

Diss. ETH Nr. 24329

**Die Entstehung der Objekte  
Überlegungen zu einer exakten Wissenschaft  
von Bewusstsein**

Abhandlung zur Erlangung des Titels  
DOKTOR DER WISSENSCHAFTEN  
(Dr. sc. ETH)

vorgelegt von  
Dr. sc. Robert Prentner  
MSc & MA, ETH Zürich  
geboren am 26. 07. 1984  
von Österreich

Angenommen auf Antrag von  
Prof. Dr. Michael Hampe  
Prof. Dr. Peter Simons, Prof. Dr. Achim Stephan

2017



# Danksagung

Die vorliegende Arbeit ist als Promotionsschrift an der Professur für Philosophie der ETH Zürich und nicht ohne Hilfe von vielen geschätzten Menschen entstanden.

Ich danke Michael Hampe für die Betreuung der Arbeit sowie meinen Zweitgutachtern, Peter Simons vom Trinity College, Dublin und Achim Stephan vom Institut für Kognitionswissenschaft der Universität Osnabrück, für die Übernahme des Koferats.

Zudem danke ich Harald Atmanspacher, Don Hoffman, Norman Sieroka, Paul Skowski und Daniel Strassberg für Ihre Hilfestellungen und Diskussionen. Des Weiteren gilt mein Dank allen Mitarbeitern der Professur für Philosophie der ETH Zürich und insbesondere Victoria Laszlo für Ihre administrative Arbeit.

Folgende Publikationen sind dieser Arbeit vorausgegangen:

- Chemistry, Context and the Objects of Thought, *Foundations of Chemistry*, 19(1): 29–41, 2017.
- Erweiterter Geist – Erweitertes Bewusstsein? in: J. G. Michel, K. J. Boström und M. Pohl (Hg.), ‘Ist der Geist im Kopf? Die These des erweiterten Geistes in Philosophie und Wissenschaft’, mentis, Münster, 2016.
- A Framework for Critical Materialists, *Mind and Matter*, 12(1): 93–118, 2014.

Ich danke den Herausgebern, Verlagen und Gutachtern für ihre Hilfe und Geduld. Für ihre Unterstützung abseits des universitären Betriebs möchte ich mich bei meiner Frau Caroline, meiner Familie und meinen Freunden herzlich bedanken.

Zollikon am 20. September 2017



## Zusammenfassung

In dieser Arbeit sollen das Verhältnis von Subjekten und Objekten näher untersucht und geeignete Methoden zu dessen formaler Darstellung entwickelt werden.

Den Ausgangspunkt hierzu bildete der zeitgenössische Diskurs in der analytischen Philosophie des Geistes, der sich um den Begriff des Bewusstseins dreht und dessen zugrundeliegenden, aber oft impliziten Annahmen und die sich daraus ergebenden Konsequenzen für eine „Wissenschaft vom Bewusstsein“ näher untersucht wurden. Besonderes Augenmerk kam dabei denjenigen Strategien zu, die auf eine naturwissenschaftliche Beschreibung des Bewusstseins abzielen und bei denen der Begriff der Emergenz oftmals eine zentrale Rolle spielt. Dies wurde an einigen Beispielen aus den Einzelwissenschaften veranschaulicht, aber auch auf systematischer Ebene beleuchtet und führte auf die Ansicht, wonach Emergenzphänomene als strukturelle, muster-relative oder kontextuelle verstanden werden müssen, wie sie typischerweise in Prozessen der Selbstorganisation zu beobachten sind. Diese Lesart unterscheidet sich von einer in der philosophischen Literatur oft üblichen, die sich auf „starke“ oder „schwache“ Ausprägung des Emergenzbegriffs beschränkt.

Vor diesem Hintergrund wurde anhand einer ausführlichen Darstellung, die sich auf die Möglichkeit der Naturalisierung einiger paradigmatischer mentaler Begriffe im Rahmen der biologischen Kybernetik bezieht, und einer anschließenden Kritik an einer funktionalistischen Interpretation für die Notwendigkeit eines Übergangs zu einem zeichentheoretischen Verständnis von Bewusstsein argumentiert. Um dies präziser fassen zu können, schloss sich daran die Untersuchung der Beziehung von Teilen und Systemganzen an, wie sie auf unterschiedliche Weise etwa von Stanislaw Leśniewski, Edmund Husserl oder Nelson Goodman konzipiert wurde und in einer einheitlichen, mathematischen Sprache formuliert und axiomatisiert werden kann. Dabei wurde die Unterscheidung zwischen kompositionalen und zerlegenden Ansätzen eingeführt und ein konkretes Modell eines zerlegenden Ansatzes – „projektive extensionale Mereologie“ – entwickelt, das in Folge auf die Darstellung von Wahrnehmungsphänomenen Anwendung fand:

Empirisch motiviert und angelehnt an die Phänomenologie, wurde bewusste Wahrnehmung als Herausbildung verweisender Objekte auf der Basis einer noch nicht differenzierten Mannigfaltig von Teilobjekten beschrieben, die selbst noch nicht auf bestimmte Gegenstände außerhalb der Erfahrung verweisen.

Bewusstsein, so die daraus abgeleitete These, ist mit der Form derjenigen Prozesse zu identifizieren, die bezeichnende Objekte in der Wahrnehmung erst hervorbringen. Die zuvor entwickelten, formalen Werkzeuge können der näheren Explikation der Struktur solcher Prozesse dienen. Dies bildet den Grundstein einer naturalistisch verstandenen Emergenztheorie, welche im Gegensatz zur Ansicht steht, dass Bewusstsein selbst gegenständlich zu betrachten wäre. Somit unterscheidet sich das hier vorgestellte Bild von den meisten „Bewusstseinstheorien“, die in den Neuro- und Kognitionswissenschaften derzeit verhandelt werden.

Zum Schluss wurde die Ähnlichkeit zu Projekten diskutiert, die sich als Spielarten einer „naturalisierten Phänomenologie“ oder einer „akteursbasierten Kognitionswissenschaft“ in bio-physikalischen Systemen verstehen lassen. Ebenfalls wurden Querverweise auf die kosmologischen Philosophien von William James und Alfred North Whitehead gezogen, wobei ein Naturbegriff vorgestellt wird, der sich durch dessen „Produktivität“ auszeichnet und sich panpsychistisch interpretieren lässt.

# Abstract

In this thesis the relation between subjects and objects shall be examined, and appropriate tools shall be developed to formally represent it.

The work's point of departure was the contemporary discourse in the analytic philosophy of mind that centers around the notion of consciousness. Its foundational but often implicit assumptions and their consequences for a "science of consciousness" were investigated. Those strategies that aim at a scientific description of consciousness thereby received particular attention. Often, the concept of emergence plays a crucial role within these strategies. This was illustrated using examples from the special sciences but was also discussed more systematically, which led to an understanding of emergence as structural, pattern-relative or contextual phenomenon, as it typically occurs in process of self-organization. Such an understanding is opposed to those often found in the philosophical literature which restrict themselves to the notion of either "strong" or "weak" emergence.

Following a detailed discussion of biological cybernetics which tries to render intelligible the naturalization of some paradigmatic mental phenomena and a criticism of functionalist interpretations thereof, a shift from functionalism to a semiotic theory of consciousness was proposed against this background. In order to give a more precise statement of the theory, the relation of parts to wholes which has been investigated by scholars like Stanislaw Leśniewski, Edmund Husserl or Nelson Goodman was investigated and formalized within an axiomatic mathematical system. The distinction between compositional and decompositional approaches was thereby introduced and a concrete model of such a decompositional approach – "projective extensional mereology" – was consequently developed and applied to the phenomena of perception:

Empirically motivated and along the lines of a Phenomenological investigation, conscious perception was described by the formation of signifying objects, based on an initially undifferentiated manifold of partial objects which themselves do not signify concrete things outside experience.

Consciousness, according to the hypothesis informed by this work, is identical to the form of those processes that lead to the emergence of objects in the first place. The formal tools developed in this work could be used to further explicate the structure of such processes. This forms the starting point of a (naturalistic) account of emergence which, however, is opposed to those views that try to reify conscious-

ness. Thus, the picture put forward in this work is substantially different from the “theories of consciousness” which are currently being investigated in the (cognitive) neurosciences.

Finally, the similarity to projects that could be labeled as “naturalized phenomenology” or “agent-based cognitive science” in bio-physical system were discussed. Also, references to the cosmological philosophies of William James and Alfred North Whitehead were given, leading to the proposal of a concept of nature that emphasizes nature’s “productivity” which could be given a panpsychist interpretation.



# Inhaltsverzeichnis

<b>Danksagung</b>	<b>i</b>
<b>Zusammenfassung</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>1 Exposition</b>	<b>1</b>
1.1 Über das Lösen von Rätseln . . . . .	1
1.2 Eine kleine Übersicht über metaphysische Positionen . . . . .	8
1.3 Methodisches . . . . .	21
Anmerkungen zu Abschnitt 1 . . . . .	32
<b>2 Wider die Abbildung</b>	<b>35</b>
2.1 Schwierigkeiten für eine gegenständliche Theorie des Bewusstseins . . . . .	35
2.2 Bilder und Modelle . . . . .	41
2.3 Emergenz . . . . .	49
Anmerkungen zu Abschnitt 2 . . . . .	59
<b>3 Geist und Gehirn</b>	<b>71</b>
3.1 Das Gehirn, ein guter Regulator? . . . . .	71
3.2 Funktionen des Bewusstseins . . . . .	78
3.3 Zeichen und Struktur . . . . .	83
Anmerkungen zu Abschnitt 3 . . . . .	90
<b>4 Der Relationale Aufbau der Gegenstände</b>	<b>97</b>
4.1 Die Theorie vom Ganzem und seinen Teilen . . . . .	97
4.2 Klassische Mereologie . . . . .	102

4.3	Erweiterte Mereologien . . . . .	111
	Anmerkungen zu Abschnitt 4 . . . . .	123
<b>5</b>	<b>Im Prozess des Teilens</b>	<b>127</b>
5.1	Projektive Mereologie . . . . .	127
5.2	Objektivität . . . . .	132
5.3	Emergenz und Topologie . . . . .	139
	Anmerkungen zu Abschnitt 5 . . . . .	144
<b>6</b>	<b>Wahrnehmung im Kontext</b>	<b>147</b>
6.1	Ding und Prozess . . . . .	147
6.2	Aspekte der Wahrnehmung . . . . .	153
6.3	Aktivität und Kontext . . . . .	159
	Anmerkungen zu Abschnitt 6 . . . . .	170
<b>7</b>	<b>Résumé: Bewusstsein und Natur</b>	<b>173</b>
7.1	Eine Emergenztheorie des Bewusstseins . . . . .	173
7.2	Brücken zur Wissenschaft . . . . .	177
7.3	Metaphysik, Natur und Panpsychismus . . . . .	184
	Anmerkungen zu Abschnitt 7 . . . . .	197
	<b>Parergon</b>	<b>201</b>
	<b>Zitatnachweise und Siglen</b>	<b>207</b>
	<b>Literatur</b>	<b>209</b>
	<b>Autorenindex</b>	<b>229</b>

# 1 Exposition

... eine Erfahrungswissenschaft, welche lehrt, wie ein freyes Spiel der Kräfte möglich seye dadurch, daß die Natur neue Verbindungen bewirkt, und bewirkte Verbindungen aufhebt. (F. W. J. Schelling: Ideen zu einer Philosophie der Natur)

## 1.1 Über das Lösen von Rätseln

Manchmal wird gesagt, dass uns „im Bewusstsein die Welt in Erscheinung trete“, oder dass „im Gehirn eine subjektive Realität erschaffen werde“ – als ob es zwei Wirklichkeiten gebe: die eine außerhalb, die andere innerhalb unseres Bewusstseins. Diese Aussagen, wenngleich sie vielleicht plausibel klingen, entsprechen keinen Ausdrücken, wie sie in wissenschaftlichen Beschreibungen üblicherweise vorkommen. Wissenschaftliche Theorien handeln von Phänomenen, von deren Struktur oder von den Prozessen, mithilfe derer sie im Labor oder in Simulationen erzeugt werden. Aber entzieht sich Bewusstsein nicht genau einer jeden objektiven Beschreibungsweise oder geht dieser sogar voraus? Kann dann überhaupt sinnvoll von einer Wissenschaft des Bewusstseins gesprochen werden?

Das Verhältnis von Subjekten und Objekten, ihre gegenseitige Bedingtheit und die Rolle der exakten Wissenschaften bei der Verständlichmachung dieses Phänomens sollen daher in dieser Arbeit näher untersucht werden. Dabei soll dafür argumentiert werden, dass Bewusstsein kein gewöhnliches Objekt der Wissenschaften darstellt, sondern das Medium ist, in dem sich solche Objekte erst herausbilden – oder, wie wir später noch näher spezifizieren werden: die Form desjenigen Prozesses, der Objekte hervorbringt.

Gleichzeitig soll dabei aber nicht behauptet werden, dass es „unbedingte Subjekte“ gäbe, die außerhalb des Naturgeschehens stünden, dieses beobachten und spukhafte Wirkungen darin hervorbringen würden. In der zweiten Hälfte des 20. Jahrhunderts verfestigte sich die Idee, dass mentale Phänomene mit Hilfe von naturwissenschaftlichen Methoden erklärt werden könnten. Insbesondere von der Metapher des Geistes als Computer geleitet, wurden Denken und Wahrnehmen naturwissenschaftlich untersucht. Dem zuvor gegangen war die Ansicht, dass Sprachfähigkeit als natürliche Veranlagung und nicht mehr als außer-biologisches Charakteristikum des Menschen zu betrachten sei, und heute sind es die Neurowissenschaften, die uns versichern, alles Geistige sei auf die Aktivität des Gehirns reduzierbar: „*You're nothing but a pack of neurons*“ (Crick 1994, S. 3).

Wichtige theoretische Grundlagen stammten dabei aus der Nachrichtentechnik, die

den Begriff der Information als quantifizierbare, objektive Größe neu zu fassen versuchte, und der Kybernetik, die Prozesse der Regelung als automatisierbare, mechanische Abläufe zu verstehen begann. Geist wurde nicht mehr als Gegenstück zur Materie begriffen, sondern als Gegenstand, über den sich in derselben Weise sprechen ließ, in der die exakten Wissenschaften über andere Phänomene der Natur sprachen, und der nichts anderes darstellte als das Ergebnis komplexer physikalischer Prozesse.

Auf die Entstofflichung der Materie durch die Physik folgte eine Mechanisierung des Geistes und die Hoffnung, dass das, was früher als getrennt gedacht wurde, in einem physikalischen Monismus vereint werden könne. Gegen Ende des Jahrhunderts rückte schließlich das Bewusstsein immer öfter ins Zentrum des Interesses, was viele als Abschluss des Projekts einer allumfassenden Naturalisierung des Geistes ansahen. Doch damit schien plötzlich das naturalistische Weltbild selbst ins Wanken zu geraten.

Der Deutsche Physiologe Emil du Bois-Reymond erklärte 1872 in *Über die Grenzen des Naturerkennens* noch, dass ein naturwissenschaftliches Verständnis von Bewusstsein unmöglich sei („*ignoramus et ignorabimus*“). Auch heute gilt ein mögliches Verständnis von Bewusstsein als der umstrittenste und problematischste Aspekt einer Naturwissenschaft des Geistigen. Gleichzeitig aber führte der Aufschwung in der Hirnforschung, insbesondere wegen der Entwicklung moderner bildgebender Verfahren, zusammen mit einer bis dato unbekanntem Rechenkapazität, zur Hoffnung, menschliches Bewusstsein könne nun Objekt wissenschaftlicher Forschung werden. Neben der Hirnforschung beanspruchen aber auch noch weitere Einzelwissenschaften, so etwa die Psychologie, die Biochemie oder die Physik, notwendige Beiträge zu einem wissenschaftlichen Verständnis von Bewusstsein zu liefern.

Über die letzten Jahre entwickelten sich daher ganz unterschiedliche Ansätze, die eine Naturalisierung des Bewusstseins anstrebten. Dabei gehen neurobiologisch, kognitionswissenschaftlich oder informationstheoretisch ausgerichtete Erklärungsversuche von der Existenz eines komplexen Mechanismus aus, der Bewusstsein verursacht oder identisch zu diesem ist [1].

Dennoch herrscht völlige Unklarheit, auf welcher Ebene denn überhaupt nach Bewusstsein gesucht werden müsse. Ist es die der Organisation des Nervensystems (also dessen „kognitive Architektur“)? Sind vielleicht nur bestimmte Hirnareale für das Bewusstsein zuständig? So scheint beispielsweise das Cerebellum am Bewusstsein nicht sonderlich beteiligt zu sein (Koch 2005). Der obere Hirnstamm hingegen ist Ausgangspunkt für die Verarbeitung instinktiver Affektionen (Solms 2013), während die höheren kognitiven Fähigkeiten meist auf Prozesse im Kortex zurückgeführt wer-

den (Fuster 2005); beides scheint jedoch für das Bewusstsein ein große Rolle zu spielen, und es ist daher schwer entscheidbar, inwieweit Bewusstsein als kortikales, sub-kortikales oder überhaupt globales neuronales Phänomen verstanden werden muss.

Oder geht es gar nur um einzelne Zellverbände, Neurotransmitter, Gerüst-Proteine oder Quantenfluktuation? Funktionale Kriterien wie etwa „hohes  $\Phi$ “ (Tononi 2008) sind eventuell in ausgezeichneten Zellverbänden aller Art realisiert, deren Interaktion den „dynamischen Kern“ (Edelmann & Tononi 2000) bei der Erzeugung des Bewusstseins bilden würde. Auch der Einfluss von neurochemisch aktiven Substanzen wie Halluzinogenen (Vollenweider 1994, Prentner 2004), Anästhetika (Hameroff 2001) oder Neurotransmittern (Perry *et al.* 2002) darf einigen Forschern zufolge nicht vernachlässigt werden, wenn wir Bewusstsein verstehen wollen. Dies würde eher gegen die Identifizierung des Bewusstseins mit einem rein systemisch-neuronalen Phänomen sprechen und auf einen molekularen Mechanismus hindeuten. Noch unklarer ist, wie diese verschiedenen Felder, die allesamt den Anspruch erheben, etwas zum Verständnis des Bewusstseins beizutragen, miteinander in Verbindung stehen.

Die obigen Redeweisen lassen vielleicht vermuten, dass man nach *dem* Bewusstsein suchen könne wie nach seinen verlegten Autoschlüsseln, dass ein bestimmter Mechanismus *fürs* Bewusstsein zuständig wäre wie das Passamt für die Ausstellung von Reisepässen oder dass es darum ginge, die Stelle im Kopf zu finden, aus der *das* Bewusstsein strömt wie ein Fluss aus seiner Quelle.

Doch möglicherweise ist Bewusstsein kein „Ding“ im Kopf, sondern subjektives Erleben, das sich zwar in den komplexen, biologischen Strukturen entfaltet, aber nicht von diesen erzeugt wird? Bewusstsein wäre möglicherweise überall anzutreffen, beginnend bei den elementaren Phänomenen von Energieübertragung bis hin zu den komplexen Mechanismen höherer Organisationsformen von Materie.

Oder sind es die Sprache und das Soziale, die uns erst zu den bewussten Wesen werden lässt, die wir sind? Reden wir uns gar nur ein, es gäbe so etwas wie Bewusstsein? Ist Bewusstsein nichts weiter als eine hartnäckige Fiktion, ähnlich wie die Vorstellung, es gäbe *die eine* Natur, *die eine* Welt oder *die eine* Zeit? Schließlich hat noch niemand das Bewusstsein je zu Gesicht bekommen!

In Hinblick auf die Philosophie sind diese Überlegungen bemerkenswert. Dort ist Bewusstsein schon seit langem ein prominentes Thema, welches gerade in der vermeintlich objektiven und auf Klarheit bedachten analytischen Philosophie zum Schauplatz eines Kampfes zwischen verschiedenen metaphysischen Auffassungen wurde. Dabei ist es nicht überraschend, dass es einerseits Philosophen gibt, die einem szientistischen Verständnis von Bewusstsein im Rahmen eines reduktiven Materialismus das

Wort Reden, und dass andererseits wissenschaftliche Erklärungsversuche einer fundamentalen Kritik unterzogen werden.

In jüngerer Zeit wurde in diesem Zusammenhang häufig vom „*hard problem of consciousness*“ (Chalmers 1995) gesprochen. Dabei handelt es sich um die These, dass die Existenz von Bewusstsein nicht im Rahmen einer materialistischen Theorie erklärbar sei. Dem zuvor gegangen sind einflussreiche Arbeiten der neueren Angelsächsischen Philosophie: Die Schwierigkeit, subjektives Erleben durch die objektiven Aussagen der Wissenschaften verständlich werden zu lassen, wurde bereits durch Thomas Nagel (1974) expliziert; die Feststellung einer *Lücke* zwischen der Erklärung physikalischer (neurobiologischer) Vorgänge und dem damit verbundenen Erleben findet sich bei Joseph Levine (1983), um nur zwei Beispiele zu nennen.

Das *hard problem* vollzieht nun eine Synthese von diesen und anderen philosophischen Positionen. Es lässt sich folgendermaßen lesen:

**Das *hard problem*:**

- Phänomenales Bewusstsein ist die Art und Weise, wie wir Situationen erleben oder „wie es sich anfühlt“, eine Erfahrung zu machen.
- Es ist nicht verständlich, wie und warum phänomenales Bewusstsein aus einer Anordnung von Materie hervorgeht.

Auf dieses Problem wird üblicherweise auf zwei Arten reagiert. Die erste Strategie leugnet die Existenz von phänomenalen Bewusstsein und erklärt, warum manche Lebewesen der Illusion eines phänomenalen Bewusstseins erliegen und worin diese Illusion eigentlich besteht. So argumentiert etwa Daniel Dennett (1991a) gegen das phänomenale Bewusstsein wie folgt: Es *scheint so*, also ob es phänomenales Bewusstsein gäbe. Daraus folgt aber nicht, dass es so etwas tatsächlich gibt. Phänomenales Bewusstsein ist möglicherweise nur eine Illusion, die durch kognitive Mechanismen erzeugt wird.

Doch dieser Gedanke ist selbst mit einigen Schwierigkeiten behaftet. Denken wir an eine optische Illusion. Zwei Geraden sind parallel, erscheinen uns aber als gebogen. Wir können einsehen, dass es hier eine Illusion gibt, etwa indem wir mithilfe eines Lineals unsere Wahrnehmung mit dem Blatt Papier vergleichen. Die Illusion wird aber durch diese Erkenntnis nicht verschwinden. Wir können vielleicht sogar erklären, warum es zu dieser Illusion kommt, etwa weil unser Gehirn die Gegenstände, die wir wahrnehmen, unbewusst immer mit deren Umgebung abgleicht.

Wir können aber dadurch nicht erklären, warum wir überhaupt etwas wahrnehmen, egal ob illusionär oder nicht. Denn anders als bei allen anderen Beispielen von Il-

lusionen geht es bei Dennetts kognitiver Illusion um das Bewusstsein selbst – und nicht etwa um krumme Geraden, Geldstücke hinterm Ohr oder Kaninchen im Hut. Es geht also nicht darum, dass uns ein Gegenstand als ein anderer erscheint, sondern darum, dass uns *überhaupt* etwas erscheint. Dennetts Vorschlag ist also nur auf den ersten Blick eine handstreichartige Lösung des *hard problems*. Bei genauem Hinsehen ergeben sich Probleme der Selbstbezüglichkeit, die mindestens ebenso schwer wiegen: Widerspricht sich die Feststellung, dass das Bewusstsein nichts als eine Ansammlung kognitiver Mechanismen sei („a bag of ordinary tricks“; Dennett 2003, S. 8), nicht selbst? Vielmehr ist Bewusstsein doch eine Bedingung dafür, überhaupt einer Illusion erliegen zu können. Falls aber die Möglichkeit einer Illusion auf der Existenz bewusster Wahrnehmung beruht, kann diese nicht mit Verweis auf Illusionen „wegerklärt“ werden [2].

Vielleicht meint Dennett aber etwas anderes mit Illusion? Vielleicht meint er, dass wir uns *nur so verhalten, als ob* wir zwei gebogene Linien sehen würden, wo wir doch in Wahrheit zwei parallele Geraden vor uns haben. Und vielleicht verhalten wir uns so, *als ob* wir ein Bewusstsein hätten, wo doch aber eigentlich gar kein Bewusstsein existiert. Doch kann die Art und Weise, wie wir einen Gegenstand (bewusst) wahrnehmen, durch das Verhalten, welches wir (oder unsere Gehirne) dabei an den Tag legen, ersetzt werden? Das Illusionsargument hinge also davon ab, ob wir Bewusstseinsphänomene durch etwas anderes *ersetzen* können. Dann aber wäre Bewusstsein nicht erklärt, sondern verschwunden!

Die zweite Strategie akzeptiert hingegen, dass Bewusstsein einen sinnvollen Teil einer naturalistischen Philosophie des Geistes bilden kann, und argumentiert nun für oder gegen dessen prinzipielle Erklärbarkeit durch die Naturwissenschaften. Kommt man dabei zum Schluss, dass eine Erklärung von Bewusstsein aus diesen oder jenen Gründen unmöglich im Rahmen einer naturwissenschaftlichen Theorie erfolgen könne, dann ist das *hard problem* nicht nur schwer, sondern unlösbar.

Im Folgenden soll hingegen dafür argumentiert werden, dass die Redeweise von einem *hard problem* zuallererst missverständlich ist. Es stellt – anders als die Fragen, wie Radioaktivität erklärbar ist, was die Träger von Erbinformation sind oder welche Anregungsmuster im Gehirn zu den verschiedenen Phasen des Schlafes korrelieren – kein eigentliches *Problem* dar, das durch die Wissenschaft zu lösen wäre. Eine solche Redeweise ist vielmehr nur ein Artefakt eines ganz bestimmten Verständnisses von Bewusstsein und der Funktionsweise wissenschaftlicher Tätigkeit vor dem Hintergrund spezifischer (aber oft impliziter) metaphysischer Annahmen. Bewusstsein stellt die Wissenschaft vor ein *Rätsel*, das es noch zu entwirren gibt, bevor mit dem Lösen von Problemen angefangen werden kann.

Als Beispiel dafür, wie sich philosophische Grundsatzfragen manchmal als vermeintliche Probleme für die Wissenschaft ausgeben, sei kurz ein Beispiel aus der Physik diskutiert. Die Lehre, die man daraus ziehen sollte, ist dass wir unsere Auffassung von Natur(wissenschaft) immer kritisch hinterfragen müssen, damit wir nicht Gefahr laufen, bestimmte Phänomene fälschlicherweise als extra-natürlich abzutun oder für prinzipiell unerklärlich halten. In einer einflussreichen Arbeit kritisierten Einstein, Podolsky & Rosen (1935) die damals noch recht junge Quantenmechanik und merkten an, dass im Rahmen eines hypothetischen Szenarios, welches von der Existenz „verschränkter Zustände“ handelt, die Quantenmechanik als wesentlich unvollständige Theorie ausgewiesen werden müsse. Viele Physiker wie Einstein hielten es damals für selbstverständlich, dass eine fundamentale wissenschaftliche Theorie folgende drei Kriterien erfüllen sollte [3]:

**Einsteins Forderungen an eine gute wissenschaftliche Theorie:**

1. Sie sollte *realistisch* sein, d.h. dass die Ergebnisse von Messungen die objektiven Eigenschaften der Dinge darstellen, falls diese ohne Störung des zu messenden physikalischen Systems vorhergesagt werden können,
2. sie sollte *vollständig* sein, d.h. dass jedes Element der Wirklichkeit eine Entsprechung im theoretischen Formalismus findet, und
3. sie sollte *lokal* sein, d.h. dass ein System nicht durch Messungen instantan beeinflusst wird, die an einem anderen, räumlich getrennten System stattfinden;

Da nun von Einstein *et al.* in ihrem Gedankenexperiment gezeigt wurde, dass die Quantenmechanik eine Interpretation als lokal-realistisch und vollständig nicht zulässt, folgerten sie – entsprechend der impliziten Überzeugung, wonach eine finale wissenschaftliche Theorie dem Realismus und der Lokalität genügen würde –, dass die Quantenmechanik unvollständig sein müsse: Es gebe Elemente der Wirklichkeit, die gerade nicht im Formalismus der Quantenmechanik behandelt würden. Eine vollständige Theorie hingegen würde über „verborgene Parameter“ verfügen, welche die Schwierigkeiten der Quantenmechanik auflösen könnten. Erst viele Jahre später konnte von John Bell (1964) theoretisch und dann von Alain Aspect *et al.* (1981) auch experimentell gezeigt werden, dass die Existenz solcher verborgener Parameter unmöglich mit der Lokalität in Einklang gebracht werden könne, was die Einstein'sche Überzeugung in schwere Erklärungsnot gebracht hätte.

Einen anderen Ansatz verfolgte der Physiker David Bohm [4], der eine Theorie



entwarf, die verborgene Parameter kennt, dabei allerdings auf das Kriterium der Lokalität verzichtete. Es gebe demnach instantane Effekte zwischen räumlich getrennten Teilchen. Auf diese Weise kann an einer realistischen und vollständigen Theorie festgehalten werden, jedoch auf Kosten der Lokalität.

Die dritte Möglichkeit, auf Einsteins Gedankenexperimente zu reagieren, ist diesbezüglich zwar sparsamer, lässt die Natur aber gemessen an unserer alltäglichen Erfahrung als höchst sonderbar erscheinen. Sie wurde etwa von Niels Bohr (1935) im Anschluss an die Kopenhagener Deutung der Quantenmechanik vertreten: Es sei der Realismus, der modifiziert werden müsse. Die Quantenmechanik würde eben nicht mehr der „naiven“ Vorstellung entsprechen, wonach eine vollständige Theorie die objektiven Eigenschaften von Systemen unabhängig von deren Messung widerspiegeln. Überspitzt formuliert, wird die physikalische Wirklichkeit, von dessen objektiver Existenz Einstein so überzeugt war, erst durch die Messung *erzeugt* [5].

Heute ist die Lage, was die Interpretation der Quantenmechanik angeht, noch einmal komplexer geworden [6], aber es hat sich gezeigt, dass die Effekte, die von Einstein *et al.* ursprünglich zur Kritik einer Theorie vor dem Hintergrund spezifischer (naturphilosophischer) Annahmen erdacht wurden, experimentell bestens bestätigt sind – genauso wie die Quantentheorie selbst, welche als vielleicht bestbestätigste empirisch-quantitative Theorie aller Zeiten gilt.

Man kann nun einwenden, dass die Quantenmechanik ja für eine Erklärung des Gehirns (oder des Bewusstseins) gar keine Rolle spielen würde. Unabhängig davon, ob diese Aussage empirisch überhaupt zutreffend ist, verkennt dies die eigentliche Moral, die wir aus der Einstein'schen Kritik ziehen sollten. Es wäre wohl kaum jemand auf die Idee gekommen, im Zuge der Arbeiten von Einstein *et al.* von einem „schweren Problem der Nicht-Lokalität“ zu sprechen und zu fragen wie es denn sein könne, dass Phänomene wie die der Verschränkung aus einer Anordnung lokalisierter Teilchen hervorgehen könnten. Noch weniger würde daraus gefolgert werden, dass hier ein metaphysischer Graben zwischen verschiedenen Seinsweisen vorliegen würde, der als Art Grundlagenproblem eines neuartigen philosophisch-naturalistischen Forschungsprogrammes aufzufassen sei.

Vielmehr scheint uns die Arbeit von Einstein *et al.* darauf aufmerksam zu machen, dass die Quantenmechanik Effekte vorhersagt, die uns als völlig widersinnig erscheinen, *aber nur, falls wir bestimmte Annahmen über das Wesen der wissenschaftlichen Naturbeschreibung treffen*. Eine Möglichkeit, mit dieser Situation umzugehen, wäre es, die Effekte als „spukhafte“ Phänomene auszuweisen, die auf eine Unstimmigkeit der Theorie deuten würden. Das Beispiel der Quantenverschränkung zeigt jedoch, dass es in einigen Fällen fruchtbarer ist, unser wissenschaftliches Naturverständnis zu ändern, was nicht dazu führen würde, dass die Phänomene „verschwinden“, son-

dern im Gegenteil erst als bedeutsame Effekte erkannt werden.

Allerdings ist Bewusstsein nichts, das derzeit von irgendeiner Theorie vorausgesagt oder erklärt würde. Vielleicht erscheint es aus diesem Grund als attraktive Option, dessen Existenz einfach zu leugnen. Dass dies von einigen mit religiösem Eifer getan wird, lässt allerdings fragen, vor welchem Übel sie uns eigentlich beschützen wollen. Oder geht es ihnen nur um ein besseres, da vermeintlich wertfreies, Naturverständnis? Doch sollten wir dann nicht erst unsere Hintergrundannahmen hinterfragen, bevor wir Bewusstsein als Aberglauben abtun? So könnte das *hard problem* möglicherweise nur deswegen so aussichtslos erscheinen, weil wir versuchen, es vor dem Hintergrund des *Physikalismus* – also einer *metaphysischen und keiner wissenschaftlichen* Position – zu verstehen [7].

## 1.2 Eine kleine Übersicht über metaphysische Positionen

Allerdings steckt im *hard problem* tatsächlich eine Schwierigkeit, mit der jede wissenschaftliche Auseinandersetzung mit dem Bewusstsein konfrontiert ist. Doch bevor wir darauf näher zu sprechen kommen, sollen traditionelle substanz-metaphysische Positionen, die relevant für den aktuellen Diskurs rund um das Bewusstsein sind, kurz erwähnt werden. Dies hat neben dem einführenden Charakter auch den Zweck, dass wichtige Argumente betrachtet werden sollen, auch wenn diese gar nicht explizit vom Bewusstsein handeln.

Beispielsweise soll dafür argumentiert werden, dass bestimmte Voraussetzung für Wissenschaft, wie sie im kartesischen Dualismus ihren Ausdruck finden, nicht einfach abgelegt werden können. So gilt manchen die Annahme eines freien Geistes, der außerhalb der zu beschreibenden Welt steht, vielleicht als fragwürdig oder gar als „falsch“, sie bildet aber eine implizite Grundvoraussetzung wissenschaftlicher Praxis.

Ein anderes Beispiel betrifft die „intrinsische Natur“ materieller Gegenstände. Diese bezeichnet alle Eigenschaften der Gegenstände, die ihnen unabhängig von allen Relationen zu anderen Gegenständen zukommen. Der Begriff der Masse, wie er in der klassischen Physik verwendet wird, liefert uns hier eine Veranschaulichung: Die Masse eines Teilchens ist unabhängig davon, in welchem Schwerfeld es sich gerade befindet; dies unterscheidet schließlich die *Masse* eines Teilchens von dessen *Gewicht*. Es könnte vermutet werden, dass es sich bei der Masse eben um eine solche „intrinsische Eigenschaft“ von Gegenständen handeln würde. Es besteht nun die Versuchung, in Analogie auch Bewusstsein vorschnell mit einer rein intrinsischen Eigenschaft physikalischer Gegenstände zu identifizieren, was aber wissenschaftliche Aussagen über das Bewusstsein zu erschweren oder gar zu verunmöglichen scheint.

Letztendlich geht es in diesem Abschnitt weniger um das Abwägen von Argumenten, die für oder gegen eine bestimmte metaphysische Position sprechen, sondern vielmehr darum, diese als Ausdruck der Art und Weise, wie wir uns reflexiv auf unsere Erfahrungen beziehen, zu verstehen. Es muss uns also auch darum gehen, die Annahmen, die dabei implizit mitschwingen, wieder explizit werden zu lassen.

## Materialismus und Physikalismus

Die derzeit vielleicht gängigste naturphilosophische Position ist der Physikalismus. Seine Vertreter behaupten, dass alles Seiende letztlich als Produkt derjenigen Gegenstände zu denken sei, welche die Physik als fundamental denkt – identifiziert man diese nun mit den Elementarteilchen des Standardmodells, mit den Hypostatierungen irreduzibler Darstellungen der fundamentalen Symmetriegruppe oder mit den Vibrationen von Strings. In diesem Sinne kann gesagt werden, dass alles letztlich *physikalisch* ist.

Mit dieser ontologischen Einstellung sind eine Menge epistemologischer Konsequenzen verbunden, etwa die Überzeugung, dass neben allen naturwissenschaftlichen auch alle kulturellen und gesellschaftlichen Phänomene, sofern wir sie als wirkliche und nicht nur als bloß illusorische betrachten, letztlich auf physikalische Wirksamkeit, und nur auf diese, zurückgeführt werden können – selbst wenn dies die kognitiven Fähigkeiten von uns Menschen überschreiten würde. Der Physikalismus lässt keinen Raum für weitere Gesetze als die der modernen Physik: Auch die Evolutionstheorie muss, wenn sie nicht nur als Beschreibung einer zufälligen, empirischen Regelmäßigkeit gelten will, als abgeleiteter oder zumindest aus der Physik hervortretender Gesetzeszusammenhang verstanden werden können. Genauso müssen psychologische, ökonomische oder soziologische Gesetze und ethische Normen letztlich auf die Physik rückführbar sein. Und auch eine Theorie vom Bewusstsein wird entweder Teil einer „neuen Physik“ oder Ergebnis einer auf Physik basierenden Wissenschaft sein. Der Physikalismus ist ein universalistisches Projekt.

Auch eine wichtige Strategie zur Immunisierung dieser Doktrin soll kurz erwähnt werden: Alles, was jemals erkannt werden könne, müsse demnach *im Prinzip* Gegenstand der Physik sein können. Eine vollständige Beschreibung von Wirklichkeit würde letztlich in einer finalen Physik aufgehen. Wirklich wäre dann genau dasjenige, was Gegenstand einer solchen finalen Wissenschaft ist, und daher wäre – per Definition – alles Seiende physikalisch. Diese Strategie ist nun einerseits selbst von der physikalistischen Überzeugung geleitet, weswegen sie als Argument für den Physikalismus als zirkulär angesehen werden muss. Andererseits würde sie alles, was sich der derzeitigen physikalischen Beschreibung entzieht, als Gegenstand einer zukünftigen

(finalen) Physik postulieren. Aus diesem Grund kann dem Physikalismus gar kein ernst zunehmendes Phänomen entgegengesetzt werden, was ihn aber zur Trivialität verkommen ließe [8]. Es scheint fast so, als ob der Physikalismus mehr mit einem Weltbild gemein hätte, als mit kritischer Wissenschaft.

Kaum vom Physikalismus zu unterscheiden, ist eine verkürzende Interpretation des Materialismus, wonach Materialismus und Physikalismus zusammenfallen würden – ohne Rücksicht auf historische Bedeutungswechsel, die der Begriff der Materie durchlaufen hat, z.B. der atomistische Materialismus bei Demokrit, die mathematische Naturphilosophie nach Newton oder das Materieverständnis der modernen Naturwissenschaften. Sofern angenommen wird, dass Materie einfach dasjenige sei, was exklusiver Gegenstand der Physik ist, so verweist jede Erklärung, die auf das Verhalten materieller Gegenstände verweist, letztlich auf Physik.

Ein ähnliches Schicksal ist dem Begriff des Naturalismus beschieden. Statt an die vielfältigen Naturvorstellungen zu denken, wie sie sich etwa noch bei den sogenannten „Naturvölkern“, den Vorsokratikern oder Goethe fanden, wird der Naturalismus heute meist als synonym zum Physikalismus verstanden und reduziert sich auf die Vorstellung, dass jegliches Sein natürlich ist, im Gegensatz zur Idee, dass bestimmte Phänomene auf eine übernatürliche Entität verweisen würden. Als paradigmatisch für übernatürliche Wesenheiten sind die unterschiedlichen Gottesvorstellungen zu nennen, aber auch ein transzendentes Subjekt, das jegliche Erkenntnis letztbegründet, ohne dabei selbst in irgendeiner Weise natürlich bedingt zu sein.

Ein Problem für diesen so verstandenen Naturalismus ist die Existenz von Bewusstsein. Wie kann es sein, dass neben den Gegenständen, die von den Naturwissenschaften beschrieben werden, so etwas wie Bewusstsein existiert, das sich einer solchen Beschreibung entzieht? Üblicherweise wird dieses Problem entweder ignoriert oder mit Verweis auf ein „kartesisches Theater“ (Dennett 1991a) für eigentlich sinnlos erklärt: Einen solchen Gegenstand gibt es gar nicht!

