwuchs überhaupt lebensfähig ist, gemessen an den allgemeinen Standards, die für Lebewesen der betreffenden biologischen Gattung, Familie oder Klasse gelten.

Dass man wissen muss, wann etwas als Nachkomme von Lebewesen einer biologischen Art zählen würde, bedeutet also noch nicht, dass man bereits wissen muss, wann etwas als Vertreter dieser biologischen Art zählen würde. Denn es ist vorstellbar, dass bei der Paarung zweier Lebewesen ein Lebewesen entsteht, das zwar zur selben Gattung, Familie oder Klasse, aber zu einer anderen Art gehört. Zum Beispiel können sich Pferde und Esel untereinander paaren, oder Löwen und

Tiger. Die Lebewesen, die dabei entstehen, Maultiere und Liger, gehören nicht zur selben Art wie ihre Erzeuger, sondern bestenfalls zur selben Gattung.

Allerdings bedeutet das nicht, dass das eben vorgebrachte Argument ganz hinfällig ist. Es gilt immer noch, dass man eine Vorstellung davon haben muss, welche nicht-relationalen Merkmale etwas aufweisen muss, um als Nachwuchs von Lebewesen einer bestimmten Art zu gelten. Um den biologischen Artbegriff anzuwenden, muss man dann wenigstens über einen typologischen Gattungs-, Familien- oder Klassenbegriff verfügen. Darüber hinaus lässt sich zeigen, dass man Fortpflanzung als einen Vorgang definieren muss, der typischerweise zum Erhalt der Art führt, und nicht etwa nur zum Erhalt der Gattung. Das bedeutet nicht, dass es keine untypischen Fälle geben kann. Um aber zu verstehen, was Fortpflanzung überhaupt ist, muss man eine Vorstellung davon haben, was es für zwei Lebewesen derselben biologischen Art hieße, weitere Vertreter genau dieser Art zu erzeugen. Man stelle sich zum Beispiel einmal vor, dass eine Population von Pferden und Eseln nur Maultiere, aber keine Pferde oder Esel hervorbringt. Dann würde nicht gelten, dass die betreffende Population sich erhalten hat. An ihre Stelle wäre eine Population von Lebewesen anderer Art getreten. Dann wäre aber fraglich, ob man sagen soll, dass wirklich Fortpflanzung stattgefunden hat. Kann es sein, dass alle Mitglieder einer Population sich erfolgreich fortpflanzen, aber die Population selbst erlischt? Wir sind zwar geneigt, in solchen Fällen immer noch von Fortpflanzung zu sprechen, weil ja hier immerhin durch Zeugung und Geburt Lebewesen einer Gattung aus anderen Lebewesen derselben Gattung entstehen. Es handelt sich aber offenbar nicht um einen klaren Fall von Fortpflanzung, und zwar deswegen nicht, weil die Lebewesen, die entstehen, nicht zur selben Art gehören wie ihre Erzeuger. Solche Fälle müssen eine Ausnahme bleiben, denn nicht alle Fälle von Fortpflanzung könnten von dieser Art sein. Also scheint es das beste zu sein, den Begriff der Fortpflanzung anhand einer Referenzklasse zu bestimmen: Fortpflanzung ist, was einem klaren Fall erfolgreicher Fortpflanzung hinreichend ähnelt, und klare Fälle erfolgreicher Fortpflanzung sind solche, bei der zwei Angehörige einer biologischen Art von Lebewesen weitere Angehörige derselben Art erzeugen. Wenn das so ist, dann muss man wissen, was es für zwei Lebewesen bedeutet, derselben biologischen Art anzugehören, um zu verstehen, was Paarung ist.

Übrigens bestimmt auch Aristoteles Fortpflanzung als den Akt, durch den ein Lebewesen ein weiteres Lebewesen derselben Art erzeugt. Er sagt, dass es für ein Lebewesen „das natürlichste“ sei, weitere Lebewesen derselben Art zu produzieren (*De Anima* II 4, 415a26-b1; vgl. Lennox, 2001, 137). Das impliziert, dass es auch unnatürliche Fälle geben kann, bei denen etwa ein Pferd und ein Esel einen Maulesel hervorbringen (Cho, 2010, 302-4). Solche Fälle müssen aber, wie Aristoteles in der *Metaphysik* Z 8 betont, Ausnahmen bleiben (1033b29-33). Also geht auch Aristoteles davon aus, dass man den Begriff der Fortpflanzung nur verstehen kann, wenn man weiß, was es für Lebewesen einer gegebenen Art bedeutet, Lebewesen genau dieser Art zu erzeugen.

Das eben vorgebrachte Argument könnte man vielleicht missverstehen. Man könnte meinen, dass der biologische Artbegriff deshalb kritisiert wird, weil es in Einzelfällen zu Hybridbildungen kommen könne, also etwa Maultieren und Ligern, und es deshalb keine klaren Grenzen zwischen biologischen Arten gebe. Ein solches Argument wäre natürlich überzogen, und darauf hat Mayr wiederholt hingewiesen (z.B. Mayr, 1982, 274 und 283-4; Mayr und Ashlock, 1991, 34; Mayr, 1992, 228). Die Artgrenze steht nicht schon in Frage, wenn sie manchmal schwimmt. Wenn es also in wenigen Fällen vorkommt, dass Pferde und Esel gemeinsam Nachkommen zeugen, ist das noch kein Grund, nicht mehr zwischen Pferden und Eseln zu unterscheiden – denn das bedeutet ja noch nicht, dass sich ganze Populationen miteinander vermischen. Auch wenn es Grenzfälle gibt, bleiben doch die Arten als solche unterschieden. Da Mayr Arten als reproduktiv isolierte Populationen definiert, kommt es nur darauf an, dass es Faktoren gibt, die eine weitgehende Vermischung zweier Arten verhindern. Obwohl es in Einzelfällen zu Maultieren und Ligern kommen kann, sorgen die jeweiligen Isolationsmechanismen offenbar doch dafür, dass das nicht beliebig so weiter geht. Zum Beispiel dadurch, dass Hybride unfruchtbar sind oder keine Partner finden, oder etwa dadurch, dass sie wieder Nachkommen haben, die klar zu einer der ursprünglichen Arten gehören. Solange das so funktioniert, spricht die Möglichkeit von Hybriden nicht gegen den biologischen Artbegriff; wenn die Isolationsmechanismen aber allgemein versagen, dann wird es auch richtig sein, die beiden betreffenden Arten nicht mehr zu unterscheiden (Mayr und Ashlock, 1991, 101-2).

Umgekehrt zeigt all das aber gerade, dass Fortpflanzung nur dann wirklich erfolgreich ist, wenn die Art erhalten bleibt. Denn gerade

wenn wir daran festhalten wollen, dass es sich bei Hybridbildungen um Grenzfälle handelt, müssen wir auch daran festhalten, dass bei der Fortpflanzung im Normalfall die Art erhalten bleibt, und nicht etwa nur die Gattung. Wenn alle Nachkommen einer Art Hybride wären, dann wäre die Fortpflanzung dieser Art fürs erste gescheitert.

## 8. Noch einmal der typologische Artbegriff

Nun geht nicht alles den Bach runter, wenn man einfach zugesteht, dass man Paarungsfähigkeit nur dann feststellen kann, wenn man bereits weiß, was im Erfolgsfall bei der Paarung herauskommen würde. Es soll ja nicht gelten, dass zwei Lebewesen nur dann zur selben Art gehören, wenn sie tatsächlich zusammen ein Lebewesen derselben Art erzeugen. Es gilt nur, dass sie zur selben Art gehören, wenn sie dazu in der Lage sind, zusammen ein Lebewesen derselben Art zu erzeugen. Und das ist ein wesentlicher Unterschied. Denn um zu wissen, ob zwei Lebewesen dazu in der Lage sind, etwas hervorzubringen, muss man nicht erst warten, bis sie Erfolg damit haben. Es reicht, dass sie etwas tun können, das dann erfolgreich wäre, wenn ein weiteres Exemplar ihrer Art dabei entstünde. Um das zu wissen, muss man eigentlich gar nicht wissen, was tatsächlich bei irgendeinem konkreten Paarungsversuch herauskommt. Man muss nur die Erfolgskriterien dessen kennen, was geschieht, wenn der Versuch stattfindet.

Ein glatter Zirkel würde entstehen, wenn man in der Definition von Paarungsfähigkeit Merkmale angeben müsste, die etwas aufweisen muss, um überhaupt als Nachwuchs zu gelten. Denn alle Lebewesen einer Art sind ja auch Nachkommen von Lebewesen ihrer Art, und wenn alle und nur die Nachkommen einer Art bestimmte Merkmale aufweisen müssten, dann müsste das auch für alle Instanzen dieser Art überhaupt gelten. Daher scheint es dann so, als hätte man sich den Umweg über einen Fokus und eine serielle Relation auch sparen können. Man hätte die Art einfach als Klasse der Wesen definieren können, die die fraglichen Merkmale haben.

Dieser glatte Zirkel ergibt sich aber nicht. Denn die Merkmale, anhand derer man Paarungsfähigkeit definieren muss, befinden sich in einem modalen Kontext. Es handelt sich nicht um Merkmale, die ein Mitglied der Art tatsächlich aufweisen muss, um als solches zu gelten. Bei der

Zeugung von Nachwuchs kann nämlich beliebig viel schief gehen, ohne dass man gleich daraus schließen müsste, dass gar keine Paarung vorliegt. Wenn man beschreibt, was bei der Paarung zweier Lebewesen herauskommen soll, beschreibt man nur den Erfolgsfall, von dem die tatsächlich vorkommenden Fälle abweichen können.

Also kann man sich den Umweg über einen Fokus und die Relation der seriellen möglichen Paarungsfähigkeit nicht sparen. Eine Art ist nicht einfach eine Klasse von Lebewesen, die die Merkmale der Produkte erfolgreicher Paarung haben, sondern sie ist eine Klasse von Lebewesen, die aus Prozessen resultieren, die bestimmten Erfolgskriterien unterliegen. Sie müssen diese Erfolgskriterien nicht erfüllen. Ob Erfolgskriterien gelten und ob sie erfüllt werden sind zwei ganz verschiedene Fragen.

Um den erforderlichen Umweg über einen Fokus und die Relation der seriellen möglichen Paarungsfähigkeit machen zu können, muss man also zwar angeben, wann etwas als gelungener Nachwuchs der betreffenden Art gilt, und wenn man das kann, dann kann man auch sagen, wann etwas als typisches und wohlgeformtes Exemplar der Art gilt. Das bedeutet aber nicht, dass wir auf den Artbegriff angewiesen sind, den Mayr als typologisch bezeichnet. Wenn man sagt, welche Merkmale ein typisches Exemplar einer Art hat, gibt man wie gesagt nicht etwa Merkmale an, die alle und nur die Instanzen der Art aufweisen. Man sagt lediglich, welche Merkmale etwas aufweisen muss, um als Resultat eines klaren Falls erfolgreicher Reproduktion dieser Art zu gelten. Man sagt nur, welche nicht-relationalen Merkmale *typische* Exemplare einer Art aufweisen würden.

## 9. Schluss

Wenn man biologische Arten anhand des Begriffs der Fortpflanzung definieren will, dann muss man wissen, was es für zwei Vertreter einer gegebenen Art bedeuten würde, weitere Vertreter dieser Art zu erzeugen. Dazu muss man eine Vorstellung von den nicht-relationalen Merkmalen haben, die für Lebewesen dieser Art typisch sind. Denn man muss angeben können, welche Merkmale etwas aufweisen müsste, um in diesem Fall als Resultat eines klaren Falles erfolgreicher Fortpflanzung zu gelten. Das bedeutet, dass man den typologischen Artbegriff



nicht einfach durch den biologischen ersetzen kann. Es bedeutet nicht, dass man schon genau wissen muss, welche Merkmale ein typischer Vertreter einer Art haben muss, *bevor* man weiß, was innerhalb dieser Art als Fortpflanzung gelten kann. Es bedeutet nur, dass die beiden Fragen nach den typischen Merkmalen von Vertretern einer Art und danach, was in Bezug auf diese Art als Fortpflanzung gelten kann, in einem Zug geklärt werden müssen. Die Fragen lassen sich nicht voneinander trennen. Das bedeutet wiederum auch nicht, dass Mayrs biologische Artbegriff unsauber oder unbrauchbar ist, sondern eben nur, dass man den typologischen Artbegriff nicht vollends abschaffen und durch den biologischen Artbegriff ersetzen kann. Denn der biologische Artbegriff, so richtig und brauchbar er ist, setzt eine Variante des typologischen Artbegriffes voraus. Diese Variante des typologischen Artbegriffes bleibt gerade auch dann im Rennen, wenn der biologische Artbegriff das Rennen gewinnt.

Die Frage, ob etwas ein Vertreter einer Art ist oder nicht, muss man von der Frage unterscheiden, ob es die für diese Art typischen Merkmale aufweist. Es ist denkbar, dass etwas die typischen Merkmale einer Art aufweist, obwohl es gar keine Instanz dieser Art ist. Die Frage, ob etwas zu einer Art gehört, klärt man also in der Tat am besten ungefähr so, wie es Mayr vorschwebt: Man sieht nach, ob es mit anderen Exemplaren der Art verwandt ist. Der Punkt, den Mayr übersieht, ist nur, dass man das nicht kann, ohne sich implizit auf den Artbegriff zu stützen, der eigentlich den Namen „typologisch“ verdient; nämlich den, den man durch Angabe des für eine Art Typischen angibt. Ich hatte anfangs darauf hingewiesen, dass Aristoteles' Formalursache in dem Sinne als Vorbild bezeichnet werden kann, als es sich um das handelt, was ein Ding werden würde, wenn es sich seiner Natur entsprechend entwickeln würde. Es ist nun leicht zu sehen, dass Aristoteles damit genau die Variante des typologischen Artbegriff trifft, die vom biologischen vorausgesetzt wird.<sup>21</sup>

Um eine biologische Art zu definieren, muss man tatsächlich nicht angeben, welche Merkmale alle und nur die Instanzen dieser Art aufweisen. Man muss nur sagen, unter welchen Umständen Instanzen dieser Art typisch für diese Art wären. Diese Umstände müssen artspezifisch sein, d.h. ihre Beschreibung sollte für keine zwei verschiedene Arten gleich sein. Es kann aber Instanzen einer Art geben, die nicht typisch sind, und deswegen kann es auch sein, dass es tatsächlich kei-

ne Merkmale gibt, die alle und nur die Instanzen einer Art tatsächlich aufweisen. Das kann sogar so sein, wenn alle Instanzen der Art tatsächlich vollkommen typisch sind, denn es ist ja denkbar, dass verschiedene Instanzen auf verschiedenartige Weise vollkommen typisch für ihre Art sind. Weibliche Löwen sind zum Beispiel typischerweise größer als männliche. Um zu sagen, was für Löwen typisch ist, muss man also zwischen weiblichen und männlichen Löwen unterscheiden. Das hat Aristoteles so nicht gesehen, aber daran ist sein Artbegriff nicht schuld.<sup>22</sup>

Elliott Sober zeigt, dass die moderne Biologie innerhalb einer Art nicht mehr zwischen Normalfällen und Abweichungen unterscheidet, sondern Variation als normal und wesentlich für eine Spezies ansieht (1980, 374). Wenn wir beschreiben, was bei der Paarung zweier Lebewesen typischerweise herauskommen sollte, müssen wir demnach eine sehr flexible Beschreibung geben, die reichlich Spielraum für Variation lässt. Das bedeutet jedoch nicht, dass die Vorstellung eines typischen Vertreters einer Art keinen Sinn mehr macht.<sup>23</sup>

Ob sich so bestimmte Arten evolutionär entwickeln können, möchte ich wie gesagt offen lassen. Ich habe allerdings nicht den Eindruck, dass etwas dagegen spricht. Gasking schlägt zum Beispiel vor, Arten als verkettete Cluster von Formen zu beschreiben (1996, 178). Was er damit meint, lässt sich ungefähr wie folgt skizzieren. Man kann für gegebene Zeitspannen ziemlich genau angeben, welche Merkmale für die Vertreter einer Art typisch sind. Gasking bezeichnet die Populationen, für die man das kann, als Formen. Wenn sich die Art entwickelt, ändern sich möglicherweise die Standards, denen ihre Instanzen unterliegen. Wenn das geschieht, muss man zu gewissen Zeitpunkten zugeben, dass eine bisher untypische Erscheinung jetzt typisch ist, und eine bisher typische Erscheinung einen Atavismus darstellt. Auch dann ist aber stets klar, dass die neue Form mit der alten verwandt ist. Denn die Vertreter der Form, für die die neuen Maßstäbe gelten, sind seriell Nachkommen von Vertretern der Form, für die ihrerzeit die alten Maßstäbe galten. Dazwischen mag es eine Grauzone geben, aber das macht nicht viel aus, solange sich irgendwann eine neue Form stabilisiert. Gasking schlägt vor, die ganze Art als Cluster von Formen zu definieren, die seriell durch Nachkommenschaft aus anderen Formen in diesem Cluster hervorgehen. Was den Aristotelischen Artbegriff angeht, scheint dem nichts im Wege zu stehen (vgl. Sober, 1980, 356; Cho, 2010, 312).

## Anmerkungen

- 1 Gegen Platon hegt Mayr eine tief verwurzelte Abneigung (Winsor, 2006, 160–61), und obwohl er Aristoteles offenbar mehr schätzt, sieht er keine Unterschiede zwischen den Artbegriffen von Platon und Aristoteles, die für Biologen relevant wären (1969b, 200). Mir geht es hier nicht darum, Mayr in diesen Hinsichten zu korrigieren, sondern um die beiden Artbegriffe, die er gegeneinander ausspielt.
- 2 Wenn Peter van Inwagen (1990) meint, dass es eigentlich nur Lebewesen und „simples“ gebe, dann ist das missratener Aristotelismus. Er meint, dass die Unterscheidung, die Aristoteles im Bereich des Seienden trifft, eigentlich eine Unterscheidung zwischen Seiendem und nicht Seiendem sei.
- 3 Vgl. zu dieser Stelle Corcilius, 2008, 326–32.
- 4 Vgl. *Physik* VIII 6, 259b1–20, sowie Wieland, 1992, 234 und Sauv  Meyer, 1994, 68.
- 5 Ich verwende hier „Form“ im Sinne von „das, was ein Ding ist,“ also gleichbedeutend mit „Art.“ Man k nnte einwenden, dass ich hier zwei ganz verschiedene Dinge gleichsetze: (1) den Gegenbegriff zur Materie eines Dinges und (2) die Art, zu der dieses Ding geh rt. Und in der Tat verbirgt sich hinter meiner Redeweise eine These, die ich im n chsten Abschnitt zu begr nden beginne: Dass die Formalursachen, von denen Aristoteles in der *Physik* spricht, dasselbe sind wie die zweiten Substanzen, von denen er in der *Kategorienschrift* spricht. Diese These bedarf einer gr ndlicheren Diskussion, als ich sie hier bieten kann; vgl. ansatzweise Hennig, 2010.
- 6 Man k nnte auch den folgenden Widerspruch sehen. Wenn Aristoteles sagt, dass Lebewesen den Ursprung von Bewegung und Ruhe in sich haben, scheint er zu sagen, dass es an ihnen liegt, ob sie sich bewegen oder nicht. In *De Motu Animalium* 8 scheint Aristoteles aber zu sagen, dass es Tieren nicht immer freisteht, ihre Bewegungen auch zu unterlassen, denn sie werden ja mechanisch von dem angezogen, was sie erstreben. Dieser Widerspruch ergibt sich jedoch nicht, wenn man die Rede vom „Ursprung der Bewegung und Ruhe“ wie folgt versteht: Wenn ein Tier sich bewegt, liegt der Ursprung seiner Bewegung in ihm, und wenn es sich nicht bewegt, liegt der Ursprung seiner Bewegungslosigkeit in ihm; egal, ob es sich auch anders verhalten k nnte.
- 7 Zumal Aristoteles in der *Metaphysik* Z16 (1040b8–10) sagt, dass die Elemente eigentlich gar keine Substanzen seien. Vgl. Gill, 1989, 235–242.
- 8 Frede und Patzig schreiben, dass das Menschsein insofern als Eigenschaft bezeichnet werden k nne, als man es von etwas aussagen kann, dass es aber keine Eigenschaft sei im Sinne von etwas, das einem gegebenen Ding zukommt (1988, *Einleitung*, 44). Ich verwende das Wort „Eigenschaft“ in diesem zweiten Sinn.
- 9 Die obige Formulierung findet sich genauer bei Su rez, *Disputationes Metaphysicae* XXXIII 1.1: „substantia enim ita substat accidentibus ut non indigeat ipsa simili sustentaculo.“
- 10 Vgl. Moravcsik, 1974; Johnson, 2005, Kap. 2; Hennig, 2009.