## Dualismus

Einige, die der letzten Aussage nicht zustimmen wollen, berufen sich stattdessen auf den Dualismus. Als Urheber der neuzeitlichen dualistischen Philosophie gilt gemeinhin René Descartes. Descartes unterscheidet zwischen einer materiellen, ausgedehnten Substanz („*res extensa*“) und einer immateriellen, denkenden Substanz („*res cogitans*“). Ein Problem für diese Auffassung ergibt sich aus dem Hinweis auf psychophysische Wechselwirkungen: Einerseits erscheint es uns als unhintergehbare Tatsache, dass Mentales einen Einfluss auf Physisches hat (wenn ich etwa den Arm hebe, weil ich das will), gleichzeitig ist auch vom umgekehrten Fall auszugehen

(schlucke ich Drogen, üben diese einen Einfluss auf mein Bewusstsein aus). Andererseits wird diese Wechselwirkung nur schwer innerhalb der substanzdualistischen Auffassung verständlich.

Mit diesem Vorwurf musste sich schon Descartes auseinandersetzen, als ihn Prinzessin Elisabeth von Böhmen dazu aufforderte, begreiflich zu machen, wie denn eine immaterielle Substanz einen kausalen Einfluss auf einen ausgedehnten Körper ausüben könne. Descartes Modell der Kausalität war rein mechanistisch, d.h. jegliche Verursachung ist auf Stöße zwischen ausgedehnten Körpern zurückzuführen. Dies scheint jedoch einen kausalen Einfluss zwischen den Substanzen per Definition zu verunmöglichen. Descartes postulierte zu diesem Zweck eine mögliche Wechselwirkung zwischen der Seele und dem Körper, die primär über die Zirbeldrüse des Gehirns stattfindet. Dabei spielten die „Lebensgeister“, ein Fluidum, das Descartes (1996, §10) mit kleinen Körpern identifizierte, „die sich sehr schnell bewegen, so wie die Teile der Flamme, die einer Fackel entsprühen“, eine zentrale Rolle: Diese können in die Muskeln eindringen und sie zur Tätigkeit veranlassen, was einerseits rein reflexartig, andererseits aber auch durch willentliche Drehung der Zirbeldrüse verursacht werden kann. Gleichzeitig führen Bewegungen der Zirbeldrüse dazu, dass die Seele Sinneseindrücke gewinnt oder von der äußeren Welt beeinflusst werden kann. Allerdings ist dies heute weder physiologisch noch physikalisch sonderlich plausibel. Zwei Alternativen, die ganz ohne einen solchen Wechselwirkungsmechanismus der beiden Substanzen auskommen, sind Malebranches Okkasionalismus, welcher die scheinbare Interaktion zwischen Körper und Geist auf eine ständige göttliche Einwirkung zurückführt, oder Leibniz' Parallelismus von Körper und Geist, der durch die Metapher zweier perfekt aufeinander abgestimmter Uhren veranschaulicht werden kann. Alle drei Positionen, psychophysische Interaktion, göttliches Einwirken und prästabilisierte Korrelation, erscheinen vielen Philosophen heute als suspekt, da sie auf vermeintlich widerlegte, unüberprüfbare oder übernatürliche (meist göttliche) Effekte verweisen.

Solche Einwände gegen den Dualismus werden in praktisch jedem modernen Lehrbuch zur Philosophie des Geistes erwähnt. Was oft weniger Beachtung findet, ist die Tatsache, dass der Dualismus unser Verständnis von Wissenschaft maßgeblich prägte. Descartes ging es dabei anscheinend vor allem um die *Tätigkeiten*, welche der immaterielle Geist ausübt. Die Frage nach der Beschaffenheit des (phänomenalen) Bewusstseins rückt in den Hintergrund und wird nur selten thematisiert, wobei der Begriff „Bewusstsein“ selbst sogar erst später bei John Locke auftaucht. Es geht also um Tätigkeiten des Geistes, die aber, so Descartes (2009) in der vierten Erwiderung an Arnauld, am ehesten dem entsprechen, was wir heute als „Aktivitäten des Bewusstseins“ oder als „Art des Denkaktes“ (Kemmerling 1996, S. 89) bezeichnen

würden. Auf der einen Seite zeigt sich also, dass Descartes dem Epiphänomenalismus höchst abgeneigt gewesen wäre, auf der anderen Seite spielt das, was von der Philosophie heute als *Qualia*-Problem verhandelt wird, gar keine große Rolle.

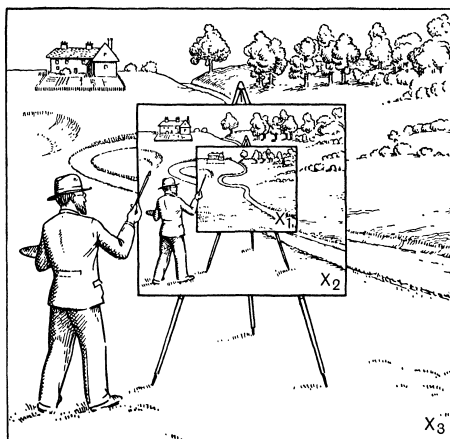
Beim Dualismus handelt es sich auch nicht um eine vor-theoretische Form der Alltagspsychologie, auch wenn eine Gleichsetzung des *ego cogitans* mit dem psychischen Ich dies nahelegt; er betrifft allerdings unser Selbstverständnis als handlungsfähige Wesen: Versteht man nämlich Wissenschaft in erster Linie als Tätigkeit, die eine Person (oder eine Gruppe von Personen) ausübt, um empirische Phänomene zu erklären, zu manipulieren oder vorhersehbar zu machen, ist ein solcher „kartesischer Schnitt“ (Primas 1994, S. 611f.) zwischen einem materiellen Objekt, welches den Naturgesetzen folgt, und einem immateriellen Bewusstsein, welches über einen Willen verfügt und bestimmte Zwecke verfolgen kann, konzeptuell kaum zu vermeiden, ist vielleicht sogar Vorbedingung von Wissenschaft überhaupt. Sowohl im Labor als auch in der Rechenstube des Theoretikers wird implizit davon ausgegangen, dass ein bewusstes, tätiges Subjekt extern zu den erzeugten Phänomenen und der Theorie existiert: Ohne die prinzipielle Freiheit des Experimentators, seinen Versuchsaufbau so oder anders zu wählen, und ohne die Freiheit des Theoretikers, seine Darstellungen daran anzupassen, wäre Wissenschaft, wie wir sie heute verstehen, gar nicht denkbar.

Dies macht insbesondere auch die Quantenmechanik und die Rolle, die dabei dem Experimentator zukommt, deutlich: Dessen Bewusstsein ist zwar selbst nicht Teil der theoretischen Beschreibung, allerdings muss seine Wahl bestimmter Observablen, d.h. des genauen experimentellen Aufbaus, berücksichtigt werden, um zu einer widerspruchsfreien Beschreibung zu gelangen [9].

Manche Materialisten schenken dieser dualistischen Voraussetzung der wissenschaftlichen Praxis keine Beachtung. Doch wenn ein empirisches Phänomen wissenschaftlich untersucht wird, handelt es sich um ein Objekt, das von jemanden erforscht wird; eine Theorie wird von jemanden dazu benutzt, eine Beschreibung eines Phänomens zu liefern oder eine Vorhersage zu tätigen. Den Dualismus im Rahmen einer materialistischen Metaphysik „zu überwinden“, kann dann gegebenenfalls so verstanden werden, dass dieser „jemand“ durch dieselbe Theorie beschrieben wird, mit der auch seine Außenwelt beschrieben wird – also etwa durch Physik und Biologie. Dies ganz naiv als wahrheitsgetreue Abbildung der Wirklichkeit zu verstehen, würde allerdings in einen Regress führen, wie an folgender Metapher verdeutlicht werden soll (angelehnt an Dunne 1934):

Nachdem eines Tages ein Insasse einer Anstalt für Geistesgestörte seinem Gefängnis entkommen war, beschloss dieser an einem schönen Sommertag ein Bild der vor ihm liegenden Landschaft zu malen. Das Gemälde zeigte jede Blume und jeden Grashalm,

es war voll lebendiger Farben und reich an Texturen. Dennoch fühlte der Verrückte den unwiderstehlichen Drang, die Landschaft nicht nur so abzubilden wie sie sich *seinem* Auge darstellte, sondern wie sie wirklich und an-sich war, wie ein allwissender Gott sie aus dem Nirgendwo sehen würde. Er begann, sein Werk noch exakter und detailreicher werden zu lassen, aber nach einiger Zeit hatte er das Gefühl, dass etwas fehlte. Plötzlich erkannte er, dass er sich selbst, den Beobachter und Erschaffer dieser Szenerie nicht abgebildet hatte. Und so beschloss er, sich selbst zu malen, wie er die sich ihm präsentierte Landschaft auf einer Leinwand festhielt. Er malte einen Maler, der eine Landschaft malt. Und auch damit gab er sich noch nicht zufrieden. Schließlich, so dachte er sich, würde sein Bild zwar ausdrücken, dass dort ein Maler zu sehen sei, der eine schöne Landschaft malt, aber nicht, dass es sich hier um einen *bewussten* Maler handle, welcher von sich selbst weiß und den Wunsch verspürt, die Wirklichkeit, wie sie nun einmal ist, abzubilden. Aus diesem Grund malte der Verrückte einen weiteren Maler, der einen Maler malt, der eine Landschaft abbildet (Abbildung 1.1)...



**Abbildung 1.1:** Das Gemälde eines Malers, der einen Maler malt, wie er eine Landschaft abbildet; nach Dunne (1934, S. 31).

Vergleicht man diese Szene mit den beiden metaphysischen Positionen, dem Materialismus und dem Dualismus, so erkennen wir, dass das ursprüngliche Gemälde eine Metapher für eine wissenschaftliche Theorie darstellt, die von einem Wissenschaftler, dessen Bewusstsein selbst nicht Teil der Theorie ist, entworfen wird. Der Metaphysiker, der Bewusstsein als „Bild von“ versteht, gleicht dem verrückten Künstler, der glaubt, er könne seine eigene Erfahrung durch das Malen eines Malers sichtbar machen. Wie würde der Künstler vorgehen? Er könnte etwa versuchen, seine Freude abzubilden und würde ein Lächeln auf das Gesicht des Malers malen. Und er würde versuchen, die Motive des Künstlers an der Art und Weise, wie dieser vor seiner Leinwand steht und die Natur abbildet, sichtbar zu machen. Und vielleicht würde

er sogar der Meinung sein, dass es außer den Gesichtszügen und der Haltung des Künstlers gar nichts anderes gäbe, das man abbilde könnte.

Doch verstünde man die Inhalte des Bewusstseins lediglich als Abbildung einer äußeren Wirklichkeit, würde die Metapher vom irren Maler nur verdeutlichen, dass es sich dabei um eine Auffassung handeln würde, die den Dualismus voraussetzt und ihn nicht überwindet. Beispielsweise wäre sein Wissen, dass *er* es ist, der die Landschaft abbildet, dadurch noch gar nicht dargestellt. Dunne (1934) lässt seinen Maler aus diesem Grund einen weiteren Maler malen, der den Maler malt, wie er die Landschaft abbildet. Doch kann er sich dem Dualismus auf *diese Weise* wirklich entziehen? Falls nein, dann würde der Versuch, das erschaffende und tätige Subjekt als Abbildung in das bereits abgebildete Objekt zu verschieben, in einen infiniten Regress von Abbildungen führen. Falls doch, dann müsste der Materialist zeigen können, wann und warum dieser Regress abbricht.

### Pluralismus ohne Substanz

Der umgekehrte Weg, nämlich die Anzahl der Substanzen zu *erhöhen*, wurde seltener besprochen, etwa durch Karl Popper (1978). Hier wird neben einer materiellen und geistigen Welt ein weiteres Reich von Gedankeninhalten (Theorien oder Ideen) postuliert. Laut Popper (1978) gibt es drei Welten: eine Welt der empirischen Dinge (die „Außenwelt“), eine Welt des Bewusstseins und, davon unabhängig, eine Welt der Ideen, wie sie etwa in Theorien, Bauplänen, Gedichten oder Mythen ausgedrückt werden. Somit beziehen sich Subjekte (Gegenstände aus Welt 2) auf physikalische Objekte (Gegenstände aus Welt 1) mithilfe von Ideen (Gegenständen aus Welt 3) und umgekehrt werden Ideen durch physikalische Gegenstände für Subjekte ausgedrückt.

Für viele ist Poppers 3-Welten-Lehre ein Kuriosum seiner Spätphilosophie, doch es ist an dieser Stelle instruktiv, Poppers Argumente kurz zu diskutieren. Für Popper ist die Realität und Existenz eines Gegenstandes, gleich welcher Welt er entstammt, an dessen Wirksamkeit geknüpft. Theorien, aber auch Symphonien, Erzählungen oder Pläne (alles Gegenstände, die Popper in Welt 3 verortet) bewirken, dass sich Menschen nach Ihnen richten: Eine Erzählung kann uns in unseren Wertvorstellungen und Handlungen beeinflussen; ein Bauplan wird in der Regel die Art und Weise bestimmen, wie jemand sein Haus baut und darin lebt; und die Aufführung einer Symphonie führt in der Regel dazu, dass Leute ihr Geld für Eintrittskarten ausgeben. Radikalisiert man den Popper'schen Gedanken, so führt dies zu folgender Auffassung: Alles, was eine Wirkung ausüben kann, müsste als real existierendes Seiendes bezeichnet werden. Gleichzeitig wäre nichts davon als Substanz zu bezeich-



nen – verstünde man unter einer Substanz dasjenige, was unabhängig von allem anderen existieren kann. Denn wenn Existenz an Wirkung gebunden ist, existiert kein Seiendes unabhängig von den Dingen, die es bewirken. Gebrauchsanweisungen bewirken, dass Menschen ihnen entsprechend handeln und führen auf diese Weise zu einer Veränderung in der physischen Welt. Somit wäre die physische Welt aber in ihrer tatsächlichen Existenz nicht unabhängig von der Gebrauchsanweisung. Umgekehrt ist jede Gebrauchsanweisung, damit sie überhaupt wirksam werden kann, auf das Papier, auf dem sie abgedruckt wird, und das Bewusstsein, welches sich ihrer bedient, angewiesen, usw.

Der Monismus müsste demgegenüber zeigen können, dass alle Dinge, die in diesem Sinne als existierend gelten können, eigentlich nichts anderes wären als Modifikation einer grundlegenden Substanz. Alle Wirkung würden immer nur von dieser einen Substanz ausgehen. Poppers Drei-Welten-Lehre kann hingegen so gelesen werden, dass es drei Substanzen, eine konkrete physische, eine konkrete psychische und eine abstrakte ideale, benötigt, um einen angemessenen Rahmen zu schaffen, innerhalb dessen die (dann nur scheinbare) Pluralität der Phänomene erklärt werden kann, und die bis zu einem gewissen Grad unabhängig voneinander sind [10].

Die Alternative ist, dass es gar keine substantiellen Dinge gibt, sondern nur je-relativ stabile Muster von Wirksamkeit. Dass wir „hinter“ diesen Mustern noch etwas hinzudenken, ist eine Eigenart unseres Denkens, verweist aber sonst auf nichts.

### **Neutraler Monismus und Panpsychismus**

Eine Position, welche den Monismus der Substanzen mit einem Dualismus der Eigenschaften zu vereinen sucht, ist der neutrale Monismus. Falls die dabei beschriebene Substanz sowohl materielle als auch geistige Eigenschaften aufweist, ist dieser Monismus verwandt zum Panpsychismus, also der Vorstellung, dass alle Dinge im Universum beseelt sind. Eine solche Position wurde in Abwandlungen von zahlreichen Forschern und Gelehrten über die Jahrhunderte vertreten (Skrbina 2005). Allerdings gibt es auch einige Unterschiede zwischen den verschiedenen Spielarten des neutralen Monismus und des Panpsychismus. Im Allgemeinen ist es daher ratsam, die Pluralität der Auffassungen im Auge zu behalten. So lässt sich mindestens zwischen Formen des Aspekte-Monismus, die geistige und materielle als gleichberechtigte, fundamentale Eigenschaften der Materie betrachten, und Varianten des neutralen Monismus unterscheiden, die geistige und materielle Eigenschaften als abgeleitete Erscheinungsweisen einer zugrunde liegenden „psycho-physisch neutralen“ Ebene ansehen. Während erstere oft den Verdacht erwecken, dass es sich hier eigentlich um eine getarnte Form des Dualismus handelt – passend hierzu die Selbstbezeichnung

als „naturalistischer Dualist“ von David Chalmers (1996, Kap. 4) –, so ist fraglich, inwiefern eine als psycho-physisch neutral postulierte Substanz tatsächlich weder als psychisch noch als physisch angesehen wird – so verwenden etwa William James oder Ernst Mach jeweils psychologisch aufgeladenes Vokabular („*pure experience*“ und „Empfindungen“) zur Beschreibung der (vermeintlich?) neutralen Ebene [11]. Eine dezidiert *panpsychistische* Position, welche die Existenz neutraler Elemente zurückweist, wird seit einigen Jahren von Galen Strawson vertreten. In mehreren Veröffentlichungen beschreibt er diese als „*Real Materialism*“, „*Realistic Monism*“ oder als „*Real Physicalism*“ (Strawson 2008). Darin argumentiert er wie folgt:

**Vom Physikalismus zum Panpsychismus:**

1. Alles Seiende ist physikalisch. (Physikalismus)
2. Es gibt zwei mögliche Weisen des Seins: experientiell und nicht-experientiell.
  - 2.1. Nur von der Existenz des experientiellen Seins kann ich mir sicher sein. (Bewusstsein ist real.)
- ∴ Bewusstsein ist physikalisch  $\Leftrightarrow$  Zumindest einige physikalische Gegenstände sind experientielle.

3. Es gibt gar keine wissenschaftlichen Gründe, die dafür sprechen, dass es überhaupt nicht-experientiell Sein in der Natur gäbe. (Eddingtons Zweifel)
4. Es ist unmöglich, dass experientiell Seiendes aus nicht-experientiell Seienden hervorgeht. (Unmöglichkeit der Emergenz von Bewusstsein)
5. Unsere Ontologie sollte möglichst sparsam sein. (Ockhams Rasiermesser)
- ∴ Alle Gegenstände sind physikalisch und verfügen über experientielle Eigenschaften; der („wahre“) Physikalismus fällt mit einer Form von Panpsychismus zusammen. Die letzte Substanz ist experientiell.

Die Ansicht, derzufolge Bewusstsein einen physikalischen Gegenstand bezeichne, wird auch von vielen („nicht-reduktiven“) Materialisten geteilt, die ebenfalls die beiden ersten Prämissen anerkennen. Hingegen ist es die Postulierung eines Panpsychismus, an dem sich die meisten wohl stoßen würden. Aus diesem Grund soll an dieser Stelle näher auf die Prämissen 3-5 eingegangen werden. In Strawson & Freeman (2006) sind eine Reihe von Kommentaren abgedruckt, welche sich mit einigen problematischen Annahmen oder Konsequenzen von Strawsons Argumentation auseinandersetzen:

- Intrinsic Naturen:

Die 3. Prämisse, der Zweifel an der Existenz nicht-experienteller Gegenstände, kann laut Strawson darauf zurückgeführt werden, dass physikalisches Wissen von Gegenständen immer nur relativ zu anderen Gegenständen, also etwa relativ zu den Zeigern von Messinstrumenten, erlangt wird. Dies entspricht einer ganz bestimmten Auffassung von Naturwissenschaft, wonach diese eben gar nicht von den „intrinsic Eigenschaften“ der Dinge handle, sondern immer nur Aussagen relativ zu anderen mache. Neben Arthur Eddington (1928) wird diese Ansicht vor allem Bertrand Russell (1927) zugeschrieben und scheint aus einer Beobachtung der naturwissenschaftlichen Erkenntnisweise zu folgen und metaphysisch unschuldig zu sein.

Um nun von dort zu einem Argument für den Panpsychismus zu gelangen, müssen zuerst intrinsic von „relationalen“ Eigenschaften näher unterschieden werden (Seager 2006). Intuitiv sind die intrinsic Eigenschaften eines Gegenstandes all jene, die auch in Abwesenheit aller anderen Gegenstände vorhanden wären, im Gegensatz zu den relationalen Eigenschaften. Beispielsweise könnte vermutet werden, die Masse eines Teilchens wäre eine intrinsic Eigenschaft eines Teilchens, im Gegensatz zu dessen Gewicht. Allerdings ist die Rede von intrinsic Eigenschaften mit grundsätzlichen Schwierigkeiten konfrontiert, denn unter der Masse eines Objekts sei immer nur dasjenige zu verstehen, was die Rolle des Ausdrucks „Masse“ spielt, wie er in physikalischen Theorien vorkommt. Wenn wir etwa sagen, ein Teilchen habe diese oder jene Masse, bedeutet dies nicht viel mehr, als dass es dasjenige ist, was in bestimmten Relationen (etwa dem Gravitationsgesetz) auf eine ganz bestimmte Art und Weise vorkommt (z.B. als radial anziehend; als äquivalent zur Trägheit oder dem Krümmungsverhalten der Raumzeit etc. [12]). Was Masse darüber hinaus ist – z.B. „wie es für das Teilchen ist, eine Masse zu haben“ –, darüber geben uns physikalische Theorien keine Auskunft. Folgt man der Eddington-Strawson'schen Auffassung, dass nicht nur Massen sondern *alle* Eigenschaften, die in physikalischen Theorien vorkommen, rein relational sind, woher dann

überhaupt die Annahme, dass es neben diesen noch intrinsische Eigenschaften gäbe? Zum einen, falls die Rede von Relationen notwendig auf Individuen verweisen würde, welche diese Relationen eingehen. Zum anderen, falls die Existenz von Bewusstsein selbst darauf hinweisen würde, dass es neben den relationalen Eigenschaften noch andere gibt (Seager 2006, S. 143f.): Immerhin ist es doch vorstellbar, dass ich über Bewusstsein selbst dann noch verfüge, wenn alle Gegenstände der Außenwelt verschwunden wären – ein Echo des kartesischen Zweifels.

Ersteres scheint eine spekulative Annahme der Metaphysik zu sein (für eine kritische Diskussion siehe etwa Esfeld 2003). Letzteres eine zumindest anzweifelbare Intuition, die im Verdacht steht, dasjenige schon vorauszusetzen, wofür argumentiert werden soll.

Eine weitere Konsequenz, die vielleicht wenig erwünscht ist, wenn eine naturwissenschaftliche Theorie des Bewusstseins angestrebt werden soll, muss hier kurz erwähnt werden: Wenn sich jegliche Form von Wissenschaft darin erschöpft, Aussagen über relationale Eigenschaften zu machen, Bewusstsein aber eine intrinsische Eigenschaft ist, liegt es nahe, von vornherein zu bezweifeln, dass so etwas wie eine Wissenschaft vom Bewusstsein überhaupt möglich ist. Zu behaupten, dass es etwas außerhalb der Wissenschaft Stehendes gäbe, das dennoch (empirisch) real ist, würde jedoch wiederum den Physikalismus (und somit Strawsons erste Prämisse) stark in Frage stellen.

- **Emergenz:**

Es gibt mehrere Möglichkeiten, für die 4. Prämisse zu argumentieren. Zum Beispiel können zuerst einige paradigmatische Fälle von Emergenz diskutiert werden. Ein häufig verwendetes Beispiel ist dasjenige des Entstehens einer Flüssigkeit aus einer Ansammlung von Molekülen. Strawson argumentiert nun dafür, dass dieses Beispiel prinzipiell unzureichend sei, um die Entstehung des Bewusstseins zu veranschaulichen. Es sei keine „Analogie der rechten Größe“; bei der Emergenz des Bewusstseins würde es sich nämlich um einen „begrifflich heterogenen“ Übergang handeln und nicht, wie etwa bei der Emergenz von Wasser um einen „begrifflich homogenen“ (Strawson 2006, S. 15) [13]. Flüssigkeiten haben zwar andere Eigenschaften als (unstrukturierte) Ansammlungen von Molekülen, jedoch ist die Flüssigkeit selbst ein gleichartiger physikalischer Gegenstand, der z.B. über eine Masse oder eine Lage im Raum verfügt. Eigenschaften wie Transparenz oder Viskosität folgen aus den (elektromagnetischen und mechanischen) Eigenschaften der Moleküle und sind in diesem Sinne nichts fundamental Neuartiges (oder Irreduzibles). Anders sieht es beim Bewusstsein aus: Dort wäre die Emergenz einer völlig andersartigen

Eigenschaft (Experientialität) postuliert, die nicht in den bereits vorhandenen (nicht-experientiellen) Eigenschaften angelegt ist.

Eine anderes Argument ergibt sich aus der 3. Prämisse: Laut Seager (2006) folgt nämlich die 4. Prämisse von der Nicht-Emergenz des Bewusstseins aus dem Argument intrinsischer Naturen, aber nur, falls zusätzlich eine Position vertreten wird, die Seager (2006, S. 131) als „Prinzip der Reduzierbarkeit relationaler Eigenschaften“ bezeichnet. Demnach sind alle relationalen Eigenschaften eines Gegenstandes auf die intrinsischen Eigenschaften reduzierbar (starke Version) oder supervenieren zumindest auf diesen (schwache Version) – und nicht umgekehrt. Wenn nun Bewusstsein eine solche intrinsische Eigenschaft darstellt, und alle gewöhnlichen physikalischen Eigenschaften relationale sind, so ist nicht ersichtlich wie Bewusstsein als emergentes Phänomen betrachtet werden könnte; im Gegenteil wäre es ja der „Urgrund“ aller physikalischen Eigenschaften schlechthin. Etwas wird hier deutlich, was vielleicht im normalen Anti-Emergenz-Argument weniger explizit ist: Bewusstsein (oder „experientielle Eigenschaften“) sind fundamentaler als alle materiellen Eigenschaften, was die Gleichwertigkeit materieller und bewusster Eigenschaften zugunsten des Bewusstseins aufhebt. Der *Panpsychismus* ist letztlich weniger neutral als vielleicht anfangs vermutet.

- Mikropsychismus:

Glaubt man, dass alle physikalischen Gegenstände aus kleinen, fundamentalen Bausteinen (Atomen, Elementarteilchen, Strings...) zusammengesetzt sind, dann folgt mit Prämisse 4, dass zumindest einige, dieser Bausteine über experientielle Eigenschaften verfügen („*micropsychism*“; Strawson 2006, S. 25). Aus der 5. Prämisse folgt, dass die sparsamste Ontologie diejenige ist, die nicht zwischen experientiellen und nicht-experientiellen Teilchen unterscheidet – was sollte denn auch zeigen können, dass z.B. Elektronen experientiell sind, Quarks hingegen nicht? –, dass also alle materiellen Bausteine (eine primitive Form von) Bewusstsein haben, genauso, wie sie auch alle eine Masse haben. In einer Ontologie, wie Strawson sie fordert, sollen letztlich die Dinge als identisch zur Summe ihrer Eigenschaften angesehen werden. Dies führt dann dazu, dass die unterschiedlichen *Qualia*, allein oder in Bündeln mit anderen Eigenschaften, die letzten Bestandteile der Wirklichkeit bilden. Ein Vorwurf, der sich hier aufdrängt, lautet, dass Strawson (recht un kreativ) die experientiellen Eigenschaften zu den Elementen der fundamentalen Physik „hinzuaddiert“ (Smart 2006, S. 158). Smarts Grund, hier besonders skeptisch zu sein, ist vor allem methodologisch: Die Existenz von Erlebnisgehalten folge aus introspektiver Betrachtung, während wir über die Eigenschaften der physikalischen

Elementarteilchen durch mathematische Theorie und Experimente informiert werden. Diese beiden Arten des Wissenserwerbs scheinen diametral entgegengesetzt zu sein und darum bleibt der Strawson'sche Physikalismus auch so suspekt, umso mehr, wenn er implizit zugleich seine Nichtüberprüfbarkeit im Rahmen der „gewöhnlichen“ Physik postuliert (vgl. auch das Argument der Intrinsic Naturen). Auch Goff (2006) sieht Schwierigkeiten darin, die unmittelbare Erkenntnis der Introspektion sinnvoll in das Gebäude der Mikrophysik einzugliedern: Gerade weil wir meinen, unmittelbar die Natur des Bewusstseins zu erfahren, schließen wir, ähnlich wie schon Descartes, auf eine metaphysische Kluft, die experientielle von nicht-experientialen Eigenschaften trennt, und begnügen uns nicht mit der Feststellung einer bloß epistemologischen Diskrepanz zwischen deren Beschreibungsweisen. Aber eine ähnliche Kluft würde auch zwischen den mikropsychischen Eigenschaften und der Psyche eines Organismus bestehen. Dann wäre aber auch die Entstehung des Bewusstseins eines Lebewesens ausgehend vom Mikropsychismus letztlich unverständlich, so Goff.

Die obigen Annahmen führen potentiell auf zwei Schwierigkeiten, die man als *Qualia*-Atomismus und Verdoppelung der Wirklichkeit bezeichnen kann. Unter *Qualia*-Atomismus soll die These verstanden werden, wonach sich Erlebnisgehalte, speziell die Inhalte von Wahrnehmungen, aus vielen atomaren *Qualia* zusammensetzen, ähnlich wie sich ein impressionistisches Gemälde aus distinkten Pinselstrichen zusammensetzt. Die Einheit des Bewusstseins ist dabei approximativ, gegebenenfalls neuronal vermittelt, ähnlich wie sich der Eindruck einer einheitlichen Szene erst ergibt, wenn das Bild aus etwas Entfernung betrachtet wird. Im Falle der *Qualia* führt dies jedoch auf ein Problem, das als „Kombinationsproblem“ des Panpsychismus bezeichnet wird (Seager 1995) und auf William James zurückgeht: Betrachtet man Bewusstsein als zusammengesetzt aus *Qualia*, mit je eigener Subjektivität und phänomenalen Charakter, bleibt unklar, wie aus dieser Summe phänomenaler Einzelergebnisse ein einziges, sich selbst als Einheit wahrnehmendes Bewusstsein entstehen soll. Das Kombinationsproblem umfasst sowohl die Schwierigkeit, mehrere Subjekte zu einem einzelnen zusammenzufassen, als auch die Schwierigkeit, zu verstehen, wie einzelne Sinneseindrücke (etwa rot, flach und rau) zu einer Einzelwahrnehmung verschmelzen (eine rote, raue Fläche) und wie phänomenale Strukturen (etwa Melodien oder räumliche Distanzen) aus strukturlosen Einzelwahrnehmungen entstehen kann (Chalmers 2016).

Dass es sich bei den Beispielen, welche das Kombinationsproblem veranschaulichen sollen, primär um Wahrnehmungsphänomene handelt, ist dabei kein Zufall. Ginge man zusätzlich davon aus, dass *alle* mentalen Prozesse, die introspektiv zugänglich

sind, über einen je eigenen phänomenalen Charakter verfügen, würden sich die Schwierigkeiten noch vervielfachen. Das Kombinationsproblem ist dabei nicht auf den Panpsychismus beschränkt. Es stellt sich für *alle* Positionen in der Philosophie des Geistes, die versuchen Bewusstsein als Summe mehrerer Teilaspekte zu verstehen. Es betrifft somit nicht nur den Panpsychismus, sondern auch den Materialismus, wenn dieser versucht, die Einheit des Bewusstseins durch synchrone Anregung verschiedener, lokalisierter Hirnfunktionen zu erklären. Das Kombinationsproblem, auch wenn es erst im Panpsychismus zu einem offensichtlich *metaphysischen* Problem wird, verdeutlicht eine strukturelle Schwierigkeit mit der wir konfrontiert sind, wenn wir versuchen, ein einheitliches Bewusstsein aus einer Vielheit von Einzelerfahrungen oder Teilzuständen zu verstehen.

Kann das Kombinationsproblem nicht gelöst werden und identifiziert man die *Qualia* mit der (intrinsischen) Natur der Dinge, führt dies jedoch ungewollt zu einer Verdopplung der Wirklichkeit. Einerseits existieren experientielle Eigenschaften auf der Mikroebene, andererseits existiert ein Bewusstsein, das sich aber nicht als Summe der mikropsychischen Eigenschaften erklären lässt. Der Erlebnisgehalt ist dann nicht die Art und Weise, wie *physikalische* Eigenschaften dargestellt werden, wofür etwa Dretske (1995) argumentiert, sondern entspricht einer „objektiven“ Eigenschaft der Wirklichkeit, die schließlich im Bewusstsein wahrgenommen wird.

Doch das ursprüngliche Ziel, aufgrund dessen die *Qualia* verdinglicht wurden, wird dadurch unterlaufen. Es sollte die Kluft zwischen zwei Wirklichkeiten, einer wahrgenommenen (empfundenen) und einer real existierenden, überwunden werden. Dabei wurde jedoch wiederum eine Verdoppelung als an-sich existierend und als wahrgenommen postuliert – diesmal allerdings bezogen auf die Instanziierung der *Qualia*. Wie beim Kombinationsproblem sind sowohl der Materialismus als auch der Mikropsychismus wieder mit den (strukturell) selben Schwierigkeiten konfrontiert. Die Ironie dabei ist, dass beide Positionen vorgeben, metaphysische Probleme der jeweils anderen aus dem Weg zu räumen.

### 1.3 Methodisches

Hier sollen in drei Schritten kurz die wichtigsten methodologischen Erwägungen diskutiert werden, die für den weiteren Verlauf dieser Arbeit von zentralem Interesse sind: Eine Darstellung von Bewusstseinsphänomenen sollte eine *mathematisch* exakte sein; sie sollte *systemisch-kontextuell* sein und sie sollte letztlich *produktiv* sein.

## Über die Rolle von Mathematisierungen

Es gibt mehrere Motive, aus denen heraus die Formalisierung empirischer Gegenstandsbereiche erfolgen kann. Neben der Rechtfertigung einer gewissen Erkenntnispraxis dienen Formalisierungen oftmals als methodisches Handwerkszeug: Dies gilt insbesondere mit Hinblick auf die Ingenieurwissenschaften, die daran interessiert sind, dass die Maschinen, die sie bauen, bestimmte Funktionen zuverlässig erfüllen, oder mit Hinblick auf naturwissenschaftliche Verfahren zur Vorhersage, etwa von Wetterphänomenen oder Sternkonstellationen.

Doch wozu bräuchte es überhaupt eine Formalisierung von *Bewusstseinsphänomenen*? Entzieht sich nicht gerade das Bewusstsein jeglicher Formalisierung? Ist nicht gerade erst Bewusstsein die Voraussetzung dafür, dass überhaupt Formalisierungen getroffen werden können?

Eine Motivation für den Versuch, eine formale Wissenschaft des Bewusstseins zu etablieren, liegt in der Hoffnung auf medizinische, besonders psychiatrische, Anwendbarkeit, aber es lässt sich kritisch fragen, ob hierfür gerade der Weg der Mathematisierung der erfolgversprechendste ist, auch wenn einige Arbeiten aus dem Umfeld der Naturwissenschaften dies vielleicht nahelegen. So versuchen etwa Carhart-Harris & Friston (2010) die Psychoanalyse auf ein formalisiertes, neurobiologisches Fundament zu stellen. Die informationstheoretische Bewusstseinstheorie von Tononi (2008) soll klären können, wann und ob Patienten in einem komatösen Zustand bei Bewusstsein sind und auch in der Anästhesie ist man daran interessiert, seltene Phänomene wie unerwünschtes Aufwachen oder die Erinnerung an einen operativen Eingriff zu verhindern [14]. Eine molekulare und wesentlich formalisierte Theorie des Bewusstseins würde sicherlich dazu beitragen (Hameroff 2001).

Dem gesellt sich jedoch noch ein weiteres Motiv hinzu: Eine Tradition, die seit den Anfängen der Philosophie in der Antike einen großen Einfluss auf das wissenschaftliche Denken ausübt, sieht in der Formalisierung eines Gegenstandsbereichs einen wichtigen Schritt zum Verständnis desselben. Gerade in den physikalisch-chemischen Grundlagenwissenschaften ist es oft dieses Motiv, welches die Verwendung komplexer mathematischer Verfahren zur Erforschung der Natur rechtfertigte. Wenn Galileo feststellt, dass im Buch der Natur mit mathematischen Zeichen geschrieben wird, so ist dies nicht bloß als rein pragmatische Rechtfertigung für die Verwendung von Mathematik zu verstehen, sondern Ausdruck der Überzeugung, dass echtes Naturverständnis erst durch die Klarheit und Systematizität der Mathematik erreicht werden kann. Wenn wir nun Bewusstsein als natürlichen Gegenstand verstehen wollen, bedingt dies notwendigerweise eine Formalisierungsleistung; nicht, um dadurch das Phänomen technologisch operational werden zu lassen, sondern, um es in eine



bestimmte Tradition der Erkenntnis einzureihen.

Ein grundlegender Einwand gegen die Idee, über Bewusstsein ließe sich in mathematischer Weise sprechen, geht davon aus, dass eine mathematische Darstellung mit einer „quantitativen“ Erfassung gleichzusetzen wäre. Dann nämlich erschiene es widersinnig, auf diese Weise über ein scheinbar „rein qualitatives“ Phänomenen wie das Bewusstsein sprechen zu wollen. Verstehen wir aber unter Mathematisierung in erster Linie die formale Darstellung von Beziehungen, ist nicht klar, warum denn Bewusstsein zwangsweise außerhalb der Tradition mathematischer Naturerkenntnis stehen sollte. Allerdings folgt daraus wiederum nicht, dass es nicht einige ganz grundsätzliche Schwierigkeiten gibt, die mit diesem Formalisierungsvorhaben zusammenhängen. Eine besondere Aufgabe, der im Laufe dieser Arbeit nachgekommen werden soll, ist daher das Sichtbarmachen eben jener Schwierigkeiten.

Aus dieser Perspektive betrachtet, liegen dem *hard problem* zwei gegenläufige Überzeugungen zugrunde: Einerseits gilt den Vertretern der Bewusstseinswissenschaften aus Abschnitt 1.1 nämlich, dass alles, was wir zu wissen meinen, darauf hindeutet, dass Bewusstsein als empirisches Phänomen zu verstehen sei. Andererseits scheinen wir aber nicht mit denselben formalen Mitteln über Bewusstsein sprechen zu können, mit denen wir uns auf alle anderen empirischen Phänomene beziehen.

Was wäre dann die Form, die eine Theorie von Bewusstsein haben müsste? Die Vorstellung, dass es sich beim Bewusstsein um ein „inneres“ undifferenziertes Wahrnehmen „äußerer“ Prozesse handelt, steht der alltäglichen Beobachtung gegenüber, wonach Bewusstsein seiner Phänomenologie nach höchst differenziert ist – man denke etwa an die „von Außen nach Innen dringenden“ Sinneswahrnehmungen, die Gedanken, welche uns „im Kopf passieren“ oder, allgemein gesprochen, die räumliche und zeitliche Strukturiertheit der Inhalte unseres Erlebens. Und handelt es sich bei der Innen/Außen-Dichotomie überhaupt um eine essentielle Eigenschaft des Bewusstseins oder lediglich um ein „dualistisches Vorurteil“ (Velmans 2009, S. 121)? Letztlich scheint es unzweifelhaft, dass Bewusstsein eng mit physischen Prozessen in Verbindung steht. Es ist nun für eine Theorie des Bewusstseins wichtig, dass die möglichen Wechselwirkungen oder *gegenseitigen Bedingtheiten* zwischen Bewusstsein und jenen physischen Prozessen aufgeklärt werden können. Dies impliziert aber, dass es hier um wesentlich *relationale* Phänomene geht.

Eine Unterscheidung, die in der Reflexion des kartesischen Dualismus bereits angeklungen ist und deren Bedeutung im Verlaufe der Arbeit noch weiter herausgearbeitet werden soll, ist die zwischen einer Theorie *des Bewusstseins* und einer Theorie von *Bewusstsein*. Letztere interessiert sich primär dafür, in welche Aktivitäten und Prozesse Bewusstsein eingebunden ist, während erstere vor allem daran interessiert ist, deutlich zu machen was das Bewusstsein nun „eigentlich sei“ – Bewusstsein wird

also das eine Mal als Art und Weise, wie sich ein bestimmter Prozess vollzieht, das andere Mal als Gegenstand, der zum Objekt wissenschaftlicher Beschreibung wird, verstanden.

Die meisten Ansätze aus der zeitgenössischen Bewusstseinswissenschaft versuchen diesen letzten Aspekt aufzuklären, etwa wenn sie Bewusstsein mit einer bestimmten Hirnfunktion oder einem bestimmten neuronalen Muster gleichsetzen. Genau hier setzt dann aber auch das *hard problem* an, wenn es diesen Übergang grundsätzlich in Frage stellt mit Verweis darauf, dass eben der „qualitative Aspekt“ des Bewusstseins niemals in neuronaler Erregung gefunden werden könne.