- 11 Oft wird zwischen dem αἴτιον und der αἰτία unterschieden, welche Letztere eine Erklärung ist, warum und in welcher Weise das αἴτιον für etwas verantwortlich ist. Wie Michael Frede anmerkt, scheint Aristoteles diese Unterscheidung nicht zu beachten (1987, S. 129–30). Mit beiden Worten bezeichnet er Faktoren, die für etwas (im weiten Sinne) verantwortlich sind, und nicht etwa Erklärungen oder Antworten auf Warum-Fragen.
- 12 In dem Sinne, in dem Heidegger schreibt, dass die Frage nach der φύσις „bei der Bewegtheit dieses Seienden anfragen“ müsse (*Vom Wesen und Begriff der Φύσις, Gesamtausgabe* 9, S. 245).
- 13 Vgl. Balme, 1962, 87 und Lennox, 2001, 160 Anm.
- 14 Er tut dies übrigens nur an der oben zitierten Stelle (*Physik* II 3, 194b26–27), sowie an der Parallelstelle *Metaphysik* Δ 2 (1013a26–27). Vgl. Owen, 1965, 128.
- 15 Vgl. zur Interpretation der Wendung Frede und Patzig, 1988, *Kommentar* 34–5, sowie die dort genannte Literatur.
- 16 In der *Metaphysik* Z 13 betont Aristoteles, dass Substanzen nichts Allgemeines seien, und das scheint der These zu widersprechen, das, was die Definition eines Dinges bezeichnet, sei stets etwas Allgemeines. Der Widerspruch lässt sich möglicherweise auflösen, wenn man Z 13 ungefähr so versteht: Substanzterme sind solche, anhand derer etwas Einzelnes als Exemplar einer bestimmten Art herausgegriffen wird. Sie unterscheiden sich also von Termen, anhand derer etwas Allgemeines über eine Sache gesagt wird. Dennoch bleiben sie in einem bestimmten Sinn allgemein, denn den Substanzterm, anhand dessen man Sokrates als Exemplar der Art Mensch herausgreifen kann, kann man ebenso verwenden, um Kallias als Exemplar der Art Mensch herauszugreifen. Vgl. Hennig, 2010, Abschn. 6.6.
- 17 Mayrs Typologie ist offenbar ein bloßer Strohmännchen; Vgl. die lesenswerte Darstellung von Winsor, 2006, 159.
- 18 Mayr schreibt, dass der typologische Artbegriff zum morphologischen wird, wenn man ihn in der Biologie anwendet (1955, zitiert nach Mayr, 1976, 485).
- 19 An einer Stelle schreibt Mayr, dass nichtgeschlechtliche Arten „typologisch umgrenzt“ werden (2002, 99). Da der typologische Artbegriff Mayr zufolge praktisch ein morphologischer ist (siehe die vorhergehende Fußnote) bedeutet das, dass hier morphologische Kriterien zum Einsatz kommen. Das widerspricht jedoch dem, was Mayr andernorts sagt: Dass Arten nichtgeschlechtlicher Organismen gar nicht morphologisch definiert werden, sondern vielmehr anhand ihrer ökologischen Nische (1982, 283).
- 20 Das ist deswegen nur eine Notlösung, weil gar nicht alle Lebewesen, die von artverschiedenen Lebewesen abstammen, unfruchtbar sein müssen. (Aristoteles geht davon aus, dass Hybride in der Regel fruchtbar sind; vgl. *De Generatione Animalium* II 8, 746b12–16.) Eine bessere Lösung ist, nur ein negatives Kriterium aufzustellen: Zwei Lebewesen gehören *nicht* zur selben Art, wenn sie zwar überhaupt paarungsfähig sind, aber nicht seriell möglicherweise miteinander.

- 21 Daraus, dass wir nicht darauf verzichten können, mit typologischen Artbegriffen zu arbeiten, folgt kein „Aristotelischer Essentialismus“ (was auch immer das ist). Ein essentialistischer Artbegriff wird hier also nicht verteidigt.
- 22 Aristoteles scheint anzunehmen, dass die Instanzen einer Art nur auf eine einzige Weise typisch sein können. Daraus schließt er, dass weibliche Lebewesen grundsätzlich untypisch sind (vgl. *De Generatione Animalium* I 3, 737a27-28).
- 23 Dass der Begriff des Typischen in der Biologie wichtig bleibt, betonen neben Thompson, 2004 und 2008, auch Johnston, 1981, und Love, 2009.

### Literatur

- Balme, David M., 1962: Γένος and Εἶδος in Aristotle's Biology. In: *Classical Quarterly* 12(1), S. 81–98.
- Cho, Dae-Ho, 2010: Beständigkeit und Veränderlichkeit der Spezies in der Biologie des Aristoteles. In: Föllinger, Sabine (Hg.), *Was ist ‚Leben‘? Aristoteles' Anschauungen zur Entstehung und Funktionsweise von Leben*. Stuttgart: Franz Steiner Verlag.
- Corcilus, Klaus, 2008: *Streben und Bewegung. Aristoteles' Theorie der animalischen Ortsbewegung*. Berlin: Walter De Gruyter.
- Frede, Michael, 1987: *Essays in Ancient Philosophy*. Minneapolis: University of Minnesota Press.
- Frede, Michael; Patzig, Günther, 1988: *Aristoteles' Metaphysik Z*. München: C. H. Beck.
- Gasking, Douglas, 1996: *Language, Logic, and Causation*. Carlton South: Melbourne University Press.
- Gill, Mary Louise, 1989: *Aristotle on Substance. The Paradox of Unity*. Princeton: University Press.
- Hennig, Boris, 2009: The Four Causes. In: *The Journal of Philosophy* 106(3), S. 137–160.
- Hennig, Boris, 2010: *The Four Causes*. Habilitationsschrift, Universität Leipzig.
- Hull, David L., 1976: Are Species Really Individuals? In: *Systematic Zoology* 25(2), S. 174–191.
- Inwagen, Peter van, 1990: *Material Beings*. Ithaca, London: Cornell University Press.
- Johnson, Monte R., 2005: *Aristotle on Teleology*. Oxford: Clarendon Press.

- Johnston, Timothy D., 1981: Species-typicality: Can Individuals Have Typical Parts? In: *Behavioral and Brain Sciences* 4, S. 291–2.
- Lennox, James G., 2001: *Aristotle's Philosophy of Biology. Studies in the Origins of Life Science*. Cambridge: University Press.
- Love, Alan C., 2009: Typology Reconfigured: From the Metaphysics of Essentialism to the Epistemology of Representation. In: *Acta Biotheoretica* 57, S. 51–75.
- Mayr, Ernst, 1955: Karl Jordan's Contribution to Current Concepts in Systematics and Evolution. In: *Transactions of the Royal Entomological Society of London* 107(1–14), S. 45–66.
- Mayr, Ernst, 1969a: The Biological Meaning of Species. In: *Biological Journal of the Linnean Society* 1(3), S. 311–320.
- Mayr, Ernst, 1969b: Footnotes on the Philosophy of Biology. In: *Philosophy of Science* 36(2), S. 197–202.
- Mayr, Ernst, 1976: *Evolution and the Diversity of Life. Selected Essays*. Cambridge, Mass.: Harvard University Press.
- Mayr, Ernst, 1982: *The Growth of Biological Thought. Diversity, Evolution, and Inheritance*. Cambridge, Mass.: Harvard University Press.
- Mayr, Ernst, 1992: A Local Flora and the Biological Species Concept. In: *American Journal of Botany* 79(2), S. 222–238.
- Mayr, Ernst, 1996: What Is a Species, and What Is Not? In: *Philosophy of Science* 63 (2), S. 262–277.
- Mayr, Ernst, 2002: Comments by Ernst Mayr. In: *Theory in Biosciences* 121(1), S. 99–100.
- Mayr, Ernst; Ashlock, Peter D., 1991: *Principles of Systematic Zoology. Second Edition*. New York: McGraw-Hill.
- Mayr, Ernst; Linsley, Gordon E.; Usinger, Robert L., 1953: *Methods and Principles of Systematic Zoology*. New York: McGraw-Hill.
- Moravcsik, Julius, 1974: Aristotle on Adequate Explanation. In: *Synthese* 28(1), S. 3–17.
- Owen, G. E. L., 1965: The Platonism of Aristotle. In: *Proceedings of the British Academy* 50, S. 125–150.
- Sauvé Meyer, Susan, 1994: Self-Movement and External Causation. In: Gill, Mary Louise Gill; Lennox, James G. (Hg.), *Self-Motion from Aristotle to Newton*. Princeton: University Press.
- Sober, Elliott, 1980: Evolution, Population Thinking, and Essentialism. In: *Philosophy of Science* 47(3), S. 350–383.
- Thompson, Michael, 2004: Apprehending Human Form. In: O'Hear,

- Anthony, (Hg.), *Modern Moral Philosophy*. Cambridge: University Press.
- Thompson, Michael, 2008: *Life and Action*. Cambridge, Mass.: Harvard University Press.
- Waterlow [Broadie], Sarah, 1982: *Nature, Change, and Agency in Aristotle's Physics*. Oxford: Clarendon Press.
- Wieland, Wolfgang, 1992: *Die Aristotelische Physik*. Göttingen: Vandenhoeck und Ruprecht.
- Winsor, Mary P., 2006: The Creation of the Essentialist Story: An Exercise in Metahistory. In: *History and Philosophy of the Life Sciences* 28(2), S. 149–174.
- Winston, Judith E., 1995: *Describing Species*. New York: Columbia University Press.

Vera Spillner

## A Bundle Definition of Scientific Understanding and its Application to Quantum Physics

### Zusammenfassung

Der vorliegende Artikel präsentiert einen Bündelbegriff ‚wissenschaftlichen Verstehens‘, mit dem sich die empirisch äquivalenten Interpretationen der Quantenmechanik bezüglich des durch sie ermöglichten Verstehens bewerten lassen. Dieser Verstehensbegriff beruht auf einem hinreichenden und notwendigen Kriterium, sowie einem Bündel von Bedingungen – wobei gilt, dass eine Theorie als umso verständlicher gilt, je mehr dieser Bedingungen durch die Theorie erfüllt werden. Somit basiert der Verstehensbegriff einerseits auf der objektiven Anzahl erfüllter Bedingungen, die andererseits durch Individuen gewichtet werden. Bei Anwendung des vorgeschlagenen Verstehensbegriffes auf drei Interpretationen der Quantenmechanik, stellt sich die Interpretation von David Bohm als die objektiv am besten verstehbare Deutung der Quantenmechanik heraus, gefolgt von der Deutung von Tim Maudlin und schließlich der so genannten Kopenhagener Deutung.