### Systemtheorie?

Da es scheinbar relationale Aspekte sind, die für ein Verständnis von Bewusstsein entscheidend sind, ist es für viele naheliegend, dass es sich bei einer Theorie des Bewusstseins um eine Art Systemtheorie (Thompson & Varela 2001, Juarrero 1999), eventuell kombiniert mit einigen metaphysischen Annahmen (Silberstein 2014), handeln müsse.

Angelehnt an Überlegungen von Ludwig von Bertalanffy (1968), gibt es hauptsächlich drei Gründe, die für die Einbeziehung systemischer Überlegungen sprechen, wie sie etwa in der Chaos- und Systemtheorie zu finden sind:

+ Die *epistemologische* Sichtweise:

Geistige Phänomene, insbesondere jene des Bewusstseins, sind komplex, so dass eine detaillierte und reduktionistische Beschreibung nicht möglich ist und somit eine systemische Beschreibungsebene gewählt werden muss.

Ein Vergleich zu Reaktionsnetzwerken, wie sie etwa in der Chemie, den Umwelt- oder den Ingenieurwissenschaften vorgefunden werden, soll dies verdeutlichen: Auch wenn wir in der Lage sind, komplexe Systeme in einfache Teilsysteme zu zerlegen, welche wir im Einzelnen mittels Differentialgleichungen exakt beschreiben können, so ist es möglich, dass das Gesamtsystem ein nicht-lineares Verhalten zeigt, das so sensitiv gegenüber den Anfangsbedingungen ist, dass sich keine genauen, quantitativen Prognosen im Einzelfall machen lassen.

Wir können jedoch wissenschaftlich exakte Aussagen über das Systemverhalten tätigen, indem wir einen systemischen Standpunkt einnehmen, aus dem dann z.B. die Beschreibung eine Klasse möglicher Dynamiken hin zu einem Attraktor folgt.

Die epistemologische Sichtweise bezieht sich dabei also in erster Linie auf unser

Unvermögen. Unser Unwissen fordert uns dazu auf, einen anderen, eben den systemischen, Standpunkt zu wählen.

+ Die *system-realistische* Sichtweise:

Es existieren Regelmäßigkeiten (und Brüche) zwischen den Verhältnissen von Teilen und Ganzheiten auf verschiedenen Beschreibungsebenen. Dies legt die Untersuchung der entsprechenden Strukturen nahe, die dann als Systemeigenschaft interpretiert werden können.

Dieser Ansatz wird oft mit der epistemologischen Sichtweise verwechselt. Anders als dort wird hier aber davon ausgegangen, dass es Strukturähnlichkeiten auf unterschiedlichen Beschreibungsebenen gibt, die nicht in unserem prinzipiellen Unvermögen zu einer detaillierten Analyse begründet sind, sondern tatsächliche Eigenschaften natürlicher Systeme widerspiegeln – ein beliebtes Beispiel in den 1970er und 1980er Jahren war die Selbstähnlichkeit fraktaler Objekte.

+ Die *transzendente* Sichtweise:

Beim Nachdenken über das Bewusstsein sind wir mit den Strukturen unseres Denkens selbst konfrontiert: Wir erfahren nur dann etwas über das Bewusstsein, wenn wir gleichzeitig etwas darüber erfahren, wie wir überhaupt Erfahrungen machen können. Dies scheint auf den ersten Blick zirkulär zu sein, bedeutet aber lediglich, dass der Erkenntnisgegenstand gleichzeitig (Teil der) Bedingung der Möglichkeit von Erkenntnis ist. Eine wissenschaftliche Analyse würde dann diesen Gegenstand nicht einfach abbilden („es existiert ein Ding, mit diesen und jenen Eigenschaften“), sondern würde aufzeigen, wie diese erzeugt werden kann („es existiert eine Tätigkeit, welche selbst Erfahrungen über diese Tätigkeit sammeln kann“).

Theorien, die beschreiben, wie sich bestimmte Entitäten selbst hervorbringen (etwa die Theorie der „*Autopoiesis*“ von Maturana & Varela 1980), scheinen hier einen ersten Ansatz zu liefern, insbesondere wo es diese explizit mit dem Thema der Selbstbezüglichkeit zu tun haben (Varela 1975, Kauffman 2003).

Allerdings gibt es auch einige kritische Einwände gegen Systemtheorien, vor allem wenn sie das Bewusstsein betreffen sollen.

– Kontextualität:

Fast alles, was auf einer Ebene als geschlossenes System angesehen werden kann, muss auf einer anderen Ebene als Komponente eines größeren Systems behandelt werden. In einer anatomischen Untersuchung der einzelnen Zelle

kann das Neuron als Ganzheit betrachtet werden, in einem anderen Kontext muss sein Verhalten als Teil des Nervensystems betrachtet werden. Aber ob wir etwas als Systemganzes oder als Teil betrachten ist relativ zu einer Beschreibung, die auf einen intentionalen Kontext verweist. Ähnlich merkte bereits W. Ross Ashby (1962) an, dass der Begriff der Organisation nicht ausschließlich auf die „intrinsic Eigenschaften der Gegenstände“ verweisen würde, sondern wesentlich abhängig von der Beziehung zwischen Beobachter und den Gegenständen sei.

Dies veranschaulicht ein übergeordnetes Problem der Kontextualität *aller* Theorien von Bewusstsein. Nicht nur in Systemtheorien, sondern, wie wir später noch sehen werden, in Theorien der Information, der Emergenz oder in der Psychologie von William James entscheidet der Kontext, welcher Art ein Gegenstand ist, also ob er einen semantischen Gehalt trägt oder nicht, ob er ein emergentes oder fundamentales Phänomen darstellt, respektive ob er das Subjekt oder das Objekt eines Gedankens bildet. Kontextualität, so könnte vermutet werden, ist aber wesentlich ein Phänomen, welches ein (unterscheidendes) Bewusstsein bereits voraussetzt [15].

– Bewusstsein versus Kognition:

Bis jetzt wurde eine Unterscheidung ausgeklammert, die in der aktuellen Diskussion einen prominenten Platz einnimmt, nämlich die zwischen dem (phänomenalen) Bewusstsein und der Kognition. Dies hängt sehr eng mit dem *hard problem* zusammen: Während Kognition, dargestellt als Mechanismus, Teil einer objektiven Naturbeschreibung sein könne, entziehe sich das phänomenale Bewusstsein einer solchen Beschreibung, *und zwar auch dann, wenn diese eine systemtheoretische wäre.*

Allerdings ist unklar, wo überhaupt die Grenze zwischen Bewusstsein und Kognition verläuft: Sind Denkprozesse, sofern sie bewusst ablaufen, nicht auch in einer charakteristischen Weise phänomenal? Fühlt es sich nicht auch irgendwie an, eine Sache zu verstehen? Sind die Inhalte des (phänomenalen) Bewusstseins vielleicht sogar eng an kognitive (und physiologische) Prozesse gekoppelt, ohne auf diese reduzierbar zu sein (Prentner 2016)? Geht es, wie Gallagher & Zahavi (2008) vermuten, nicht vielmehr um Subjektivität, welche das Bewusstsein gegenüber der (unbewussten) Kognition auszeichnet? Falls eine dieser Fragen bejaht würde, spricht dies aber eher *für* eine (erweiterte) Systemtheorie, die wesentlich darum bemüht ist, die unterschiedlichen (phänomenalen) Aspekte von Erfahrung zu berücksichtigen und Bewusstes und Unbewusstes als zusammenhängende Glieder innerhalb einer natürlichen Abfolge zu verstehen.

- Empirische und nicht-empirische Voraussetzungen:

Eine Kritik an der Theorie der Autopoiesis stammt von Alfred Locker (1981) und gilt stellvertretend für alle Ansätze, die Selbstbezüglichkeit als rein materielles Phänomen darstellen wollen. Dabei ist Lockers Unterscheidung von Bedingtheiten und Vorannahmen („preconditions“ vs. „presuppositions“; Locker 1998) zentral, der auf den Unterschied zwischen *empirischen* Vorbedingungen, welche im Rahmen des Formalismus der Autopoiesis dargestellt werden, und den *transzendentalen* Vorannahmen, die einen meta-theoretischen Standpunkt erfordern, zielt. Dass die Natur spontan geschlossene und selbstbezügliche Systeme bilde, sei, so Locker, unerklärlich ohne dabei die Existenz einer produzierenden Tätigkeit schon vorauszusetzen; um selbstbezügliche Systeme als solche erkennen zu können, müsste bereits ein selbstbezügliches Bewusstsein vorhanden sein. Dies scheint ein Echo Kants zu sein, der in der Kritik der Urteilskraft bereits davon sprach, dass „das Prinzip der Zweckmäßigkeit der Natur [...] ein transzendentales Prinzip [sei]“ [16].

Systemtheorie muss also mit Vorsicht genossen werden und stellt weniger eine fundamentale Theorie zur Erklärung „von eigentlich allem“ dar, als vielmehr ein punktuell nützliches Werkzeug, um eine Reihe von Phänomenen überhaupt erst sichtbar zu machen.

## Objekte als Resultate von Prozessen

Eine formal-wissenschaftliche Theorie von Bewusstsein zu etablieren, stößt anscheinend auf größere Schwierigkeiten als es der Ansatz, Bewusstsein einfach als Objekt analog zu Elektronen, Zahnrädern oder Genen zu betrachten, vermuten ließe [17].

Wir werden im weiteren Verlauf dieser Arbeit zuerst zwei Arten von Theorien besprechen, die man als „*bottom-up* und *top-down* Theorien des Bewusstseins“ bezeichnen kann. Letztere sind mit den Schwierigkeiten konfrontiert, die sich immer dann ergeben, wenn wir versuchen, ein „intuitiv oder unmittelbar erkanntes“ Bewusstsein mit bestimmten physikalischen Gegenständen (oder mit Vorgängen innerhalb oder Beziehungen zwischen physikalischen Gegenständen) zu identifizieren.

In der Literatur wurden Einwände gegen ein solches Vorgehen an vielen Stellen vorgebracht. Wir werden dies exemplarisch am Argument der invertierten *Qualia* diskutieren und an einem aktuellen Beispiel aus der Bewusstseinsforschung veranschaulichen. Dabei geht es weniger darum, ob das Argument invertierter *Qualia* ein gutes philosophisches Argument ist oder wie man allfällige dialektische Hintertürchen schließen könnte, sondern darum, welche Einschränkungen für eine ma-

thematische Theorie des Bewusstseins aus ihm folgen. Diese Einschränkungen, so scheint es, sind aber dergestalt, dass das Vorhaben, derartige Theorien des Bewusstseins zu errichten, als sinnlos erscheinen, weil sie ihr zentrales Versprechen gar nicht einlösen *können*.

Daneben gibt es aber auch die raffinierteren *bottom-up* Theorien des Bewusstseins, die jedoch zur unerfreulichen Konsequenz führen, dass sie Bewusstsein gar nicht erklären, sondern scheinbar obsolet machen: Dort nämlich, wo eine bestimmte Funktion, etwa mit Verweis auf die Arbeitsweise des Gehirns, erklärt werden kann, scheint Bewusstsein gar keine Rolle mehr zu spielen. Diese Konsequenz ist umso unerfreulicher, als dass es ja genau darum geht, einen theoretischen Ansatz dafür zu liefern, wie vom Bewusstsein als wirklichem und nicht bloß als illusionärem Gegenstand gesprochen werden kann.

*Top-down* und *bottom-up* Theorien können sehr allgemeinen anhand von formalen, strukturalistischen Modellen diskutiert werden, die davon ausgehen, dass in diesen Theorien eine gegenständliche und prinzipiell erfahrbare Wirklichkeit abgebildet oder modelliert würde. Die aussichtsreichsten Ansätze zu einer Theorie des Bewusstseins berufen sich dabei auf den Begriff der Emergenz: Emergent sind systemische Eigenschaften, die als neuartig in Erscheinung treten und dabei nicht bereits als Eigenschaften der Komponenten vorliegen. Ein Versprechen vieler sog. Emergenztheorien ist es nun, dass alles, was wir kolloquial als Eigenschaft des Bewusstseins bezeichnen, als emergente Eigenschaft materieller (neuronaler) Systeme umgedeutet werden kann, kurz: dass „Bewusstsein“ eben nichts anderes als der Name für einen Komplex emergenter physikalischer Phänomene ist.

Solche Emergenztheorien können anhand verschiedener Kriterien klassifiziert werden, wobei eine Unterscheidung von schwachen und starken Auffassungen sehr geläufig ist. Die vielleicht interessanteste Klasse von Emergenztheorien bilden hingegen diejenigen, welche gerade *nicht* in dieses Schema fallen und z.B. durch Topologien (Primas 1998, Bishop & Atmanspacher 2006), Musterbildung (Dennett 1991*b*) oder die (diachrone) Neuartigkeit von Strukturen (Stephan 1999, Kap. 4) näher expliziert werden können. Die (emergenten) Objekte dieser Theorien können nun allerdings nicht mehr als Abbilder unabhängig existierender Gegenstände „da draußen“ angesehen werden. Den Schwärmen, den Gleitern und Seglern oder den klassischen Molekülgestalten – um beliebige Beispiele der jeweiligen Emergenztheorien aufzugreifen – entsprechen keine eigenständigen physikalischen Gegenstände mehr, sondern eben nur noch Muster oder strukturierte Ganzheiten. Sofern sie aber als einfache und selbstgenügsame Objekte aufgefasst werden, sind sie relativ zu einer Beschreibung zu verstehen, in der diese als Zeichen vorkommen. Was letztere auszeichnet und wie es zu deren Erscheinen kommt, ist jedoch – anders als in einigen mentalistischen

Deutungen, in denen ein bezeichnender Geist einer bezeichneten Materie kategorisch gegenübersteht – nicht willkürlich.

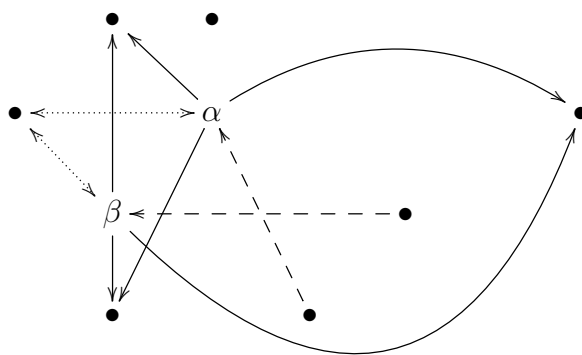
Uns interessiert hier vor allem die Bedeutung, die dies für eine Theorie von Bewusstsein und Geist hat. Paradigmatisch sei daher die Wahrnehmung diskutiert, einerseits im Kontext einer zeitgenössischen neuronalen Theorie und andererseits als intentionaler Prozess, wobei hierfür ein geeignetes formales Vokabular entwickelt werden soll. Dazu soll an die formale Darstellung der Beziehung von Teilen und Systemganzen angeknüpft werden, wie sie auf unterschiedliche Weise etwa von Stanislaw Leśniewski, Edmund Husserl oder Nelson Goodman konzipiert wurde und in einer einheitlichen, mathematischen Sprache formuliert werden kann (Simons 1987). Eine solche „Mereologie“ verspricht neue Ansätze zur Beschreibung relationaler Phänomene, welche in Folge ausgearbeitet und schließlich angewandt werden sollen. So wird bewusste Wahrnehmung als Herausbildung verweisender Objekte auf der Basis einer noch nicht differenzierten Mannigfaltigkeit von Teilobjekten, die selbst noch gar nicht auf bestimmte Gegenstände außerhalb des Bewusstseins verweisen, beschrieben werden. Zur „Wahrnehmung *von*“ korreliert also die Emergenz eines „Objektes *für*“. Dies kann Thema einer naturalistisch gefassten Emergenztheorie sein, was jedoch nicht gleichbedeutend dazu ist, dass Bewusstsein als „emergent“ zu bezeichnen wäre.

Dies lässt sich an einer weiteren Metapher gut veranschaulichen. Das Entstehen von Objekten im Bewusstsein, lässt sich mit dem Aufwachen vergleichen: Während wir uns müde aus dem Bett erheben, während uns vielleicht der Duft des morgendlichen Kaffees in der Nase liegt oder während die ersten Sonnenstrahlen unsere Augen reizen, formieren sich allmählich die ersten konkreten Gedanken in unserem Kopf – seien es die Sorgen des letzten Tages, die langsam in unser Bewusstsein zurückkehren, oder seien es die ersten Objekte, die wir bewusst wahrnehmen, während wir unsere tägliche Routine starten. Dabei kann höchstens im übertragenen Sinn die Rede davon sein, dass hier unser „Bewusstsein entstände“. Viel eher handelt es sich um einen Übergang, der sich in unserem Bewusstsein vollzieht, ähnlich wie unsere Gedanken während des Einschlafens in die nächtlichen Träume abgleiten.

Die Struktur dieses Vorganges – wie Objekte im Bewusstsein entstehen – gilt es nun, explizit zu machen. Dabei kann nicht lediglich auf die Gegenstände außerhalb des Bewusstseins verwiesen werden, da, erstens, so gar nicht verständlich würde, warum eine Abfolge solcher Gegenstände überhaupt zu so etwas wie einer bewussten „Wahrnehmung von“ führen sollte, und da, zweitens, die Gegenstände selbst zuallererst als Objekte des Bewusstseins gegeben sind und erst dann zu Gegenständen außerhalb des Bewusstseins *abstrahiert* oder *hypostasiert* werden. Bewusstsein durch eine Konfiguration von Gegenständen außerhalb des Bewusstseins erklären zu wollen, hieße

somit, das Pferd von hinten aufzuzäumen.

Der Gedanke, dass Objekte im Bewusstsein entstehen, ließe sich weiter verallgemeinern, indem nicht mehr davon ausgegangen wird, dass eine Sorte von Dingen (die physischen) einer mit dieser *per Definition* unvereinbaren Sorte von Dingen (den psychischen) gegenübersteht. (Sollten wir uns wundern, dass eine Beschreibung, die für die eine Sorte so erfolgreich zu sein scheint, für die andere Sorte unbrauchbar ist?) Anstelle von (substantiellen) Dingen, die einem (substantiellen) Subjekt erscheinen, würde in deflationärer Weise lediglich von erfahrbaren Regelmäßigkeiten gesprochen werden, die in Prozessen realisiert sind und fallweise zur Emergenz von Objekten führen, was gleichbedeutend zu deren Wahrnehmung ist (vgl. Abb. 1.2).



**Abbildung 1.2:** Dinge als Knoten in Netzwerken. Die Objekte  $\alpha$  und  $\beta$  seien allein durch ihre Position innerhalb eines Netzwerke von Beziehungen (Verweisen, Reaktionen, Wechselwirkungen) definiert (nach Prentner 2017). Ihre Eigenschaften sind daher nicht *intrinsisch* sondern *relational*. Der Wahrnehmung und den Formen der bewussten Kognition entsprechen spezifische Konfigurationen eines solchen Netzwerkes.

Ein Argument für die Relationalität von Objekten (angelehnt an Rovelli 1996) lautet, dass

1. aus der Tatsache, dass es unterschiedliche (korrekte) Wahrnehmungen desselben Objektes gibt, und
2. aus der Überzeugung, dass es keine ausgezeichneten Objekte gibt (= Ablehnung eines Dualismus), folgt
3. dass die Eigenschaften von Objekten immer relativ zum entsprechenden Wahrnehmungskontext und nicht „absolut“ oder intrinsisch sind.



Andernfalls müssten wir uns auf einen absoluten Standpunkt begeben können, den „Blick von nirgendwo“ auf die Dinge richten und sie so abbilden, „wie sie nun einmal sind“. Wir erkennen die Dinge jedoch immer nur so, wie sie uns als Objekte gegeben sind.

Wenn wir nun aber gleichzeitig daran festhalten wollen, dass eine Bewusstwerdung dieser Objekte nichts Übernatürliches ist, so stellt sich folgende Frage: Welcher Naturbegriff liegt überhaupt vor, wenn er so sein soll, dass er Bewusstsein zulässt? Die vorläufige Antwort soll lauten, dass es einer ist, der von einer grundsätzlichen **Produktivität** ausgeht: Die Entstehung von unvorhersagbar Neuartigem ist möglich. *Die eine* Natur löst sich auf in eine Vielzahl scheinbar emergenter Prozesse. Der Glaube an eine natürliche Tätigkeit ist kein Anthropomorphismus; es ist die Vorstellung von Untätigkeit, welche die Folge unseres endlichen Denkens ist.

## Anmerkungen zu Abschnitt 1

[1] Für rein neurobiologische Erklärungsansätze siehe die Arbeiten von Crick & Koch (2003) und Koch (2005); ein kognitionswissenschaftliches Bewusstseinsmodell ist die *Global Workspace*-Theorie von Baars (1988) und deren neuropsychologische Ausarbeitung durch Dehaene & Naccache (2001); rein informationstheoretisch formuliert ist die *Integrated Information Theory of Consciousness* von Tononi (2008), welche Bewusstsein mit einer höherrangigen Form von Information in (kortikalen und anderen) Netzwerken gleichsetzt. Demgegenüber stehen Versuche, Bewusstsein im Rahmen der Fundamentalphysik erklären zu wollen. Insbesondere geschieht dies etwa bei Hameroff & Penrose (2014).

[2] In den Worte Galen Strawsons: „[T]here must be experience for there to be appearance.“ (Strawson 2006, S. 26).

[3] Die ersten beiden Annahmen wurden von Einstein *et al.* explizit als Kriterien der Realität und der Vollständigkeit erwähnt. Die dritte Annahme bleibt implizit, ist aber zentral für das Verständnis der gesamten Arbeit (Fine 2007) und ist bis heute Gegenstand eingehender experimenteller Untersuchungen, etwa von Gröblacher *et al.* (2007) oder Hensen *et al.* (2015).

[4] Bohm (1952a, 1952b). Die Bohm'sche Interpretation der Quantenmechanik ist unter praktizierenden Physikern (zu Recht oder zu Unrecht) nicht mehrheitsfähig, vermutlich da sie Entitäten wie das Quantenpotential postuliert, die vielen als zu *ad-hoc* erscheinen.

[5] Etwas weniger dramatisch ausgedrückt, gilt, dass bestimmte, natürlich erscheinende Annahmen über den wissenschaftlichen Realismus aufgrund der Vorhersagen der Quantenmechanik zurückgewiesen werden müssen (Gröblacher *et al.* 2007); als Beispiel sei hier die Auffassung von Esfeld (2008) erwähnt, der die Prinzipien der Separabilität, Lokalisiertheit und Individualität von Quantenobjekten zurückweist, welche aber unser intuitives Verständnis vom Realismus wesentlich prägen. Zudem ist anzumerken, dass natürlich auch mehrere der obigen Annahmen verletzt sein können; so könnte die Quantentheorie nicht-lokal *und* unvollständig sein oder es könnte, wie im Falle einer Variante der Kopenhagener Interpretation, sowohl das Kriterium der Realität als auch der Lokalität (nicht aber der Unmöglichkeit von Signalübertragung mit Überlichtgeschwindigkeit) verletzt sein.

[6] Für einen Überblick siehe z.B. Jaeger (2009).

[7] Galen Strawson (2006, S. 6f.) drückt dies wie folgt aus: „They [Dennett und die eliminativen Materialisten, Anm. RP] are prepared to deny the existence of experience, more or less (c)overtly, because they are committed to physicalism [...] This particular denial is the strangest thing that has ever happened in the whole history of human thought [...]

Next to this denial every known religious belief is only a little less sensible than the belief that grass is green.“

[8] In einem Aufsatz fragen sich Crane & Mellor (1990), ob Geistiges, wie es die Psychologie beschreibt, ein Phänomen darstellt, an dem sich der Physikalismus überprüfen ließe. Dabei kommen sie zum Schluss, dass mentale Phänomene mit keinem guten Grund als unphysikalisch betrachtet werden können. Somit existiert aber auch nichts, an dem sich der Physikalismus überhaupt messen lassen könnte. Der Physikalismus sei die falsche Antwort auf eine simple Frage: „[S]aying that minds are all physical no more helps to explain how some physical things can think than saying that all flesh is grass helps to explain the difference between carnivores and vegetarians“ (Crane & Mellor 1990, S. 206).

[9] In der Quantenmechanik kommt, je nach Interpretation, eine weitere begriffliche Trennung hinzu, und zwar diejenige zwischen System, Umgebung und Messapparatur („Heisenberg Schnitt“). Hierbei handelt es sich um einen weiteren Schnitt, *zusätzlich* zum kartesischen. Selbst wenn eine Theorie vorläge, die den Heisenbergschnitt wegerklären könnte, wie das etwa Ziel der Dekohärenztheorie ist, wäre damit der kartesische Schnitt noch nicht aufgehoben (Primas 1994). Dies zeigt sich zum Beispiel darin, dass eine Dekohärenztheorie immer noch auf zusätzliche Interpretationen, etwa die Viele-Welten-Theorie Everetts oder Zureks Quantendarwinismus, angewiesen ist, um das Messproblem zu lösen (vergleiche auch Schlosshauer 2008, Wallace 2012).

[10] Ein Argument, welches Popper für die Notwendigkeit anführt, die abstrakten Ideen (aus Welt 3) von konkreten Bewusstseinsprozessen (aus Welt 2) zu unterscheiden, ist dass jene oft Beziehungen untereinander aufweisen, die nicht Teil des Bewusstseins sind, welches sie hervorgebracht hat: Theorien haben z.B. logische Konsequenzen, die sich aus ihrer mathematischen Formulierung ergeben, die aber nicht bereits im Bewusstsein, welches sie entwickelt hat, angelegt waren. Einsteins Formulierung der speziellen Relativitätstheorie hat die Äquivalenz von Masse und Energie zur Folge. Dieser Relation war sich Einstein bei der Formulierung der speziellen Relativitätstheorie vermutlich gar nicht bewusst (Popper 1978, S. 158).

Ein ähnliches Argument lässt sich aber nicht nur für Ideen und Theorien anführen, sondern für alle möglichen Objekte. So wurde etwa die Dampfmaschine erfunden, um Wasser in einem Bergwerk zu heben. Dass jene letztlich im großindustriellen Maßstab Verbreitung fand, war nicht bereits Teil der Intentionen des Erfinders Thomas Newcomen, sondern hat vielmehr mit den Kontingenzen von Infrastruktur, Wartungs- und Produktionskosten zu tun. Umgekehrt scheint es jedoch genauso unplausibel zu sein, die Ideen von Newcomen lediglich als Ergebnis der materiellen Bedingtheiten des 19. Jahrhunderts zu betrachten.

[11] Weitere Kriterien, die mehr auf die strukturellen Eigenarten monistischer Theorien eingehen, finden sich bei Atmanspacher (2014).

[12] Und dass ein Teilchen neben einer Masse noch andere relationale Eigenschaften oder Dispositionen hat, unterscheidet es wiederum von anderen Teilchen, welche zufällig über die gleiche Masse verfügen. Genau da, wo eine solche Charakterisierung nicht mehr möglich ist, endet dann auch die Individuierbarkeit als „Teilchen“.

[13] Etwas anders ausgedrückt hat dies Godehard Brüntrup (2011), wenn er „intra-attributive“ von „inter-attributiver“ Emergenz unterscheidet. Im ersten Fall handelt es sich um Übergänge innerhalb derselben ontologischen Kategorie (Moleküle → Flüssigkeit), im zweiten Fall jedoch um problematische Übergänge zwischen verschiedenen Kategorien (Nicht-Experientielles → Experientielles).

[14] Dabei ist gar nicht so klar, wie selten diese Phänomene tatsächlich sind. Umstrittenen Schätzungen zufolge liegt die Wahrscheinlichkeit für ein unerwünschtes Aufwachen während der Narkose in den USA sogar bei 0.1 - 0.2%, siehe (Sebel *et al.* 2004).

[15] Ein ähnliches Argument wurde unlängst von Evan Thompson (2007) vorgebracht, wonach eine Anwendung der Theorie Dynamischer Systeme bereits ein verkörpertes Bewusstsein voraussetze.

[16] KU, B XXIX/ A XXVII; vgl. auch §75 der KU, B 333 - 338/A 329 - 335.

[17] Was einige Argumente gegen den Physikalismus, die in der zeitgenössischen analytischen Philosophie des Geistes von zentralem Interesse sind, auf elegante Weise verständlich machen würde: Da Bewusstsein und physikalische Objekte gar keine Entitäten derselben Art sind, folgt, dass keine *analytischen* oder *begrifflichen* Beziehungen zwischen ihnen hergestellt werden können – anders als „Blitze“, die mit atmosphärischer elektrischer Entladung identifiziert werden können, oder „Flüssigkeiten“, die bezüglich einer (strukturierten) Klasse von Molekülen emergent sind. Dies stellt aber den eigentlichen Kern der Irreduzibilitätsthese des Bewusstseins dar (vgl. Crane 2001*a*, Kap. 30).

## 2 Wider die Abbildung

*Früher habe ich einmal gesagt, der Satz sei ein Bild der Wirklichkeit. Damit könnte man zwar eine sehr nützliche Betrachtungsweise der Sprache ins Bild bringen, aber es besagt nicht mehr als 'Ich will den Satz als Bild der Welt betrachten'. (Ludwig Wittgenstein: Vorlesungen 1930-1935)*

### 2.1 Schwierigkeiten für eine gegenständliche Theorie des Bewusstseins

#### Das Mobiliar der Welt

Da, wo Wissenschaft mit dem *Gebrauch* von Begriffen und Kalkülen zur Erklärung und Formalisierung empirischer Phänomene gleichgesetzt wird, wird Philosophie oft als das *Überdenken* eben dieser Begrifflichkeiten angesehen [1]. So wird jede Auseinandersetzung mit den Problemen des Bewusstseins, gerade auch dann, wenn sie in Teilen auf eine wissenschaftliche Erklärung hinauslaufen soll, immer auch ein gewisses Maß an Philosophie beinhalten müssen und zwar nicht nur im reflexiven Sinne einer „Bestandsaufnahme“, sondern auch im revolutionären Sinne einer Herstellung alternativer Verknüpfungsweisen von Phänomenen.

Zu der im vorigen Abschnitt festgestellten Unbestimmtheit der disziplinären Zugehörigkeit von Bewusstseinsphänomenen gesellt sich nun die Einsicht, dass auch auf der Ebene der Begrifflichkeit noch völlige Unklarheit darüber besteht, wie mit dem Bewusstsein wissenschaftlich umzugehen sei. Aus diesem Grund ähneln die Versuche der gegenwärtigen Philosophie, ein materialistisches, panpsychistisches oder doch dualistisches Bild des Geistes zu zeichnen, den Versuchen aus dem 18. Jahrhundert, ohne die Begriffe „DNA“ oder „Selektion“ die Phänomene des Lebens zu ergründen. Anstelle einer inhaltlichen Analyse der Einzeldebatten über Zombies oder über metaphysische Notwendigkeit soll daher im Folgenden der konzeptionelle Rahmen abgesteckt werden, in dem eine Theorie des Bewusstseins entwickelt werden kann.

Eine gemeinhin als „natürlich“ geltende Einstellung, die im Lichte einer empirischen Wissenschaft des Geistes, überdacht werden muss, ist die des *Abbildungsrealismus* und dessen Konsequenzen für eine Theorie von Bewusstsein. Dabei soll unter einem *Abbild* eine wahrheitsgetreue und wesentlich vollständige Darstellung verstanden werden. *Realistisch* ist sie dann, wenn sie sich auf eine objektive (subjektunabhängige) und empirisch zugängliche Wirklichkeit bezieht, deren Entschlüsselung

als das Ziel von Wissenschaft angesehen werden kann.

Spätestens wenn Bewusstsein ins Zentrum wissenschaftlichen Fragens rückt, muss die Angemessenheit einer solche Einstellung hinterfragt werden [2]. Eine konzeptuelle Vorstufe zum Abbildungsrealismus bildet die Vorstellung, dass im Bewusstsein eine Abbildung oder Spiegelung einer objektiven Wirklichkeit vorliegen würde. Und auch wenn wir nicht mehr daran glauben, dass die Eigenschaften, die wir wahrnehmen – etwa die Farben der Gegenstände –, wirkliche Eigenschaften der Dinge sind, gehen wir meist davon aus, dass sich unsere Wahrnehmung dennoch auf unabhängig existierende Gegenstände beziehen, die z.B. eine bestimmte Ausdehnung, Lage im Raum oder Masse besitzen. In der Wahrnehmung „erscheinen“ uns dann diese Gegenstände, wobei wir im Nachhinein zwischen den wahrgenommenen Eigenschaften und ihren tatsächlichen unterscheiden.

Eine verwandte Idee ist die Vorstellung, dass einige sprachliche Ausdrücke Abbildungen von „inneren“ psychischen Zuständen sind. Auch hier gilt wiederum, dass wir nicht alle diese Ausdrücke wörtlich nehmen – unsere Rede ist zumeist metaphorisch –, aber dennoch gehen wir in der Regel implizit davon aus, dass sie sich auf einen vorhandenen Gegenstand beziehen, ein introspektiver Bericht scheint ja von einem bestimmten Vorgang zu handeln, der wirklich, wenn auch nur „im Kopf“ stattfindet.

Diese Gedanken lassen sich verallgemeinern, indem wir nicht mehr auf die Dinge „da draußen“ oder „im Kopf“ selbst verweisen, sondern auf die Beziehungen, Formen oder Strukturen, in welche diese eingebettet sind. Letztlich gelangt man auf diese Weise vielleicht zum Schluss, dass es wenigstens die besten Theorien sein sollten, die auf irgendeine Weise Abbilder der (Struktur der) Wirklichkeit darstellen. Wirklich wäre nur, was sich objektiv über die Elemente dieser Theorien (respektive deren Struktur) sagen ließe; das *Mobilier der Welt* könnte an unseren besten Theorien abgelesen werden.

Selbst wenn uns also weder unsere Intuitionen, noch unser Alltagsverstand oder unsere (inneren wie äußeren) Wahrnehmungen ein Bild der Wirklichkeit, „wie sie nun einmal ist“, präsentieren, so finden wir vielleicht Trost in der Vorstellung, dass uns zumindest die Wissenschaft eine subjekt-unabhängige Darstellung der Wirklichkeit liefern könnte.

Gegen die Vorstellung, unsere Theorien würden die (Struktur der) Gegenstände abbilden, wie sie außerhalb und unabhängig von unserem Bewusstsein existieren, sprechen die inhärenten Schwierigkeiten, die sich aus dieser Auffassung ergeben, wenn etwas wie eine Theorie von Bewusstsein angestrebt wird. Dies soll nun in mehreren Schritten weiter problematisiert werden, bevor in den späteren Abschnitten dieser Arbeit eine alternative Sichtweise vorgestellt wird, bei dem gerade der Aspekt der

Verobjektivierung im Bewusstsein betrachtet werden soll, allerdings nicht in einer phänomenologisch-narrativen Weise, sondern mithilfe einer geeigneten mathematischen Darstellung: Wenn wir davon sprechen, dass etwas als *Phänomen* im Rahmen einer Naturbeschreibung zu begreifen sei, dann stellt sich sofort die Frage, was es denn überhaupt bedeuten soll, dass hier etwas in „Erscheinung“ ( $\varphi\alpha\nu\acute{o}\mu\epsilon\nu\omicron\nu$ ) tritt.

### Inversion und Integration

Eine Möglichkeit, eine empirische Theorie zu entwickeln, die Bewusstsein zum Gegenstand hat, kann als abwärtsgerichtetes oder *top-down* Vorgehen bezeichnet werden. Dabei wird so vorgegangen, dass zuerst Bewusstsein analog zu gewöhnlichen empirischen Phänomenen analysiert wird und dass anschließend diejenigen Eigenschaften, die seinen „Kern“ ausmachen, mit Eigenschaften von physikalischen Zuständen oder deren Struktur identifiziert werden.

Ein Beispiel aus der aktuellen Bewusstseinsforschung soll diese Vorgehensweise veranschaulichen. Einer Hypothese zufolge, ist Bewusstsein identisch zu „integrierter Information“ (Tononi 2008), deren Wert  $\Phi$  für physikalische Systeme prinzipiell (de facto aber nur schwer) berechenbar und messbar ist. Integrierte Information  $\Phi$  kann demzufolge als Maß von Bewusstsein betrachtet werden: Wenn ein System einen hohen  $\Phi$ -Wert aufweist, ist es bei „hohem Bewusstsein“; wenn es nur durch niedrige Werte von  $\Phi$  charakterisiert wird, dann verfügt es entsprechend über „wenig Bewusstsein“. Es lassen sich somit sowohl Photodioden (kleines  $\Phi$ ) als auch Menschen im Wachzustand (hohes  $\Phi$ ) auf diese Weise betrachten. Dies legt den Verdacht nahe, dass es sich bei der Theorie der integrierten Information um eine moderne Form von Panpsychismus handeln würde, allerdings unter der Einschränkung, dass es Systeme gibt, die *nicht* bei Bewusstsein sind und dass Bewusstsein als graduelle Eigenschaft betrachtet werden sollte (Tononi & Koch 2015).

Dass eine solche Theorie notwendig mit hausgemachten Schwierigkeiten konfrontiert ist, dafür soll nun im Folgenden argumentiert werden. Exemplarisch stellt das Gedankenexperiment der „invertierten *Qualia*“ ein Problem für all jene Theorien dar, die eine Identität zwischen Bewusstseinsinhalten und physikalischen Strukturen postulieren, und geht in seinen Grundzügen bereits auf John Locke und dessen *Essay* zurück. Das vielleicht bekannteste Beispiel für invertierte *Qualia* betrifft die Farbwahrnehmung: Wäre es nicht denkbar, dass „mein Rot“ genau „deinem Grün“ entspricht, obwohl wir dasselbe Objekt betrachten und sich unsere neuronalen Zustände dabei gleichen? Dies wäre ausgeschlossen, wenn eine Theorie existieren würde, die eine Klasse von Hirnzuständen einem ganz bestimmten Wahrnehmungszustand zuordnet. Aber auch andere, etwas ausgefeiltere Beispiele, die auf ähnlichen

Überlegungen beruhen, lassen sich konstruieren [3]. Wichtig dabei ist, dass die Existenz einer derartigen Vertauschung von mentalen (phänomenalen) Zuständen gar nicht tatsächlich empirisch realisiert sein muss. Das Szenario der Inversionen stellt ein Problem für reduktive Auffassungen („Bewusstsein ist nichts anderes als...“) dar, sofern jenes Szenario auch nur „logisch möglich“, d.h. sofern es nicht von vornherein auszuschließen ist. Um dieses Szenario näher zu untersuchen, kann das Argument der Inversionen formalisiert werden (Hoffman 2006), wobei sich zeigt, dass ein beliebiger Einwand gegen die Möglichkeit solcher Inversionen – nämlich ein Verweis auf die Asymmetrie der Struktur bewusster Erfahrung – nicht greift [4].

Wird aber tatsächlich eine Theorie *über Bewusstsein* (und nicht nur über das zu ihm korrelierende Verhalten) angestrebt, sollten solche Inversions-Szenarien ausgeschlossen sein. Zwar würde aus diesen gar nicht folgen, dass wir keine intersubjektiven Aussagen über unsere Erfahrungen machen könnten – im Gegenteil folgt sogar, dass sich zwei Personen in verschiedenen subjektiven Zuständen auf die exakt gleiche Weise ausdrücken können; Alice und Bob können also beide über dasselbe Objekt sprechen, auch wenn es Alice rot und Bob grün erscheint. Allerdings wäre impliziert, dass dabei keine verbindlichen Aussagen über den phänomenalen Gehalt ihrer subjektiver Zustände gemacht werden können – etwa „wie es sich für Alice anfühlt, rot zu sehen“ oder „wie es sich für Bob anfühlt, grün zu sehen“.

Eine analoge Situation, so ein kritischer Einwand, würde auch für *alle* physikalischen Theorien zutreffen, da wir immer nur bis auf Isomorphismen sagen können, wie sich die Entitäten physikalischer Theorien zu den empirisch-realen Gegenständen verhalten. Aus dieser Tatsache sei aber kein fundamentales Problem abzuleiten; physikalische Theorien seien auch dann noch möglich und sinnvoll.

Der zentrale Unterschied, der jedoch zwischen Theorien von Bewusstsein und gewöhnlichen physikalischen Theorien besteht, ist folgender: Während es für letztere gar nicht darauf ankommt, ob die darin vorkommenden Entitäten neben den (funktionalen) Beziehungen, in denen sie auftreten, noch andere (nicht-funktionale) Eigenschaften besitzen, ist es ja gerade das Ziel einer Theorie von Bewusstsein, verständlich zu machen, inwiefern ein bestimmtes Erlebnis diese oder jene Eigenschaft (etwa den Farbeindruck) aufweist, indem sie Beziehungen zwischen physikalischen Objekten mit den („intrinsischen“) Eigenschaften des Bewusstseins identifiziert oder darauf abbildet.