### Summary

This paper presents a bundle definition for ‚scientific understanding‘ through which the empirically equivalent interpretations of quantum mechanics can be evaluated with respect to the understanding they generate. This definition of understanding is based on a sufficient and necessary criterion, as well as a bundle of conditions – where a theory can be called better understandable than others, whenever it fulfills a higher number of bundle criteria. Thereby the definition of understanding is based on the one hand on the objective number of criteria a theory fulfills, as well as, on the other hand, on the individual’s preference of bundle criteria. Applying the definition onto three interpretations of quantum mechanics, the interpretation of David Bohm appears as most understandable, followed by the interpretation of Tim Maudlin and the Copenhagen interpretation.



Quantum Mechanics is thought to be one of the most successful theories of physics, despite of the fact that its interpretations are manifold and have been subject of intensive discussions for almost a century. The large variety of interpretations and the problems which the different interpretations entail have led some physicists like Richard Feynman to think that quantum mechanics can only be used for predictions of measurement results, but cannot be ‚understood‘. As Feynman prominently remarked: „I think I can safely say that nobody understands quantum mechanics.“ In this article, we argue that producing understanding is one of the central goals of the scientific enquiry. We will therefore suggest a bundle definition of scientific understanding by which quantum mechanical interpretations can be classified according to the scientific understanding they help to achieve. The definition of scientific understanding presented here does not depend on any temporal or social context and allows for an objective – i. e., non-individualistic, context-independent – evaluation of how much understanding a quantum mechanical interpretation produces.

## Scientific Explanation and Understanding

In his 1974 essay ‚Explanation and Scientific Understanding‘ in the *Journal of Philosophy* (Friedman, 1974, 5–19), Michael Friedman points out, that the definition of ‚scientific explanation‘ has been widely discussed, whereas the notion of ‚scientific understanding‘ still lacks a clear definition:

„Some philosophers [...] have relatively precise proposals as to the nature of the explanation relation, but relatively little to say about the connection between their proposals and the scientific understanding, i. e., about what it is about the relation they propose that gives us understanding of the world.“ (Friedman, 1974, 5)

He furthermore formulates two central conditions a theory of understanding is supposed to fulfill: First, the notion of understanding should be objective – i. e., it should be independent of individuals and historical context – and second, the connection between scientific explanation and understanding should become obvious. The theory of understanding offered in this paper claims to fulfill these conditions and there-

by claims to be a good candidate for a satisfactory theory of scientific understanding.

## Problems of current Theories of Understanding

The here presented definition of scientific understanding will comprise three of the most prominent theories of explanation. These are the deductive-nomological (DN) model of explanation (Hempel, 1965), the Unificationist approach (Kitcher, 1989, 410-505) and the causal model of explanation (Salmon, 1984). Each of these theories defines what counts as an explanation and then assumes that understanding arises as their model of explanation is fulfilled by a scientific theory. In this article, we argue, that this assumption is problematic for various reasons and can therefore not be called very satisfactory. We claim that a better model of understanding, which will include the best known theories of scientific explanation, can be formulated.

As stated above, many philosophers assume that scientific understanding follows directly as their model of scientific explanation is fulfilled. In his book *Der Schluss auf die beste Erklärung* (Klärner, 1974), Holger Klärner describes the connection between explanation and understanding that C. G. Hempel (Hempel, 1965) formulated using the DN-model: There, understanding of a phenomenon is achieved by the fact that through nomological deduction – which is the basis of scientific explanation in the DN-model – the phenomenon was to be expected.

The DN-model, however, has been criticised, for example by Michael Friedman in ‚Explanation and Scientific Understanding‘ (Friedman, 1974), who argues that if one correctly predicts – and thus expects – a storm using a barometer, one still has not understood the phenomenon of the storm as such. Thereby, the DN-model can not offer a sufficient condition for scientific understanding.

In the second prominent model, the causal model of explanation as formulated by Wesley Salmon (Salmon, 1984), a scientific explanation of an event E is achieved by uncovering the causal interactions and processes leading to E. The explanation eventually shows, how an event fits into a causal web. In his Stanford Library essay, James Woodward (Woodward, 2003) offers a visual example for this causal model of explanation: A billiard stick makes one billiard ball move forward and

doing so, it attaches a piece of blue chalk to it (a ‚mark‘, which is then transferred by the next ball, identifying the ball as a ‚causal process‘, i. e., the relevant factor in the explanation of the change in the state of motion of the second ball). Some of the blue chalk is furthermore passed on to the second ball, which marks that a ‚causal interaction‘ has taken place. Finally, scientific understanding in the causal model is achieved through listing all causal processes and interactions.

Again, this model of scientific understanding cannot be regarded as satisfactory. As argued by Christopher Hitchcock (Hitchcock, 1995, 304–320) in 1995, the identification of causal processes in Salmon’s model, which is a relevant part of achieving scientific explanation, becomes problematic, since the causal agents are not always clearly identifiable. In the example of the billiard balls, we find that it cannot be clearly identified, what ‚caused‘ the change in the state of motion of the second ball: both the momentum of the first ball and equally the blue chalk make up possible causal agents (both being transferred ‚marks‘). Since we agree that the blue chalk is a bad candidate, we must conclude that a ‚mark‘ transferred by a process does not necessarily have to be the causal reason for an event E. The result is, that Salmon’s model cannot clearly distinguish between relevant and irrelevant explanations. As a reply to this criticism, Salmon modified his model (Salmon, 1997, 461–477) from the transfer of a ‚mark‘ to the transfer of a non zero amount of a conserved quantity. Yet, it still has remained unclear whether the momentum, the angular momentum or possibly a charge (various in the collision process conserved quantities) had ‚caused‘ the state in motion of the second billiard ball (Spillner, 2011, 42). Thereby the causal agent – which needed to be identified for a scientific explanation – still remains not uniquely identifiable.

Identifying causal agents however is an important part of Salmon’s definition of scientific understanding:

„[...] causal processes, causal interactions, and causal laws provide the mechanisms by which the world works; to understand why certain things happen, we need to see how they are produced by these mechanisms. [...] Underlying causal mechanisms hold the key to our understanding of the world [...] (Salmon, 1984, 260)

Since identifying causal agents is essential for Salmon’s theory of scientific understanding, the lack of a unique identification poses a problem for his definition.

In this paper, we argue that the principle of causality is neither sufficient nor necessary for understanding and therefore possesses no absolute standard for understandability or scientific understanding. Some philosophers, like Henk de Regt and Dennis Dieks, have assumed that the understandability of theories depends on historical context:

„[The] intelligibility of scientific theories [...] is essentially contextual, [...] [it] depends on contextual factors, and can change in the course of time.“  
(Regt, Dieks, 2005, 137–170)

In this paper, we argue that the basis of what is called ‚scientific understanding‘ does not depend on time and context but remains constant over historical periods. The only effect the historical context has on the definition of scientific understanding presented here, lies in individual’s weighting of bundle criteria for understanding, as we will see below. This, however, will not effect the absolute – and context-independent – definition of understanding that we will present in the following.

Finally, the most prominent alternative to the causal model of explanation is the unificationist approach as formulated for example by Philip Kitcher (Kitcher, 1989, 410–505). Exponents of the unificationist approach assume that scientific explanation consists in producing a unified picture of a number of different phenomena, which previously counted as unconnected. Scientific explanation thereby consists of obtaining a common explanatory basis.

The main criticism lies in the question whether unification and explanation are uniquely connected. Intuitively we feel that summarizing various physical phenomena as stemming from the same physical cause indeed is connected to explanation. However, we also feel that other unifications do not have the same explanatory power: a common mathematical description for various physical phenomena – as using Lagrange equations in mechanics and electrodynamics – will not entail a unification of Newton’s gravity and electrodynamics. In fact, a common mathematical description – which very closely meets the unificationists’ concept of what is meant by ‚unification‘ – does not give us any insight into the causal factors responsible for the different phenomena, hence a common mathematical formalism does not automatically entail a common physical explanation – and might possibly not entail anything that we would call a ‚scientific explanation‘.

The unificationist definition of scientific understanding is based on the concept of explanation. As summarized by Michael Friedman:

„Science increases our understanding of the world by reducing the total number of independent phenomena that we have to accept as ultimate or given. A world with fewer independent phenomena is, other things equal, more comprehensible than one with more.“ (Friedman, 1974, 15)

Friedman's model of Unification was criticised and revised by Kitcher (Kitcher, 1981, 507–531), Schurz and Lambert (Schurz, Lambert, 1994, 109) into a theory, in which so called argument patterns play a central role:

„Understanding the phenomena is not simply a matter of reducing the fundamental incomprehensibilities but of seeing connections, common patterns, in what initially appeared to be different situations. [...] Science advances our understanding of nature by showing us how to derive descriptions of many phenomena, using the same patterns of derivation again and again [...]“ (Kitcher, 1989, 432)

The revised model has some advantages, one of which is its more general applicability. But still, the above raised question remains whether unification always entails understanding. For example, considering Bohr's atomic theory we find that the unification of spectral analysis and of the principle of quantization did not lead to a deeper understanding of the atom: many physicists to this day evaluate the theory as ‚not well understood‘, since it violates important principles like the continuity of motion.

It seems that the DN-model, the causal model of understanding and the unificationist approach are important approaches to a concept of scientific explanation – whereas they fail to prove necessary or sufficient for the production of scientific understanding. Neither the principle of causality nor that of unification defines a unique or absolute standard for the understanding of scientific theories.

In this paper, we do not want to single out one approach to scientific explanation as uniquely responsible for the production of scientific understanding. Rather, we want to propose, that the theories of scientific explanation together enable ‚understanding‘. Thereby neither the causal nor the unificationist approach can claim any privileged status, but they will both appear as pillars (bundle conditions) which can enable a scientist to make predictions and thereby understand a phenomenon.

We emphasize that the final notion of understanding presented here will in the end not depend on historical context or on the individual.

## A Bundle Theory of Scientific Understanding

We here assume that all scientific theories that we shall furthermore discuss concerning understanding shall already comply with all logical, empirical and methodological conditions of the scientific conduct. In this sense, a fortune teller, who supposes to ‚understand‘ future developments, will not ‚understand‘ these phenomena by our definition, if he doesn't base his predictions on a scientific theory.