Nun würden die meisten Physikalisten sofort einwenden, dass es eben eine Sache der Empirie sei, inwieweit „intrinsische Eigenschaften des Bewusstseins“ als identisch zu Struktureigenschaften physikalischer Systeme angesehen werden können oder nicht. Genau dies ist ja z.B. das Ziel der Theorie integrierter Information, die eine Funktion  $\Phi$  postuliert, die bewusste Systeme von nicht-bewussten unterscheiden kann. Aber



wie können wir im Einzelfall überprüfen, ob denn  $\Phi$  überhaupt adäquat ist? Wir haben keine Möglichkeit zu bestimmen, ob dieser oder jener Zustand eines anderen Organismus tatsächlich bewusst ist, da wir (wenn überhaupt!) nur Aussagen über die Existenz unseres je eigenen Bewusstseins machen können. Zudem werden uns andere funktionale (behaviorale) Kriterien nicht weiterhelfen, da diese – wiederum verstanden als Zuordnung mentaler Zustände zu einer geeigneten Menge physikalischer Zustände – selbst dem Szenario der Inversion unterliegen; ob jemand auch meint, was er sagt, kann nicht dadurch entschieden werden, dass er sagt, was er meint [5].

Zwei Organismen könnten demnach behavioral identisch sein, obwohl der eine bei Bewusstsein ist, während es der andere nicht ist. Wie sollen wir dann aber entscheiden, welcher von beiden nun bei Bewusstsein ist und welcher nicht?

### Intuitionen

Eine Möglichkeit wäre es, auf unsere Intuitionen bezüglich der „intrinsischen“ (im Gegensatz zu den funktionalen oder relationalen) Eigenschaften mentaler Zustände und auf die Vorannahmen, wonach diese durch eine bestimmtes  $\Phi$  am besten wiedergegeben werden, zu verweisen [6]. Hier wird deutlich, warum in einer Wissenschaft des Bewusstseins im Vergleich zu anderen Wissenschaften unsere Intuitionen nicht nur den Startpunkt der Theorieentwicklung darstellen: Angenommen, wir würden bei unseren Intuitionen über das Bewusstsein beginnen und diese in die Sprache einer funktionalistischen (mechanistischen) Theorie übersetzen. Dadurch ist es uns vermeintlich möglich, die Träger von Bewusstsein zu bestimmen. Die Fragen, auf die wir Antworten erwarten, wären etwa folgende:

Sind Menschen im Traume bei Bewusstsein?

Wie und warum schwindet das Bewusstsein im Tiefschlaf?

Zu welchem Zeitpunkt in der Ontogenese „entsteht“ Bewusstsein? Sind Neugeborene oder Föten bei Bewusstsein?

Sind nicht-menschliche Tiere bei Bewusstsein? Pflanzen? Kristalle? Atome? Elektronen? Usw.

Doch wie überprüfen wir letztlich, ob unsere Theorie adäquat ist? Wir verweisen wiederum auf Intuitionen, wonach dieser und jener Gegenstand bewusst ist oder eben nicht, und eine Theorie von Bewusstsein sollte mit möglichst vielen dieser Intuitionen übereinstimmen [7].

Es kann nun argumentiert werden, dass wissenschaftliche Theorien genau dazu die-

nen sollen, unsere alltäglichen Intuitionen herauszufordern, ähnlich wie Kopernikus und Galilei die intuitiven Annahmen zur Ruhelage der Erde herausforderten und Einstein die Absolutheit des Raumes und der Zeit in Frage stellte. Doch hier stoßen wir auf eine Schwierigkeit: Die (impliziten) Intuitionen, die wir über das Bewusstsein haben, würden, wenn sie nicht recht in das Bild passten, das unsere Theorie zeichnet, ja nicht ohne Weiteres abgetan werden – etwa mit Verweis darauf, „dass wir unser Bewusstsein eben nicht richtig verstünden“.

Vielmehr kann uns die Berichtigung unserer Intuitionen durch Wissenschaft im Nachhinein sogar nur dadurch plausibel werden, dass wir zwar an den Intuitionen als *vorthoretische*, *natürliche* Einstellung festhalten können (die Erde erscheint mir immer noch als das Zentrum des Universums, wenn ich in den Himmel blicke; der Raum erscheint mir immer noch wie ein unendlicher Behälter), dahinter aber eine „eigentliche Wirklichkeit“ vermuten, mit der es sich anders verhält, und zwar eben so, wie uns das die Wissenschaft sagt.

Hingegen wäre eine Theorie, die aussagen würde, wie uns, im Gegensatz zu unserer tatsächlichen Wahrnehmung, die Welt „wirklich“ erscheinen müsste, vieles, aber keine Theorie unseres Bewusstseins. Wir können zwar Illusionen aufdecken (also eine postulierte äußere Realität mit deren Wahrnehmung vergleichen) und vielleicht sogar erklären, warum wir diese Illusionen haben (etwa weil wir unbewusst die Größe von Objekten immer mit deren Erscheinungskontext abgleichen); wir würden aber nicht die Wahrnehmung selbst, also dass uns hier überhaupt etwas erscheint, erklärt haben.

Wir sehen also, dass eine Theorie von Bewusstsein notwendig an unsere Intuitionen gebunden bleibt, sofern wir jene für unfehlbare Einsichten in das Wesen des Bewusstseins halten. Des Weiteren wären diese Intuitionen theorieunabhängig. Bloß: Erheben wir unsere Intuitionen zum alleinigen Richter über eine wissenschaftliche Theorie, dann sind wir nicht mehr in der Lage unsere Theorie an einem unabhängigen Kriterium zu testen; und zwar nicht nur weil wir, wie im Falle der Stringtheorie, *de facto* keine Möglichkeit zu einer empirischen Überprüfung hätten, sondern weil uns eine solche aus prinzipiellen Gründen verwehrt bliebe.

Allerdings liegt es nahe, dass zumindest einige unserer Intuitionen, anders als ursprünglich vermutet, gar nicht die „intrinsischen Eigenschaften“ mentaler Zustände (falls es so etwas überhaupt gibt) widerspiegeln, sondern vielmehr historisch gewachsene und religiös (oder kulturell) vermittelte Überzeugungen, Wünsche oder Hoffnungen darstellen.

Neben der bereits erwähnten Intuition, wonach Bewusstsein nur ganz bestimmten Lebewesen zugesprochen werden dürfe, ließen sich z.B. die Auffassungen, wonach Bewusstsein *privat*, *subjektiv*, *intern*, *jenseits von Raum und Zeit* oder *nicht-materiell*

sei, nennen [8]. Bei alle diesen Punkten scheinen wir vor dem folgenden Dilemma zu stehen: Entweder es handelt sich hier um eine „intrinsische Eigenschaft“ des Bewusstseins, die uns intuitiv gegeben ist – dann ist diese für eine Theorie von Bewusstsein jedoch nur bedingt zu gebrauchen –, oder wir betrachten jene als bloß historisch gewachsen und kontingent – dann aber müssen wir dazu bereit sein, uns wieder von ihr zu verabschieden, allerdings immer mit der Frage im Hinterkopf, ob wir dann wirklich noch über Bewusstsein sprechen. Theorien, die sich ausschließlich auf unseren Intuitionen über Bewusstsein gründen, sind also letztlich unüberprüfbar oder sie stehen gar im Verdacht, nicht so sehr vom Bewusstsein zu handeln, sondern nur von unseren Annahmen über Bewusstsein.

## 2.2 Bilder und Modelle

### Theorien und Strukturen

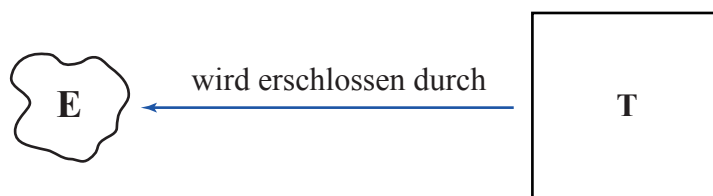
Im Gegensatz zu einer solchen Vorgehensweise entspricht es vielleicht eher dem materialistischen Temperament, zuallererst eine physikalische Theorie der „objektiven Welt“ anzustreben – etwa eine biologische Theorie des Gehirns oder eine vollständige Quantentheorie –, aus der sich schließlich eine Theorie des Bewusstseins ableiten ließe. Diese Hoffnung wird durch die Annahme gestützt, dass Bewusstsein ja ein empirisches Phänomen sei, das, wenn auch besonders komplex, letztlich wie andere Phänomene auch durch eine wissenschaftliche Theorie beschreibbar wäre. Ein solches Vorgehen kann als „*bottom-up*“-Strategie bezeichnet werden.

Wir können dies wie folgt veranschaulichen: Eine Mannigfaltigkeit von Gegenständen oder objektiven Ereignissen,  $\mathbf{E}$ , wird durch eine Menge von Aussagen oder Theorien,  $\mathbf{T}$ , dargestellt. Wenn wir annehmen, dass Bewusstsein als (Eigenschaft von einem) Gegenstand objektiv „in der Welt vorkommt“, und dass alle objektiv existierenden Gegenstände durch Theorien abbildbar sind, dann müsste eine vollständige Theorie der Gegenstände letztlich auf eine Theorie oder eine Beschreibung des Bewusstseins führen.

1. Wir wollen an dieser Stelle keine metaphysischen Annahmen hinlänglich Einheit oder Pluralität von  $\mathbf{E}$  treffen: ob es sich bei  $\mathbf{E}$  z.B. um eine „Welt“ im metaphysischen, starken Sinne, also um einen einheitlichen Zusammenhang empirischer Regelmäßigkeiten, oder ob es sich um eine Vielzahl individueller und wesentlich unabhängiger Einzelereignisse handelt.

2. Wir orientieren uns an einer Darstellung  $\mathbf{T}$ , die zwar möglichst allgemein bleiben soll, aber dennoch an den theoretischen Beschreibungsweisen der mathematischen Wissenschaften orientiert ist. (Dies impliziert, dass viele Arten der Verobjektivierung oder des Ordnen von Erfahrung ausgeschlossen werden oder nur schwer zu berücksichtigen sind, etwa die Kunst, Teile der Philosophie oder die Religion.)
3. Die Beantwortung von Fragen gelte als Metapher für die Beziehung zwischen  $\mathbf{E}$  und  $\mathbf{T}$ , aber auch das Messen oder die Gewinnung von Information:  $\mathbf{T}$  entspricht den Antworten, die wir auf Fragen über  $\mathbf{E}$  erhalten; sie entspricht den Ergebnissen von Messungen oder den Informationen, die wir über  $\mathbf{E}$  sammeln können.

Nun kann versucht werden, die Beziehung zwischen den Bereichen  $\mathbf{E}$  und  $\mathbf{T}$  aufzuklären, etwa in einer Strukturanalyse [9]; es geht dann nicht mehr, wie noch zuvor, um das Identifizieren von Bewusstsein mit anderen, objektiven Gegenständen, sondern um das Auffinden mathematischer Strukturen, die in einem Ähnlichkeitsverhältnis zur gegenständlichen Welt stehen (vgl. Abb. 2.1). Dies kann z.B. in Anlehnung an Heinrich Hertz' Bildtheorie verstanden werden, die dieser in seinen *Prinzipien der Mechanik* entwickelte.



**Abbildung 2.1:** Theorien  $\mathbf{T}$  beziehen sich auf eine Mannigfaltigkeit von Ereignissen in  $\mathbf{E}$ . Aus der Strukturiertheit dieser Bereiche sollen Rückschlüsse auf eine Wissenschaft des Bewusstseins gezogen werden.

Wie genau ist diese Ähnlichkeitsbeziehung jedoch zu verstehen? Eine erste Vermutung legt nahe, einen Isomorphismus zu vermuten, also eine vollständig strukturerhaltende und umkehrbare Abbildung zwischen unseren Theorien und (einer Beschreibung) der gegenständlichen Welt. Es ist instruktiv, zuerst eine Reihe von paradigmatischen Beispielen aus den physikalischen Wissenschaften zu betrachten.

**Beispiel 1:** Es beschreibe  $\{(q, p) \in \mathbb{R}^2\}$  den Phasenraum eines Punktteilchens der Masse  $m$  in einer Dimension;  $(q_0, p_0)$  bezeichne den Ort und Impuls zum

Zeitpunkt  $t = 0$ . Dann gilt, dass zu jedem Zeitpunkt  $t > 0$  der entsprechende Punkt im Phasenraum mit Hilfe der newtonschen Bewegungsgleichungen berechnet werden kann, wobei eine zusätzliche externe Kraft, die auf das Punktteilchen wirkt, durch einen Parameter oder eine vorgegebene Funktion  $a$  repräsentiert wird. Die Zeitentwicklung ist durch

$$\delta_a : \mathbb{R}_+ \rightarrow \mathbb{R}^2 : t \mapsto \left( \begin{array}{l} q(t) = \frac{1}{m} \int_0^t p(t') dt' \\ p(t) = m \int_0^t a(t') dt' \end{array} \right) \quad (2.1)$$

beschrieben. Die Funktion  $\delta_a$  beschreibt alle möglichen Zeitentwicklungen des Systems via Gl. (2.1). Diese funktionale Abhängigkeit innerhalb des theoretischen Formalismus entspricht unter der Annahme, dass Theorie und Welt in einem Abbildungsverhältnis stünden, der „kausalen Struktur“ der gegenständlichen Welt..

**Beispiel 2:** Es beschreibe die Funktion  $\vec{E} = (E_x, E_y, E_z)$  das 3-dimensionale elektrostatische Feld im Vakuum. Die Feldgleichungen der Elektrostatik legen  $\vec{E}$  zu einer gegebenen Ladungsverteilung  $\rho$  fest:

$$\nabla \cdot \vec{E} = \frac{1}{4\pi\epsilon_0} \rho \quad (2.2)$$

$$\nabla \times \vec{E} = 0. \quad (2.3)$$

Da es sich um zeitunabhängige Phänomene handelt, folgt eine jede mögliche „Abfolge“ von Zuständen innerhalb der theoretischen Beschreibung aus dem statischen elektrischen Feld, welches die Struktur der gegenständlichen Welt spezifiziert [10].

Eine Weiterentwicklung dieser Auffassung führt auf den derzeit populären ontologischen Strukturrealismus in der Wissenschaftsphilosophie (siehe z.B. Ladyman & Ross 2007, Esfeld 2008), wonach die letzten Elemente der Wirklichkeit eben nur Strukturen sind, die aus den Darstellungen, die sich in den Theorien der modernen Physik finden, abgelesen werden können.

Dabei wurden einige sehr starke Annahmen getätigt (Isomorphie von Theorie und Erfahrung, eindeutige Unterscheidbarkeit zwischen theorieunabhängigen Daten und ihren theoretischen Beschreibungen etc.), denen man vielleicht auf den ersten Blick zustimmen möchte, die aber bei näherem Hinsehen als zu starke Idealisierung erkannt werden müssen.

**Anmerkung:** Wissenschaftliche Erklärungen handeln oft davon, wie verschiedene Aussagen über Beobachtungen miteinander *korrelieren*. Dies wird üblicherweise dadurch dargestellt, dass verschiedene Wahrscheinlichkeitsverteilungen,

$p(X)$  und  $p(Y)$ , untersucht werden. Dies wird in Abschnitt 3 noch ausführlicher an einem Beispiel diskutiert werden.

Dabei scheinen auf den ersten Blick Korrelationen zwischen zwei empirischen Gegenstandsbereichen,  $X$  und  $Y$ , untersucht zu werden, etwa wenn Geburtswahrscheinlichkeiten mit Storchenpopulationen oder erst-personale Beschreibungen von Wahrnehmungsprozessen mit fMRI-Daten verglichen werden. Nun gilt es jedoch zu berücksichtigen, dass solche Darstellungen oft verarbeitete sind, d.h., dass sie wesentlich theoretisch vorstrukturierte Objekte darstellen. Auch wird manchmal eine der beiden Verteilungen als „verborgene“ bezeichnet, d.h. als empirisch nicht direkt zugängliche und somit (theoretisch) zu erschließende. Statt von der Korrelation zwischen empirischen Gegenstandsbereichen  $X$  und  $Y$  zu sprechen, ist es dann oft adäquater, von Korrelationen zwischen gegenständlichen und theoretischen Bereichen zu sprechen.

Ein potentieller Fehlschluss ist dabei zu vermeiden. Es gilt nämlich, dass zwischen den beiden kein Kausalverhältnis bestehen *muss*: Korrelationen sind nicht dasselbe wie Ursachen, auch wenn Verursachung zu Korrelation führt. Der Schluss von Korrelation auf Ursächlichkeit ist daher kein logisch-deduktiver sondern ein hypothetischer Schluss. Das Fehlen einer umkehrbaren, strukturhaltenden (isomorphen) Abbildung zwischen den Ereignissen aus  $\mathbf{E}$  und den Beschreibungen aus  $\mathbf{T}$  lässt den Schluss von theoretischen Regelmäßigkeiten auf empirische „Naturnotwendigkeiten“ nur bedingt zu. Als solches müssen Aussagen, die von Korrelationen auf Kausalbeziehungen schließen, mit einer gewissen Skepsis betrachtet werden. Dies betrifft insbesondere auch die „neuronalen Korrelate des Bewusstseins“ (Chalmers 2000).

Andere, pluralistische Lesarten, welche die (produktive!) Unterbestimmtheit der von Hertz eingeführten „Bilder“ betont scheinen daher adäquater zu sein (vgl. etwa Sieroka 2007). Statt wahrheitsgetreuen Abbildern, scheinen wir uns nur noch (ev. einander widersprechende) *Modelle* zu machen und dürfen höchstens hoffen, dass diese vereinzelt Regelmäßigkeiten beschreiben, von denen wir annehmen, dass sie auch „objektiv“ in der gegenständlichen Welt gegeben sind [11].

Nun sollen jedoch Rückschlüsse auf eine mögliche Theorie des *Bewusstseins* gezogen werden, etwa dadurch, dass eine Theorie des Gehirns auf dasjenige theoretische Bild führen würde, in welchem nicht mehr länger von „mentalenen“ Phänomenen die Rede ist, sondern nur noch von den Funktionen oder der Dynamik neurobiologischer Systeme. Allgemein scheint eine solche Auffassung aber immer dort problematisch zu werden, wo es spezifischer um eben jenes Bewusstsein geht – sei es, wenn von

der (Phänomenalität der) Wahrnehmung, vom (freien) Treffen von Entscheidungen, oder vom (bewussten) Handeln nach Gründen die Rede ist.

### **Kausale Geschlossenheit und prinzipielle Reduzierbarkeit**

Bevor auf einen viel bemühten Vorschlag eingegangen wird, wie diese Schwierigkeiten umgangen werden könnten, sollen nun zwei Diskussionen aus der zeitgenössischen Philosophie kurz aufgegriffen werden, die damit im Zusammenhang stehen. Zum einen geht es um die Annahme der kausalen Geschlossenheit des Physischen und um einen konkreten Vorschlag, wie wir am Prinzip der kausalen Geschlossenheit festhalten und dennoch nach Gründen handeln können. Zum anderen geht es um die Idee, dass Beschreibungen von Prozessen in allen Fällen auf objektive, physikalische Beschreibungen reduziert werden können.

Beide Thesen, die der kausalen Geschlossenheit des Physischen und die der prinzipiellen Reduzierbarkeit aller begrifflich beschreibbaren Vorkommnisse in der Welt drücken starke regulative Ideen der Wissenschaften aus, sind aber als Universalausagen letztlich nicht empirisch überprüfbar. Beide Thesen gehen über das enge Feld der Bewusstseinsphilosophie hinaus und betreffen die Philosophie des Geistes im Allgemeinen. Da wo jedoch Rückschlüsse auf das Bewusstsein gezogen werden sollen, führen diese zu bemerkenswerten aber meist unerwünschten Bedingungen, denen eine Theorie desselben genügen müsste. Dies lässt sich nun zweierlei deuten: einmal so, dass hier spezifische Probleme angesprochen würden, die mit der „Natur des Bewusstseins“ zusammenhängen; ein andermal aber auch so, dass eine breitere philosophische Problematik vor allem dann akut würde, wenn wir sie vor dem Hintergrund des Diskurses über Bewusstsein betrachten. In beiden Fällen würde eine Theorie von Bewusstsein die aufgeworfenen Rätsel vermutlich nicht lösen können, allerdings würde sie wesentlich zu einer neuen Perspektive auf diese beitragen.

Das *Prinzip der kausalen Geschlossenheit* des Physischen besagt, dass jeglicher kausaler Einfluss auf physikalische Zustände selbst immer nur von physikalischen Zuständen ausgehen kann [12]. Anders als in seiner physikalischen Realisierung kann ein Subjekt keinen Einfluss auf die Vorgänge in der (physikalischen) Welt nehmen; weder als Experimentator im Labor (vgl. den in Abschnitt 1.2 erwähnten Wissenschaftler, der im Labor Phänomene erzeugt), noch als Teilnehmer einer Gemeinschaft (etwa als Stimmbürger, der über die Umsetzung eines politischen Vorhabens entscheidet).

Obwohl das Prinzip der kausalen Geschlossenheit vermeintlich an einer strengen Wissenschaftlichkeit orientiert ist, übersieht man leicht – gerade auch weil man sich de-

zidiert von allem „Außernatürlichen“ abwendet – dass die Bedingungen, unter denen wissenschaftliche Erkenntnis produziert wird, das Prinzip der kausalen Geschlossenheit zu verletzen scheinen: Wenn wir dem Experimentator nicht zugestehen würden, dass er frei ist, seinen Versuchsaufbau so zu wählen, dass die Materie entsprechend seinen Absichten manipulieren wird, müssten wir die Vorstellung, Wissenschaft sei (freie und zielgerichtete) Aktivität einer Gruppe von Subjekten, aufgeben. Insbesondere gilt dies auch für eine experimentelle Überprüfung des Prinzips der kausalen Geschlossenheit selbst. Betrachten wir jenes Prinzip nämlich als empirisches, ähnlich wie etwa den zweiten Hauptsatz der Thermodynamik, so würden wir in einen pragmatischen Zirkel geraten, da wir eine – experimentell zu überprüfende – Aussage tätigen, welche aber die Möglichkeit ihrer eigenen Überprüfung unterminiert. Das Prinzip der kausalen Geschlossenheit kann also höchstens als *metaphysisches* Prinzip Geltung beanspruchen.

Nun scheint es auf den ersten Blick aber tatsächlich so zu sein, dass wir den Verweis auf ein Subjekt zur Erklärung einer Situation immer weiter hinausschieben können: Warum stellt der Wissenschaftler im Labor seine optischen Geräte so und nicht anders auf? Weil er ein bestimmtes Phänomen künstlich erzeugen will oder weil seine Muskeln entsprechend kontrahieren? Warum tun sie dies? Weil die Gründe, die er hat, eben jene Aktivität herbeiführen oder weil seine Muskeln einen Impuls aus dem Zentralnervensystem erhalten haben? Wie kam es zu diesem Impuls? Weil das Bewusstsein des Wissenschaftlers diesen verursacht hat oder weil er die Folge der neuronalen Verarbeitung eines Reizes ist? Usw.

Sehen wir nun aber etwas genauer hin, scheinen wir es hier mit zwei verschiedenen Arten von Ursächlichkeit zu tun zu haben, die der Abbildungsrealist in Einklang bringen will, indem er entweder einfach ein „bewusstes Subjekt“ als objektiven Gegenstand postuliert, der das letzte (oder zumindest ein ausgezeichnetes) Glied einer Kausalkette bildet, oder indem er (subjektive) Gründe restlos durch (objektive) Kausalursachen ersetzt.

Ein möglicher Ausweg aus diesem Dilemma, der zwar das Handeln aus Gründen mit der Verursachung durch physikalische Ereignisse vereinbart, aber nicht auf eine Reduktion (oder auf die Hypostasierung von Epiphänomenen) hinausläuft, stammt von Donald Davidson. Von all den möglichen Gründen für eine Handlung hebt Davidson den „primären Grund“ hervor, für den gilt, dass er i) eine Einstellung (ein Wollen, Begehren, eine flüchtige Laune oder einen stabilen Charakterzug) gegenüber eines bestimmten zukünftigen Ereignisses ausdrückt und ii) die Überzeugung beinhaltet, wonach die Handlung zum Eintreten dieses Ereignisses (je relativ zu einer bestimmten Beschreibung) führen würde. Solche Gründe könne man dann auch als Ursache einer Handlung im engeren Sinne ansehen (Davidson 1990, S. 20).



Gleichzeitig wendet sich Davidson aber gegen reduktionistische Versuche, die nun, motiviert dadurch, dass (primäre) Gründe ja auch Ursachen sind, teleologische Erklärungen durch Kausalerklärungen ersetzen wollen. In einer Position, die oft als „anomaler Monismus“ bezeichnet wird, geht er davon aus, dass zwar mentale und neurophysiologische Zustände im *Einzelfall* (dieser bestimmte Gehirnzustand  $x$  ist identisch zum geistigen Zustand  $y$ ), aber nicht des Typus nach identisch sind (Gehirnzustände der Art  $X$  sind nicht identisch zu geistigen Zuständen der Art  $Y$ ). Dies impliziert, dass bestimmte mentale Einzelereignisse mit neurophysiologischen Ursachen identisch sind und somit als Ursache für ein bestimmtes Verhalten angesehen werden können – ganz ohne dabei das Prinzip der kausalen Geschlossenheit des Physischen zu verletzen.

Kausalerklärungen scheinen jedoch die Möglichkeit von Allgemeinaussagen zu bedingen, etwa wenn wir Erklären mit dem Subsumieren von erklärungsbedürftigen Einzelaussagen unter Naturgesetze und Randbedingungen gleichsetzen. Eine Ersetzung der intentionalen Erklärungsweise durch eine reine neurophysiologisch, kausale würde daher der Gleichheit von Gehirn- und Bewusstseinszuständen des Typus nach bedürfen. Aber das Aufstellen von Brückengesetzen zwischen diesen sind, so der *anomale* Monismus, nicht möglich und somit auch eine Identität des Typus nach nicht [13].

Dies scheint ein philosophisch raffinierter Schachzug zu sein, wirklich befriedigend ist er in Hinblick auf eine formale Wissenschaft von Bewusstsein aber nicht. Zwar wird der anomale Monismus allen drei Intuitionen (kausale Geschlossenheit des Physischen, Existenz mentaler Verursachung, Irreduzibilität des Mentalen) gerecht, wendet sich aber explizit gegen die Möglichkeit psychophysischer und psychologischer Gesetzmäßigkeiten, hat deren Unmöglichkeit sogar notwendig zur Folge [14].

Eine zweite These betrifft die *prinzipielle Reduzierbarkeit* aller Tatsachen auf die Aussagen der fundamentalen Physik. Selbst wenn dies im Einzelfall nicht durchführbar ist, so denkt es sich der Physikalist, ändere dies nichts daran, dass *im Prinzip* eine solche Reduktion machbar sein müsste. Im Laufe des letzten Jahrhunderts wurden mehrere Versionen des Reduktionismus vorgeschlagen [15]. Im folgenden soll kurz auf eine bestimmte Spielart eingegangen werden, die man als „übersetzenden Reduktionismus“ bezeichnen könnte, der im Rahmen der Identitätstheorie des Geistes eine wichtige Rolle spielte [16]. Historisch markierte diese Diskussion einen Ausgangspunkt für die Entwicklungen einer modernen naturalistischen Philosophie des Geistes. Dabei wird angenommen, dass Sätze über Erfahrungen (Empfindungen) in einem ersten Schritt aus einem psychologisch-phänomenalen Vokabular („Ich sehe einen roten Apfel“) in eine neutrale Beobachtungssprache übersetzt werden können („Unter

diesen und jenen Lichtverhältnissen, wird ein apfelartiges Bild auf der Netzhaut eines Versuchsteilnehmers abgebildet und verarbeitet“). In einem zweiten Schritt kann dieser neutral formulierte Ausdruck, so die Prognose der Identitätstheorie, mit einem neurologischen Prozess identifiziert werden. Kritiker haben früh angemerkt, dass der Übersetzungsschritt, anders als ursprünglich gefordert, keinen bedeutungserhaltenden Schritt darstellen kann, da er immer aus „dualistisch vorbelastetem“ Vokabular erfolgt (Feyerabend 1963). Daraus kann dann gefolgert werden, dass eine Identitätstheorie des Geistes, die auf einem solchen Reduktionismusverständnis beruht, nur schwer vertretbar ist. Eine Konsequenz, die man nun ziehen könnte, ist die Verfolgung eines „eliminativen Materialismus“ (Churchland 1981), der gar nicht mehr daran interessiert ist, zu zeigen, dass unser erst-personales Vokabular verlustfrei übertragen werden kann, sondern dieses durch ein neues Vokabular ersetzen möchte.

Dies mag vielen an dieser Stelle vielleicht als gar zu radikal erscheinen. Allerdings, und das ist der wesentliche Punkt dieser Diskussion, ist das Erste-Personenvokabular, von dem weiter oben die Rede war, nicht so unschuldig, wie es oft scheint, und die ersten Schwierigkeiten der „Übersetzung“ ergeben sich bereits an dieser Stelle: Der Satz „Ich sehe einen roten Apfel“ ist ja selbst eine reflexive Bezugnahme auf den Akt der Wahrnehmung. Dabei ist es prinzipiell zu hinterfragen, ob dies überhaupt dem Abbilden eines (Wahrnehmungs-)Zustandes entspricht. In der Tat: Sofern sich das Verständnis des Sprechaktes auf eine Position ähnlich dem Abbildungsrealismus stützt, bleiben sowohl eine naiv-materialistische als auch eine „*ordinary language*“-Philosophie des Geistes eher fragwürdig.

Insofern, als dass sie immer nur Aussagen (weg) von Erfahrung macht, ist eine jede sprachliche Darstellung (die wissenschaftliche, aber auch die alltägliche) bereits Abstraktion; höchstens die Vorstellung, dass es sich hier um Abbildungen objektiv existierender „innerer Dinge“ handle, täuscht uns vielleicht vor, wir würden auch etwas über das Wesen oder die Essenz unseres Bewusstseins aussagen können. Es ist eben mehr als fraglich, ob es überhaupt einen Isomorphismus (oder eine Modellbeziehung) zwischen Erfahrung und sprachlichem Ausdruck gibt, der (psychische) Gegenstände mit sprachlichen Ausdrücken korrelieren kann und zudem die Struktur zwischen beiden Bereichen erhält [17].

Ein wichtiges Korollar betrifft die Auffassung der analytischen Sprachphilosophie, intentionale Zustände seien als propositionale Einstellungen darstellbar. Warum sollte (intentionales) Bewusstsein als Einstellung gegenüber einer Proposition zu verstehen sein? Warum sollte der psychische Zustand des „Aua“ eine Einstellung gegenüber der Tatsache, „dass ich Zahnschmerzen habe“, darstellen? Die philosophischen Analysen, denen man (propositionale) Sätze unterwerfen kann, machen höchstens dann

eine Aussage über das Bewusstsein, wenn dieses zu jenen Sätzen i) identisch ist oder ii) in einer strukturgleichen Beziehung steht, d.h. wenn ich davon ausgehen kann, dass Bewusstsein und Sprache derselben Struktur unterliegen. Von Smith & Mulligan (1983, S. 79) wurde daher die Forderung vorgetragen, die propositionalen Umschreibungen mentaler Zustände sollten dieselbe „ontologische Struktur“ aufweisen, wie die mentalen Zustände selbst (oder zumindest deren entsprechenden nicht-propositionalen Umschreibungen.). Um die Adäquatheit der Analyse propositionaler Einstellungen beurteilen zu können, muss besagte ontologische Struktur expliziert werden, was das Ziel vieler „formal-ontologischer“ Arbeiten darstellt.

## 2.3 Emergenz

### Starke und schwache Ausprägung

Als Reaktion auf die oben diskutierten Schwierigkeiten wird oft vom Begriff der *Emergenz* gesprochen, womit die Vorstellung verbunden ist, dass neuartige Eigenschaften einer Ganzheit (eines Systems) aus den Eigenschaften ihrer Teile hervorgehen können, ohne dass dabei die Dynamik der Ganzheit auf die der Teile reduzierbar wäre.

So ließe sich etwa vermuten, dass mentale Verursachung, Bewusstseinsinhalte oder das Handeln nach Gründen in einem hierarchischen System durch die Einführung einer oder mehrerer „emergenter“ Ebenen beschrieben werden können. Gerade da, wo es um Bewusstsein als *explanandum* einer naturalistischen Theorie geht, ist der Verweis auf Emergenz schon beinahe Teil der Orthodoxie einer „nicht-reduktiven“ materialistischen Auffassung vom Geist (Crane 2001b) und der Emergentismus wird als vielversprechende Erklärungsstrategie zur Lösung des „Geist-Materie-Problems“ betrachtet.

Das Gehirn, dieser Denkweise zufolge, ist als materielles System anzusehen, das aus Millionen informationsverarbeitenden Subsystemen, den Neuronen, besteht und eine Vielzahl an komplexen Funktionen hervorbringt, die das Verhalten des Gesamtorganismus steuern. Wenn davon die Rede ist, dass einige dieser Prozesse „*bewusst* ablaufen“, bedeutet dies, dass Bewusstsein eine weitere, möglicherweise übergeordnete physikalische Eigenschaft ist, die aus der Dynamik der interagierenden Nervenzellen hervorgeht. Die Redeweise von Emergenz impliziert dabei auch, dass Bewusstsein eine Eigenschaft des Gesamtorganismus ist, und keinesfalls schon in den Nervenzellen selbst vorliegen würde, und distanziert sich damit von panpsychistischen Überlegungen.

Bewusstsein als emergentes Phänomen zu betrachten, ist ein Versuch, an dessen

Wirklichkeit und Irreduzibilität festzuhalten und gleichzeitig nicht davon ausgehen zu müssen, dass es überhaupt etwas Nicht-Physikalisches gäbe. Viele versprechen sich daher von Emergenztheorien, den Graben zwischen dem Materialismus und dem Dualismus zu überwinden und dabei das Beste beider Welten zu vereinen. Gleichzeitig macht dies aber auch die Schwäche von Emergenztheorien aus: Kritiker könnten einwenden, dass es ja zu gar keiner konzeptionellen Änderung gegenüber den bisher vorgestellten metaphysischen Positionen zum Thema gekommen sei und daher Emergenztheorien im Großen und Ganzen mit denselben Schwierigkeiten konfrontiert wären.

Die Befürworter von Emergenztheorien sehen sich hingegen dadurch bestätigt, dass eine Vielzahl an Beispielen für emergente Eigenschaften aus anderen Naturwissenschaften bekannt sind, so etwa der Magnetismus, die thermischen Eigenschaften makroskopischer Systeme und viele weitere komplexe Phänomene, die in Chemie oder Biologie (Luisi 2002, Boogerd *et al.* 2005) vorgefunden werden. Wenn die vermutete Emergenz des Bewusstseins analog zu den Emergenzphänomenen in den Einzelwissenschaften verstanden werden kann, würde das vielleicht für die Ansicht sprechen, dass Bewusstsein als ein gewöhnlicher wissenschaftlicher Gegenstand unter vielen aufzufassen sei: Bewusstsein würde aus dem Gehirn hervorgehen wie Flüssigkeit aus Molekülen.

Doch kann die Frage nach der Angemessenheit des Emergenzbegriffs nicht so einfach beantwortet werden. Betrachtet man die Literatur – sowohl die philosophische als auch die populärwissenschaftliche – zum Begriff der Emergenz etwas genauer, zeigt sich schnell, dass sich hinter diesem Schlagwort eine Vielzahl verschiedener Definitionen und Verständnisse verbirgt. Dabei lässt sich grob zwischen zwei Ausprägungen von Emergenz unterscheiden (Stephan 1999, Kap. 4). Eine *schwache* Form von Emergenz kann durch folgende drei Eigenschaften näher charakterisiert werden:

- (i) Alles, was existiert, ist *physikalisch*, d.h. aus rein physikalischen Bausteinen zusammengesetzt.
- (ii) Es gibt *systemische* Eigenschaften, die nicht bereits als Eigenschaften der Teile vorliegen.
- (iii) Die Eigenschaften des Systems stehen mit den Eigenschaften der Teile in einer naturgesetzlichen Beziehung; die Systemeigenschaften sind durch diese *determiniert*.

Eine *starke* (und synchrone) Form von Emergenz, manchmal auch als radikale Emergenz bezeichnet, erfordert zusätzlich zu den Bedingungen (i) - (iii) noch:

- (iv) Die Systemeigenschaften sind *irreduzibel* bezüglich der Teileigenschaften, entweder weil die üblichen wissenschaftlichen Methoden der Analyse versagen, oder weil die Teileigenschaften selbst irreduzibel sind.

Sind die Teileigenschaften eines Systems irreduzibel, hat dies zwei mögliche Ursachen: Entweder sie sind ihrerseits bereits als systemisch-emergent zu betrachten, oder es liegt eine Form von abwärtsgerichteter Kausalität vor, d.h. dass das System einen kausalen Einfluss auf seine Bestandteile ausübt. Dies impliziert umgekehrt, dass eine kategorische Trennung in System- und Komponentenverhalten nicht mehr durchgeführt werden kann. Wegen der Supervenienzanforderung iii) gilt, dass eine Änderung im Systemverhalten Änderungen im Verhalten der Bestandteile implizieren. Gleichzeitig gilt aber auch beim Vorhandensein einer abwärtsgerichteten Ursächlichkeit, dass sich das Verhalten der Komponenten in einem System mitunter stark vom Verhalten der isoliert vorliegenden Komponenten unterscheiden kann.

Mit Hinblick auf eine Emergenz von Bewusstsein ist es wichtig, klar zu machen, von welcher Art von Emergenz hier die Rede ist. Ist damit eine starke Form von Emergenz gemeint? Dann handelt es sich bei der Emergenz des Bewusstseins aber um einen fundamental anderen Prozess als in den Beispielen aus den Einzelwissenschaften, bei welchen es sich anscheinend nur um Formen von schwacher Emergenz handelt [18]. Emergentisten könnten dann nicht mehr behaupten, dass der Mechanismus, der das Entstehen von Bewusstsein verständlich machen würde, analog wäre zum Mechanismus, der die Ausbildung magnetischer Phänomene erklärt. Gleichzeitig legen metaphysische Überlegungen nahe, dass starke Emergenz überhaupt nicht intelligibel sei. Es würde, so eine Kritik am starken Emergentismus, doch seltsam anmuten, falls in einem System Bewusstsein entstünde, ohne dass irgendeine Basis für Bewusstsein in der Materie selbst vorhanden wäre (vgl. die Diskussion zum Panpsychismus in Abschnitt 1.2.).

Geht man aber nur von der schwachen Form von Emergenz aus, so scheint dies im Widerspruch zu vielen Annahmen über das Bewusstsein zu stehen, etwa dass Bewusstsein einen kausalen Einfluss auf den Körper haben kann (eine Form der abwärtsgerichteten Kausalität?), dass es sich bei Bewusstsein um etwas anderes als ein Produkt raumzeitlich lokalisierbarer Eigenschaften materieller Gegenstände handle, oder dass (phänomenale) Erlebnisgehalte unmöglich aus der Dynamik von Nervenzellen folgen können, egal wie komplex diese angeordnet sind.

Die folgenden zwei Unterabschnitte sollen die Problematik des Emergenzbegriffs für eine Wissenschaft des Bewusstseins noch weiter vertiefen – zuerst anhand konkreter Beispiele aus den Einzelwissenschaften, schließlich anhand grundsätzlicher Überlegungen.

## Komplexität, Information und Selbstorganisation

Einige Wissenschaftler versuchen nun, den oft als schwammig empfundenen Begriff der Emergenz auf andere, besser verstandene und (manchmal nur vermeintlich) exaktere Konzepte aus den jeweiligen Einzelwissenschaften zurückzuführen. Dabei ist öfters von Begriffen wie Selbstorganisation, der Entstehung von Ordnungsphänomenen, Protektoraten, oder der Brechung von Symmetrien die Rede. Insbesondere soll dadurch die Irreduzibilitätsthese vermieden werden, die als bedrohlich für das wissenschaftliche Unternehmen angesehen wird. Im Folgenden wird auf drei Strategien eingegangen, die den Begriff der *schwachen Emergenz* (oder davon abgeleitete) durch (i) *Komplexität und Information*, (ii) *Symmetrie* oder durch Verweis auf die (iii) *Selbstorganisation* materieller Systeme plausibel machen wollen.

Der Begriff der Emergenz wird in den Wissenschaften immer auch da angetroffen, wo von „Komplexität“ die Rede ist. Den Vertretern der Komplexitäts- und Systemwissenschaften zufolge, führen komplexe Systeme aufgrund ihrer inneren Strukturiertheit oder Organisation zur Emergenz von neuen Eigenschaften. Trotz jahrelangen Forschungsanstrengungen ist bis heute jedoch nicht geklärt, wie der Begriff der Komplexität genau zu fassen ist. Es gibt eine Unzahl an „Komplexitätsmaßen“ wobei jedes auf unterschiedliche Intuitionen verweist, was Komplexität eigentlich sei [19]. Ohne weitere Einschränkung, entstünde hier das Gefühl, ein *schlecht gefasstes explanandum* – Emergenz – würde durch ein *schlecht verstandenes explanans* – Komplexität – ersetzt werden. Fragen, die eine Theorie der Komplexität beantworten müsste, bevor sie eine tragende Rolle in einer Theorie von Bewusstsein spielen könnte, sind beispielsweise: Wie lässt sich der diffuse Begriff der Komplexität quantitativ fassen? Wie skaliert Komplexität mit der Größe eines Systems bzw. dem Größenverhältnis des Systems und seiner Teilsysteme? Wann kann ein System als einfach betrachtet werden, wann muss es als komplex angesehen werden? Wie verhält sich die Komplexität eines Teilsystems zur Komplexität des Ganzen? Wie lassen sich die Komplexitätsmerkmale, die eventuell im Gehirn realisiert sind, mit den evolutionären Bedingungen, unter denen diese entstanden sind, in Verbindung bringen? U.a.m.

Zur Lösung der ersten Frage wurde vorgeschlagen, den Informationsgehalt oder damit zusammenhängende Größen als Maß der Komplexität, und davon abgeleitet: des Bewusstseins (Balduzzi & Tononi 2008), zu verwenden. Es muss allerdings beachtet werden, dass „Information“ mehrere Bedeutungen hat und es unklar bleibt, ob es überhaupt einen einheitlichen Informationsbegriff geben kann. Die Möglichkeit eines einheitlichen Informationsbegriffs ist dabei sehr umstritten und wurde etwa

vom „Vater der Informationstheorie“, Claude Shannon (1993, S. 180), bezweifelt, findet aber auch Fürsprecher. Fruchtbarer als der Versuch, dies abschließend beurteilen zu wollen, ist es, auf die Vielfalt von Informationsbegriffen und -konzepten hinzuweisen: Was wir unter Information im Alltag verstehen, unterscheidet sich von dem, was als Information in der Nachrichtentechnik verstanden wird, und auch von dessen Generalisierungen in Physik oder Biologie.

Dabei lässt sich prinzipiell zwischen drei Aspekten von Information unterscheiden: dem syntaktischen, dem semantischen und schließlich dem pragmatischen (Küppers 1990, Burgin 2010). Syntaktische Auffassungen betrachten Information als Funktion der relativen Häufigkeit von (im Grunde bedeutungslosen) Zeichenketten.

Ein semantischer Informationsbegriff hingegen ordnet den Zeichen (oder den aus ihnen geformten „Nachrichten“) eine Bedeutung innerhalb eines Kontextes zu. Die meisten Versuche, Bedeutung zu naturalisieren, bringen die Bedeutung eines Zustandes mit dessen kausaler Rolle in Zusammenhang: In der Biologie, ein Feld, in dem der syntaktische Informationsbegriff mit großem Erfolg mit dem Begriff der „Erbinformation“ verschmolzen wurde, folgen aus den nur 4 Nukleotidpaaren der DNA nahezu unendlich viele ( $\approx 4^{4 \times 10^6}$ ) mögliche Kombinationen, das Genom eines E.coli Bakteriums zu spezifizieren (das Beispiel stammt von Küppers 1990). Allerdings sind nur einige dieser möglichen Kombinationen relevant für den Organismus und tragen in diesem Sinne Bedeutung innerhalb eines Kontextes. Dabei muss beachtet werden, dass hier oft von einer Bedeutung *für* den Organismus gesprochen wird. Kritiker wie Searle (1992) werfen solchen und ähnlichen Ansätzen an dieser Stelle vor, dass diese, anstatt eine echte Naturalisierung von Bedeutung liefern zu können, immer nur von Bedeutung relativ zu einem sinnverstehenden Wesen sprechen können. Folglich müssten wir bereits über eine Theorie des Bewusstseins verfügen, um über Bedeutung „in der Natur“ Aussagen machen zu können. Eine solche müsste aber erst entwickelt werden. Die zentrale Herausforderung naturalisierter Bedeutungstheorien besteht also darin, diesen impliziten Rückverweis auf ein verstehendes Bewusstsein durch die Angabe eines Kontextes zu ersetzen [20].

Der Informationsbegriff, der uns aus dem Alltag bekannt ist, enthält zudem eine pragmatische Komponente. Dabei führt Information zu einem spezifischen, zielgerichteten Verhalten. Eine Information, die mir zuvor telefonisch mitgeteilt wurde, führt vielleicht dazu, dass ich mich zufrieden in den Sessel lehne oder dass ich plötzlich ins nächste Krankenhaus stürme (Information induziert eine neuartige Handlungen). Der Soldat, der einen Befehl erhält, führt eine Geste aus, die seinem Vorgesetzten vermitteln soll, dass der Befehl verstanden wurde und vom Soldaten entsprechend ausgeführt wird (Information induziert Bestätigung). Weder der rein syntaktische noch der rein semantische Aspekt von Information scheint ein solches

Verhalten zu induzieren; stattdessen verweisen wir üblicherweise auf Gründe, um uns solche und ähnliche Handlungsabläufe verständlich zu machen. Dennoch scheinen der syntaktische und semantische Informationsbegriff die Randbedingungen für das Auftreten von pragmatischen Informationen zu definieren, indem sie das (physikalische) Vehikel strukturieren, mit dem pragmatische Information übertragen wird (Pattee 2013). Die Übermittlung pragmatischer Information folgt *in diesem Sinne* den Gesetzen semantischer Informationsübermittlung, deren Träger wiederum den Gesetzen syntaktischer Informationsübermittlung folgen.

Auf den ersten Blick scheint es vielleicht eine vielversprechende Idee zu sein, Bewusstsein als emergentes Phänomen auf der Basis von Information zu erklären. Allerdings wird meist mit einem rein syntaktischen Informationsbegriff gearbeitet, der zwar genau definiert ist aber ohne semantischen oder pragmatischen Gehalt zu verstehen ist, oder es wird umgekehrt stillschweigend von einem sehr reichhaltigen Begriff von Information ausgegangen, der aber üblicherweise gar nicht klar definiert ist (Pockett 2014). Beim Vorhaben, Bewusstsein im Rahmen der Informationstheorie zu erklären, muss daher gezeigt werden können, wie die höherrangigen Aspekte von Information mit syntaktischer Information zusammenhängen, etwas, was den Kern der Frage betrifft, ob Computer – also Systeme, die syntaktische Information verarbeiten – über Bewusstsein (oder „Intentionalität“) verfügen können. Es ist jedoch unklar, inwieweit mathematische Formulierungen von Information überhaupt dazu beitragen, Begriffe wie Sinn, Bedeutung oder Absicht naturwissenschaftlich zugänglich zu machen. Trotz den Hoffnungen, die sich hier im Laufe des „Informationszeitalters“ ergeben haben, ist und bleibt sich die Philosophie dabei uneins [21]. Ein umfassender Ansatz zu einer Informationstheorie, die semantische und pragmatische Aspekte berücksichtigt, wurde neuerdings von Burgin (2010) vorgelegt, und fußt letztlich auf einer semiotischen Interpretation, was einige Ähnlichkeiten zu den Darstellungsweisen psychischer Prozesse in den Abschnitten 5 und 6 aufweist. Dann kann letztlich jedoch gar nicht mehr von einer Naturalisierung im eigentlichen Sinn gesprochen werden, da es gar nicht mehr der Fall ist, dass ein bestimmtes *explanandum*, nämlich Bewusstsein, auf der Basis eines „rein natürlichen“, nicht mit dem zu-erklärenden Objekt verknüpften *explanans* verständlich würde (Seager 2012, Kap 9.3).

Alternativ zu informationstheoretischen Ansätzen kann versucht werden, Emergenz auf die Erhaltung oder Brechung von Symmetrien zurückzuführen. In diese Richtung argumentieren zum Beispiel die Physiker Philip Anderson (1972) und Robert Laughlin (2007). Besonders in der Physik spielen Symmetriebrüche eine besondere Rolle in der theoretischen Beschreibung materieller Phänomene. Der eingangs erwähnte Magnetismus, das Higgsteilchen, oder das Phänomen der Supraleitung sind gern ge-



nannte Beispiele, aber auch in den Nachbardisziplinen spielt das Konzept der Symmetrie eine zentrale Rolle, sei es in der Kristallographie oder der Molekülchemie in Form diskreter Raumsymmetrien von Gitterstrukturen oder als Austauschsymmetrie zwischen Atomen und Molekülen (Pauliprinzip und Permutationssymmetrien). Was diesen teils sehr unterschiedlichen Verwendungsweisen von Symmetrieprinzipien gemein ist, ist die Feststellung, dass mit dem Bruch einer Symmetrie jeweils neue Observablen verbunden sind. Symmetrien führen dazu, dass wir bestimmte Unterscheidungen an materiellen Systemen nicht treffen können. Erst der Bruch dieser Symmetrien ermöglicht uns beispielsweise, zwischen rechts- und linkshändigen Naturprozessen zu unterscheiden. Diejenigen Eigenschaften der Materie, die auf einer grundlegender Ebene nicht beobachtbar sind und erst relativ zu einem Symmetriebruch beobachtbar werden, können dann als emergent bezeichnet werden.

Der Physik hinreichend großer und gekoppelter Systeme, etwa der Festkörperphysik oder der Synergetik (Haken 1983), entstammt die Idee, wonach die Brechung von Symmetrien eng mit dem Phänomen der Selbstorganisation zusammenhängt, aber auch in den Bereichen der Biologie und den Neurowissenschaften findet die Einführung von Symmetriebrüchen und Selbstorganisation (und Emergenz) immer mehr Verbreitung, wenn auch vieles spekulativ bleibt. Dies wird gern mit der Theorie von Phasenübergängen oder Minimalprinzipien in Verbindung gebracht.

Ein Vorteil dieser dritten Lesart von Emergenz ist es, dass informationstheoretische und physikalische Ansätze unter ein Prinzip subsumiert werden können. Im Gegensatz zur Vorstellung, Emergenz beruhe auf dem Bruch fundamentaler Symmetrien, handelt es sich hier aber primär um eine Anwendung im Rahmen dynamischer oder komplexer Systeme. Daraus leitet sich auch die Spekulation ab, die – aus Sicht der Physik – einfachen und fundamentalen Gesetze ergeben sich erst aus der Selbstorganisation komplexer und zusammengesetzter Systeme (Laughlin 2007). Bei deren Dynamik kommt es (in manchen Fällen sogar notwendigerweise) zur Ausbildung bestimmter, unvorhersagbarer Regularitäten, welche dann als emergente Muster oder Eigenschaften identifiziert werden können. Insbesondere für das Gehirn, wenn es als chaotisches System fern des thermodynamischen Gleichgewichts verstanden wird, verspricht ein solcher Ansatz zu erklären, wie aus der neuronalen Dynamik emergente Makroeigenschaften des Gehirns hervortreten. Ein prominenter Vertreter einer solchen Auffassung ist neuerdings etwa Karl Friston (2010). In eine ähnliche Richtung gehen aber auch die Theorien von Stuart Kauffman (1995) zur Selbstorganisation in biologischen Systemen oder von György Buzsáki (2006) zu den selbstorganisierenden Oszillationen des Nervensystems.

Die drei erwähnten Strategien, den Begriff der Emergenz anhand des Verhaltens natürlicher Systeme zu beleuchten, sind in der Lage, die Rede von der „schwachen

Emergenz“ zu plausibilisieren und zu spezifizieren. Dennoch bleibt unklar, inwieweit der Letztverweis auf Bewusstsein wirklich ersetzt werden kann, wie gerade am Beispiel der Information deutlich wird.

Bis jetzt wurde Emergenz immer mit Bezug auf die Eigenschaften zusammengesetzter Systeme diskutiert. Vielleicht sollte es in der Bewusstseinsfrage aber gar nicht darum gehen, Bewusstsein als emergente *Eigenschaft* des Gehirns zu erklären, sondern darum, die *Gesetze*, mit denen das Verhalten von Materie beschrieben werden kann, mit psychologischen Regelmäßigkeiten in Verbindung zu bringen. Emergenz wäre dann eine Relationen zwischen Theorien oder Modellen und handelte nicht so sehr von Gegenständen respektive deren Eigenschaften, sondern vielmehr von der Rolle, welche diese als Referenten einer (naturgesetzlichen) Beschreibung spielen. Emergenz wäre *epistemologisch* zu verstehen: Relativ zur Theorie  $\mathbf{T}_1$  erscheinen uns bestimmte Phänomene, deren Verhalten uns nur in Anbetracht einer anderen Theorie  $\mathbf{T}_2$  verständlich ist, als emergent. Wenn wir etwa Bewusstsein als Objekt einer psychologischen Theorie auffassen, erscheint uns Bewusstsein als emergent bezüglich unseres derzeitigen neurophysiologischen Wissens, aber daraus ist kein metaphysische Argument zu stricken, welches von der Essenz des Bewusstseins ausgeht und auf eine Unergründbarkeit des Bewusstseins schließt [22].

### Topologien, Kontexte und Muster

Vielleicht hindert uns aber auch eine kategorische Trennung von starker und schwacher Emergenz daran, den Begriff der Emergenz sinnvoll in der Philosophie des Bewusstseins anzuwenden. Im Rahmen einer alternativen Klassifikationen von Emergenz kann dann vielleicht die These hinterfragt werden, wonach nur im Rahmen von starker Emergenz wirklich Aussagen zum Bewusstsein gemacht werden können. Anstelle des Verweises auf Irreduzibilität soll die Unterscheidung zwischen hinreichenden und notwendigen Bedingungen eingeführt werden: Eine Ableitung der Systemeigenschaften aus den Teileigenschaften erfordert das Vorliegen sowohl hinreichender als auch notwendiger Bedingungen für die Ausbildung der Systemeigenschaften.

Falls jedoch die Eigenschaften der Teile nur notwendige, aber nicht hinreichende Bedingungen für die Existenz systemischer Eigenschaften liefern, müssen Annahmen wie Randbedingungen oder Kontexte zusätzlich berücksichtigt werden, um die Herausbildung systemischer Eigenschaften zu verstehen. Emergenz ist dann ein „*kontextuelles*“ (Bishop & Atmanspacher 2006) Phänomen und wird nicht als Gegenbegriff zur Reduktion, sondern als komplementär zu dieser verstanden.

Ein Beispiel liefern etwa die thermodynamischen Zustandsgrößen: Die messbare Größe der Temperatur folgt aus einer statistischen Beschreibung eines Systems, wenn

neben den notwendigen Bedingungen, die durch die mechanischen Bewegungsgleichungen gegeben sind, zugleich eine Gleichgewichtsbeziehung (der „0te Hauptsatz der Thermodynamik“) zwischen Referenzzuständen realisiert ist [23].

Im Unterschied zu vielen anderen Theorien von Emergenz wird hier davon ausgegangen, dass ein externer Beobachter (bzw. ein experimentelles Verfahren oder ein mustererkennendes Instrument) eine zentrale Rolle spielt, da er den Kontext einer Untersuchung spezifiziert, was etwa in der Wahl einer geeigneten *Topologie* [24] oder durch Einschränkung der Dynamik auf bestimmte Zeit- oder Energieskalen ausdrückbar ist [25]. Statt von einer topologischen Struktur kann auch von unterschiedlichen Partitionierungen eines Phasenraumes relativ zu einer Funktion oder „Observablen“ gesprochen werden (beim Graben & Atmanspacher 2009).

Eine wichtige Eigenschaft der Theorie der kontextuellen Emergenz betrifft nun die Tatsache, dass der spezifische Kontext, der aus der (kontingenten) Wahl einer Funktion oder eines Referenzzustandes folgt, letztlich gewissen objektiven Bedingungen – sogenannten „Robustheitskriterien“ (Atmanspacher & Bishop 2007) – genügen muss. Emergente Eigenschaften sind somit als (objektive) Eigenschaften von (physikalischen) Gesamtheiten anzusehen, auch wenn ihr Hervortreten ursprünglich auf die Wahl eines „Beobachters“ zurückzuführen ist. Die Maximierung von „effektiver Information“ von Mustern, wie sie von Bedau (2008) oder Ladyman & Ross (2007) als Kriterium der objektiven Realität meso- und makroskopischer Phänomene vorgeschlagen wurde, kann als Analogon zu solch einem Robustheitskriterium verstanden werden.

Anstelle von theorie-externen Kontexten ist es manchmal adäquater von nicht-vorhersagbaren Strukturen zu sprechen, die sich spontan in dynamischen (z.B. chaotischen und komplexen) Systemen ausbilden. Die neuartigen Eigenschaften, die sich dann zeigen, sind Ergebnisse der Neuartigkeit solcher Strukturen, welche die Funktion der kontextuellen Topologien übernehmen. Im Gegensatz zu „synchronen“ Emergenzvorstellungen, spielt der Faktor Zeit eine große Rolle, weshalb manchmal auch von einer „diachronen Strukturemergenz“ (Stephan 1999, S. 69f.) gesprochen wird. Beide Ansätze, kontextuelle und synchrone sowie strukturelle und asynchrone, können als Alternativen zur Dichotomie von starker/schwacher Emergenz verstanden werden und scheinen daher die erfolgversprechendste Position für weitere Überlegungen darzustellen.

Was sich überhaupt beobachten lässt oder als Objekt unseren Theorien vorkommt, hängt also sowohl von der Beschaffenheit der Gegenstände selbst ab, als auch von unserer Fähigkeit, stabiler Muster zu erkennen. Emergente Eigenschaften sind solche, die nicht als Summe der Eigenschaften der (isolierten) Teile zu betrachten sind, aber dennoch objektiven und zugleich kontextsensitiven Mustern im Verhalten von

(physikalischen) Systemen entsprechen. Dabei können die jeweiligen Muster nicht aufeinander abbildbar sein und müssen nicht Teile einer allumfassenden Hierarchie von Beschreibungen sein, wie es die klassischen Theorien des Reduktionismus oftmals postuliert haben. Ein weiterer Vorteil dieser Lesart von Emergenz besteht darin, dass sie neben der Analyse vorhandener Beispiele auch Kriterien zum Auffinden (ev. noch unbekannter) emergenter Eigenschaften liefert [26].

Natürlich lässt sich fragen, welche Bedeutung dies denn alles für eine Beschreibung von Bewusstsein haben sollte. Bei den vorgestellten Ideen, die zu einer Wissenschaft vom Bewusstsein führen sollen, kommt es zu einem Kurzschluss, der entweder Subjektivität objektiviert und in Folge mit anderen Gegenständen der physikalischen Welt identifiziert, oder sie irgendwo „in“ ihnen findet.

Doch letztlich lösen sich die Objekte selbst im Rahmen kontextueller Emergenztheorien in Muster auf, die zwar objektiv und als Zeichen in einer übergeordneten Beschreibung vorkommen, aber immer je relativ zu unseren An- und Absichten sind (d.h. zu unseren „*ontological commitments*“; vgl. Atmanspacher & Kronz 1999). Ganz ähnlich argumentiert Seager (2012), wenn er vom „Paradoxon des Bewusstseins“ spricht, mit dem alle (nicht-radikalen) Emergenztheorien konfrontiert sind: Die emergenten Objekte sind letztlich immer nur emergente Objekte *für* ein Bewusstsein. Was soll es dann aber bedeuten, dass Bewusstsein selbst ein emergentes Phänomen ist?

Existiert nun jenseits der verschiedenen Beschreibungsebenen ein ontologisch primitiver Bereich, auf den die verschiedenen, teils-hierarchischen und emergenten Beschreibungen gegründet sind? Und falls ja, können wir ihm dann überhaupt bestimmte Eigenschaften zusprechen, oder haben wir es hier mit einer Form von negativer Ontologie zu tun: Wir wissen, dass hier etwas existiert, aber wir können darüber letztlich nichts aussagen?

Dem Ganzen ließe sich aber auch eine positive Wendung geben: Was wir als Objekt identifizieren können, ist nicht einfach „gegeben“, sondern erst Ergebnis eines Prozesses, der zwischen Kontingenz und Regelmäßigkeit vermittelt. „Emergenz“ bezeichnet die Erscheinungsweise von Objekten und verdeutlicht die je-relative Existenz aller Objekte. Die Bewusstseinsfrage ist dann weniger ein Problem der Ontologie – also der Frage, welche Objekte existieren und was Bewusstsein „eigentlich ist“ –, sondern rückt vielmehr in den Bereich der *Metaontologie*: Unter welchen Bedingungen lassen sich Regelmäßigkeiten identifizieren und welche Rolle spielt Bewusstsein dabei? Was führt überhaupt dazu, dass wir bestimmten Objekten eine gegenständliche Existenz zusprechen? Was ist der Unterschied zwischen den Kobolden in meiner Phantasie und den „echten“ Unruhestiftern da draußen?

## Anmerkungen zu Abschnitt 2

[1] Ausführlich hat sich etwa Arno Ros (2005, S. 52) mit dieser Problematik in Bezug auf das Geist-Materie-Problem beschäftigt:

Zur Wissenschaft [...] gehört sowohl die *empirische* als auch die *begrifflich-philosophische* Forschung. Beide stehen [...] in einem unauflöselichen systematischen Zusammenhang.

Leider wurde zu oft übersehen, dass gerade für *Naturwissenschaftler*, die sich produktiv mit den Problemen des Geistes beschäftigen wollen, ein Hinterfragen ihrer Begrifflichkeit fruchtbarer wäre, als es die ein oder andere philosophische Anbietung an die Neurowissenschaften vermuten ließe. Oft herrscht die Meinung vor, wonach die quasi-scholastische Tätigkeit der Philosophie sowieso wenig relevant sei für „eigentliche“ Forschung. Diese mag vielleicht für die Frage nach dem Mechanismus der Proteinsynthese gerechtfertigt sein; für eine Wissenschaft vom Geist, ist sie es nicht.

Die Rolle, welcher dabei einem „Begriff“ zukommt, entspricht dabei weniger dem einer (ontologischen) Kategorie als viel mehr dem, was von Michael Hampe (2007, Kap. 18) als „Unterscheidungsgewohnheit“ bezeichnet wurde, also als Abstraktion einer wiederkehrenden Konstellation von Einzelereignissen und Bedingung der Möglichkeit konkreter Zeigehandlungen.

[2] Die Skepsis gegenüber dem Realismus einer Abbildung scheint eng mit einer Interpretation der Quantenphysik zu korrelieren, die von der Bedingung von (subjektiver) Erfahrung handelt. Ausgehend von Überlegungen von Bohr, Heisenberg und Pauli tun dies etwa Fuchs & Schack (2014); eine damit verwandte Diskussion des „Subjekt-Objekt-Zirkels“ in der Quantenmechanik findet sich bei Bitbol (2001). Ganz unabhängig zur Quantenmechanik scheinen auch die Vertreter der radikaleren Spielarten des Konstruktivismus wie Heinz von Foerster, Humberto Maturana und Francisco Varela den Glauben an die Möglichkeit (und Nützlichkeit) realistischer Abbildungen nicht zu teilen. Für eine Theorie von Wahrnehmung, die sich *explizit gegen* den Abbildungsgedanken wendet, siehe Bennett *et al.* (1989), Mausfeld (2012) oder Hoffman *et al.* (2015).

[3] Vgl. z.B. die Szenarien bei Block (1990) und Chalmers (1996, Kap. 7). Eine behauptete Identität zwischen Bewusstsein und materiellen (oder funktionalen) Zuständen wird dabei üblicherweise im starken (metaphysischen) Sinne verstanden, also als Identität „in allen möglichen Welten“.

[4] Das *Scrambling-Theorem* lautet wie folgt: Es seien  $X, Y$  zwei Mengen, welche die mentalen Zustände zweier Personen repräsentieren, und  $\phi : X \rightarrow Z$  eine Abbildung der Zustände aus  $X$  auf eine Menge (von Äquivalenzklassen in)  $Z$ ; dies kann etwa die Menge

der „propositionalen Einstellungen“, eine Menge an Funktionen oder auch nur ein Zahlenintervall sein. Dann gibt es für jede Bijektion  $b$  eine Abbildung  $\phi' : Y \rightarrow Z$  mit der Eigenschaft  $\phi' = \phi \circ b^{-1}$ .

$$\begin{array}{ccc} \text{Person}_1 : & X & \xrightarrow{\phi} Z \\ & \downarrow b & \nearrow \phi' \\ \text{Person}_2 : & Y & \end{array}$$

Insbesondere heißt dies, dass, falls eine Bijektion zwischen mentalen Zuständen aus  $X, Y$  existiert, die Funktion  $\phi'$  existiert, welche die durch  $b$  „vertauschten“ mentalen Zustände aus  $Y$  denselben Elementen aus  $Z$  zuordnet, denen auch jene aus  $X$  zugeordnet wurden („Vertauschte Qualia“).

Wenn  $\phi$  zudem eine *messbare* Funktion darstellt, wird auch deren Struktur unter einer (stetigen) Bijektion  $b$  erhalten: Vertauschung von *Qualia* führt zu keinem *messbaren* Unterschied im Verhalten der beiden Personen, welches zu diesen *Qualia* korreliert.

[5] Für jedes weitere (behaviorale oder funktionale) Kriterium  $\kappa$  existiert gemäß des *Scrambling-Theorems* ein  $\kappa' = \kappa \circ b^{-1}$ , welches unsere vertauschten Zustände wiederum denselben Werten zuordnet, u.s.w.

[6] Dies ist etwa die Strategie von Oizumi *et al.* (2014), die ihr Kriterium der maximalen integrierten Information auf Intuitionen über das Bewusstsein (vermeintlichen „phänomenologischen Axiomen“) begründen wollen. Für eine Kritik an diesem Vorgehen siehe etwa die Arbeit von Cerullo (2015). Eine weitere Schwierigkeit besteht darin, dass die Theorie integrierter Information vorgibt, eine nicht-reduktive Theorie des Bewusstseins zu sein, da sie in der Formulierung von Oizumi *et al.* (2014) anscheinend vom Bewusstsein ausgeht, gleichzeitig aber eine *Identität* zwischen Bewusstsein und integrierter Information postuliert. Dies macht sie jedoch wiederum anfällig für Inversions-Argumente.

[7] Dabei könnte man etwa Nagel (1996, S. 233) folgen, der schreibt dass „viele Menschen um so weniger guten Glaubens sind, dass es dort so etwas wie innere Erlebnisse gibt, je weiter sie den phylogenetischen Stammbaum hinabsteigen.“

[8] Eine ähnliche Liste findet sich bei Velmans (2009, Kap. 6), wobei er diese als typisch dualistische Intuitionen über Bewusstsein bezeichnet.

[9] Diese könnte wie folgt aussehen: Die Struktur einer Menge  $M$  kann ganz allgemein als Mengensysteme  $\mathcal{M}$  zur Grundmenge  $M$  aufgefasst werden. Etwas formaler: Die *Struktur* (folgend Bourbaki 1961) auf einer Menge ist i) durch die Zuordnung von Elementen von  $M$  auf sich selbst gegeben („Struktur erster Ordnung“):

$$\varphi^{(1)} : M \times M \rightarrow M, (m_j, m_k) \mapsto m_l; \quad (2.4)$$

oder ii) durch Abbildungen von Teilmengen  $\Omega \subseteq M$ : („Struktur zweiter Ordnung“):

$$\varphi^{(2)} : \wp(M) \times \wp(M) \rightarrow \wp(M), (\Omega_j, \Omega_k) \mapsto \Omega_l. \quad (2.5)$$

Wichtige Strukturen in der Mathematik sind Mengen  $M$  mit einer (Gruppen) oder zwei (Ringe, Körper) inneren Verknüpfung, die ganz bestimmten Eigenschaften (Axiomen) genügen; daran anknüpfend dann etwa Mengenalgebren und Topologien, bei denen wir uns für Operationen auf Teilmengen zur Grundmenge  $X$  interessieren (z.B. Vereinigungen und Durchschnitte).

Es bezeichne dabei die Potenzmenge,

$$\wp(M) = \{X \mid X \subseteq M\}, \quad (2.6)$$

die Menge aller möglichen Teilmengen von  $M$ . Ein Beispiel soll dies verdeutlichen: Beinhaltet eine Menge,

$$M = \{a, b, c\}, \quad (2.7a)$$

$|M| = 3$  Elemente, so ist

$$\wp(M) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\} \quad (2.7b)$$

eine Menge, die  $|\wp(M)| = 2^{|M|} = 8$  Mengen (inklusive der leeren Menge) als Elemente beinhaltet. Jede  $n$ -wertige Relation zwischen Elementen der Menge  $M$  kann als Teilmenge von Potenzmengen zu  $M$  (also also Teilmengen von  $\wp(M)$  bzw.  $\wp^n(M) \stackrel{\text{def.}}{=} \underbrace{\wp(\wp(\dots(\wp(M))))}_{\times n}$ ) dargestellt werden.

Im Beispiel von oben sind etwa die (geordneten) binären *Paare*

$$(a, b) = \{\{a\}, \{a, b\}\}, \quad (b, c) = \{\{b\}, \{b, c\}\}, \quad (d, d) = \{\{d\}, \{d, d\}\} = \{\{d\}, \{d\}\} = \{\{d\}\} \quad (2.8)$$

jeweils eine Teilmenge von  $\wp(M)$ . (Eine analoge Definition lässt sich für beliebige Paare von Elementen einer Menge geben.) Die Menge der Teilmengen von  $\wp(M)$  ist dann diejenige Menge, welche solche Paare als Elemente enthält. Es gilt dann, dass

$$(a, b) \subset \wp(M) \Rightarrow (a, b) \in \wp(\wp(M)) = \wp^2(M). \quad (2.9)$$

Eine *Relation* lässt sich als Menge  $R$  von Paaren  $(a, b)$  verstehen. Wenn nun ein Element  $x \in M$  zu einem Element  $y \in M$  in einer Relation steht, bedeutet dies nichts anderes als dass

$$xRy \equiv (x, y) \in R, \quad (2.10)$$

d.h.  $R$  ist eine Teilmenge von  $\wp^2(M)$  und es gilt analog zu Gl. (2.9):

$$R \subset \wp(\wp(M)) \Rightarrow R \in \wp^3(M) \quad (2.11)$$

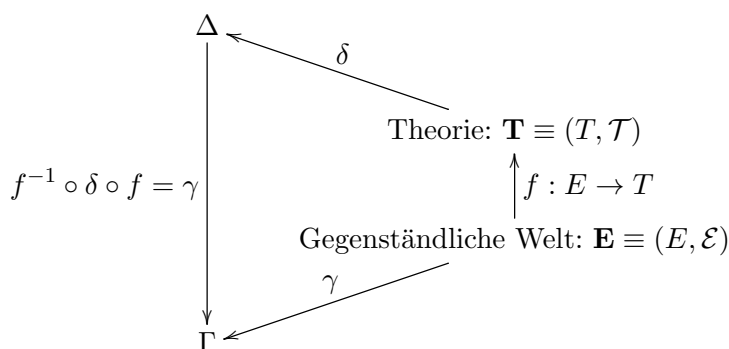
Für symmetrische (ungeordnete) Relationen wie etwa Distanzen (z.B. „Abständen im *Qualia*-Raum“) gilt sogar, dass bereits die Potenzmenge (näherungsweise) isomorph zur Menge aller möglichen solcher Relationen ist. Auch Operationen, in der Mathematik üblicherweise durch  $+$  oder  $\circ$  bezeichnet, lassen sich als mengentheoretische Konstruktion zur Grundmenge  $M$  explizieren, was vielleicht als wenig elegant oder unnötig umständlich („pathologisch“; Halmos 1968, S. 37) erscheint.

Für die philosophische Analyse des Abbildungsrealismus scheint jedoch diese Art der Darstellung geeignet zu sein, da hier nicht allzu viele zusätzliche Begriffe wie geordnete Paare, kartesische Produkte o.ä. eingeführt werden müssen. Wenn wir zudem der Prämisse folgen, wonach alle erfahrbaren Gegenstände Elemente *einer* umfassenden Menge („Natur“) sind, so können wir mit Hilfe von Teil- und Potenzmengen alle formalen Beziehungen zwischen diesen Gegenständen darstellen.

Def. 1: Es bestehe eine Theorie oder „Objektbeschreibung“  $\mathbf{T}$  aus einer „Aussagenmenge“  $T = \{t_1, t_2, \dots, t_N\}$  zusammen mit einer Menge  $\mathcal{T}$ , welche als „Aussagenstruktur“ angesehen werden kann. Zudem existiere eine Funktion  $\delta : T \rightarrow T$ , die der Menge  $T$  bestimmte Elemente aus sich selbst zuordnet.

Def. 2:  $\mathbf{T}$  bezieht sich auf eine gegenständliche Welt  $\mathbf{E}$ , welche eine Menge an erfahrbaren Gegenständen  $E = \{e_1, e_2, \dots, e_M\}$  und eine Strukturmenge  $\mathcal{E}$  umfasst. Dabei bezeichnen die  $e_i$  allgemein diejenigen Gegenstände, von welchen wir annehmen, dass sie uns direkt in der Erfahrung in Form von (empirischen) „Daten“ gegeben sind. Bei den Gegenständen kann es sich um physikalische Dinge und ihre Eigenschaften handeln, aber auch um psychische Zustände, oder Tatsachen und Sachverhalte.

Def. 3: Eine („streng realistische“) Abbildungsbeziehung zwischen diesen liegt genau dann vor, wenn gilt dass  $\mathbf{E} \stackrel{f}{\cong} \mathbf{T}$ ; d.h. falls ein Isomorphismus  $f : E \rightarrow T$  existiert, der Elementen  $e \in E$  die Elemente  $t \in T$  zuordnet und dabei die Strukturen  $\mathcal{E}$  und  $\mathcal{T}$  erhält. Wir bezeichnen dann  $f$  als „streng realistisch“. Eventuell wird man die Menge  $E$  auf Teilmengen von  $\Omega \subseteq E$  einschränken wollen; etwa wenn unsere Theorien nur von ganz bestimmten Gegenständen handeln.



**Streng realistische Abbildungen:** Sei  $f : E \rightarrow T$  eine streng realistische Abbildung zwischen empirischen Daten  $e_i$  und theoretischen Aussagen  $t_i$ ;  $\mathbf{E} \stackrel{f}{\cong} \mathbf{T}$  sind isomorph;  $f$  ist strukturerhaltend: Eine Abbildung  $\delta : T \rightarrow T$  transformiert in eine Abbildung  $\gamma : E \rightarrow E$  mit  $f^{-1} \circ \delta \circ f = \gamma$ . Die Struktur der Theorie  $\mathbf{T}$  spiegelt 1:1 die der gegenständlichen Welt  $\mathbf{E}$  wider. Jeder Folge  $\Delta$  in  $T$ , die wir mit Hilfe der Abbildung  $\delta$  konstruieren, entspricht eine Folge in  $\Gamma$  in  $E$ .



Man findet dann, dass jede Strukturbeziehung, die als Element  $R \in \mathcal{T}$  geschrieben werden kann, wegen der Isomorphie von  $\mathbf{E}$  und  $\mathbf{T}$  einer Beziehung

$$B = f^{-1}(R) \stackrel{\text{def}}{=} \{f^{-1}(r) | r \in R\} \in \mathcal{E} \tag{2.12}$$

zwischen den erfahrbaren Gegenständen entspricht. Dies gilt analog auch für Abbildungen  $\delta$ , deren Auffinden neben der Formulierung einer theoretischen Beschreibungsweise  $T$  das Geschäft der theoretischen Wissenschaften ausmacht. Auf diese Weise lässt sich eine Menge  $\Delta \subseteq T$  konstruieren, für die gilt, dass alle ihre Elemente der durch  $\delta$  beschriebenen „Dynamik“ genügen (logische oder „denknotwendige“ Folgen). Wegen  $\Delta \subseteq T$  existiert eine Menge

$$\Gamma = f^{-1}(\Delta) \subseteq E \tag{2.13}$$

sowie eine Abbildung

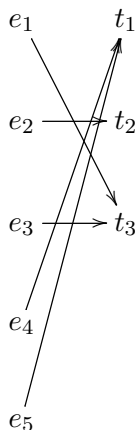
$$\gamma : E \rightarrow E, \quad \gamma = f^{-1} \circ \delta \circ f. \tag{2.14}$$

Dies bedeutet, dass wir zu jeder durch  $\delta$  generierten Folge in  $T$  eine Entsprechung in  $E$  finden können. Diese könnte man als „naturnotwendige“ oder kausale Folge bezeichnen.

[10] In beiden Beispielen wurden jeweils zwei Teilgebiete der gegenständlichen Welt  $E$  beschrieben (einmal als eindimensionale Dynamik eines Punktteilchens, das andere Mal als Elektrostatik in drei Dimensionen). Zudem wurden externe Parameter  $a, \rho$  eingeführt, von denen zwar ausgegangen wird, dass sie irgendeine Entsprechung in  $E$  besitzen, aber nicht unbedingt eine direkte Entsprechung in den betrachteten Teilmengen von  $E$  zu haben brauchen. Als solches ist bereits zu vermuten, dass sich wissenschaftliche Aussagen niemals auf  $E$  als ganzes beziehen werden, sondern jeweils auf (arbiträr festgelegte) Teilausschnitte davon. Allgemeinaussagen zu  $E$  sind daher bereits als metaphysische zu kennzeichnen. Wir können dann in Folge lediglich auf die Konsequenzen solcher Annahmen hinweisen und sie nicht (relativ zu einem bestimmten empirischen Verfahren) als wahr oder falsch ausweisen.

[11] Ein elementares Beispiel soll dies veranschaulichen: Es seien dabei  $E = \{e_1, e_2, e_3, e_4, e_5\}$  und  $T = \{t_1, t_2, t_3\}$ . Unser „Modell“  $g : E \rightarrow T$  ordnet diese wie folgt zu:

$$t_1 = g(e_4) = g(e_5), \quad t_2 = g(e_2), \quad t_3 = g(e_1) = g(e_3) \tag{2.15}$$



Zudem sei  $\mathcal{T}$  die Menge der (ungeordneten)  $n$ -Tupel  $\underbrace{\langle t_j, t_k, \dots, t_l \rangle}_n$ , mit den Eigenschaften

$$\langle \dots, t_j, t_k, \dots \rangle = \langle \dots, t_k, t_j, \dots \rangle \quad (2.16)$$

$$\langle \dots, t_j, t_j, \dots \rangle = \langle \dots, t_j, \dots \rangle \quad (2.17)$$

Daraus folgt aber, dass die Menge  $\mathcal{T}$  aller  $n$ -Tupel der Menge aller Teilmengen von  $T$ , also der Potenzmenge  $\mathcal{P}(T)$  entspricht:

$$\mathcal{T} = \left\{ \begin{array}{l} \emptyset \\ \langle t_1 \rangle, \\ \langle t_2 \rangle, \\ \langle t_3 \rangle, \\ \langle t_1, t_2 \rangle, \\ \langle t_2, t_3 \rangle, \\ \langle t_1, t_3 \rangle, \\ \langle t_1, t_2, t_3 \rangle \end{array} \right\} \cong \wp(T). \quad (2.18)$$

Folgende Eigenschaften werden von  $\mathcal{T}$  erfüllt:

$$T' \in \mathcal{T}, \text{ (mit } T' \cong T) \quad (2.19)$$

$$A, B \in \mathcal{T} \Rightarrow A \cup B = \langle x | x \in A \vee x \in B \rangle \in \mathcal{T}. \quad (2.20)$$

Das Komplement  $\bar{A} = T \setminus A$  ist definiert als

$$\bar{A} = \langle a | a \in T \wedge a \notin A \rangle \quad (2.21)$$

und liegt ebenfalls in  $\mathcal{T}$ , falls  $A \in \mathcal{T}$ . Zudem ist jede abzählbare Vereinigung  $\cup_{n=1}^{\infty} A_n$  von Mengen  $A_n \in \mathcal{T}$  selbst Menge in  $\mathcal{T}$ .