In order to approach our definition of scientific understanding, we refer to an idea of Paul Dirac's, as quoted by Richard Feynman in his ‚lectures‘ (Feynman, 1964, 2-1):

„I understand what an equation means if I have a way of figuring out the characteristics of its solution without actually solving it.“

Motivated by this idea, we suggest the following *necessary and sufficient* criterion for scientific understanding:

*K: A theory is understandable for a scientist in a certain context if he or she is able to make an empirically adequate qualitative prediction without performing exact calculations.*

With ‚context‘ we here denote a number of factors characterizing the individual scientist: as for example his individual abilities, his knowledge or personal convictions. His personal convictions finally are composed of a *number of demands* on a scientific explanation, which the individual scientist weights as especially important for his or her personal understanding – i. e., criteria a theory needs to fulfill, in order to enable him or her to ‚understand‘ a theory in the sense of *K*. A similar condition for understanding was formulated by de Regt in ‚Ludwig Boltzmann's Bildtheorie and Scientific Understanding‘ (Regt, 1999). However, there the author does not consider the criterion to be necessary or sufficient for explanation and a differing final conclusion is reached. In addition, the term ‚context‘ is defined differently.

The bundle of all known demands on a scientific theory that scien-

tists have formulated from the past until today include the principles of causality and unification and other yet to be discussed possible aspects of a theory (like the continuity of motion, symmetries etc.). The individual's weighting of any of these otherwise equally weighted criteria is an entirely personal, individual matter – such that in our theory of understanding none of the criteria ever gains an absolute – i.e., non-individual – privilege over any other criterion.

We defined the above criterion *K* to be necessary and sufficient for scientific understanding. Therefore, making an adequate qualitative prediction without performing a mathematical calculation should be a sufficient criterion for understanding, whereas at the same time understanding should be sufficient for an adequate qualitative prediction. The first direction of this implication is fulfilled since ‚not understanding a scientific theory‘ is here defined to be equivalent to not being able to make an adequate prediction using the theory under question. Vice versa, we also define that not being able to make an adequate prediction is identical to not (yet) having understood the theory under question.

It should be added here that we specifically exclude the cases in which a scientist has either just performed an experiment and remembers the result or where he or she is referring to an experiment with known and all times equivalent result. In this case, the scientists are referring to a fact rather than deducing a result from a theory as they must in order to fulfill *K*. We also want to point out that a scientist who is unable to perform mathematical calculations can still have understood a theory if he or she is able to make adequate qualitative predictions with it. And finally, a scientist who is able to perform exact calculations and who can thereby predict phenomena quantitatively must by our definition not necessarily have understood the theory. This shows that the intuitive notion often used by physicists that a theory is ‚understood‘ if exact predictions can be mathematically calculated does not coincide with the understanding defined in this paper.

Let us inquire once again whether *K* is sensibly defined and if there is no case in which a scientist could have understood a theory even if he were unable to make qualitative predictions. Let's therefore consider the example of a physicist who has been working with problems of quantum mechanics for many years. Having constructed an entirely new experiment, he is asked to make a prediction for the outcome. If he cannot use quantum theory to make a qualitative and empirically adequate predic-

tion, we here claim that this implies his lack of understanding. However, we want to mention that in the daily routine of science, scientists sometimes say that they ‚understand parts of a theory‘, if they can always predict phenomena correctly under certain limited circumstances. This objection points to the fact that the definition of a ‚scientific theory‘ is not always unambiguously definable in the sciences. Parts of a larger theory can themselves form a theory, which in application to a small number of natural phenomena allows for qualitative predictions. An example is electrostatics, which forms a closed sub-theory of electrodynamics – made up of the static phenomena of the larger theory. If an expert of electrostatics were not able to predict the consequences of an oscillating dipole, he could still be said to have understood electrostatics by the criterion *K*, if within this realm of phenomena his predictions were always adequate. However, he would not be said to have understood the larger theory of electrodynamics. This objection, however, does not endanger the adequateness of the criterion *K* with respect to the understanding of a physical theory: it only alludes to the necessity to more sharply define the range of the theory under consideration.

Let us now take a closer look at the previously mentioned *bundle of criteria* for scientific understanding – i.e., the individual demands a scientist has to a theory to enable him or her to make predictions and thereby fulfill *K*. In order to motivate this approach, let’s take a look at the often quoted famous example of the so called Boltzmann gas.

From thermodynamics it is known that there is a well-defined connection between pressure and temperature of an ideal gas (of which the constituents are point-like and do not interact). To show whether one ‚understands‘ the theory of ideal gases, one would e.g. need to derive the connection between temperature and pressure without performing an explicit calculation, thus reaching an adequate qualitative prediction.

Following Boltzmann’s interpretation, one can imagine the gas container as an ensemble of freely moving non-interacting atoms. This atomic picture allows for further conclusions: we can picture the atoms as colliding with the container walls and thereby transferring momentum. Knowing that the average kinetic energy of an ensemble of atoms corresponds to the temperature of the substance, we can infer that increasing the temperature will be identical to an increase in the average velocity of the atoms – and thereby in their average absolute momentum and the momentum transferred to the container walls. Without explic-



it calculation, we have derived a qualitative prediction: namely, that a higher net force is transferred to the container walls as the temperature increases, which implies an increase in pressure.

Boltzmann's interpretation of gases as composed of atoms lead to a qualitative and adequate description of the behavior of the gas in the sense of our criterion *K*. We now can identify various elements of the theory, which enabled the qualitative prediction: Namely assuming *individual agents*, formulating a *visualization* for the problem, as well as the usage of the principle of *causality* (atoms cause momentum transfer) and the principle of *unification* (picturing atoms like classical idealized point-like particles). None of these conditions by itself can be called sufficient or necessary for the prediction.

Assuming the equivalent weight of all bundle conditions, we thus here propose that *scientific understanding can be expressed via a bundle definition*: where we find that the more criteria are fulfilled by a theory, the higher the probability that this theory can be used by a scientist for qualitatively adequate predictions, such that *K* is fulfilled.

The reason for the assumption that a higher number of criteria enables more scientists to make predictions is the following: Let  $T_1$  be a causal theory. Our assumption is that the causal structure of the theory will enable some scientists (let us call them  $s_{11}$ ,  $s_{12}$ ,  $s_{13}$  ...) to make qualitative predictions using this theory, whereas others might not be able to use the causal structure for adequate predictions and thus will not ,understand'  $T_1$  by the criterion *K*. The reason for being able to or not being able to use the causal structure for predictions could be based in the historical and individual context of the scientist. Let us furthermore consider a second theory  $T_2$ . This theory is not causal but unifying, in the sense that it unifies to formerly accepted and ,understood' scientific theories. We claim that some scientists (let us call them  $s_{21}$ ,  $s_{22}$ ,  $s_{23}$  ...) will be able to use this theory for qualitative predictions and thereby ,understand' the theory, whereas other scientists won't be able to make predictions with it, which could be explained by different social, cultural or historical backgrounds. We finally claim that a theory *which is both causal and unifying* will allow for a larger net number of scientists to use the theory: those who by context are more used to thinking in causal structures *and* those who need unification to make predictions.

In this sense, we will call a theory ,more understandable' than another

er, when it fulfills more criteria than the competing theories. We emphasize that it is not important that all possible criteria are fulfilled for a theory to be called ‚understandable‘ or for an individual scientist to ‚understand‘ it.

Whereas the above presented causal or unificationist approaches to scientific explanation and understanding implied that their approach uniquely yielded understanding, we here argue that causality as well as unification can be helpful to enable a scientist to fulfill *K*, yet by themselves we claim them neither to be necessary nor sufficient for scientific understanding in the sense of *K*. In the following, we want to elucidate and discuss those bundle criteria we identified as central in the natural sciences and especially in physics. The roles of causality and unification in theories of scientific explanation and understanding have already been discussed. In the following, we want to present and discuss a list of further demands to physical theories, which is supposed to be as complete as possible. However, this list will not claim completeness, yet it should contain all relevant principles used in modern physical theories.

## Visualizability and Individuals

Many physicists hold the possibility of a *visualization* to be an essential condition for understanding scientific theories. Be it the field lines of electrostatics or the Feynman graphs of Quantum Electrodynamics (QED), visualization is widely considered to be a tool to support intuition and thus qualitative predictions. In the case of the Feynman graphs, in which incoming particles interact and fly apart, the impression of a spatio-temporal interaction is evoked, whereas in fact the graphs represent mathematical terms in an integral of which the result is a probability amplitude in a scattering process. Even so, the graphs still advance intuition by proposing an analogy to classical interactions of point-like elements on continuous space-time paths. This then allows for further creative thought processes, as to which additional terms should contribute to the scattering amplitude. Still, some physicists, like Murray Gell-Mann, consider the possibility of understanding nature through Feynman graphs to be an illusion:

„In QED, as in other quantum field theories, we can use the little pictures invented by my colleague Richard Feynman, which are supposed to give the illusion of understanding what is going on in quantum field theory.“ (Gell-Mann, 1980, 169–198)

We do not share this point of view and instead claim that Feynman graphs in fact achieve more than just an illusion of understanding. As an element of the bundle of conditions, visualizations of any kind can enable qualitative predictions and thus produce actual understanding in the sense of the criterion *K*. We find that the discomfort expressed in Murray Gell-Mann's words does not stem from an actual lack of understanding a theory, but rather marks the scientist's search for a ‚better‘ theory, which we identify to be a theory which fulfills more bundle criteria.

Visualizations try to approach abstract theories through classical systems, where we here mean by classical system any entity which is *individualized* in space-time and whose properties can be defined independently of the environment. In the case of the Boltzmann gas, the individuals were individual point particles by which the atoms of the gas were approximated. In the case of the Feynman graphs, the individual agents again are single point-particles. However, a model does not have to implement point particles in order for us to speak of individuals – we here use the intuitive notion of ‚individuals‘ in the sense of visualizable entities, which could also be water surfaces or complexes like larger solid bodies. We could equally have called this bundle criterium that of *separability* – which is the physical notion of system states which possess properties and property values independently of their environment.

However, neither visualization nor the concept of individual agents can claim to have a prominent status amongst the bundle criteria. Understanding can be possible without visualization. As an example we can take a mathematical function. A well-trained physicist might know where the extrema of the function will lie without performing explicit calculations. This recognition can happen without actual visualization, just caused by experience and general knowledge of mathematical rules. The same holds for solving differential equations.