Nun kann der Menge  $\mathcal{T}$  ein „Urbild“  $\mathcal{U}$  zugeordnet werden;  $g$  sei per Definition homomorph bezüglich der Tupelbildung, d.h. dass  $g(\langle e_i, e_j \rangle) = \langle g(e_i), g(e_j) \rangle$ , und mit Gl. (2.15) und (2.16) folgt beispielsweise:

$$\begin{array}{ccc} \langle e_4, e_2 \rangle & & (2.22) \\ & \searrow & \\ & \langle t_1, t_2 \rangle & \\ & \nearrow & \\ \langle e_2, e_5 \rangle & & \end{array}$$

Bei gegebenen Tupel über  $T$  ist es also nicht eindeutig, welchem Tupel über  $E$  es zugeordnet wurde. Das Tupel  $g^{-1}(\langle t_1, t_2 \rangle)$  hingegen sei gegeben als

$$g^{-1}(\langle t_1, t_2 \rangle) = \langle e_i | g(e_i) = t_1 \vee g(e_i) = t_2 \rangle \cong \{e_4, e_5, e_2\}. \quad (2.23)$$

Es bezeichne dann allgemein  $\mathcal{U} = g^{-1}(\mathcal{T}) \stackrel{\text{def.}}{=} \{g^{-1}(C) | C \in \mathcal{T}\}$  mit  $g^{-1}(C) \stackrel{\text{def.}}{=} \langle e \in E | g(e) \in C \rangle$ . Das Urbild  $\mathcal{U}$  lässt sich dann berechnen zu:

$$\mathcal{U} = g^{-1}\left(\left\{\begin{array}{l} \emptyset \\ \langle t_1 \rangle, \\ \langle t_2 \rangle, \\ \langle t_3 \rangle, \\ \langle t_1, t_2 \rangle, \\ \langle t_2, t_3 \rangle, \\ \langle t_1, t_3 \rangle, \\ \langle t_1, t_2, t_3 \rangle \end{array}\right\}\right) = \left\{\begin{array}{l} \emptyset \\ \langle e_4, e_5 \rangle, \\ \langle e_2 \rangle, \\ \langle e_1, e_3 \rangle, \\ \langle e_4, e_5, e_2 \rangle, \\ \langle e_2, e_1, e_3 \rangle, \\ \langle e_4, e_5, e_3, e_1 \rangle, \\ \langle e_4, e_5, e_2, e_1, e_3 \rangle, \end{array}\right\} = \left\{\begin{array}{l} \emptyset \\ \langle e_2 \rangle, \\ \langle e_1, e_3 \rangle, \\ \langle e_4, e_5 \rangle, \\ \langle e_1, e_2, e_3 \rangle, \\ \langle e_2, e_4, e_5 \rangle, \\ \langle e_1, e_3, e_4, e_5 \rangle, \\ \langle e_1, e_2, e_3, e_4, e_5 \rangle, \end{array}\right\}. \quad (2.24)$$

Anders als zuvor ist  $\mathcal{U}$  *nicht* mehr isomorph zur Potenzmenge  $\wp(E)$ , sondern nur zu einer Teilmenge davon. Es lässt sich jedoch Folgendes zeigen:

$$E = \{e_1, e_2, e_3, e_4, e_5\} \Rightarrow E \cong g^{-1}(\langle t_1, t_2, t_3 \rangle) \in \mathcal{U} \quad (2.25)$$

$$A \in \mathcal{U}, B \in \mathcal{U} \Rightarrow A \cup B \in \mathcal{U} \quad (2.26)$$

$$A \in \mathcal{U} \Rightarrow \bar{A} \in \mathcal{U} \quad (2.27)$$

$$A_n \in \mathcal{U} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{U}, \quad (2.28)$$

d.h., dem Bild  $\mathcal{U} = g^{-1}(\mathcal{T})$  kommt dieselbe Struktur (in unserem Fall die einer Algebra mit  $\sigma$ -Eigenschaft) zu wie  $\mathcal{T}$ . Insbesondere gilt, dass wenn  $\mathbf{T}$  ein messbarer Raum mit

$$\mu_T : \mathcal{T} \rightarrow \mathbb{R}^+ \quad (2.29)$$

einem Maß auf  $T$  ist, auch  $\mathbf{E} = (E, \mathcal{U})$  ein messbarer Raum mit

$$\mu_E : \mathcal{U} \rightarrow \mathbb{R}^+, \quad \mu_E = \mu_T \circ g, \quad (2.30)$$

ist. Für den Fall, wo die Zielmenge der Abbildung  $\mu_T$  das Intervall  $[0, 1]$  ist (mit  $\mu(T) = 1$ ,  $\mu(\bigcup A_n) = \sum \mu(A_n)$ ), spricht man von einem Wahrscheinlichkeitsmaß; den „Ereignissen“ aus  $\mathcal{T}$  kann dann eine (nicht-negative) Wahrscheinlichkeit  $p \leq 1$  zugeordnet werden. Die Abbildung  $g$  „überträgt“ somit ein Maß, das für  $T$  definiert ist, auf einen Unterbereich der empirischen Daten. Die im Beispiel behandelte Abbildung  $g$  definierte lediglich eine eindeutige aber nicht umkehrbare Zuordnungsvorschrift zwischen Elementen aus (finiten Mengen)  $E$  und  $T$ . Über die Struktur der Bereiche  $E$  und  $T$  und bzw. deren Erhaltung durch  $g$  wurde hier nicht viel gesagt, da wir die ungeordnete Tupelbildung als einzige Operation betrachtet haben.

Im konkreten Fall werden sich die Kandidaten für eine Theorie des Bewusstseins darin unterscheiden, inwieweit sie Gruppen- oder Ringstrukturen, Metriken, Topologien etc. erhalten, und es wäre ein aufschlussreiches Projekt, solche Theorien hinsichtlich ihrer mathematischen Struktur zu untersuchen und ihre etwaige Stärken und Schwächen in Abhängigkeit der Abbildungseigenschaften, die sie postulieren, zu untersuchen.

[12] Genauer kann dieses Prinzip als Konsequenz der Forderung nach Bereichsspezifität (veränderungsbezogener) Erklärungen aufgefasst werden und lässt sich auf Aristoteles zurückführen. Neben veränderungsbezogenen (Wirk-)Ursachen kennt Aristoteles noch Ursachen des Zwecks, der Form und der Materie (im aristotelischen Sinne). Bereichsspezifität geht nun davon aus, dass Erklärungen von Phänomenen, die einem Bereich entstammen, immer nur durch Verweise auf Gegenstände dieses Bereichs zu erfolgen haben – (physikalische) Bewegungen durch andere Bewegungen, Gründe durch Verweis auf andere Gründe etc.

Wenn hingegen davon ausgegangen wird, dass es in der Natur nur *eine Art* von Gegenständen geben kann, muss diese Bereichsspezifität aufgegeben oder alternativ interpretiert werden, z.B. dadurch dass sie der (epistemologischen) Irreduzibilität von Beschreibungen entspringt.

Besonders interessiert etwa die Frage, inwiefern Materielles und Psychisches als solche getrennt zu betrachtende Gegenstandsbereiche angesehen werden müssen. Der Physikalist wird dies verneinen, stellen für ihn doch Psychisches und Physisches jeweils nur ein Teilgebiet der einen Natur dar; allenfalls ist Psychisches eine Struktur (Funktion), die auf dem Physischen definiert ist. Eine Alternative, die enger an die Aristotelische Sicht angelehnt ist, wäre es, die Existenz verschiedener Arten von Verursachung auf das Verhältnis zwischen einem Systemganzen und seinen Teilen zurückzuführen (siehe dafür etwa Ros 2008).

[13] In der Schreibweise aus Anmerkung [9] gilt zwar, dass im Einzelfall alle subjektiven Zustände (wie das Haben eines Grundes) einer bestimmten neurophysiologischen Beschreibung entspricht,

$$\forall g_i \exists! t_j (g_i \Leftrightarrow t_j), \quad (2.31)$$

dass aber keine Folge  $\Delta_G \subset T$  gefunden werden kann, welche zu einer Folge  $\Gamma_G \subset E$  isomorph wäre. Somit können Erklärungen durch Gründe nicht durch ursächliche Erklärungen ersetzt werden, obwohl gewisse Gründe als Ursache von Handlungen angesehen werden können. Zudem kann die Beziehung aus Gl. (2.31) nicht mittels Brückengesetzen dargestellt werden (oder anders ausgedrückt: Es kann keine eindeutige und strukturerhaltende Zuordnungsfunktion  $f$  mit  $t = f(g)$  bestimmt werden).

[14] Allerdings bezieht sich Davidsons Analyse in erster Linie auf rationale mentale Zustände, die mithilfe von propositionalen Einstellungen umschrieben werden können. Nicht-begriffliche Zustände des Bewusstseins, wie sie vielleicht bei nicht-menschlichen Tieren oder in der Wahrnehmung auftreten, werden nicht berücksichtigt. Jedoch ist die Vorstellung, dass es zwar (a-rationale) phänomenale Zustände geben kann, diese aber kausal inert sind, nicht sonderlich befriedigend.

[15] Siehe etwa die Monographien von Nagel (1961) oder Scheibe (1997).

[16] So etwa bei J. J. C. Smart (1959); für Smarts Reduktionsmusverständnis siehe Bickle (2012). Die Diskussion dieser Spielart des Reduktionismus dient in erster Linie der

Verdeutlichung eines Punktes über die Beziehung zwischen Gegenständen und deren Beschreibung.

[17] Wissen als (mathematische, begriffliche, bildliche ...) Erkenntnis ist primär immer als vermittelt anzusehen. Der unerschütterliche Glaube an die Existenz eines unvermittelten Wissens wurde von Sellars (1956) als „Mythos des Gegebenen“ bezeichnet. Eine Konsequenz, die man im Anschluss ziehen könnte, wäre es, an der „Unmittelbarkeit des Bewusstseins“ zu zweifeln; die andere, zu hinterfragen, ob denn Bewusstsein überhaupt ein *Erkenntnisgegenstand* ist.

[18] Wobei „schwache Emergenz“ hier nur einen sehr allgemeinen begrifflichen Rahmen bildet. Achim Stephan (2006, S. 150) spricht sogar davon, dass der Begriff der schwachen Emergenz „die Natur zwar an ihren ‚Scharnieren‘ [schneide], doch diese zu zahlreich sind“, um hier vielsagende Konsequenzen aus der Tatsache, dass einige Eigenschaften im schwachen Sinne emergent sind, zu ziehen.

[19] Vgl. etwa die Arbeiten von Wackerbauer *et al.* (1994) und Ladyman *et al.* (2013) für einen Überblick.

[20] Oizumi *et al.* (2014) versuchen, diesen Aspekt von Bedeutung im Rahmen einer probabilistischen Theorie zu verstehen, die allen prinzipiell möglichen Zuständen eines gegebenen Mechanismus Wahrscheinlichkeitswerte zuordnet. Dabei schränkt der aktuelle Zustand eines System die Wahrscheinlichkeit (vergänger und zukünftiger) Zustände innerhalb des Ensembles ein. Auf diese Weise kann davon gesprochen werden, dass der Zustand eines Systems Bedeutung „für sich“ habe.

[21] Vgl. etwa die Diskussionen in (Dretske 1983) rund um den Vorschlag, wonach Intentionalität aus einem Informationsbegriff ableitbar sei (als „*nested information*“), oder in (Searle 1980) für die klassische Streitfrage, ob Bedeutung durch Syntax erklärt werden könne. Auch die Redeweise vom „*information space*“ bei Chalmers (1996, Kap. 8) oder den „*informational relations*“ bei Tononi (2008) ist Zeuge einer anscheinend noch ungebrochenen Faszination mit dem Begriff der Information. Searle nimmt innerhalb der philosophischen Szene eine Mittelposition ein, wenn er daran zweifelt, dass Formulierungen, die auf der Informationstheorie von Shannon & Weaver (1949) aufbauen, Bewusstsein erklären können, und zugleich davon überzeugt ist, dass Bewusstsein ein biologisches Phänomen ist und daher im Rahmen einer naturwissenschaftlichen Theorie verstanden werden kann.

[22] Eine wichtige Einsicht stammt von Mark Bedau (2008), der dafür argumentiert, dass es Situationen gibt, in denen uns *alle* Regeln bekannt sind, die das Verhalten eines Objektes (auf der Mikroebene) beschreiben, und dennoch Fälle von Emergenz auf der Makroebene auftreten können. Dies liegt dann aber nicht daran, dass uns eine zusätzliche Einsicht fehlen würde, da uns ja alles, was es auf der Ebene von Regeln oder Naturgesetzen zu sagen

gäbe, bereits bekannt ist. Als Beispiel nennt er die Herausbildung stabiler makroskopischer Muster und deren Verhalten in Computersimulationen (vgl. auch Dennett 1991b).

[23] Für Details siehe die Arbeiten von Sewell (2002, S. 108-126) und Atmanspacher & Bishop (2007, S. 144-146).

[24] Dies kann in der Terminologie aus Anmerkung [9] ausgedrückt werden: Die Struktur  $\mathcal{T}$  einer Theorie  $\mathbf{T} = (T, \mathcal{T})$  wird dabei so modifiziert, dass neuartige theoretische Beschreibungen möglich werden.

Es sei dazu eine Menge  $T$  zusammen mit einer Struktur (einem Teilmengensystem)  $\mathcal{T}$  gegeben, die als Topologie von  $T$  bezeichnet werden kann, genau dann wenn,

- $\mathcal{T}$  die leere Menge und  $T$  selbst beinhaltet, und
- Vereinigungen und endliche Durchschnitte von Elementen aus  $\mathcal{T}$  stets wieder in  $\mathcal{T}$  enthalten sind.

Die Topologie  $\mathcal{T}$  beschreibt die „Lagebeziehung“ zwischen den „Punkten“ (Elementen) von  $T$ . Anschaulich gesagt, „teilt“ die Topologie  $\mathcal{T}$  die Menge  $T$  in unterschiedliche „Pakete“ (Teilmengen) auf.

Wir können nun zwei Topologien  $\mathcal{T}_1$  und  $\mathcal{T}_2$  vergleichen. Für  $\mathcal{T}_1 \supset \mathcal{T}_2$  bezeichnen wir die erste als „feinere“ und die zweite als „gröbere“ Topologie. Die feinste Topologie ist die Potenzmenge  $\mathcal{T} = \mathcal{P}(T)$ , da diese alle möglichen Teilmengen von  $T$  als Elemente enthält, die gröbste besteht nur aus der leeren Menge und der Menge selbst,  $\mathcal{T} = \{\emptyset, T\}$ , welche per Definition als Teilmenge in jeder Topologie vorkommen müssen. Nicht alle Topologien lassen sich auf diese Weise ordnen:  $\{\emptyset, \{1\}, \{1, 2\}\}$  und  $\{\emptyset, \{2\}, \{1, 2\}\}$  sind je „inkommensurable“ Topologien der Menge  $\{1, 2\}$ . Im Allgemeinen wird es daher möglich sein, mehrere solche Topologien zu betrachten, die untereinander geordnet werden können, aber nicht in einer totalen Ordnungsrelation stehen, d.h. keiner strikten „Hierarchie“ von topologischen Strukturen entsprechen (was gegen die Vorstellung einer strikten Hierarchie von Beschreibungsweisen spricht).

Die extern zur fundamentalen Theorie vorzuziehenden Kontexte können nun als gröbere Topologien dargestellt werden. Das technische Verfahren, welches geeignete *kontextuelle Topologien* erzeugt, ist jedoch nicht trivial. Für die neue Theorie  $\mathbf{T}_2 = (T, \mathcal{T}_2)$  mit der kontextuellen Topologie gilt, dass in ihr Aussagen gemacht werden können, die im Rahmen der ursprünglichen Theorie nicht gemacht werden können, d.h. dass diese relativ zu  $\mathbf{T}_1$  als emergent angesehen werden müssen. Gleichzeitig gilt, dass sich  $\mathbf{T}_2$  aus der fundamentalen Theorie  $\mathbf{T}_1$  ableiten lässt, *aber nur unter Berücksichtigung des entsprechenden Kontexts*, der jedoch nicht bereits in  $\mathbf{T}_1$  gegeben ist (Primas 1998).

[25] Bishop & Atmanspacher (2006, S. 1759) sprechen hier sogar davon, dass kontextuelle Emergenz auch als Rezept zum Auffinden von Brückengesetzen zwischen Theorien interpretiert werden kann, etwas, das in der philosophischen Literatur traditionell mit re-

duktionistischen Erklärungsstrategien in Verbindung gebracht wurde (*Der locus classicus* ist Nagel 1961).

[26] Für eine mögliche Anwendung in den Neurowissenschaften siehe z.B. Atmanspacher & beim Graben (2007) und beim Graben *et al.* (2009).





## 3 Geist und Gehirn

*But to the biologist the brain is not a thinking machine. It is an acting machine; it gets information and then it does something about it.* (W. Ross Ashby: Design for a Brain)

### 3.1 Das Gehirn, ein guter Regulator?

Im Folgenden soll an einem konkreten Beispiel untersucht werden, wie i) Bewusstsein manchmal als Abbild oder Modell einer gegebenen Wirklichkeit (miss-)gedeutet wird, ii) Emergenz als Folge der Dynamik biologischer Systeme verstanden werden kann, iii) eine „Naturalisierung“ von Subjektivität und intentionalem Verhalten zu verstehen ist, und iv) unter welchen Bedingungen, die Redeweise von psychischer Aktivität (Wahrnehmung) durch Beschreibungen materieller Vorgänge ersetzt werden kann. Welche Konsequenzen dies für eine mögliche Theorie von Bewusstsein hat, darauf soll in den Abschnitten 3.2 und 3.3 dann separat eingegangen werden, wobei für die Notwendigkeit eines Übergangs von einer funktionalistischen Theorie hin zu einer semiotischen Theorie argumentiert wird. Letztlich soll dadurch auch eine begriffliche Unterscheidung zwischen „bewusst“ und „subjektiv“ vorgenommen werden. Die Untersuchung selbst-organisierender Systeme scheint in erster Linie nur auf einen naturalistisch interpretierbaren Begriff der Subjektivität zu führen, bei dem noch unklar ist, wie sich dieser zum Begriff des (intentionalen und phänomenalen) Bewusstseins verhält. Ein konkreter Vorschlag, der jene zueinander in Beziehung setzt, soll im Anschluss vorgestellt werden.

#### Biologische Kybernetik

Die Vorstellung, dass unser Bewusstsein zur gegenständlichen Welt in einer abbildenden Beziehung steht, kann vielleicht durch die Behauptung motiviert werden, dass die Aktivität unseres Gehirns eben darin besteht, solche „Weltmodelle“ zu erzeugen. So betonen z.B. immer mehr Neurowissenschaftler und -Philosophen, das menschliche Gehirn sei eine „Inferenz-Maschine“, die sich aus sensorischen Daten ein Modell der Welt erschaffe, ähnlich wie Wissenschaftler (bewusst?) Inferenzen aus experimentellen Daten ableiten [1]. Dabei geht es ihnen vor allem darum, einen plausiblen Mechanismus zu finden, der klar macht, wie sich der Prozess der Wahrnehmung als Funktion des Gehirns verstehen lässt, wobei die Redeweise von „Wahrnehmung als Prozess der Inferenz“ andeutet, dass hier bereits eine Verschiebung

von der Auffassung, dass Wahrnehmung eine (wesentlich neutrale) Beobachtung der Gegenstände darstelle, hin zu einem Verständnis von Wahrnehmung als Form einer (vor)theoretischen Konstruktion stattgefunden hat.

Da zudem Wahrnehmung oft als Paradigma einer mentalen Operation aufgefasst wird, wäre eine Naturalisierung derselben ein Indiz dafür, dass unser mentales Vokabular durch eine Beschreibung von neurobiologischer Aktivität ersetzbar wäre. Und aus der Struktur einer so naturalisierten Wahrnehmungstheorie, so könnte man im Anschluss vermuten, ließen sich letztlich Schlüsse auf das phänomenale Bewusstsein ziehen. Wie wir aber sehen werden, ist es die eine Sache, eine globale, biologisch-mathematische Theorie des Gehirns aufzustellen (was schwer genug ist); eine andere Sache ist es jedoch, daraus eine Theorie des Bewusstseins abzuleiten.

Ohne auf die Metaphorik des dabei meist vorgefunden „Neuro- und Systemsprech“ näher eingehen zu wollen, muss an dieser Stelle mehr zum Verhältnis von Gehirn, Geist und Modellen gesagt werden. Das „Theorem des guten Regulators“ von Coant & Ashby (1970) soll zeigen, dass jeder gute Regulator ein Modell des zu regulierenden Systems darstellt. Dieser Befund gilt als zentrales Postulat der biologischen Kybernetik, welche die Brücke zwischen Informationstheorie und (Neuro-) Biologie schlagen will.

Es bezeichnen dabei  $E$  und  $R$  zwei dynamische Systeme. Den Produktzuständen  $X = E \times R$  werden über die Abbildung  $\psi$  Zustände aus eine Menge  $Z$  zugeordnet:

$$\psi : X = E \times R \rightarrow Z, \psi(e_j, r_k) = z_l \quad (3.1)$$

$Z$  ist dabei eine Menge beliebiger abzählbarer Ereignisse  $z_i$  (z.B. „Ziel getroffen“ oder „Ziel verfehlt“). Unter einer *guten Regulierung* wird verstanden, dass bei gegebenen  $E$  und  $\psi$  das Objekt

$$H(Z) = - \sum_i p(z_i) \cdot \log p(z_i) \quad (3.2)$$

durch Anpassung der Zustände von  $R$  minimiert wird;  $p(z_i)$  bezeichnet dabei die relative Wahrscheinlichkeit, den Zustand  $z_i$  (nach erfolgter Regulation) zu erhalten. Ein Verschwinden von  $H(Z)$  ist gleichbedeutend zur Forderung, dass der Regulator auf Änderungen in  $E$  so reagiert, dass die Schwankungen in  $Z$  minimal sind [2].

Dabei ist ein *guter Regulator* ein solcher, der eine gute Regulierung auf möglichst einfache Weise bewerkstelligt. In welcher Weise die Ereignisse  $z_i$  von den Zuständen  $r_k$  und  $e_j$  des Regulators bzw. des externen Systems  $E$  abhängen, ist durch die Zuordnung  $\psi$  festgelegt (der Einfachheit halber sei angenommen, dass die Zustände

von  $E$ ,  $R$  und  $Z$  abzählbar sind). Nun lässt sich  $\psi$  durch die Matrix  $\langle\psi\rangle_{j,k} \equiv \psi(e_j, r_k)$  darstellen:

$$\langle\psi\rangle = \begin{matrix} & e_1 & e_2 & \dots & e_n \\ \begin{matrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{matrix} & \begin{pmatrix} z_a & z_b & \dots & z_c \\ z_d & z_e & \dots & z_f \\ \vdots & \vdots & \ddots & \vdots \\ z_u & z_v & \dots & z_w \end{pmatrix} \end{matrix}. \quad (3.3)$$

Das Ereignis  $z_b$  wird also laut Gl. (3.3) etwa dann gefunden, wenn  $e_2$  und  $r_1$  vorliegen. Dabei kann  $\psi$  im Prinzip beliebig vielen Paaren  $(e_j, r_k)$  dieselben Objekte aus  $Z$  zuordnen. Beispielsweise könnte

$$\psi(e_2, r_1) = z_b \quad (3.4)$$

sowie aber auch

$$\psi(e_4, r_7) = z_b \quad (3.5)$$

gelten. Berücksichtigt man dies, kann  $p(z_l)$  als Summe von Wahrscheinlichkeiten ausgedrückt werden

$$p(z_l) = \sum_{\psi^{-1}(z_l)} p(e_j, r_k), \quad (3.6)$$

wobei die Summe über die Wahrscheinlichkeiten aller möglichen Vorkommnisse von  $(e_j, r_k)$ , die zum Ereignis  $z_l$  führen, geht;  $\psi^{-1}(z_l)$  bezeichnet dabei die Menge  $\{(j, k)\}$  für die gilt, dass  $\psi(e_j, r_k) = z_l$ .

Falls keine Regulierung stattfindet, sind die Verteilungen von  $E$  und  $R$  unabhängig voneinander,  $p(e_j, r_k) = p(e_j)p(r_k)$ . Für einen guten Regulator  $R$  wird jedoch dessen Zustand  $r_k$  abhängig davon sein, welcher Zustand  $e_j$  des externen Systems gerade realisiert ist. *Regulation kann demnach als zweckmäßiges Reagieren auf Änderungen in der Umwelt aufgefasst werden.* Die Wahrscheinlichkeit,  $R$  im Zustand  $r_k$  und  $E$  im Zustand  $e_j$  anzutreffen, schreibt sich dann als:

$$p(e_j, r_k) = p(e_j)p(r_k|e_j). \quad (3.7)$$

$p(R|e_j)$  ist dabei die bedingte Wahrscheinlichkeitsverteilung von  $R$  bei gegebenem  $e_j \in E$ . Einsetzen in Gl. (3.6) führt auf:

$$p(z_l) = \sum_{\psi^{-1}(z_l)} p(e_j)p(r_k|e_j). \quad (3.8)$$

Ein guter Regulator ist also durch die Verteilung  $p(R|e_j)$  charakterisiert, die  $H(Z) = \sum_l p(z_l) \log p(z_l)$  minimiert und dabei möglichst einfach (effektiv) ist. Dies ist aber genau dann der Fall, so das Theorem vom guten Regulator, falls  $R$  das zu regulierenden System  $E$  wie abbildet. Einem externen Beobachter würde der Zustand

des Regulators  $R$  als der des Systems  $E$  erscheinen, wenn er ihn durch die „Brille“  $h$  betrachtet. Jedem  $e_j$  wird dann genau ein  $r_k$  zugeordnet; dies ist äquivalent zur Aussage, dass es eine Zuordnungsfunktion  $h : E \rightarrow R$  gibt, welche die bedingte Wahrscheinlichkeitsverteilung  $p(R|e_j)$  des guten Regulators beschreibt [3]. Mit anderen Worten, *der Regulator ist ein Modell des Systems*, welches er reguliert, mit

$$p(r_k|e_j) = \begin{cases} 1 & \text{für } r_k = h(e_j) \\ 0 & \text{sonst} \end{cases} . \quad (3.9)$$

Unter der Annahme, dass das Gehirn regulierend tätig ist — etwa, wenn es einen homöostatischen Prozess steuert –, können wir, so das Theorem des guten Regulators, davon sprechen, dass das Gehirn ein Modell seiner Umwelt ist. Wenn wir dabei „ins Gehirn blicken“ könnten, dann würden wir, ein Modell seiner Außenwelt (inkl. des sensorischen Bildes des eigenen Körpers) vorfinden.

Eine Kritik, die an kybernetischen Betrachtungen geäußert wurde, lautet, dass kybernetische Prinzipien in erster Linie für menschengemachte, mechanisch oder anders regulierende technische Objekte gültig sind und nicht auf lebende Systeme übertragen werden sollten. Insbesondere die Annahme eines Regulators, der „von außerhalb“ in wohldefinierter Weise auf ein System einwirkt, scheint nicht mehr auf biologische Situationen übertragbar zu sein. Es würde ja immer schon vorausgesetzt werden, dass man von der Existenz einer gegebenen Zuordnungsfunktion  $\psi$  ausgehen kann. Sollte man hier also von einer Form von Intentionalität sprechen, so wäre diese, in den Worten von John Searle (1992), immer nur eine „abgeleitete“. Man könne dann aber nicht davon sprechen, dass der Regulator über einen Geist oder ein Bewusstsein verfügte – ähnlich wie man dies auch nicht einem Thermometer zusprechen würde. Man kann allerdings versuchen, das vorherige Resultat zu verallgemeinern.

### Wahrnehmung als biologische Operation

Wir interessieren uns nun in erster Linie für selbst-organisierende Systeme, die, anders als die Untersuchungsgegenstände zuvor, nicht mehr als Artefakte gedeutet werden können, die so entworfen oder programmiert wurden, dass sie ein bereits im Vorhinein bestimmtes Verhalten an den Tagen legen. Dies entgeht dem Vorwurf, dass deren „Intentionalität“ ja lediglich abgeleitet und somit nicht „echt“ sei. In anderen Worten, produzieren solche Systeme, sofern sie ein scheinbar intentionales Verhalten zeigen, dieses aus sich heraus und ohne äußeres Zutun.

Die Redeweise von *selbstorganisierenden* Systemen kann leicht missverstanden werden: Gemeint seien hier funktionale Ganzheiten – etwa „Organismen“, die sich *zu-*

*sammen* mit ihrer „Umwelt“ entwickeln –, deren jeweiliger Zustand aber nicht im Vorhinein festgelegt ist. Einerseits liegt dem die Beobachtung zugrunde, dass der Begriff der Selbstorganisation immer nur für ein gekoppeltes System (z.B. Organismus-Umwelt) sinnvoll ist (Ashby 1962); andererseits soll dies die *relative* Autonomie der Zustandsentwicklung eines so gekoppelten Systems zum Ausdruck bringen.

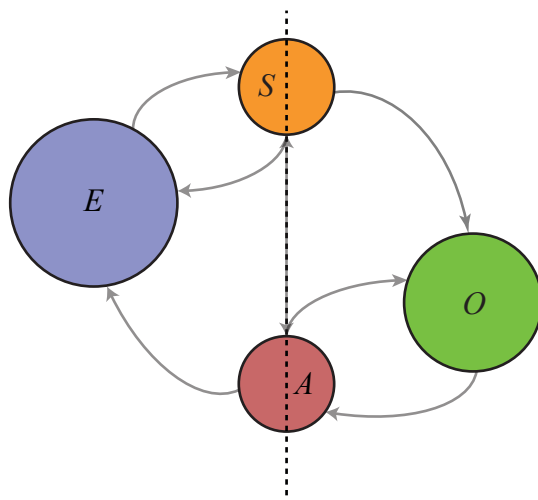
Dabei handelt es sich oft um *komplexe* Systeme, was implizieren soll, dass sich diese aus verschiedenen, funktional getrennten Komponenten (etwa ein zentrales Nervensystem, Motor- und Sinnesorgane) zusammensetzen und oft in nicht-linearer Weise gekoppelt, also chaotisch sind [4]. Die Regelmäßigkeiten, die sich in deren Dynamik ausmachen lassen, können dann als emergente Eigenschaften eben jener Systeme gedeutet werden.

Anders als bei der Diskussion des guten Regulators können wir nun nicht mehr vom Gegebensein einer Zielfunktion  $\psi$  ausgehen, welche den (immer wohl-definierten) Zuständen einer Umwelt  $E$  und den Zuständen eines davon separierbaren Regulators  $R$  eine Ereignismenge  $Z$  „zuordnet“. Allerdings findet man ein Pendant dazu in der Forderung nach der Begrenzung einiger ausgewählter „interner“ (physiologischer) Parameter oder davon abgeleiteter Größen.

Einer Hypothese von Karl Friston (2013) zufolge geschieht dies mit Notwendigkeit in einer Klasse von (ergodischen) Systemen, die über eine bestimmte, zellartige Struktur verfügen und sich fern vom thermodynamischen Gleichgewicht befinden [5]. Zu beachten ist, dass es sich nicht nur um *Zellen* im engeren Sinne handelt, sondern allgemein um Systeme, die durch eine „Membran“ gegen ihre Außenwelt abgeschirmt sind (Pearl 1988). Durch diese sollen aber Informationen (via einem „Sensorium“) ins Innere gelangen können; umgekehrt soll die Zelle auf ihre Umwelt einwirken können. Letzteres erfolgt aber nicht durch direkten kausalen Kontakt zwischen dem Zellinneren und seiner Außenwelt, sondern über einen „aktiven Bereich“, der wie das Sensorium Teil der Membran ist.

Man spricht in diesem Zusammenhang manchmal (mehr oder weniger metaphorisch) von „zirkulärer Kausalität“. Ganz ähnliche Konzepte finden sich in der Synergetik (wechselseitige Verursachung von Bestandteilen und Ordnungsstrukturen) sowie in der Kybernetik (Rückkopplungsprozesse), was noch einmal die strukturellen Analogien zwischen kybernetischen, systemtheoretischen und komplexitätswissenschaftlichen Ansätzen deutlich macht [6].

Etwas technischer gesprochen: Für die internen und aktiven Zustände des Organismus findet man im genannten Fall, dass deren Dynamik der Ableitung eines verallgemeinerten Potentials  $F$  entspricht, dessen Werte nur von den internen Zuständen selbst und denen der Membran abhängt [7]. Zudem stellt der zeitliche Mittelwert von



**Abbildung 3.1:** „Zirkuläre Kausalität“. Dargestellt ist der Zustandsraum des Gesamtsystems bestehend aus einer Umwelt ( $E$ ) und einem Sensorium ( $S$ ), dem aktiven Bereich ( $A$ ) und den internen Zuständen ( $O$ ) des Organismus; nach Friston (2013). Die Pfeile sollen die Existenz einer funktionalen Abhängigkeit der durch sie verbundenen Bereiche anzeigen. Die gestrichelte Linie soll andeuten, dass die internen Zustände des Organismus kausal von seiner Umwelt durch eine Membran  $M = S \times A$  abgeschirmt werden.

$F$  eine obere Schranke für die Dispersion  $H$  der internen, sensorischen und aktiven Zustände eines Organismus dar, d.h. dass jene von  $F$  begrenzt wird [8]:

$$\langle F(s, a, o) \rangle_t \geq H(s, a, o). \quad (3.10)$$

Anders als zuvor sind die Werte der internen Zustände  $o$  also nicht mehr durch eine vorgegebene Regulierung  $\psi(e, o)$  bestimmt (was auf  $o = h(e)$  und somit auf eine Abbildung der Umgebung in die internen Zustände führen würde), sondern folgen der Minimierung von  $F$  und begrenzen somit die entropische Tendenz des Organismus. Für eine (recht allgemeine) Klasse von Systemen findet man, dass diese Form der Selbsterhaltung auf natürliche Weise realisiert wird, *ohne* dass dabei von einem externen „Urheber“ auszugehen wäre.

Doch inwiefern kann dabei von *Wahrnehmung* gesprochen werden? Die Minimierung von  $F$  (und in Folge die Begrenzung von  $H$ ) kann prinzipiell auf zwei Arten erfolgen: Indem sich die internen Zustände des Organismus den externen anpassen, was dann so interpretiert werden könnte, dass der Organismus näherungsweise auf die äußeren Ursachen der Sinnesreizung „schließt“ (also „passiv“ bleibt); oder aber, indem der Organismus so auf seine Sinnesorgane und die Umgebung einwirkt, dass diese an seine Zustände angepasst werden („aktive Inferenz“). Dabei soll die

Redeweise, wonach ein Organismus „auf seine Umgebung einwirkt“ anzeigen, dass sich die Umwelt abhängig von denen aktiven Zuständen des Organismus entwickelt; und die Redeweise, dass ein Organismus auf „äußere Ursachen schließt“, soll anzeigen, dass die internen Zustände des Organismus (approximativ) ein Modell der Umwelt *verkörpern*, in welcher dieser eingebettet ist. Es handelt sich also nicht um einen Prozess, der abstrakt „im Geist der Zelle“ stattfindet. Wahrnehmung wird als physiologischer Prozess gedeutet, währenddessen sich Organismen graduell an ihre Umwelt anpassen. Dabei sind zwei Charakteristika herauszustreichen: Erstens handelt es sich bei der Wahrnehmung um einen dynamischen Vorgang, dessen Resultat einer Art Gleichgewichtsbeziehung (relatives Minimum von  $F$ ) entspricht. Zweitens spielen dabei sowohl aktive als auch passive Teilprozesse eine Rolle.

Bis jetzt betraf die Diskussion ein ergodisches System, das als formales Modell einer Zelle aufgefasst werden kann. Das Gehirn spielte dabei keine wesentliche Rolle und es könnte mit Recht behauptet werden, dass es sich primär nur um ein Modell „zellulärer Wahrnehmung“ handle. Dieses bildet jedoch die Grundlage, auf welcher ein komplexerer Mechanismus basiert, der zur Erklärung neuronaler Aktivität dienen soll. Dabei geht es in erster Linie um die neurophysiologische Realisierung dieses Mechanismus in hierarchischen Strukturen, die in kortikalen und neokortikalen Zentren des Gehirns gefunden werden (Friston 2010). Die Anpassung der internen Zustände findet nicht in einem Schritt statt, sondern erfolgt stufenweise und lässt sich ähnlich zur hierarchischen Datenverarbeitung in Computersystemen („*predictive coding*“) verstehen.

Dabei werden die sensorischen Informationen aus  $S$  in einem ersten Schritt mit einem Erwartungswert verglichen, der sich aus den Zuständen einer übergeordneten neurophysiologischen Ebene ableitet. Die Abweichung dieses Wertes wird in Folge an die höhere Ebene weitergereicht, welche diesen wiederum mit einem Erwartungswert gemäß der nächst-höheren Ebene vergleicht usw. Es werden also immer nur Abweichungen (bzw. Abweichungen von Abweichungen) zwischen den sensorischen Daten und den neuronal kodierten Erwartungswerten in die jeweils übergeordnete physiologische Ebene weitergereicht, bis das Gesamtsystem einen Zustand erreicht, der durch einen minimalen Wert von  $F$  charakterisiert wird.

Die internen Zustände verkörpern (gemäß eines neuronal realisierten Bayes'schen Mechanismus) ein „probabilistisches Modell“, dessen „Vorhersagegenauigkeit“ minimiert wird („*prediction error minimization*“, Hohwy 2013). Um dies zu tun, können die neuronalen Zustände entweder solange angepasst werden, bis die jeweiligen Abweichung zwischen Erwartungswerten und sensorischen Daten verschwinden, oder indem aktive Zustände entsprechend auf die Umgebung bzw. die Sinnesorgane ein-

wirken. Dies impliziert auch die Redeweise, wonach die Minimierung von  $F$  (wegen ihrer formalen Ähnlichkeit auch oft als „freie Energie“ des biologischen Systems bezeichnet) die eigentliche Funktion des Nervensystems sei – analog dem Herzen, dessen Funktion das Pumpen von Blut sei (Hohwy 2015b). Das Gehirn dient primär der Realisierung dieser Funktion.

Inbesondere die Subsumierung und Formalisierung mehrerer Ansätze zur Erforschung des Gehirns unter denselben theoretischen Rahmen (Friston 2010) scheint das hier vorgestellte Modell besonders attraktiv zu machen. Das zentrale Konzept der Minimierung der Vorhersageungenauigkeit wird als Blaupause zur Analyse mentaler Funktionen herangezogen und auf verschiedenste Themenfelder angewandt, welche dann (vermeintlich wertfrei) als naturalisiert betrachtet werden: Neben der Wahrnehmung sind das etwa Entscheidungsprozesse (Schwartenbeck *et al.* 2015), Emotionen (Seth 2013) oder Psycho-Pathologien (Hohwy 2015a); und selbst die Psychoanalyse soll unter dieses Prinzip subsumiert werden können (Carhart-Harris & Friston 2010).

Doch lässt sich daraus auch eine neurobiologisch realistische und wesentlich computationalistische Theorie des Bewusstseins ableiten?

## 3.2 Funktionen des Bewusstseins

An dieser Stelle offenbart sich eine Schwierigkeit. Die vielleicht wichtigste Konsequenz, die nach den letzten beiden Abschnitten gezogen werden kann, ist die Feststellung, dass Modellierung eben *nicht* als die kognitionsbiologische Lesart einer Abbildungstätigkeit des Bewusstseins anzusehen ist, sondern eine (emergente) Eigenschaft, die aus der „Selbstregulierung“ (kybernetisch gesprochen) bzw. aus der Selbsterhaltung eines biologischen Systems (gemäß der Theorie komplexer Systeme oder der Synergetik) folgt oder mit dieser einher geht. Oder umgekehrt: Obwohl das Gehirn „ein Modell seiner Umwelt konstruiert“, ist das Bewusstsein nicht notwendigerweise ein Abbild der Welt und auch der Verweis auf ein Gehirn, das „Inferenzen zieht“, ist in diesem Sinne missverständlich. Zwar kann vielleicht gesagt werden kann, dass die Aktivitäten des Gehirns, so aussähen, als würde es „Schlüsse und Vorhersagen mit den Sinnesdaten abgleichen“, allerdings ist dies wohl eher metaphorisch als wörtlich zu verstehen. Bewusstseinsinhalte sind nicht Elemente von Mengen, die auf eine subjekt-unabhängige Wirklichkeit abbildbar sind.