## Continuity and Analogy

Another criterion that we identified to be part of the bundle of criteria which enable an individual scientist to fulfill *K* is the *principle of continuous motion*. When a classical object propagates through space-time, it will not ‚jump‘, i. e., suddenly disappear from one space time point only to suddenly reappear in another point without having crossed the space in between. Continuous motion thus means that an object’s Minkowski world-line has no holes.

Again, one finds that the continuous description can enable a scientist to make qualitatively adequate predictions – if he or she knows the surrounding forces on an object and its initial conditions, the scientist will be able to predict a continuous path. Yet, the energy levels of an atom in Bohr’s model prove to be a counter example for any privileged status of a continuous description: we are in fact able to make qualitative predictions using Bohr’s model, even though the paths of the electrons during transitions cannot be pictured as classically continuous. On the contrary, standard interpretations assume that the electrons jump between energy levels without ever existing ‚between‘ the levels at any point in time. Therefore continuity proves to be helpful, yet neither necessary nor sufficient for a prediction.

Another bundle criterion is the possibility of finding *analogies*. Many physicists claim that analogies between physical theories are one of the main instruments to reach understanding:

„Analogy is perhaps the physicist’s most powerful conceptual tool for understanding new phenomena or opening new areas of investigation.“ (Povh, 1962, 117)

A famous example for a analogy which was successful in producing understanding is again Bohr’s model of the atom. This model was created as Niels Bohr and Ernest Rutherford compared the energy levels of electrons around an atomic nucleus to the paths of the planets around the sun. Even if this model was later substituted with the model of Schrödinger’s atomic orbitals, it still led to qualitatively adequate predictions. Another successful example of an analogy was relating the energy levels of the positronium – a bound state of an electron and its anti-matter counter part (the positron) – to the states of the Hydrogen atom, which were already well known. This analogy also enabled quali-

tative predictions and was thus successful in producing understanding. Analogies can be unifying, but they do not need to be. In our above example, the analogy between the atom and the planets did not lead to a unification of atomic theory and gravity; this is why the bundle criterion of analogy was discussed separately from the criterion of unification.

## Locality and symmetries

Salmon's definition of causality, which we find to be closest to the possibly more intuitive and less sharply defined terminology many physicists use when speaking of causality, suggests that *locality* should also be amongst the criteria of the bundle: The transfer of conserved quantities draws upon the principle of direct contact. In fact, locality is a principle held to be sacrosanct by most physicists, so much that a theory violating this principle is often neglected by the physics community. Locality enables predictions, as we know from every day experience that nothing happens without a cause that can in the end be traced back to a local contact. As quantum physics seems to prove, however, locality is again neither necessary nor sufficient for qualitative predictions and thus for understanding. Despite of a lack of locality in most quantum mechanical interpretations, qualitative predictions are still possible, as for example in the case of the EPR-experiment, which in some quantum mechanical interpretations is regarded as a non-local process.

Another principle many physicists hold to be sacrosanct is the principle of energy conservation. We want to subsume the latter under the bundle condition of *symmetry* since, as Emmy Noether has shown, every conservation law in physics can be traced back to a continuous symmetry and vice versa. In fact, every continuous symmetry implies a conservation law, as proven by Emmy Noether (Noether, 1981, 235–257). Continuous symmetries denote such symmetries like translational or rotational, yet no such symmetries like mirror symmetry. From time invariance (a temporal symmetry by which experimental outcomes should not change when performed at different times) one can derive energy conservation – from translational invariance (spatial symmetry), momentum is conserved and so forth.

Symmetries imply at least two things for a physicist: on the one hand mathematical simplicity and on the other hand the manifestation of a

‘given fact’ which doesn’t call for any further explanation – different from non-symmetric theories in which the reason for the breaking of the symmetry needs to be formulated. Instead of asking why a certain symmetry exists, physicists prefer to ask why there should be no such symmetry.

At this point we feel we ought to mention that we so far did not require all bundle criteria to be independent from one another: thus calling each conserved quantity a criterion by itself would have been possible, since each will enable certain qualitatively adequate predictions and thereby understanding. We have decided to restrict ourselves to solely calling ‘symmetry’ a bundle condition, but we feel that the discussion of the independence of bundle criteria could be an interesting aspect of a further analysis.

To show how symmetries and thus conservation laws allow for qualitative predictions, we consider two ice skaters facing each other. If they hold hands and then push hard to move apart, their initial momentum (which was zero) implies that under the assumption of equal masses they will move apart with identical and opposite velocities (which at each moment will add up to zero and thus conserve momentum). Another example draws upon rotational symmetry: physicists assume that the solutions of a rotationally symmetrical problem will themselves again be rotationally symmetric. This will yield a qualitative prediction for the shape of stars: gravity only depends on distance and not on any angles, therefore matter clouds will collapse into maximally symmetrical, i. e., spherical objects.

However, some processes of nature are asymmetric – like the incline in entropy in a closed system. Assuming temporal symmetry would then lead to a non-adequate prediction. Symmetry is also violated in quantum mechanics, when an elementary particle borrows energy from the quantum vacuum to tunnel through a potential barrier and thereby violates the principle of energy conservation. The energy it borrowed has to be returned within a small time frame, in order to fulfill Heisenberg’s uncertainty relation between time and energy. We see that symmetries again have predictive powers but are not necessary nor sufficient for understanding.

## Lorentz Invariance and Simplicity

*Lorentz Invariance*, named after Hendrik Lorentz, displays the relativistic principle by which all reference frames are assumed to be equal. This principle implies that all physical laws must be independent of reference frames, and it leads to relativistic effects like length contraction and time dilation, which enable scientists to make qualitative predictions. Knowing that time moves more slowly in systems which are moving with respect to a resting frame, scientists can predict that relativistic myon particles approaching Earth with a velocity close to the speed of light from outer space actually have a chance to hit the Earth's surface, whereas their life span in a rest frame would have them decay long before their arrival.

To give up Lorentz Invariance is held as an impossibility by many physicists, mainly because it would contradict current observations. However, many collapse-interpretations of quantum mechanics, which assume the collapse of the wave-function to be an actual 'physical process', are manifestly non-Lorentz-invariant, since an instantaneous collapse (as a physical process) happens faster than with the speed of light. However, in his book *Quantum Non-Locality And Relativity*, Tim Maudlin finds a way to formulate an instantaneous physical collapse process in a Lorentz-invariant way. However, as we argue in a following chapter, in order to do so he will violate the principle of *simplicity* – a principle by which a theory is supposed to make as little assumptions as possible and should possess the least amount of parameters, which are not deducible from theory.

Yet theories which cannot be called simple can still allow for predictions. In the later to be discussed quantum theory by David Bohm, the spatial coordinate of a system is well-defined at all times. This needs to be postulated to be true, which makes Bohmian Theory more complicated than the standard interpretation of quantum mechanics, the Kopenhagen interpretation, which does not draw upon this postulate and thus displays a smaller amount of extra assumptions. Still, Bohmian mechanics allows for qualitative predictions and we will later apply the bundle definition of scientific understanding to show that it actually produces more understanding than the Kopenhagen interpretation by the definition of understanding presented here.

## Understanding the EPR-Experiment

Let us summarize what we have achieved so far. We have introduced a sufficient and necessary criterion  $K$  and a *bundle of criteria* which enable scientists to fulfill  $K$ . We further argued that none of these criteria can by themselves be called necessary or sufficient for understanding, nor can any of them claim a privileged status. This definition of understanding is subjective in so far as the context of the scientist is responsible for his or her individual weighing of the criteria. However, the absoluteness of our evaluation of the understandability of a theory lies in the insight that the more criteria are fulfilled by a theory, the more scientists on average will be able to use this theory for qualitatively adequate predictions and thereby fulfill  $K$ .

The question remains, why many physicists, like Albert Einstein, Nathan Rosen and Boris Podolski felt that there was something missing in quantum theory, so that physicists still did not actually ‚understand‘ it – whereas the theory enabled scientists to make qualitative predictions. Using our definition of understanding, we conclude that this feeling does not arise from a lack of understanding, but is due to a misunderstanding.

The fact that Einstein, Rosen and Podolski felt the scientific community lacked understanding of quantum theory does therefore not actually lie in a lack of understandability of the Kopenhagen interpretation, which was known at the time, or of quantum theory – but much more in the conviction that a ‚better‘ theory should exist, one which recognized more conditions which these physicists considered to be important to make a theory understandable! Thus, a non causal or indeterministic theory could have appeared dissatisfactory in case one demanded a theory to be causal or continuous. Such demands on a theory would then manifest as questions: to the inner mechanisms of quantum processes, to the causes and the spatio-temporal paths or the connection to other theories. The basic motivation for these questions is to be seen in the search for a coherent world picture.

## Comparing three interpretations of quantum mechanics

We now apply the here sketched definition of scientific understanding to three interpretations of the EPR-experiment of quantum mechan-



ics: the Copenhagen Interpretation (Bohr, 1934), the interpretation by Tim Maudlin (Maudlin, 1994) and finally the interpretation by David Bohm (Bohm, 1952, 166–179). We are not considering the Many Worlds Interpretation as formulated by Hugh Everett and Bryce deWitt (Everett, 1957, 454–462), since we claim that it suffers from a fatal flaw, which needs to be clarified first before further considering this interpretation. The flaw lies at the heart of the interpretation: When a measurement is performed, all of the varying many worlds interpretations assume that in some way or another the world we know splits up into a number of worlds, such that in each of the worlds one single measurement result is found and all possible measurements are realized in at least one world. Let us now assume, that we have two possible outcomes for a certain measurement. Then the many worlds interpretation might assume that the world splits into two branches. However, these branches are not weighted, so that each of the measurement results should appear to an observer – who happens to be in one branch or the other with a probability of  $\frac{1}{2}$  each – with a probability of 50 %. Here we encounter a fatal problem. In this model it becomes impossible to explain measurements, in which one result happens with a different probability from the other result, i. e., a case in which eigenvalues (measurement results) are not equally weighted. Let us assume, one eigenvalue appears with  $\frac{1}{3}$  probability. One could then try to resolve the problem by assuming that the world splits in three branches, two realizing one eigenvalue and the third the other. However, the question remains whether the universe is this complex. It is also difficult to explain, how the world needs to split up in order to account for irrational probabilities (McCall, 1994). These yet to be unambiguously clarified questions are the reason as to why we do not further consider the many worlds approach in this paper.