Für manche sind solche Metaphern vielleicht Paradefälle einer falschen oder zumindest missverständlichen Sprechweise („das Gehirn denkt“). Andererseits zeigen sie an einem konkreten Beispiel, unter welchen Annahmen – *Repräsentation* als Korrelation zwischen internen Zuständen und systemischen Kausalbeziehungen, *Intentionalität*



als nicht-abgeleitetes und (scheinbar) zweckmäßiges Verhalten von Organismen und *Subjektivität* als spezifische Art der Darstellung ihrer Umgebung – Aussagen über den Geist durch Aussagen über materielle Systeme ersetzt werden.

Dabei ist vom Bewusstsein, verstanden als „magische Zutat“, die erst „echte“ Wahrnehmung und Intentionalität von der scheinbaren Intentionalität eines Thermometers oder der „a-mentalen“ Wahrnehmung einer Amöbe unterscheidet, jedoch nirgends die Rede. Anscheinend braucht es den Verweis auf „das Bewusstsein“ gar nicht, um bestimmte, bisher als *mental gedeutete* Phänomene zu erklären. Die Frage, die sich nun stellt, ist, ob dies für *alle* mentalen Aktivitäten möglich ist, d.h. ob Bewusstsein überhaupt eine Rolle in einer (naturalistischen) Beschreibung des Geistes spielt.

Doch entspricht es nicht einer Tatsache, dass wir uns der Welt bewusst sind? Eine Beobachtung zum Verhältnis zwischen Bewusstsein und Hirnprozess scheint zu folgender Interpretation zu verführen: Da man sich nie eines neuronalen Prozesses, sondern maximal seines Resultates bewusst ist, scheint es plausibel zu sein, dass eine Theorie des Gehirns die bewussten Endprodukte neuronaler Prozesse spezifiziert. Angewandt auf obige Überlegungen würde dies bedeuten, dass wir uns zwar nie des Prozesses der Selbsterhaltung via Inferenz bewusst sind, dennoch aber dessen Resultat – das sensorisch abgeglichene Modell der Außenwelt – „in unserem Bewusstsein tragen“. Problematisch an dieser Interpretation ist jedoch die Postulierung eines Bewusstseins, auf „dessen Bühne“ die Resultate neuronaler Prozesse erscheinen würden – eine Sichtweise, die naturalistische Ansätze ja eigentlich vermeiden sollten (Dennett 1991a).

Betrachten wir für einen Moment weniger die Produkte neuronaler Prozesse oder, alternativ gesprochen, die durch sie vermeintlich spezifizierten „Objekte des Bewusstseins“, und fragen stattdessen danach, wie denn Bewusstsein überhaupt anderes verstanden werden könnte denn als Epiphänomen. Im Rahmen einer kognitionspsychologischen Theorie bedeutet dies, zu fragen, ob dem Bewusstsein überhaupt irgendeine Funktion zukomme bzw. mit dieser zu identifizieren sei. So wird beispielsweise in der *Global-Workspace*-Theorie von Bernard Baars (1988) dem Bewusstsein die Rolle der zentralen Verfügbarmachung von Information („*global broadcasting*“) zur weiteren Verarbeitung zugeschrieben.

Nun lassen sich zwei Fälle unterscheiden: Im ersten könnte man vermuten, dass uns dessen kausale Rolle gar nicht bewusst ist. Bewusstsein wäre wesentlich opak. Eine andere Weise, dies auszudrücken, wäre es zu sagen, dass, aus einer dritt-personalen Perspektive betrachtet, Bewusstsein identisch mit einer Funktion des Gehirns ist, jedoch das Ausüben einer solchen Funktion in einer (als komplementär zu betrach-

tenden) erstpersionalen Perspektive niemals wahrzunehmen ist.

Dass dies im Rahmen des Funktionalismus auf den ersten Blick plausibel erscheint, aber bei näherem Nachdenken widersinnig ist, dafür argumentiert Max Velmans (2009, S. 259f.) wie folgt:

1. Bewusstsein ist identisch zu einer bestimmten Hirnfunktion.
  2. Solche Funktionen sind vollständig als Transformationen zwischen Sinnesreizen, Verhaltensantworten und anderen funktionalen (mental) Zuständen spezifizierbar.
  3. Ich bin mir dieser Transformationen aber gar nie bewusst, d.h. ich kann sie auch nicht bewusst steuern.
- ∴ Bewusstsein ist identisch zu einer Funktion (innerhalb einer Theorie), hat aber keine Funktion (für mich).

Im zweiten Fall hätte Bewusstsein eine Funktion, derer wir uns auch bewusst sind. Beispielsweise wird dem Bewusstsein oftmals die Funktion zugeschrieben, unsere Aufmerksamkeit selektiv auf Gegenstände oder Abläufe zu lenken: Wenn wir z.B. lernen, ein Auto zu fahren, dann achten wir auf die einzelnen Prozesse, wie das Einlegen des richtigen Ganges, das korrekte Setzen des Blinkers oder das Halten der Spur. Erst allmählich werden diese Abläufe internalisiert und finden fortan unbewusst, d.h. ohne dass wir ihnen Aufmerksamkeit schenken, statt. In der Tat ist es sogar so, dass wir im allgemeinen Sprachgebrauch oft nicht zwischen Bewusstsein und Aufmerksamkeit unterscheiden – sich eines Objekts in einer visuellen Darstellung bewusst zu sein wird manchmal synonym dazu verstanden, dass man seine Aufmerksamkeit auf eben jenes Objekt gerichtet hat.

In den letzten Jahren wurde jedoch in kognitionswissenschaftlichen Experimenten gezeigt, dass bewusste Wahrnehmung und selektive Aufmerksamkeit („*top-down attention*“) zwei zu unterscheidende Phänomene sind, die zwar oft zusammen auftreten, aber getrennt manipuliert werden können und beizeiten sogar gegenläufige Effekte tragen (Tsuchiya & Koch 2009). Auch die Unterscheidung zwischen phänomenalem „P-Bewusstsein“ und funktionalem „A-Bewusstsein“ impliziert, dass zwischen Bewusstsein und aufmerksamsbedingter Verarbeitung von Reizen zu unterscheiden ist (Block 1995, Block 2011).

Im Allgemeinen bleibt immer ein hartnäckiger Zweifel bestehen: Warum sollte es so sein, dass diese oder jene Funktion identisch zu Bewusstsein ist? Ein funktional analoger Hirnmechanismus, der ohne Bewusstsein stattfindet, ist prinzipiell immer

vorstellbar und wäre nicht auf eine nomologisch unmotivierte, mysteriöse Identität zwischen Bewusstsein und Funktion angewiesen.

**Die Metapher des Sprechens:**

Um obige Kritik am Funktionalismus zu veranschaulichen, kann folgendes Bild gebraucht werden, das den Zusammenhang von Sprechen und Bewusstsein zum Thema hat.

Man spricht manchmal davon, dass wir uns erst bewusst würden, was wir wirklich sagen wollten, nachdem wir es gesagt hätten. Schon Heinrich von Kleist berichtete „über die allmähliche Verfertigung der Gedanken beim Reden“. Dabei beobachtete er, dass einem ein Gedanke oft erst dann klar würde, wenn man ihn zu artikulieren sucht. Doch inwiefern kann dann überhaupt das Sprechen als „Funktion des Bewusstseins“ bezeichnet werden? Wie sollten wir „unsere Gedanken in Sprache kleiden“, wenn uns deren Inhalte erst durch das Sprechen bewusst würden? Es ließe sich dann ja nur schwerlich behaupten, dass gemäß dieser Inhalte das Sprechen (bewusst) gesteuert oder verursacht worden wäre.

Nun, vielleicht sollten wir uns einfach damit zufrieden geben, dass die Inhalte des Bewusstseins identisch zu den Endprodukten unseres Sprechens sind? Im Idealfall wären wir plappernde Automaten, die ihr Plappern wahrnehmen können. Doch warum sollte dann überhaupt noch vom Bewusstsein gesprochen werden, wenn dieses sowieso keine Rolle mehr spielt?

Interessant an dieser Stelle ist jedoch die Feststellung, dass, phänomenologisch betrachtet, der Unterschied zwischen der Bewusstwerdung eines Prozesses und der Bewusstwerdung seines Endproduktes ein *relativer* ist: Gegenstand einiger Meditations- oder Achtsamkeitstechniken ist das Gewahrwerden des Entstehens psychischer (emotionaler) Muster, noch bevor oder während sich diese entfalten. Es geht also um die (vorzeitige) Einsicht in einen Prozess, von dem wir normalerweise nur den Endzustand erfahren, auch wenn man sich dabei freilich nicht der neurobiologischen Aktivität bewusst wird, die einen solchen Prozess begleitet.

Ein anderes Beispiel liefern die sinnlichen Wahrnehmung. Diese erscheinen uns ja wesentlich als „transparent“ (Metzinger 2003b), d.h. dass wir uns für gewöhnlich nur der Eigenschaften der wahrgenommenen Objekte bewusst sind, nicht aber des Prozesses der Wahrnehmung: Wir sehen durch den Wahrnehmungsapparat „hindurch“. So ist uns die Entstehung einer dreidimensionalen Szene aus dem zweidimensionalen Bild auf der Netzhaut und den daraus abgeleiteten „Entwürfen“ (Marr 1982) introspektiv nicht zugänglich. Andererseits gibt es Techniken, die uns (auf der phänomenalen Ebene) die Entstehung von Dreidimensionalität bewusst werden lassen. Dieser Effekt lässt sich etwa durch Stereogramme erzielen. Die Entstehung von Dreidimensionalität, wie wir sie in Stereogrammen erfahren, lässt sich dabei wiederum kognitions-

wissenschaftlich untersuchen (z.B. Poggio & Poggio 1984).

Auch die Gewissheit, wonach es sich bei den Gegenständen unserer Wahrnehmung um reale Dinge handelt, ist Resultat eines Prozesses. In den Techniken der *Virtual Reality* werden wir vermehrt mit der Tatsache konfrontiert, dass Realität als „internes Attribut“ der Wahrnehmung zugeschrieben wird und nicht in den Sinnesreizen selbst zu suchen ist (Mausfeld 2012). Die Aufgabe der Wahrnehmungspsychologie, so Mausfeld in Anlehnung an die Gestaltpsychologie, sei es nun die „Eingengesetzlichkeit“ dieser Realitätszuschreibung in die Charakterisierung von Wahrnehmungsphänomenen einfließen zu lassen und nicht mehr als Ergebnis eines grundsätzlich opaken Informationsverarbeitungsprozesses zu betrachten, der lediglich auf den sensorischen Daten beruht, die zu den wahrgenommenen Objekten korrespondierend.

Wir können also zumindest davon sprechen, dass wir uns die *Form* dieser Prozesse bewusst machen können und dies gilt auch dann, wenn uns die dazu *korrelierende Hirnaktivität* verborgen bleiben.

Noch einen Schritt weiter geht die Vorstellung, dass Bewusstsein selbst bestimmte Effekte in der physischen Welt bewirken kann [9]. Neben alltäglichen Erfahrungen, wonach Bewusstseinsakte willentlich Bewegungen einleiten können (wenn ich etwa meinen Arm hebe) und der früher diskutierten Situation, wonach wir im Rahmen wissenschaftlicher (experimenteller) Forschung immer schon ein intentionales Bewusstsein voraussetzen, lassen sich noch weitere Fälle anführen, in denen Bewusstsein, eine empfundene Ursächlichkeit und physiologische Effekte zumindest korrelieren. Ein schönes Beispiel dafür liefert das Neurofeedback, eine Technik, bei der Probanden lernen, mittels einer akustischen oder visuellen Darstellung die eigene Hirntätigkeit (global oder lokal) zu beeinflussen (deCharms *et al.* 2005). Andere Beispiele liefern die Entwicklungen auf dem Gebiet der Prothetik, bei der Patienten via Gehirn-Computer-Schnittstellen bestimmte Motorfunktionen willkürlich steuern können (Hochberg *et al.* 2006).

Aus der Beobachtung, dass es anscheinend ein enges Verhältnis, vielleicht sogar einen Wirkungszusammenhang, zwischen Bewusstsein und physiologischen Phänomenen gibt, folgt lediglich, dass diejenigen Positionen, die dem Bewusstsein einzig den Status eines stillen Beobachters zusprechen, verdächtig sind. Selbst der Physikalismus scheint, zumindest auf den ersten Blick, sowohl mit der Vorstellung, dass Bewusstsein kausal inert als auch kausal wirksam ist, vereinbar zu sein (vgl. Perry 2001, S. 78f.), wenngleich damit über den etwaigen Mechanismus noch gar nichts gesagt wurde.

Was sich allerdings neben der weit verbreiteten Meinung, wonach der Funktionalismus die Phänomenalität des Bewusstseins nicht erklären könne, feststellen lässt,

ist, dass im Funktionalismus die Struktur der Gehirn-Geist-Beziehung problematisch bleibt: Funktionalisten landen schnell bei widersinnigen Aussagen („Bewusstsein ist eine Funktion, hat aber keine“) oder, bestenfalls, beim Epiphänomenalismus, also der Vorstellung, dass ein (phänomenales) Bewusstsein zwar existiere, aber keinerlei Wirksamkeit besitze.

### 3.3 Zeichen und Struktur

#### Komplementarität statt Korrespondenz

Wie kann die vermutete Beziehung zwischen den Inhalten unseres Bewusstseins und der physischen Aktivität von Nervenzellen dargestellt und vielleicht sogar empirisch zugänglich gemacht werden? Wie wir eben gesehen haben, kommt eine funktionalistische Theorie dabei schnell in Erklärungsnot. In diesem Zusammenhang wurden unlängst die Prinzipien der *Strukturellen Kohärenz* (Chalmers 1996, Kap. 6) bzw. der *Gleichartigkeit von Repräsentations- und Informationstrukturen* (Velmans 2009, Kap. 13) vorgeschlagen. Diese beruhen auf der Vorstellung, dass *zwei* („epistemischen“) Perspektiven der *einen* („ontischen“) Wirklichkeit zugrunde liegen, die als nicht-näher explizierte Struktur von Informationsflüssen gedeutet wird. Der Erhalt dieser Kohärenz wird dabei als zentrales theoretisches Kriterium einer Wissenschaft des Bewusstseins angesehen. Hauptmotivation ist eine angenommene Korrespondenz zwischen erst- und dritt-personalen Aussagen und der Schluss auf eine Repräsentationsbeziehung zwischen ihnen und einer „psycho-physisch neutralen“ Welt. Im Gegensatz dazu wollen wir für die Notwendigkeit argumentieren, die beiden Perspektiven als kontextuelle und komplementäre Sichtweisen zu verstehen. Dies soll im Rahmen einer Zeichentheorie des Bewusstseins geschehen. Eine solche wird sich unter anderem auf die folgenden Punkte stützen:

1. Objekte werden zu *Zeichen*. Die Zustände eines Regulators sind immer bezogen auf das System, welches reguliert werden soll. Die Beziehung zwischen dem Regulator und seiner „Umwelt“ ist jedoch spezifisch – nur das kann als Umwelt gelten, was auch tatsächlich dem Einfluss der Regulierung ausgesetzt ist und umgekehrt [10]. Im Falle des Regulators kommt es dabei zu einer *expliziten* Darstellung der Welt, verkörpert in den Zuständen des Regulators. Dies ist für ein selbstorganisierendes System nicht mehr der Fall. Von einem Organismus kann beispielsweise höchstens gesagt werden, dass einem externen Beobachter die interne Zustände des Organismus als (probabilistisches) Modell der Beziehungen zwischen ihm und seiner Umwelt erscheinen würde. Aus

der Sicht des Organismus geht es jedoch primär nur darum, seinen eigenen (strukturellen) Erhalt zu garantieren. Somit muss die Vorstellung, dass zwischen seinem „Inneren“ und seiner Außenwelt eine Beziehung der Abbildung existieren würde, modifiziert oder sogar ersetzt werden.

Dies kann dadurch geschehen, dass fortan Wahrnehmung als Zeichenprozess verstanden wird, welcher dem Organismus ein scheinbar zweckmäßiges Handeln ermöglicht. Unter einem „Zeichen“ soll dabei ein jedes Objekt verstanden werden, das auf einen Gegenstand verweist, der nicht unmittelbar erfahren werden kann (etwa weil er hinter einer „Membran“ verborgen ist) und eine Änderung im System (etwa in einem Organismus) bewirkt, welches von ihm Gebrauch macht, also dessen „Interpretant“ (Peirce 1983, §4) er ist. Um in ihrer Umwelt bestehen zu können, benutzen Organismen Zeichen, die deren Verhalten steuern oder zur Kommunikation mit anderen Organismen dienen. Ein solcher Zeichenbegriff ist *prima facie* nicht „mentalistisch“.

Doch wie ließe sich dann überhaupt noch vom Bewusstsein sprechen? Einerseits dadurch, dass *Bewusstsein* selbst nur zu einem Zeichen unter vielen würde, so etwa, wenn wir von „dem Bewusstsein“ sprechen als würde es auf einen Gegenstand verweisen, welcher in den Köpfen von Menschen herumspukt und in näher zu spezifizierender Art und Weise auf andere Gegenstände einwirkt. Diese Redeweise ist, wenn sie anders als metaphorisch gebraucht wird, mit so vielen Problemen behaftet, dass sie am besten ganz aufgegeben wird. Andererseits lässt sich *Bewusstsein* aber auch so verstehen, dass es ein bestimmtes Moment oder eine bestimmte Vollzugsweise dieses Zeichenprozesses benennt. Eine wichtige Unterscheidung ist daher die zwischen Bewusstsein (verstanden als Form des Zeichenprozesses) und den verweisenden Objekten selbst (verstanden als dessen Inhalte):

2. Diejenigen Zeichen, die wir *Bewusstseinsinhalte* nennen, sind wesentlich strukturiert. Andernfalls würden wir unser Bewusstsein als undifferenziertes Empfinden erleben, das beispielsweise als instantaner Nebeneffekt einer autonom ablaufenden Hirnaktivität in Erscheinung träte und auf phänomenaler Ebene weder über eine zeitliche noch über eine räumliche Struktur verfügen würde. Dies erscheint vielleicht noch für eine simple Motorleistung wie das Strecken des rechten Handgelenks plausibel, nicht aber für komplexes, zielgerichtetes Verhalten wie die Durchführung eines Experiments oder die *feedback*-gesteuerte Beeinflussung neuronaler Aktivität.

Einen anderen Hinweis, der für eine Strukturiertheit von Bewusstseinsinhalten spricht, liefern deskriptive Verfahren, die über die Beobachtung der Alltagserfahrung hinausgehen. So geht etwa die Phänomenologie selbst auf der Ebene

von einfachen Wahrnehmungen von einer reichhaltigen Struktur der damit verbundenen Erfahrung aus. Folgende weiterführende Beobachtungen können an dieser Stelle zur fundamentalen Strukturiertheit der Bewusstseinsinhalte angeführt werden:

- Ein Charakteristikum, das in Abschnitt 6.2 noch ausführlich behandelt werden wird, ist die sogenannte „Einheit des Bewusstseins“. Dabei mag auf den ersten Blick gar nicht klar sein, dass dies weiterer Erklärungen bedürfte, sind wir es doch aus der alltäglichen Wahrnehmung gewohnt, dass die Welt von einheitlichen und wesentlich selbstständigen Objekten bevölkert ist. Doch warum ist das so? Ein naheliegender Grund wäre, dass dies nun einmal die (objektive) Existenzweise von (makroskopischen) Dingen widerspiegeln würde. Doch auch dann stehen wir vor dem Rätsel, wie denn räumlich und zeitlich getrennte Prozesse innerhalb (des Gehirns) eines Organismus zu einer einheitlichen Wahrnehmung verschmelzen.
- Sowohl Körper als auch Bewusstsein scheinen in doppelter Weise gegeben zu sein: Der Körper kann als Objekt betrachtet werden – etwa wenn ich an mir herab blicke und die Bewegung meiner Füße beobachte –, aber gleichzeitig wird er immer auch als *empfundener* Leib wahrgenommen – wenn ich die Bewegung meiner Füße spüre. Bewusstsein wird oft als zu erklärendes Objekt betrachtet – etwa wenn ich es zum Gegenstand einer biologischen Theorie machen will –, bildet gleichzeitig aber immer auch den subjektiven Hintergrund einer solchen Aktivität: Die Objekte der Theorie, die *ich* anwenden *will*, scheinen *Objekte für mein* Bewusstsein sein. Körper und Bewusstsein sind in doppelter Weise gegeben: einerseits als bezeichnendes Objekt innerhalb eines Prozesses, andererseits aber auch als Grundlage der Möglichkeit und Realisierung des Prozesses.
- Die Inhalte meiner Wahrnehmung sind immer schon geformt: Die Dinge sehe ich *im* Raum und *außerhalb* meines Leibes, die Berührungen spüre ich *auf* der Haut und mein Denken findet vermeintlich *im* Kopf statt. Auch deren zeitliche Struktur ist auffällig. Mentale Zustände werden oft als eindeutig individuierbare aber zeitliche ausgedehnte und „fließende“ Ereignisse beschrieben. Entweder wir glauben daher, dass unsere Erfahrung selbst strukturiert ist, oder dass diese äußerlich strukturiert wird.

Sowohl Punkt 1 als auch Punkt 2 legen nun nahe, eine Korrespondenztheorie durch eine Komplementaritätsauffassung zu ersetzen. Punkt 1 behauptet schließlich, dass eine natürlich erscheinende Abbildungsbeziehungen nur die Folge der Wechselwirkung evolutionär oder „intentional“ aufeinander bezogener Systeme ist. Die Rede-

weise von einer doppelten Sichtweise auf denselben Gegenstand erscheint aus dieser Perspektive als irreführend [11].

Punkt 2 spricht gegen eine Korrespondenztheorie, da die Einheit des Bewusstseins nicht ohne weiteres als isomorph zur Existenzweise physikalischer Gegenstände gedacht werden kann. Die einheitliche Struktur der Objekte ergibt sich ja erst aus einer vermittelnden Wechselwirkung – sowohl gemäß der transzendentalen als auch der physikalistischen Lesart, wenn auch unter jeweils anderen Vorzeichen. Ein weiteres Argument, das gegen eine Isomorphie zwischen der Existenzweise der physikalischen Gegenstände und der des Bewusstseins spricht, ist die bereits angesprochene Zeitlichkeit: eine parametrische Struktur auf der einen, eine ausgedehnte Momentstruktur auf der anderen Seite.

Zudem gibt es außer unserer metaphysischen Überzeugung gar keinen Grund anzunehmen, dass die zweifache Art des Gegebenseins eine kategorische sei. So verweisen die erst- und dritt-personale Perspektive auf keine Dichotomie, wie es etwa die Rede von „Innen“ und „Außen“ vermuten ließe und die es irgendwie „zu überwinden“ gäbe: Berichten wir aus der ersten Person über unsere Erfahrung, greifen wir dazu immer bereits auf eine objektive Ausdrucksform, d.h. auf die Beschreibungen der dritt-personalen Perspektive zurück. Eine Schilderung aus der erst-personalen Perspektive ist also streng genommen nicht „rein subjektiv“, sondern wesentlich das Resultat eines Prozesses – gegebenenfalls kann sie als erstarrtes Ergebnis eines solchen Prozesses gedacht werden [12]. Die Art und Weise, wie sich Bedeutung unter Rückgriff auf das (objektive) System der Sprache ausdrückt, wird paradigmatisch etwa bei Eugene Gendlin (1997) behandelt. Dabei wird sogar davon gesprochen, dass Bedeutung aus dem Prozess der (subjektiven) Erfahrung *entsteht* („*creation of meaning*“).

Umgekehrt sind die Beschreibungen der dritten Person wiederum aus (erst-personalen) Erfahrungen abgeleitet. Historisch kann dieser Prozess der Verobjektivierung, der durchaus in „Brüchen“ verläuft (Bachelard 1987), etwa an der Entwicklung der Thermodynamik im 19. Jahrhundert illustriert werden, bei der sich die (objektiven) Bezeichnungen wie Wärme oder Arbeit aus den (subjektiven) Wahrnehmungen und lebensweltlichen Praktiken der beteiligten Akteure heraus entwickelten, um schließlich ihren subjektiven Charakter abzulegen und als rein objektive Größen zu erscheinen. Besonders anschaulich wird dies am vermeintlich so abstrakten Begriff der Entropie, der in seiner Übersetzung nichts anderes als den „Wandlungsinhalt“ (ἐν τροπία) bezeichnet.



## Die zwei Tische

Das dualistische Fehltrium, dass diese zwei Perspektiven auf eine gänzlich verschiedene *Seinsweise* verweisen würden, ist dadurch bedingt, dass die genetischen Abhängigkeiten zwischen den Perspektiven übersehen werden und stattdessen auf eine zugrundeliegende ontologische Differenz geschlossen wird.

Die Probleme, die uns eine solche Dichotomie beschereu würde, werden in einer Parabel Sir Arthur Eddingtons veranschaulicht. Am Beginn seines Buches *The Nature of the Physical World* erzählt Eddington von zwei Tischen, dem einem an dem er sitzt, den er als farbig, ausgedehnt und substantiell wahrnimmt und der andere, den die Physik beschreibt und der zum größten Teil aus Leere und vereinzelt, farblosen Teilchen besteht. Wenn Eddington fragt, wie sich die Metapher der beiden Tische auflösen lässt, um zu dem Bild der *einen* Welt zu gelangen, steht er vor einem Rätsel, denn „der Vorgang, durch den die externe Welt der Physik transformiert wird in eine Welt vertrauter Bekanntschaft im menschlichen Bewusstsein, *liegt außerhalb der Physik*“ [13].

Gibt es einen Mechanismus, der uns diese Transformation plausibel machen könnte? Von der Existenz *zweier* Tische auszugehen, erschiene suspekt, einerseits weil dies dem Ockham'schen Gebot der Denkökonomie zuwider laufen würde, andererseits weil ja beide Gegenstände auf intimste Weise miteinander verbunden sind: Die Farbe von Tisch<sub>1</sub> korreliert zu den (gegebenenfalls neuronal verarbeiteten) Wellenlängen der von Tisch<sub>2</sub> reflektierten Photonen unter typischen Beleuchtungsbedingungen, seine Härte der zellulären und kristallinen Struktur, seine Substanzhaftigkeit den Wechselwirkungen zwischen den Rezeptoren meines Körpers und den Molekülen von Tisch<sub>2</sub> usw.

Folgende Auflösung des Eddington'schen Dilemmas liegt nahe: Die ganze Verwirrung besteht ja nur darin, dass wir uns mit unterschiedlichem „Vokabular“ auf *denselben* Tisch beziehen, einmal unmittelbar und subjektiv als Wahrnehmende, das andere mal mittelbar und objektiv als Wissenschaftler. Dies tun wir mit jeweils zueinander korrespondierenden Begriffen (Konzepten, Kategorien,...), was schließlich im Studium des Gehirns seinen Höhepunkt findet, schließlich kreuzen sich dort Erkennen und Erkanntes, Wahrnehmen und Wahrgenommenes, Denken und Gedachtes: Als der Teil der objektiven Welt, der das Bewusstsein *erzeugt*, müsste das Gehirn (die Erkenntnis von) sich selbst erzeugen [14]. Doch ist man einmal hier angelangt, kann die „bloß subjektive“ Art des Gegebenseins nicht mehr weiter verschoben werden: Wenn wir die (subjektiven) Eigenschaften der erkannten Gegenstände tatsächlich aus den (objektiven) Eigenschaften des Erkenntnisapparats rekonstruieren könnten, was wäre mit den subjektiven Eigenschaften des Erkennens – etwa, dass es *meine*

Erkenntnis ist und nicht die eines Menschen, der zufällig meinen Namen trägt – zu tun?

Alternativ könnte man einwenden, dass wissenschaftliche Aussagen ja sowieso nur von einem je-anderen Gehirn und dessen als mental *gedeutete* Funktionen zu tätigen sind. Um ein Wort Carnaps aufzugreifen, wären dann nur Aussagen über die Fremdpsyche und nicht die Eigenpsyche zulässig; Wissenschaft wäre *Objektwissenschaft*. Dies ist solange unproblematisch, wie nicht gleichzeitig davon ausgegangen wird, dass wissenschaftliche Theorien, die Welt in ihrer Ganzheit abbilden und dabei Aussagen über Bewusstsein ermöglichen. Es stellten sich dann nämlich folgende Schwierigkeiten:

1. Ein Teil dieser Welt, nämlich unser eigenes Bewusstsein, ist nicht Teil der wissenschaftlichen Beschreibungsweise. Die Theorien der Wissenschaft sind also, entgegen aller physikalistischen Überzeugung, *unvollständig*,
2. Die Bedingungen, unter denen jene Theorien entstehen und gerechtfertigt werden (z.B. unser Denken und Wahrnehmen), können nicht zufriedenstellend erklärt werden. Was vielleicht auf den ersten Blick als eher philosophisches Problem erscheint, bedroht die physikalistische Einstellung noch mehr. Anders als zuvor, wo wir die Wirklichkeit in zwei Bereiche aufteilen mussten – in einen „objektiven“, von dem unsere Theorien handeln, und in einen „subjektiven“, über den wir letztlich wissenschaftlich gar nichts aussagen können, scheint hier die Existenz eines unphysikalischen Bewusstseins postuliert zu sein: Wann immer wir wissenschaftliche Aussagen machen, müssen wir einen diesen Aussagen übergeordneten Geist dazu denken, der jedoch selbst nicht Teil der Wirklichkeit ist, die von unseren Theorien beschrieben wird.

Eine empirisch orientierte Wissenschaft von Bewusstsein *kann* demnach also nicht aus einer reinen Wissenschaft vom Gehirn abgeleitet werden, wenn sie mehr sein will als eine unvollständige, immer auf eine letzte Instanz verweisende Sammlung gesetzmäßiger Zusammenhänge. Selbst Ansätze zur Untersuchung des Bewusstseins, die auf den ersten Blick rein neurobiologische sind – also die davon ausgehen, Bewusstsein würde sich lediglich mit Verweis auf die Dynamik des Gehirns verstehen lassen – sind immer bereits auf psychologische oder phänomenologische Vorannahmen angewiesen. Die Suche nach den neuronalen Korrelaten des Bewusstseins in den Bildern neuronaler Aktivität setzt das Vorhandensein von Bewusstsein (und im Idealfall eine Kenntnis dessen Inhalts) in den untersuchten Subjekten bereits voraus und erklärt ja gar nicht, wie Bewusstsein entsteht, sondern nur wie es zu Gehirnaktivität korreliert. Ein stures Festhalten am Physikalismus führt also nicht dazu,

dass sich das Rätselhafte am Bewusstsein letztlich erschließen würde. Das *hard problem* wäre unter diesen Voraussetzungen nicht nur schwer (dies impliziert, dass es ja doch irgendwie gelöst werden könnte), sondern unlösbar. Es kann höchstens so weit aufgeschoben werden, bis es uns nicht mehr interessiert.

Doch ließe sich überhaupt eine Theorie formulieren, die nicht versucht, so vom Bewusstsein zu sprechen wie von anderen Objekten „in der Welt“, die mit unterschiedlichem Vokabular, das aber letztlich aus zueinander korrespondierenden Begriffen aufgebaut ist, beschrieben werden? Auf die Frage nach der Rolle von Bewusstsein antwortet Velmans (2009, S. 347) am Ende seines Buches *Understanding Consciousness* mit der Bemerkung, dass die wesentliche Funktion von Bewusstsein darin bestünde, „subjektive Realität“ zu erschaffen: Keine Beschreibungsweise würde uns verständlich machen, wie es sich anfühlt, einen Sachverhalt zu *realisieren*. Doch was könnte damit gemeint sein und lässt sich hier vielleicht ein etwas weniger mystisches Bild zeichnen?

Die Beantwortung dieser Fragen erfordert eine neue, *komplementäre* Perspektive, die wir in den folgenden Abschnitten einnehmen wollen, wobei wir nicht länger fragen wollen, wie Bewusstsein als empirischer Gegenstand aufzufassen ist, sondern wie Bewusstsein als Form eines subjektiven Prozesses, der Objekte hervorbringt, verstanden werden kann.

Der Ansatz, der in den folgenden Abschnitten gewählt wird, versucht, Subjektivität als Zeichenprozess zu beschreiben. Deutlich zu machen, wie Zeichenprozesse strukturiert sind, wäre ein erster Ansatz, in exakter Weise vom Bewusstsein als Form einer natürlichen Tätigkeit zu sprechen. Letzteres soll Thema der nun folgenden Abschnitte sein.

## Anmerkungen zu Abschnitt 3

[1] Dieser Gedanke wird oft mit den sinnesphysiologischen Untersuchungen von Hermann von Helmholtz in Verbindung gebracht. Für eine zeitgenössische Rezeption siehe etwa Dayan *et al.* (1995), Friston *et al.* (2006) oder Hohwy (2013).

Ein ähnlicher Gedanke – Wahrnehmung als Prozess des Schließens – wurde Ende des 19. Jahrhunderts auch von Charles Sanders Peirce (1991, §11) expliziert, allerdings in einem ganz anderen Kontext.

[2] Einige Eigenschaften der „Entropiefunktion“  $H(Z)$ :

Aus  $0 \leq p(z_i) \leq 1$  folgt  $H(Z) \geq 0$ , da jeder Summand in Gl. (3.2) kleiner oder gleich 0 ist. Falls nur ein Element  $z_j$  aus  $Z$  realisiert werden kann, gilt  $\forall i : p(z_i) = \delta_{i,j}$  und somit  $H(Z) = 0$  (wegen  $\lim_{x \rightarrow 0} x \log x = 0$ ). Dies ist das globale Minimum von  $H(Z)$ .

Eine Gleichverteilung der  $z_i$  hingegen führt auf das globale Maximum von  $H(Z)$ : Die Variation zweier Elemente  $p(z_j)$  und  $p(z_k)$  um einen beliebig kleinen Faktor  $\epsilon > 0$  führt auf eine Abnahme von  $H(Z)$ , falls die Variation zu einer größeren Ungleichheit von  $p(z_j)$  und  $p(z_k)$  führt.

Beweis: Seien  $p'(z_j) = p(z_j) + \epsilon$  und  $p'(z_k) = p(z_k) - \epsilon$  mit  $p(z_j) > p(z_k) > 0$ . Einsetzen in die Entropiefunktion (3.2) gibt:

$$H'(Z) = - \sum_i p'(z_i) \log p'(z_i) \quad (3.11)$$

$$= - \sum_{i \neq j,k} p(z_i) \log p(z_i) - \underbrace{\sum_{i \in \{j,k\}} p'(z_i) \log p'(z_i)}_{\equiv \delta_H(\epsilon)} \quad (3.12)$$

mit

$$\delta_H(\epsilon) = -(p(z_j) + \epsilon) \log (p(z_j) + \epsilon) - (p(z_k) - \epsilon) \log (p(z_k) - \epsilon). \quad (3.13)$$

Ableitung nach  $\epsilon$  ergibt:

$$\frac{d\delta_H(\epsilon)}{d\epsilon} = -\log (p(z_j) + \epsilon) + \frac{p(z_j) + \epsilon}{p(z_j) + \epsilon} + \log (p(z_k) - \epsilon) - \frac{p(z_k) - \epsilon}{p(z_k) - \epsilon} \quad (3.14)$$

$$= \log \left( \frac{p(z_k) - \epsilon}{p(z_j) + \epsilon} \right). \quad (3.15)$$

Da  $p(z_j) + \epsilon > p(z_k) - \epsilon$ , ist die Ableitung negativ. Für infinitesimale  $\epsilon$  gilt insbesondere, dass

$$H'(Z) - H(Z) = \delta_H(0) - \delta_H(\epsilon) = \epsilon \cdot \frac{d\delta_H(\epsilon)}{d\epsilon} < 0, \quad (3.16)$$

und somit

$$H'(Z) < H(Z). \quad (3.17)$$

Dies bedeutet umgekehrt, dass die Gleichverteilung  $p(z_i) = 1/|Z|$ ,  $\forall i$ , mit  $|Z|$  der Anzahl Elemente in  $Z$ , die Funktion  $H(Z)$  maximal werden lässt:

$$H(Z) = -|Z|(1/|Z| \cdot \log 1/|Z|) \quad (3.18)$$

$$= -\log 1 + \log |Z| = \log |Z|. \quad (3.19)$$

Optimale Regulierung ( $H(Z) = 0$ ) bedeutet also, dass das System genau einen bestimmten Zustand  $z_i$  (mit  $p(z_i) = 1$ ) einnehmen würde, während der schlechtest denkbare Fall von Regulierung eine statistische Verteilung aller mögliche Zustände ( $H(Z) = \log |Z|$ ) wäre. Realistischerweise wird es nicht immer möglich sein, eine Verteilung der Form  $p(z_i) = 1$ ,  $p(z_{j \neq i}) = 0$  zu erzwingen, aber näherungsweise wird ein minimaler Wert von  $H(Z)$  dadurch erhalten, dass ein möglichst kleines Ensemble in  $Z$  mit nicht-verschwindenden Wahrscheinlichkeiten erzeugt wird: Es bezeichne  $G \subset Z$  die Menge der „möglichen Regulierungen“ in  $Z$ . Für die Entropie folgt, dass

$$H(Z) = - \sum_{z \in G} p(z) \log(p(z)) - \sum_{z \in Z \setminus G} p(z) \log(p(z)) \equiv H_G + H_{\setminus G}. \quad (3.20)$$

Es gilt nun analog zu zuvor, dass innerhalb von  $G$  jede Ungleichverteilung die Entropie verringert, mit  $H_G \geq 0$ .

In vielen Fällen entspricht ein niedriger Wert von  $H$  der erhöhten Präzision einer Messung. Gute Regulierung kann dann auch als Optimierung der Präzision einer Messungen aufgefasst werden. Dass eine Messungen aber einen bestimmten Absolutwert liefert, bedingt zusätzlich noch eine vormalige Kalibrierung, was jedoch nicht vom Theorem des guten Regulators beschrieben wird.

[3] Es bezeichne  $\pi$  die Menge derjenigen Verteilungen  $p(R|E)$ , die bei gegebenen  $p(E)$  zu einer Wahrscheinlichkeitsverteilung  $p(Z)$  führen, die  $H(Z)$  minimiert. Der gute Regulator ist dann durch ein Element aus  $\pi$  charakterisiert. Der Beweis des Theorems nach Conant & Ashby verläuft nun in zwei Schritten:

1. Für jedes  $p(r_k|e_j) \neq 0 \in \pi$  gilt, dass  $\psi(r_k, e_j) = z_l$ , d.h. bei gegebenem  $e_j$  wird jedem Regulatorzustand, falls dieser überhaupt besetzt ist, dasselbe Ereignis  $z_l$  zugeordnet: Der Regulator „reagiert“ auf eine Änderung in seiner Umgebung auf eine ganz bestimmte Weise.

Andernfalls gäbe es mindestens zwei positive  $p(r_1|e_j)$ ,  $p(r_2|e_j)$  mit  $\psi(r_1, e_j) = z_1 \neq \psi(r_2, e_j) = z_2$ . Mit Gl. (3.7) existieren dann  $p(e_j, r_1) = p(e_j)p(r_1|e_j) > 0$  sowie  $p(e_j, r_2) = p(e_j)p(r_2|e_j) > 0$ , die zu verschiedenen  $p(z_1), p(z_2)$  beitragen, die laut Annahme die Entropie  $H(Z)$  minimieren. Durch Variation der Regulatorverteilung  $p(R|e_j)$  könnte nun  $H(Z)$  jedoch weiter minimiert werden (vgl. Anmerkung [2]), was aber ein Widerspruch ist. Daraus folgt, dass  $z_1 = z_2$  und alle  $p(r_k|e_j) > 0$  zur selben Zuordnung  $\psi(r_k, e_j) = z_l$  gehören müssen.

2. Teilen wir  $\psi^{-1}(z_l)$  in die unterschiedlichen „Äste“ abhängig von  $e_j$  auf, so folgt:

$$p(z_l) = \sum_j \left( p(e_j) \sum_{\psi_j^{-1}(z_l)} p(r_k|e_j) \right). \quad (3.21)$$

Es sei nun  $e_h$  ein beliebiger Zustand des externen Systems, dem eine bestimmte Regulatorverteilung  $p(R|e_h) \in \pi$  zugeordnet ist. Da gemäß (1.) jedem  $e_h$  genau ein Wert  $z_m$  zugeordnet ist, trägt  $p(R|e_h)$  nur zur Wahrscheinlichkeit  $p(z_m)$  bei:

Wegen der Normierung

$$\sum_{j,k} p(e_j, r_k) = \sum_j p(e_j) \left( \sum_k p(r_k|e_j) \right) = 1 \quad (3.22)$$

und weil alle  $p(r_k|e_h) > 0$  genau auf einem Ast  $\psi_h^{-1}(z_l)$  liegen, folgt:

$$\sum_{\psi_h^{-1}(z_l)} p(r_k|e_h) = 1. \quad (3.23)$$

Umschreiben von Gl. (3.21) führt auf:

$$p(z_l) = p(e_h) \underbrace{\sum_{\psi_h^{-1}(z_l)} p(r_k|e_h)}_{=1} + \sum_{i \neq h} p(e_i) \sum_{\psi_{i \neq h}^{-1}(z_l)} p(r_k|e_i). \quad (3.24)$$

Im einfachsten Fall, gilt also für genau ein  $(k, h) \in \psi_h^{-1}(z_l)$ , dass

$$p(r_k|e_h) = 1, \quad (3.25)$$

mit  $p(r_{k'}, e_h) = 0$  für alle  $k' \neq k$ .