## A Physical Realization of the EPR-Experiment

In the summer of 1935, A. Einstein, N. Rosen and B. Podolski published a paper carrying the title ‚Can Quantum Mechanical Description of Reality be Considered Complete?’ (Einstein, Rosen, Podolski, 1935, 777–780). In their publication, the authors presented a thought experiment which was supposed to show that quantum theory could not yet

count as a complete theory. The flaw they thought to have detected lay in parts of a spatially separated experiment which seemed to instantaneously communicate with one another – an effect, which is sometimes called ‚non local‘ or ‚action at a distance‘. For the sake of completeness of our argument, we here want to quickly sketch one of the modern realizations of an EPR-experiment.

Physically, the experiment can for example be realized by having a pion decay, which is a particle composed of a quark and an anti-quark. When it decays, it can emit two electrons. Because of conservation of momentum, where in the rest system of the pion total momentum is zero, one electron will fly to the right and the other to the left, the sum of their momenta being zero. The entanglement now manifests itself in the fact that the spins of these electrons will always turn out to be opposite of one another when measured. This is due to spin conservation with the pion having zero spin. So if one of the electrons' spins is measured in a certain direction (here: in x-direction) and is found to point up in x-direction, the other electron's spin will be found to point down in x-direction.

Here a problem arises. A variant of Heisenberg's uncertainty principle states that different spin components of a system (which can point in x-, y- or z-direction) can not be simultaneously well-defined, i. e., an electron cannot point up in x-direction and simultaneously have any well-defined spin either in y- or z-direction, etc. In this, Einstein, Podolski and Rosen formulated the core problem of the EPR experiment. If quantum theory was right, then all components of either electrons truly were undefined until a measurement (say on the left electron, measuring spin in x-direction) had taken place, at which point the spin of both particles changed from being undefined in all directions to being well-defined in x-direction. However, an observer could equally have conducted a measurement of the spin in y- or z-direction. In every case this implies that the remote electron would be instantaneously well-defined along the same axis of measurement.

Since the authors excluded the possibility action at a distance from the measured electron to its remote partner, the information about a measurement having taken place and the axis chosen needed to be either sent with a faster-than-light signal – a possibility the authors excluded – or the spin components of the second electron had to have been well-defined from the start of the electron's flight.

If the latter were true, quantum theory would be an incomplete theory, since it does not describe particles having a well-defined spin outside of a measurement context. Furthermore, an electron for which all spin-components are simultaneously well-defined violates Heisenberg's uncertainty relation for spin components.

## The Kopenhagen Interpretation of the EPR-experiment

Different interpretations of quantum mechanics tried to deal with the riddles of the EPR-experiment and its inexplicable connection between remote particles, i.e. the entanglement structure. Of these interpretations, the Kopenhagen interpretation was the first formally recognized approach to understand the physics beyond the mathematical equations. Niels Bohr, Werner Heisenberg, Max Born and Erwin Schrödinger are some of the physicists who contributed to this interpretation.

A central building block of the Kopenhagen interpretation is Bohr's idea that actual experiments need to be conducted in order for physical properties to be at all definable. Only through the conduction of actual experiments, the conditions would be set to assign certain values to physical properties. Concerning the EPR experiment, individual states for particles could therefore not be assumed without performing a measurement – and the particles could therefore not be thought of as separate entities having influences on one another.

Having said this, we see that the Kopenhagen Interpretation violates separability. The collapse of the wave function, which explains the transition from a superposition to a well-defined eigenstate of a physical observable – which many physicists to this day might consider to correspond to an actual physical process – in the frame of the Kopenhagen interpretation is no physical concept at all. This interpretation denies any visual or physical reality for the collapse. The only possible statements concern the prediction of outcomes – of measurement results, which the formalism is able to produce with high precision. This way, the Kopenhagen interpretation conserves locality: no non-local physical process takes place and all that is spoken of are measurement results – which are determined locally. Thus, whereas the bundle criterion of separability is violated, that of locality is conserved.

## Tim Maudlin's Interpretation of the EPR experiment

The American philosopher Tim Maudlin did not want to accept the Copenhagen interpretation, but was rather convinced that the collapse of the wave function had to correspond to an actual physical process. This belief can be motivated by considering that certain quantum mechanical experiments prove the existence of superposition states, whereas all experiment only ever register eigenstates of systems – a difference in physical realizations of states, which some feel must be gapped by a physical process as well.

As Tim Maudlin argues in his book ‚Quantum Non-Locality and Relativity‘, this physical – and as he argues: causal process of the collapse can not be described to be a local process, i.e it cannot happen with a velocity smaller than that of light. On the contrary, the connection between the EPR-particles must be instantaneous in order to explain the correlations, which can exist even over space-like distances.

Since the concept of instantaneity in relativistic physics has become relative to the reference frame, Maudlin calls for a radically new concept of causality, in which the collapse does not happen instantaneously in one certain frame of reference. Instead, he draws on the Hyperplane-Theory, which was first formulated by Gordon Fleming (Fleming, 1986, 574–575). This theory assumes that the collapse is an instantaneous process in all possible hyperplanes which can be drawn through an event –, an idea, which allows for a Lorentz invariant description of the collapse of the wave-function and thus for a connection to Special Relativity Theory.

Despite of the fact that the Maudlin encounters problems within this interpretation – with the order of events and thus truth values about object properties depending on the hyperplane under consideration – he still achieves a Lorentz invariant and causal description of the EPR-experiment. On the other hand, locality and separability are violated: locality is violated since all collapses in all hyperplanes are physical processes which happen instantaneously (with a velocity larger than the speed of light), and separability does not hold since properties of particles depend on the space-like, flat hyperplane (whether a particle is ‚already‘ measured and well-defined in the measured property or whether it is ‚still‘ in a superposition and thus not well-defined in a certain property, is something different observers will not agree on).

## The Interpretation of David Bohm

The theory by David Bohm (Bohm, 1952, 166–179) is a hidden variable theory. In it, the position observable is well defined at all times – it is the ‚hidden‘ variable of the interpretation, since it is not manifestly described in the quantum formalism. However, the interpretation is not excluded by Bell’s inequality (Bell, 1987, 160), since the latter only excludes the possibility of local hidden variable theories – whereas Bohm’s theory is manifestly non-local. Despite of insistent claims in the scientific community that Bohm’s theory cannot be formulated relativistically, relativistic generalizations of it have been presented and are being developed, for example in *The Quantum Theory Of Motion* by Peter Holland (Holland, 1993).

In Bohm’s theory the wave function is interpreted as a physical field, which ‚guides‘ a point particle on a continuous quantum path. Thereby particles move in a fully deterministic way, following the force the quantum field exerts on them – which can be determined through the deterministic Schrödinger equation – its path then being uniquely fixed by initial conditions and deterministic equations.

The Bohmian Theory can be criticised for not being maximally simple: it needs to assume the position observable to be well-defined at all times – and thereby to be singled out. Also, locality is violated: the non-locality in the EPR experiment enters through the guiding equations for the particles, which explicitly depend on the position of the – possibly even space-like separated – entangled partner particles. However, Sheldon Goldstein remarks that the non-locality in Bohm’s theory is not worse than that of other quantum mechanical interpretations:

„It should be emphasized that the nonlocality of Bohmian mechanics derives solely from the nonlocality built into the structure of standard quantum theory, as provided by a wave function on configuration space, an abstraction which, roughly speaking, combines or binds distant particles into a single irreducible reality.“ (Goldstein, 2001)

## Evaluation of the interpretations

To evaluate how much understanding a quantum mechanical interpretation produces with respect to the EPR-experiment, we apply our above

offered definition for scientific understanding. Thus we need to identify the number of bundle criteria which are fulfilled by each interpretation.

We find, that the Kopenhagen interpretation fulfills two bundle criteria: it conserves locality (1) and can be called simple (2) – since it only assumes a collapse of the wave function, which is not part of the mathematical formalism, but otherwise uniquely sticks to the mathematical frame of quantum theory.

We furthermore find that Tim Maudlin's interpretation of the EPR-experiment fulfills three bundle criteria: it allows for a causal description (1), since the collapse is interpreted as a physical process, it offers a Lorentz invariant formulation (2) of the EPR-experiment and last but not least, the interpretation is unifying (3) since Special Relativity and quantum mechanics are united via Fleming's Hyperplane Theory.

Finally, we see that Bohm's interpretation fulfills most bundle criteria and thus will be evaluated as best suited for understanding quantum mechanics. It uses the concept of individual agents (1) – all times well defined point particles –, which move on continuous paths (2), it enables a clear visualization (3) of the particles moving on the wave function like dust grains on a water surface, the description is causal (4) and unifying (5) since classical mechanics and quantum mechanics are united mathematically (the equation of motion for a particle is a Hamilton Jacobi (classical) equation with an additional quantum term, which vanishes for large quantum numbers). Additionally, an analogy (6) is made possible, between classical probabilities and quantum probabilities, with many point particles creating an ensemble of actual paths. And finally, the Bohmian approach can still be called simple (7), since it does not need to postulate a collapse (no collapse ever happens in Bohmian mechanics), and only assumes the position observable to be special, drawing equal in extra assumptions with the Kopenhagen interpretation. But Bohm's theory is thereby not drawing equal with Maudlin's interpretation in simplicity. On the contrary, we claim that it is actually superior because of the multitude of hyperplanes and the alternative concept of 'properties' which Maudlin's interpretation entails.

## Summary

Even though quantum mechanics is one of the most successful physical theories, it raises puzzling questions which have produced many widely varying interpretations. These interpretations imply the same observations and thus are empirically equivalent and often simply chosen as matter of taste. In this essay, we formulated a possible answer to Richard Feynman's statement: „I think I can safely say that nobody understands quantum mechanics.“ We identified this as a misunderstanding or a misuse of the meaning of ‚scientific understanding‘. We find that quantum mechanics can be understood via a *necessary and sufficient criterion K*, which is fulfilled if an individual scientist can make a qualitatively adequate prediction without performing mathematical calculations. We furthermore formulated a *bundle of conditions*, of which each element is neither sufficient nor necessary for scientific understanding, but by which a theory can be called ‚better suited for understanding‘ than another if it fulfills more bundle criteria than its competitor. Amongst the bundle criteria we identified and discussed the following principles of physical theories: *Causality, Unification, Continuity, Individuals, Visualization, Simplicity, Locality, Analogies, Symmetries and Lorentz-Invariance*.

We discussed three interpretations of quantum mechanics with which qualitative predictions for the EPR experiment are possible. Of these, the Kopenhagen Interpretation proved to fulfill two bundle criteria, Tim Maudlin's interpretation fulfilled four criteria, whereas Bohm's interpretation fulfilled eight criteria and thereby was identified as the interpretation which is best suited to understand quantum mechanics, since it will on average enable most physicists to make qualitative predictions.

The seeming contradiction to identifying Bohm's theory as best suited for understanding whereas physicists usually prefer the Kopenhagen interpretation can be explained by the realization that in the scientist's daily routine being able to calculate results has a higher priority than making qualitative predictions – whereby theories which stick to a maximally simple mathematical frame gain popularity.

Our definition of scientific understanding has proven to be a means to differentiate between theories which imply the same observations and which otherwise might therefore be considered to be equivalent.

We find this to be an answer to the critique formulated against empirically equivalent theories as having a negative implication for the pursuit of science: we claim that empirically equivalent theories are no argument against the efficacy of the pursuit of science. Rather, we claim that formulating new scientific theories, even if empirically equivalent to existing theories, is an actual progress, whenever the theories are also evaluated with respect to the understanding they produce.

Thus, finding more empirically equivalent theories eventually leads to a continuous improvement of man's understanding of the world. The search for new physical theories is therefore not only justified through empirical superiority of a new theory, but also through the consequent improvement of our understanding of the phenomena around us.

### *Literature*

- Bell, John, 1987: *Speakable and Unspeakable in Quantum Mechanics*. Cambridge: Cambridge University Press.
- Bohr, Niels, 1934: *Atomic Theory and the Description of Nature*. Cambridge: Cambridge University Press.
- Bohm, David, 1952: A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables. In: *Physical Review* 85, p. 166–179.
- DeWitt, Bryce, 1973: *The Many-Worlds Interpretation of Quantum Mechanics*. Princeton: Princeton University Press.
- Einstein, Albert; Podolski, Boris; Rosen, Nathan, 1935: Can quantum-mechanical description of physical reality be considered complete? In: *Physical Review* 47, p. 777–780.
- Everett, Hugh, 1957: Relative State Formulation of Quantum Mechanics. In: *Review of Modern Physics*, 29, p. 454–462.
- Feynman, Richard, 1964: *The Feynman Lectures on Physics, Volume II*. Massachusetts: Addison-Wesley Publishing Company.
- Fleming, Gordon, 1986: Towards a Lorentz Invariant Quantum Theory of Measurement. In: Greenberger, Daniel (ed.): *New Techniques and Ideas in Quantum Measurement Theory*. New York: New York Academy of Sciences, p. 574–575.
- Friedman, Michael, 1974: Explanation and Scientific Understanding. In: *The Journal of Philosophy* 71, p. 5–19.
- Gell-Mann, Murray, 1980: Questions for the Future. 8th Wolfson Lec-



- ture. In: Mulvey, John (ed.): *The Nature of Matter*. Oxford: Clarendon Press, p. 169–198.
- Goldstein, Sheldon, 2001: *Bohmian Mechanics*. Stanford: Stanford Library of Philosophy.
- Hempel, Carl G., 1965: *Aspects of Scientific Explanation and other Essays in The Philosophy of Science*. New York: New York Free Press.
- Hitchcock, Christopher, 1995: Discussion: Salmon on Explanatory Relevance. In: *Philosophy of Science* 62, p. 304–320.
- Holland, Peter, 1993: *The Quantum Theory of Motion*. Cambridge: Cambridge University Press.
- Kitcher, Philip, 1989: Explanatory Unification and the Causal Structure of the World. In: Kitcher, Philip; Salmon, Wesley (ed.): *Minnesota Studies in the Philosophy of Science, Volume XIII*. Minneapolis: University of Minnesota Press, p. 410–505.
- Kitcher, Philip, 1981: Explanatory unification. In: *Philosophy of Science* 48, p. 507–531.
- Klärner, Holger, 2003: *Der Schluss auf die beste Erklärung*. Berlin / New York: Gruyter.
- Maudlin, Tim, 1994: *Quantum Non-Locality and Relativity*. Oxford: Blackwell Publishing.
- Noether, Emmy, 1918: Invariante Variationsprobleme. In: *Nachrichten von der königlichen Gesellschaft der Wissenschaften Zu Göttingen*, p. 235–257.
- McCall, Storrs, 1994: *A Model of the Universe*. Oxford: Oxford University Press.
- Regt, de Henks; Dieks, Dennis, 2005: A Contextual Approach to Scientific Understanding. In: *Synthese* 144, p. 137–170.
- Salmon, Wesley, 1984: *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
- Salmon, Wesley, 1997: Causality and Explanation: A Reply to Two Critiques. In: *Philosophy of Science* 64, p. 461–477.
- Schurz, Gerhard; Lambert, Karel, 1994: Outline of a theory of scientific understanding. In: *Synthese* 101, p. 109.
- Spillner, Vera, 2011: Verstehen in der Quantenphysik. In: *Online-Publikation an deutschen Hochschulen*. Bonn: Univ. Diss. URN: urn:nbn:de:hbz:5-24475, p. 42.
- Povh, Bogdan, 2004: *Particles and nuclei: an introduction to the physical concepts*. Berlin / Heidelberg: Springer-Verlag GmbH.

Woodward, John, 2003: *Scientific Explanation*. Stanford: Stanford Library Of Philosophy, <http://plato.stanford.edu/entries/scientificexplanation/>

## Verzeichnis der Autoren

Prof. Dr. Monika Betzler  
Institut für Philosophie  
Universität Bern  
Länggassstrasse 49a  
CH-3000 Bern 9

Dr. Boris Hennig  
Hospitalstr. 61  
22767 Hamburg

Dr. Dr. Norman Sieroka  
Oberassistent  
Professur für Philosophie  
ETH Zürich  
RAC F18  
CH-8092 Zürich

Dr. Vera Spillner  
Institut für Philosophie  
Universität Bonn  
Am Hof 1  
53113 Bonn

## PHILOSOPHIA NATURALIS

Eingereichte Beiträge dürfen weder schon veröffentlicht worden sein noch gleichzeitig einem anderen Organ angeboten werden. Mit der Annahme des Manuskriptes zur Veröffentlichung in der *Philosophia naturalis* räumt der Autor dem Verlag Vittorio Klostermann das zeitlich und inhaltlich unbeschränkte Nutzungsrecht im Rahmen der Print- und Online-Ausgabe der Zeitschrift ein. Dieses beinhaltet das Recht der Nutzung und Wiedergabe im In- und Ausland in körperlicher und unkörperlicher Form sowie die Befugnis, Dritten die Wiedergabe und Speicherung des Werkes zu gestatten. Der Autor behält jedoch das Recht, nach Ablauf eines Jahres anderen Verlagen eine einfache Abdruckgenehmigung zu erteilen.

### *Richtlinien zur Manuskriptgestaltung*

Bitte jeden Beitrag mit *Titelblatt* abgeben, das folgende Angaben enthält: Name und Vorname des Autors / der Autorin (mit akad. Titel), Titel des Beitrags, vollständige Adresse (inkl. Telefon-Nummer), nähere Bezeichnung der Arbeitsstätte.

Die *Manuskripte* sollten 3-fach und als WORD-File auf Diskette oder CD eingereicht werden und ein deutsch- und englischsprachiges Abstract enthalten. Das Manuskript sollte einen breiten Rand haben.

Der *Umfang* (einschließlich Anmerkungen und Bibliografie) soll bei den Aufsätzen nicht mehr als 30 maschinengeschriebene Seiten (ca. 2.000 Anschläge, 2-zeilig) betragen.

Für *Abbildungen* im Text bitte die Originalvorlage einreichen. Abbildungen müssen nummeriert und mit Autorennamen versehen sein.

*Zitate* im Text sollten vom Haupttext durch eine Leerzeile abgehoben werden. Nach dem zitierten Text stehen Name des zitierten Verfassers, Erscheinungsjahr und Seitenangaben in Klammern, z. B.: (Elkana 1974, S. 34). Bei mehreren Autoren werden die jeweiligen Namen durch Schrägstriche getrennt, z. B.: Krantz/Luce/Suppes/Tversky 1971, S. 8). Wird auf mehrere Publikationen desselben Autors im selben Erscheinungsjahr verwiesen, so sollen sie nummeriert werden: (Ludwig 1970 a) bzw. (Ludwig 1970 b).

Die *Anmerkungen* sind im Manuskript fortlaufend zu nummerieren; sie stehen am Schluss des Beitrags in numerischer Reihenfolge.

Für das anschließende *Literaturverzeichnis* in alphabetischer und chronologischer Reihenfolge gilt folgendes Muster:

- Elkana, Y., 1974: *The Discovery of the Conservation of Energy*. London: Hutchinson.  
Clausius, R., 1850: Über die bewegende Kraft der Wärme. In: *Annalen der Physik und Chemie*, 79, S. 500–524.  
Klein, M.J., 1978: The Early Papers of J. Willard Gibbs: A Transformation of Thermodynamics. In: E. G. Forbes (Hg.): *Human Implications of Scientific Advance*. Edinburgh: University Press, S. 330–341.

*Korrekturen*: Die Autoren erhalten vom Verlag die Fahnen ihres Beitrags mit der Bitte, die korrigierten Fahnen *innerhalb von zwei Wochen* an den Herausgeber zu schicken. In den Fahnen sollen nur noch Satzfehler berichtet werden.

Nach Erscheinen des Heftes erhalten die Autoren 3 Belegexemplare des jeweiligen Heftes.

*philosophia naturalis*

Located at the crossroads between natural philosophy, the theory and history of science, and the philosophy of technology, JOURNAL FOR THE PHILOSOPHY OF NATURE has represented for many decades – not only in the German speaking countries but internationally – a broad range of topics not addressed by any other periodical.

The journal has a highly interdisciplinary focus. Articles with systematic as well as historical approaches are published in German and English. Their quality is assured by a strict peer review policy.

*philosophia naturalis*

Inhaltlich an der Schnittstelle zwischen Naturphilosophie, Wissenschaftstheorie, Wissenschaftsgeschichte und Technik-Philosophie angesiedelt, vertritt die Zeitschrift

JOURNAL FOR THE PHILOSOPHY OF NATURE seit mehreren Jahrzehnten nicht nur im deutschen Sprachraum, sondern auch im internationalen Vergleich, einen weiten Themenbereich, der von keinem anderen Publikationsorgan vertreten wird. Die Zeitschrift ist ausgesprochen interdisziplinär ausgerichtet. Sie veröffentlicht Aufsätze in deutscher und englischer Sprache, die sowohl systematisch als auch historisch orientiert sind. Deren Qualität wird durch ein besonders strenges Begutachtungsverfahren gesichert.