[4] Ein Diskussion und Klassifikation von komplexen Systemen findet sich beispielhaft bei Ladyman *et al.* (2013).

[5] Hier und in der folgenden Anmerkung stützen wir uns größtenteils auf die Darstellungen von Haken (1983) und Friston (2013). In die Sprache dynamischer Systeme übersetzt, lässt sich der Zustandsraum, der zuvor noch durch das Produktsystem  $R \times E$  charakterisiert wurde und auf dem die Funktion  $\psi$  definiert war, nun wie folgt darstellen:

- Es existiere ein als *Sensorium*  $S$  bezeichnetes Teilsystem, dessen Aufgabe es ist, Änderungen in der Umwelt  $E$  zu registrieren und an die internen Zustände  $O$  weiterzuleiten.
- Es existiere eine Klasse von *aktiven Zuständen*  $A$ , mithilfe derer der Organismus in die Dynamik von  $E$  eingreifen kann.
- Die weder zu  $A$  noch  $S$  gehörenden Zustände des Organismus bilden dessen *interne Zustände*  $O$ .

- Zusätzlich sollen zufällige Fluktuationen  $\omega \in \Omega$  einen Einfluss auf die Dynamik von  $S$  und  $E$  haben.

Sowohl das Sensorium  $S$  als auch die aktiven Zustände  $A$  sind nicht Teile von  $O$  oder  $E$ , sondern bilden selbst einen je irreduziblen Teil des gesamten dynamischen Systems  $X$ . Gesamthaft gilt also, dass

$$X = E \times \underbrace{S \times A \times O}_R. \quad (3.26)$$

Wir können  $S$  und  $A$  als relativ stabile Bereiche, die  $O$  von  $E$  trennen, betrachten, falls die Existenz von  $S$  und  $A$  zur Entkoppelung der Dynamiken des Organismus und dessen Umwelt führt, z.B. wenn die Zustandsentwicklung von  $O$  im Vergleich zu jener von  $E$  langsam verläuft. Dies entspricht den physiologisch realistischen Situationen, in denen ein Organismus durch einige Variablen (z.B. Temperatur, Blutdruck oder Teile der neuronalen Aktivität) charakterisiert werden kann, deren Änderungen gegenüber den ständigen Fluktuationen in  $E$  vergleichsweise klein, ja nahezu konstant sind,

$$\delta_O \ll \delta_E. \quad (3.27)$$

[6] Dabei ließe sich vermuten, dass es sich hier um eine Version der oft postulierten „abwärtsgerichteten Kausalität“ handelt. Gemäß Stephan (1999, S. 232 ff.) beruht dies allerdings auf einem kategorischen Fehlschluss, falls man hier von (einer Approximation) der mathematischen Theorie auf eine Kausalbeziehung schließen würde, was wiederum die Abbildungsthese aus Abschnitt 2.2 aufgreift, wonach funktionale Abhängigkeiten aus  $\mathbf{T}$  der kausalen Struktur in  $\mathbf{E}$  entsprechen würden. Zudem sei noch einmal betont, dass ein Emergenzbegriff, wie er sich aus der Theorie selbstorganisierender Systeme ableitet, im Allgemeinen *kein* starker Emergenzbegriff (im ontologischen Sinne) ist.

[7] Um dies verständlich zu machen, kann die Zeitentwicklung eines dynamischen Systems,  $x \in X$ , betrachtet werden. Diese ist charakterisiert durch den Fluss  $\mathbf{f}(x)$ ,

$$\dot{x} = \mathbf{f}(x) + \omega, \quad x = (e, s, a, o). \quad (3.28)$$

Für die (ergodische) Wahrscheinlichkeitsdichte  $p(x)$  unter  $\delta$ -korrelierten Fluktuationen,  $\langle \omega_j(t) \omega_k(t') \rangle = \mathbf{W}_{jk} \delta(t - t')$ , lässt sich die sog. Fokker-Planck-Gleichung definieren (vgl. z.B. Haken 1983, Gl. (6.95) – (6.98)),

$$\dot{p}(x) = -\nabla \cdot \mathbf{f}(x)p(x) + \nabla \cdot \frac{1}{2} \mathbf{W} \nabla p(x). \quad (3.29)$$

Für beliebige, stetige Vektorfelder  $\mathbf{f}$  gilt die Helmholtz-Zerlegung:

$$\mathbf{f}(x) = \nabla \phi(x) + \nabla \times \mathbf{A}(x). \quad (3.30)$$

Ein Ansatz für eine Gleichgewichtskonfiguration ( $\dot{p}(x) = 0$ ) ist

$$p(x) = \exp(-L(x)) \Leftrightarrow L(x) = -\ln[p(x)]. \quad (3.31)$$

Wegen  $\nabla p(x) = -p(x)\nabla L(x)$  lässt sich die Fokker-Planck-Gleichung zu

$$\dot{p}(x) = -\nabla \cdot (\mathbf{f}(x)p(x) + \frac{1}{2}\mathbf{W}p(x)\nabla L(x)) \quad (3.32)$$

umschreiben. Einsetzen von Gl. (3.30) liefert

$$\dot{p}(x) = -\nabla \cdot \underbrace{(\nabla\phi(x)p(x) + (\nabla \times \mathbf{A}(x))p(x) + \frac{1}{2}\mathbf{W}\nabla L(x)p(x))}_{\equiv \mathbf{k}} \quad (3.33)$$

Dies ist aber genau dann gleich 0, wenn die Divergenz von  $\mathbf{k}$  verschwindet,

$$\dot{p}(x) = \nabla \cdot \mathbf{k} = 0 \Leftrightarrow \begin{cases} \nabla\phi(x) = -1/2\mathbf{W}\nabla L(x) \\ \nabla \times \mathbf{A}(x) = -\mathbf{R}\nabla L(x) \end{cases}, \quad (3.34)$$

mit  $\mathbf{R}$  einer schiefssymmetrischen Matrix,  $\mathbf{R}^T = -\mathbf{R}$ .

Es folgt also, dass

$$\mathbf{f}(x) = -(\mathbf{W}/2 + \mathbf{R})\nabla L(x). \quad (3.35)$$

Eine Mittelung über eine beliebige Wahrscheinlichkeitsverteilung  $q(E|o)$  der externen Zustände  $e$  mit

$$\int_E q(e|o)de = 1 \quad (3.36)$$

liefert:

$$\langle L(x) \rangle_q = \int_E q(e|o)L(x)de \quad (3.37)$$

$$= -\int_E q(e|o)\ln[p(x)]de \quad (3.38)$$

$$= -\int_E q(e|o)\ln[p(x)]de + \underbrace{\int_E q(e|o)(\ln[q(e|o)] - \ln[q(e|o)])de}_{=0} \quad (3.39)$$

$$= \underbrace{-\int_E q(e|o)\ln[p(x)/q(e|o)]de}_{\equiv F(s,a,o)} + \underbrace{\int_E q(e|o)\ln[q(e|o)]de}_{\equiv H[q(e|o)]}. \quad (3.40)$$

$$\Rightarrow \langle L(x) \rangle_q = F(s, a, o) + H[q(e|o)]. \quad (3.41)$$

Entsprechend der gewählten Partition gilt für die Zustände  $a$  und  $o$ , dass diese nicht explizit von den externen Zuständen  $e$  abhängen. Dies lässt sich mit  $\lambda = \{a, o\}$  wie folgt durch den Fluss  $f_\lambda$  ausdrücken:

$$\langle \dot{\lambda} \rangle_t = f_\lambda(s, a, o), \quad (3.42)$$

mit:

$$f_\lambda(s, a, o) \propto -\nabla_\lambda F(s, a, o), \quad (3.43)$$

vgl. auch Friston (2013, Lemma 2.1). D.h. die Zustände  $a$  und  $o$  folgen den Ableitungen der Funktion  $F(s, a, o)$  nach den entsprechenden Variablen.



[8] Aus der rechten Seite von Gl. (3.40) und der Identität

$$p(x) = p(e, s, a, o) = p(s, a, o)p(e|s, a, o) \quad (3.44)$$

folgt mit, dass

$$F(s, a, o) = - \int_E q(e|o) \ln[p(s, a, o)p(e|s, a, o)/q(e|o)]de \quad (3.45)$$

$$= - \int_E q(e|o) \ln[p(s, a, o)]de - \int_E q(e|o) \ln[p(e|s, a, o)/q(e|o)]de = (3.46)$$

$$= \underbrace{- \ln[p(s, a, o)]}_{=L(s,a,o)} + \underbrace{\int_E q(e|o) \ln[q(e|o)/p(e|s, a, o)]de}_{\equiv D_{\text{KL}}[q(e|o)||p(e|s,a,o)]}. \quad (3.47)$$

Der letzte Term wird oft als „relative Entropie“ oder Kullback-Leibler-Divergenz bezeichnet und ist streng größer 0, außer für den Fall, wo

$$q(e|o) = p(e|s, a, o) \Leftrightarrow D_{\text{KL}}[q(e|o)||p(e|s, a, o)] = 0. \quad (3.48)$$

Mit Gl. (3.47) und der Ergodizitätsannahme folgt daher:

$$\langle F(s, a, o) \rangle_t \geq - \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \ln[p(s, a, o)]dt = - \int_X p(x) \ln[p(s, a, o)]dx \quad (3.49)$$

$$\geq - \int_{S \times A \times O} p(s, a, o) \ln[p(s, a, o)]dz = H[p(s, a, o)]. \quad (3.50)$$

Zu beachten ist, dass weder die Redeweise von einer „freien Energie“,  $F$ , noch von einer „Entropie“,  $H$ , streng im thermodynamischen Sinne genommen werden darf, da es sich eben nicht um abgeschlossene Systeme im thermodynamischen Gleichgewicht handelt. Es handelt sich um verallgemeinerte Begriffe, welche aber grob den Intuitionen hinter den thermodynamischen Begriffen folgen.

[9] Bei der Diskussion solcher Beispiele ist es wichtig anzumerken, dass die Annahme einer kausalen Wirksamkeit von Bewusstsein nur dann in der klassischen dualistischen Position aufgeht, falls davon ausgegangen wird, dass eine solche „mentale Verursachung“ als Interaktionen zwischen einer geistigen und materiellen Substanz aufzufassen sei, also wenn wir ein substantielles Bewusstsein denken, das einem Körper gegenübersteht und äußerlich auf diesen einwirkt.

[10] So kann ebenfalls gesagt werden, dass die Umwelt die „Dimension des Modells“ bestimme – vgl. dazu Ashbys (1956) *law of requisite variety*.

[11] Ein möglicher Einwand an dieser Stelle würde lauten, dass unter *Repräsentation* ja etwas ganz anderes zu verstehen sei als eine raffinierte Form von Abbildung, da Repräsentation immer bereits ein Element der Bezugnahme oder Interpretation beinhaltet. Doch üblicherweise bildet diese Annahme nicht den Ausgangspunkt in den Wissenschaften vom Bewusstsein – man denke z.B. an die „neuronalen Korrelate des Bewusstseins“, die im Idealfall punktuell zu den „Inhalten des phänomenalen Bewusstseins korrelieren“

(Chalmers 2000) oder an die Definition der Repräsentationsbeziehung bei Fred Dretske (1995, S. 2), wonach ein System  $S$  genau dann die Eigenschaft  $F$  repräsentiert, wenn dessen Zustände  $\{s_1, s_2, \dots, s_n\}$  zu den Werten  $\{f_1, f_2, \dots, f_m\}$  korrespondieren. Auf den ersten Blick scheint dies eher für ein Verständnis von Repräsentation als (raffinierte Form von) Abbildung zu sprechen als für irgendeine Form der aktiven Bezugnahme oder Interpretation. Am Ende von *Naturalizing the Mind* qualifiziert Dretske seine Definition der Repräsentation allerdings weiter, indem er schreibt, dass nur da von *mentaler* Repräsentation (bzw. von bewusster, sensorischer Wahrnehmung) gesprochen werden könne, wo Repräsentation auch zu (biologisch relevanten) *Wirkungen* führe (Dretske 1995, S. 165 f.), da sie sonst nicht Gegenstand natürlicher Selektion sein könne. Dies impliziert, dass bewusste Wahrnehmung bei Dretske letztlich *nicht* als Abbildungsbeziehungen verstanden werden kann, auch wenn dies scheinbar den Ausgangspunkt seiner Untersuchung darstellt.

[12] Eine interessante Spekulation betrifft nun die Frage, ob der Prozess der zwischen erst- und dritt-personaler Perspektive vermittelt, derselbe ist, der die einheitliche Struktur des Bewusstseins begründet, vgl. dazu Abschnitt 7.

[13] Zitiert nach Scheibe (2006, S. 38f.); Hervorhebung RP.

[14] Hierin spiegelt sich also auch das Rätsel einer doppelten Bedingtheit des Geistes wieder. Ein Wissenschaftler-Philosoph, der versucht hat, diese doppelte Bedingtheit durch eine „Philosophie des Kreisgangs“ zu überwinden, war Carl Friedrich von Weizsäcker. Siehe für diesen Aspekt der Philosophie Weizsäckers etwa die Ausführungen von Schüz (1986) oder, insbesondere für das Schelling'sche Erbe bei Weizsäcker, von Sieroka (2009): Geist bedingt Erkenntnis, welche die Natur beschreibt, die den erkennenden Geist erst hervorbringt.

## 4 Der Relationale Aufbau der Gegenstände

[Das Kontinuum ist] ein Gebilde, das fortgesetzter Teilungen fähig ist, ohne daß man je auf letzte, nicht mehr teilbare Elemente stößt. (Hermann Weyl: Riemanns Geometrische Ideen, ihre Auswirkung und ihre Verknüpfung mit der Gruppentheorie)

### 4.1 Die Theorie vom Ganzem und seinen Teilen

#### Elemente, Relationen und die Mereologie

Ein erste Herangehensweise wäre es, zu versuchen, Bewusstseinsaktivität in möglichst stabile Einzelereignisse, in *Grundelemente* zu analysieren. Auf diese Weise ließe sich Einsicht in *fundamentale Prozesse der Psyche* erlangen – angelehnt an die Methode der modernen Wissenschaft, materielle Prozesse durch Zerlegung in möglichst elementare und wechselwirkende Systeme („Elementarteilchen“) aufzuklären.

Dass ein solcher *elementaristischer* Ansatz jedoch mit grundsätzlichen Schwierigkeiten behaftet ist, dafür argumentierte etwa Barry Smith (1995, S. 5-7). Dabei bezieht er sich auf das Beispiel des (wahrgenommenen) Kontinuums und dessen Rekonstruktion als Menge von Punkten („simuliertes Kontinuum“), die er aus folgenden Gründen als problematisch erachtet:

- Aus der Reflexion über Wahrnehmungsphänomene ergibt sich ein „qualitativer“ Begriff des Kontinuums, das subjektiv empfunden wird und quantitativ, also zum Beispiel durch Nummerierungen, Metriken, usw., nur unzulänglich strukturiert werden kann.
- Eine mengentheoretische Konstruktion würde implizieren, dass unausgedehnte Elemente (Punkte) letztlich ein ausgedehntes Ganzes konstituieren. Es ist jedoch völlig unklar, welche „Urelemente“ den Ausgangspunkt einer solchen Konstruktion bilden sollten. Weder sind solche Urelemente direkt gegeben, noch ist es überhaupt klar, dass sich unteilbare Elemente, d.h. atomare, phänomenale Objekte überhaupt isolieren lassen.
- Eine mengentheoretische Konstruktion, die letztlich auf Punkte rekurriert, scheint nicht natürlicherweise mit räumlicher oder zeitlicher Begrenztheit umgehen zu können, da solche Grenzen „immer und notwendigerweise Grenzen von etwas Ausgedehntem [sind].“

Ein alternativer Ansatz kann als *relationaler* bezeichnet werden. Hier wollen wir uns auf eine ganz bestimmte Relation, nämlich auf die Beziehung zwischen Teilen und Ganzheiten, beschränken. Diese stellt vielleicht das grundlegendste relationale Prinzip dar, welches formalisiert werden kann.

Die Lehre, welche das Verhältnis eines Ganzen zu den es konstituierenden Teilen zu klären sucht, bezeichnet sich als *Mereologie*. Ausgangspunkt vieler mereologischer Systeme war der Versuch, geometrische oder abstrakte Teil-Ganzesbeziehungen in einem Axiomensystem auszudrücken. Mereologie wurde dabei etwa als topologische (Whitehead, PNK) oder nominalistische Fundierung der Mathematik (Leśniewski 1916/1992, Tarski 1935, Tarski 1956) sowie als Formalontologie (Smith & Mulligan 1983, Simons 1987, Simons 2003, Schaffer 2010) verstanden. Wichtige Vorüberlegungen entstammen aber auch der Psychologie des späten 19. Jahrhunderts und finden sich prominent etwa bei Carl Stumpf und Franz Brentano (Smith & Mulligan 1982, S. 25-35), und dann insbesondere in der 3. Logischen Untersuchung von Edmund Husserl. Diese können als metaphysische Betrachtungen, aber auch als frühe, wenn auch nur wenig formalisierte Darstellungen der Struktur bewussten Erlebens oder geistiger („intentionaler“) Tätigkeit angesehen werden.

Insbesondere diese letzte Deutung ist es, in deren Tradition die folgenden Abschnitte stehen. Dabei kann dafür argumentiert werden, dass Mereologie die Struktur eines Verweissystems beschreibt, mithilfe dessen eine Reihe von Symbolsystemen konstruiert werden können, die *dem Anschein nach* Abbilder oder Modelle einer gegenständlichen Wirklichkeit liefern, die aber in Wahrheit gar nicht als unabhängig vom Akt des Bezeichnens gedacht werden können.

Folgende zentrale Interpretation soll an dieser Stelle eingeführt werden: Die Objekte der Mereologie sind Verweise auf Gegenstände, Mereologie handelt in diesem Sinne von Verweisen auf konkrete [1] Gegenstände und nicht wie die (naive) Mengenlehre von der Zugehörigkeit zu einer (logischen) Klasse. Welcher Art sind diese Gegenstände? Dies kann auf verschiedene Weise beantwortet werden, etwa wie folgt:

- Naiver Realismus: Die Gegenstände, auf die verwiesen wird, entsprechen den Gegenständen des Alltags, etwa Stühlen, Fahrzeugen oder Kaffeetassen. Ein solcher Realismus wird etwa von Searle (1991, Kap. 2) in Bezug auf intentionale Objekte vertreten.
- Kritischer Realismus: Die Gegenstände, auf die verwiesen wird, entsprechen scheinbar den oben genannten, aber bei genauer Analyse lösen sich diese auf, z.B. in Gruppierungen von Elementarteilchen (Physikalismus), welche dann als „hinter den Verweisen stehend“ betrachtet werden können,

- Prozessmetaphysik: Die Gegenstände, die hinter diesen Verweisen stehen, sind in Wahrheit die Referenten einer Verkettung von Zeichen (Peirce 1983), Ereignisse (Whitehead, PR), Netzwerke interagierender Subjekte (Hoffman 2008) oder symbolverwendende Einzelwesen (Hampe 2014).

Mereologische Beschreibungen gehen dabei von einem Teileinsbegriff aus, der einerseits möglichst allgemein bleiben soll, andererseits häufig auf konkrete Intuitionen verweist. Eine Kritik, die von Casati & Varzi (1999) an der mereologischen Vorgehensweise geäußert wurde, lautet, dass Mereologie selbst nichts darüber aussagt, was denn „eigentlich“ unter einem Ganzen zu verstehen sei. Genauso wenig, möchte man hinzufügen, gibt sie darüber Auskunft, was denn eigentlich unter einem Teil zu verstehen sei. Gleichzeitig liegt es nahe, dass Fragen nach der Natur der Teil-Ganzes-Beziehung *a priori* gar nicht beantwortet werden können (Simons 2006).

Richten wir unseren Blick dazu kurz auf die empirischen Wissenschaften, die ja ebenfalls als strukturierte Systeme von Zeichen, die sich auf eine gegenständliche Welt beziehen, aufgefasst werden können. In einer neueren Arbeit versuchen Findlay & Thagard (2012) zu zeigen, wie eine vollständige mereologische Aufschlüsselung aller Wissenschaftsbereiche von Teilchenphysik bis Soziologie mit Hilfe eines einzigen Schemas möglich ist. Dadurch erhoffen sich die Autoren, „ein einheitliches Verständnis der Struktur, Funktion und Dynamik von Organisationsprozessen in Physik, Biologie, den Kognitions- und Sozialwissenschaften“ [2] zu befördern. Im Gegensatz zu den formalen Ausführungen dieses Abschnittes arbeiten Findlay & Thagard vor allem an einer empiristischen Katalogisierung der Einzelwissenschaften. Zudem wenden sie sich explizit gegen ein Verständnis von Mereologie als Teilgebiet der analytischen Metaphysik. Dies könnte als empiristischer Gegenentwurf zu einer als „klassisch“ zu bezeichnenden Mereologie angesehen werden.

Doch würde gerade die Verbindung von mathematisch-logischem Formalismus und empirischer Adäquatheit einen wichtigen Zug in der Weiterentwicklung des klassischen Systems darstellen. Dessen größtes Manko ist weder die schwere Zugänglichkeit noch das axiomatisches Vorgehen – dies hätte sie ja mit anderen Gebieten der Mathematik gemein –, sondern ihre Orientierung an einer Trivialempirie des Alltags (wenn es überhaupt eine Orientierung an irgendeiner Form von Empirie gibt). Da, wo dies lediglich zur Illustration formaler Konzepte gebraucht wird, ist wenig dagegen einzuwenden. Oftmals dient es aber gerade dazu, bestimmte Intuitionen bezüglich der Teil-Ganzes-Beziehung zu befördern und axiomatisch zu verewigen.

Umgekehrt liefert die formale Analyse aber erst genau jenen Grad an begrifflicher Schärfe, der für eine Analyse der Verfahrensweise der empirischen Wissenschaften als angemessen erscheint (für eine Versöhnung zwischen formalen und empirischen Zugangsweisen zur Mereologie siehe auch Hovda 2014). Dass z.B. naturwissenschaft-

liche Theorien unter mereologischen Gesichtspunkten betrachten werden können, fand erst kürzlich wieder vermehrte Aufmerksamkeit, wobei sich zeigt, dass den unterschiedlichen empirischen Teilgebieten durchaus mereologische Axiomatisierungen zugeordnet werden können (etwa bei Calosi & Tarozzi 2014). Jene stellen dann „Modelle“ eines bestimmten mereologischen Systems dar, dies allerdings nicht immer *eindeutig* (Harré & Llored 2011, Llored & Harré 2014, Prentner 2017). Dies impliziert, dass uns eine Reflexion gar keinen bestimmten Standpunkt bezüglich des Verhältnisses zwischen unserer abstrakten Darstellung und dessen empirischer Realisierung aufzwingt.

### Mereologische Operatoren

Anstelle sich weiter diesen systematischen Fragen zu widmen, soll jedoch zuerst auf die formale Struktur mereologischer Beschreibungen eingegangen werden [3]. Mereologie wird dabei als formales, axiomatisch aufgebautes Notationssystem verstanden, ähnlich der Mengenlehre, allerdings mit einigen Unterschieden. Als Referenz hierfür soll der erste Teil des Buchs *Parts* von Peter Simons (1987) dienen. Historisch wichtige Kalküle stammen etwa von Leśniewski (1916/1992) („Mereologie“) sowie von Leonard & Goodman (1940) („*calculus of individuals*“). Im Vergleich zur mengentheoretischen Elementrelation ( $\in$ ) tritt hier die Teilseinsrelation ( $<$ ) [4]. Die verwendeten Variablen  $x, y, z$  etc. können dann gemäß obiger Deutung als Platzhalter für Verweise auf einen konkreten Gegenstand gelesen werden. Einen solchen bezeichnen wir als das „Individuum“, auf welches von mereologischen Objekten verwiesen wird. Zusätzlich können nun *mereologische Operationen* auf diesen Verweisen definiert werden. Eine kurze Übersicht über einige mereologische Operatoren soll nun folgen:

$$x < y, \tag{4.1}$$

bedeutet, dass  $x$  einen Teil des Objektes  $y$  bezeichnet und auf einen Gegenstand verweist, der in einer analogen Beziehung zum Gegenstand steht, auf welchen von  $y$  verwiesen wird. Dabei können  $x$  und  $y$  auch identisch sein, also auf denselben Gegenstand verweisen. Ist letzteres nicht Fall, spricht man von *echtem Teilsein*:

$$x \ll y \Leftrightarrow x < y \wedge x \neq y \tag{4.2}$$

Von *mereologischem Überlapp* spricht man, wenn zwei Objekte  $x$  und  $y$  einen gemeinsamen Teil haben,

$$x \circ y \Leftrightarrow \exists z (z < x \wedge z < y). \tag{4.3}$$

(Wenn  $x < y$ , dann überlappen sie trivialerweise.) Ist dies nicht der Fall, sind  $x$  und  $y$  *mereologisch disjunkt*:

$$x \wr y \Leftrightarrow \neg \exists z (z < x \wedge z < y). \quad (4.4)$$

$$= \neg(x \circ y). \quad (4.5)$$

Sind  $x$  und  $y$  disjunkt, kann daher auch keine Teilseinrelation zwischen ihnen bestehen.

Die *binäre Summe*,

$$x + y, \quad (4.6)$$

bezeichnet das Objekt, welches mit allen Objekten überlappt, die mit  $x$  oder  $y$  überlappen. Die binäre Summe umfasst  $x$  und  $y$ : Der Besen ist die mereologische Summe aus Stiel und Kopf. In einigen Axiomatisierungen wird davon ausgegangen, dass die mereologische Summe *immer* existiert. Mereologische Summen können dabei auch potentiell aus nicht-zusammenhängenden (disjunkten) Objekten generiert werden. Zu beachten ist ferner, dass  $x + y$  einen Ausdruck bezeichnet, der selbst auf einen Gegenstand (Individuum) verweist, also nicht als rein abstraktes Konzept oder Menge von Elementen betrachtet werden sollte.

Zudem scheinen einige Teile essentiell für das Ganze zu sein: Entfernen wir den Kopf des Besens, hört dieser auf ein Besen zu sein; selbiges gilt jedoch nicht für einen (echten) Teil des Stiels. Umgekehrt können Teile auch abhängig sein von ihrem Ganzen: Lautstärken ohne Töne machen genauso wenig Sinn wie Töne ohne Lautstärken. Diese und ähnliche Situationen geben oft Anlass zu Fragen bezüglich der Zulässigkeit von Summation. Plakativ gesprochen, raubt die Summation der Mereologie ihre ontologische Unschuld.

Abgeleitet von der Summe bezeichnet das *binäre Produkt*,

$$x \cdot y, \quad (4.7)$$

die Summe derjenigen Objekte, die sowohl Teile von  $x$  als auch von  $y$  sind.

Schließlich bezeichnet die *mereologische Differenz*,

$$x - y, \quad (4.8)$$

den (größten) Teil von  $x$ , der nicht mit  $y$  überlappt (also disjunkt zu  $y$  ist).

Am einfachsten illustriert man diese Operationen mithilfe mengentheoretischer Darstellung (Tab. 4.1). Es gilt aber einige Feinheiten zu beachten, da Mereologie und Mengenlehre vom logischen (und semiotischen) Standpunkt aus prinzipiell verschieden sind: So bezeichnet etwa die Schnittmenge zweier disjunkter Mengen eine leere Menge; in der Mereologie ist das binäre Produkt zweier disjunkter Objekte nicht-existent, da mereologische Objekte (laut unserer Interpretation) immer auf konkrete

Gegenstände verweisen; auch die Null-Menge, also eine Menge, mit dem Maß 0, das aber möglicherweise unendlich viele Elemente besitzt oder die einelementige Menge, finden keine direkten Entsprechungen in der Mereologie.

Die binäre Summe kann verallgemeinert werden zur *Fusion*

$$\sum x(Fx), \quad (4.9)$$

wobei hier eine Summe über alle Objekte  $x$  genommen wird, welche die Eigenschaft  $F$  aufweisen, etwa die mereologische Summe aller Kreise. Analog kann das binäre Produkt zum *Kern oder Nukleus* verallgemeinert werden:

$$\prod x(Fx). \quad (4.10)$$

In einem geometrischen Beispiel würde der Nukleus denjenigen Punkt bezeichnen, in dem sich eine (möglicherweise unendliche) Menge an Geraden schneidet.

Zudem ist es hilfreich, folgendes Objekt im Auge zu behalten:

- Das „Universum“  $U$  wird als die mereologische Summe aller Objekte (inklusive deren möglichen mereologischen Summen) bezeichnet. Gegeben die Elemente  $a$  und  $b$ , bezeichnet  $U$  also die mereologische Summe aus  $a$ ,  $b$ ,  $a + b$ .

Im Unterschied zur Menge (oder Klasse)  $\bigvee$  aller Objekte,  $\bigvee = \{a, b, a + b\}$ , verweist  $U$  als mereologische Summe auf einen konkreten Gegenstand (auf ein Individuum). Auch die Russell'sche Klasse, also die Menge aller Mengen, die sich selbst *nicht* enthalten, haben kein natürliches Pendant in der Mereologie, was insbesondere für Leśniewski einen großen Vorteil der Mereologie gegenüber der (naiven) Mengenlehre darstellte.

## 4.2 Klassische Mereologie

### Mereologische Axiome

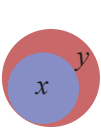
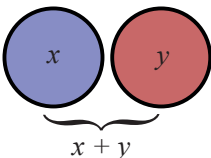
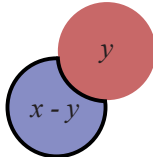
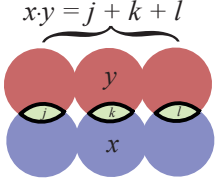
Wichtige Eigenschaften der (echten) Teilseinsrelation sind:

$$\text{M1: Transitivität: } x \ll y \wedge y \ll z \rightarrow x \ll z$$

$$\text{M2: Asymmetrie: } x \ll y \rightarrow \neg(y \ll x)$$

Transitivität und Asymmetrie können als mereologische Axiome betrachtet werden (die Nummerierung der Axiome folgt dabei Simons 1987, Kap. 1); aus Ihnen folgt,



Mereologie	Mengenlehre
$x \ll y$	
$x \circ y$	
$x + y$	
$x - y$	
$x \cdot y$	
?	Leere Menge $\{\}$
?	Einelementige Menge $\{i\}$
Universum $U$	?
?	Russell'sche Klasse, Null-Menge

**Tabelle 4.1:** Mengentheoretische Veranschaulichung einiger mereologischer Konzepte und Objekte, die keine eindeutige Entsprechung in beiden Kalkülen besitzen.

dass die (echte) Teil-Ganzes-Beziehung irreflexiv ist: Angenommen  $x \ll x$  wäre wahr, dann würde nach M2 folgen, dass  $\neg(x \ll x)$  ebenso wahr wäre, was einen Widerspruch darstellen würde.

Allerdings würden diese Axiome allein, eine Intuition bezüglich der Teilseinsrelation nicht widerspiegeln: Etwas hat mindestens zwei echte Teile (oder gar keinen). Ist  $x \ll y$ , dann bleibt ein „Rest“, der disjunkt zu  $x$  und gleichzeitig ein Teil von  $y$  ist (und im Falle, dass dieser Rest auf ein einzelnes Objekt verweist, wäre er identisch zu  $y - x$ ). Diese Intuition kann durch ein zusätzliches mereologisches Axiom (das *Schwache Supplement*) ausgedrückt werden:

$$\text{M3: } x \ll y \rightarrow \exists z(z \ll y \wedge z \wr x) [5].$$

Ähnlich wie Zermelo (1908) mit dem Aussonderungsaxiom verhinderte, dass die Russell'sche Klasse in Zermelos Mengenlehre als echte Menge bezeichnet werden durfte, können wir weitere scheinbar pathologische Fälle von Teilsein durch das *Starke Supplement* ausschließen. Jenes hat zudem die wichtige Konsequenz, dass keine zwei Objekte aus denselben (echten) Teilen bestehen, gleichzeitig aber verschieden sein können. In anderen Worten: Objekte mit denselben (echten) Teilen sind identisch,

$$\text{M5: } \neg(x < y) \rightarrow \exists z(z < x \wedge z \wr y).$$

M5 lässt sich z.B. so lesen: Wenn  $x$  über keinen Teil  $z$  verfügt, der disjunkt zu  $y$  ist – wenn z.B. alle echten Teile von  $x$  echte Teile von  $y$  sind –, dann ist  $x$  entweder identisch zu  $y$  oder  $x$  ist echter Teil von  $y$ . Wenn gleichzeitig aber alle echten Teile

von  $y$  auch echte Teil von  $x$  sind, bleibt nur noch die Option  $x = y$  offen (wegen der Asymmetrie der Teilseigenschaft).

Das axiomatische System, bestehend aus M1-M3 und M5, wird manchmal als „extensionale Mereologie“ (EM), etwa von Casati & Varzi (1999) bezeichnet. Simons (1987) bezeichnet dies als „minimale extensionale Mereologie“ (MEM), mit der zusätzlichen Forderung, dass es im Falle des Überlappens von Objekten,  $x \circ y$ , einen größtmöglichen gemeinsamen Teil (ein mereologisches Produkt  $z = x \cdot y$ ) geben müsse:

$$x \circ y \rightarrow \exists z \forall w (w < z \Leftrightarrow w < x \wedge w < y) \quad (4.11)$$

Alle weiteren Axiome betreffen vor allem die Zulässigkeit von Summen- und Produktbildung bzw. deren garantierte Existenz. So könnte die Existenz einer finiten mereologischen Summe daran geknüpft sein, dass die Objekte, aus denen sie gebildet wird, mindestens überlappen müssen. (Man denke etwa an Summen aus überlappenden Raumzeitgebieten.) Tun sie dies, so erscheint es plausibel, ein gemeinsames Ganzes  $z$  zu postulieren:

$$\text{M7: } x \circ y \rightarrow \exists z (x < z \wedge y < z).$$

Ein Objekt  $z$ , das sowohl  $x$  als auch  $y$  umfasst, existiert also dann, wenn  $x$  und  $y$  überlappen. Umgekehrt folgt, dass  $x$  und  $y$  *nicht* überlappen, wenn es kein umfassendes  $z$  gibt. Eine etwas stärkere Forderung ist die Existenz einer (einzigen) mereologischen Summe:

$$\text{M9: } x \circ y \rightarrow \exists!(x + y),$$

für die gilt, dass jedes  $w$ , das mit ihr überlappt auch mit  $x$  und  $y$  überlappt:

$$\forall w (w \circ (x + y) \Leftrightarrow w \circ x \vee w \circ y) \quad (4.12)$$

Wenn die Summe  $x + y$  existiert, existiert auch mindestens ein  $z$ , das  $x$  und  $y$  umfasst (nämlich  $x + y$ ). Der umgekehrte Fall muss nicht immer gelten: Aus der Existenz eines gemeinsamen Überlappenden folgt nicht unbedingt die Existenz der Summe  $x + y$  [6]. Alternativ könnte, wie bei Casati & Varzi (1999), M9 durch ein anderes Axiom ersetzt werden: Sollte ein  $x$  und  $y$  umfassendes Objekt  $z$  existieren, so existiert auch die mereologische Summe  $x + y$ :

$$\text{M9': } \exists z (x < z \wedge y < z) \rightarrow \exists!(x + y),$$

Folgendes Beispiel dient der Veranschaulichung: Der Körper umfasst die Finger der rechten Hand. Nach M9' würde dann folgen, dass auch eine mereologische Summe

existiert, die nur aus den Fingern der rechten Hand besteht.

Die unbedingte Existenz einer (einzigen) finiten mereologischen Summe, unabhängig davon, ob  $x$  oder  $y$  überlappen oder umfasst werden, kann durch

$$\text{M14: } \exists!(x + y)$$

ausgedrückt werden.

Die Existenz eines Universums (= die mereologische Summe aller Elemente inklusive deren Summen) muss zusätzlich postuliert werden als

$$\text{M16: } \exists U \forall y (y < U).$$

Eine interessante Einschränkung geht dabei auf den dritten Teil der PNK von Whitehead zurück: Fordert man, dass jedes Objekt *echter* Teil eines anderen ist, ist die Existenz eines Universums nicht mehr gegeben, da das Universum  $U$  selbst echter Teil eines anderen Objektes wäre, was aber gegen die Definition von  $U$  verstößt, wonach  $U$  auf die Summe *aller* Objekte (mitsamt deren mereologischen Summen) verweist. Eine Diskussion über den Status des Universums findet sich bei Simons (2003), der dafür argumentiert, dass die Existenz eines Universums (im Sinne eines Verweises auf ein Individuum) nur dann sinnvoll ist, wenn zugleich angenommen werden kann, dass alle Objekte derselben Kategorie entspringen.

Noch stärkere Axiomensysteme postulieren die Existenz unendlicher mereologischer Summen (Fusionen):

$$\text{M24: } \exists x Fx \rightarrow \exists! \sum x(Fx).$$

Es gilt dann, dass, sobald auch tatsächlich ein Objekt existiert, welches das Prädikat  $F$  erfüllt, eine mereologische Fusion existiert. Die Fusion ist definiert, als dasjenige  $\zeta$ , für welches gilt:

$$\forall y (y \circ \zeta \Leftrightarrow \exists z (Fz \wedge y \circ z)). \quad (4.13)$$

In Worten: Für alle  $y$  gilt, dass, wenn sie mit  $\zeta$  überlappen, auch mit (mindestens einem)  $z$ , das die Eigenschaft  $F$  aufweist, überlappen; vgl. auch die Definition der binären Summe [7]. Wenn lediglich ein Objekt  $x$  die Eigenschaft  $F$  realisiert, ist  $z = x$ . Dies entspricht z.B. dem trivialen Fall:

$$x = \sum z(z = x). \quad (4.14)$$

Für die Existenz unendlicher Produkte (Nuklei)  $\prod$  muss kein zusätzliches Axiom angenommen werden; der Nukleus ist dann definiert als die Fusion derjenigen Objekte (falls es solche gibt), die Teile aller  $x$  sind, für die  $Fx$  gilt. Ein Axiomensystem,

dass M1-M3, M24 beinhaltet, wird als „klassische extensionale Mereologie“ (KEM) bezeichnet. Alle anderen vorgestellten Axiome lassen sich aus diesen vier ableiten und nach ihrer Mächtigkeit ordnen (ungefähr in der Reihenfolge, in der sie vorgestellt wurden). Zusätzlich können eine Reihe mereologischer Theoreme abgeleitet werden, wie sie z.B. Simons (1987, S. 38-40) auflistet.

Eine wichtige Klasse von Prädikaten  $F$  erfüllt folgende zwei Eigenschaften:

- Distributivität: Jeder Teil von  $y$  verfügt über die Eigenschaft  $F$ , falls  $y$  darüber verfügt:

$$(x \ll y \wedge Fy) \rightarrow Fx \quad (4.15)$$

- Kumulativität: Gilt ein Prädikat für  $x$  und  $y$ , gilt es auch für deren mereologische Summe:

$$(Fx \wedge Fy) \Leftrightarrow F(x + y), \quad (4.16)$$

und allgemeiner für Fusionen:

$$\forall x (Fx \rightarrow Gx) \Leftrightarrow G\left(\sum x(Fx)\right), \quad (4.17)$$

d.h., falls die Summe  $\sum x(Fx)$  existiert (und nach M24 muss daher auch  $\exists x Fx$  gelten), erfüllt die Summe jedes Prädikat  $G$ , genau dann, wenn für alle Summanden  $Gx$  aus  $Fx$  folgt.

Eine Frage, die sich nun stellt, ist, welche Art von Objekten oder Strukturen die Axiome der KEM überhaupt erfüllen. Insbesondere die Prinzipien der unbedingten Existenz mereologischer Summen (M14 bzw. M24) erscheinen in vielen Fällen fraglich. Wenn man die mereologische Summe aus den vier Rädern eines (bestimmten) Autos bildet, folgt nicht unbedingt, dass diese Summe auf einen konkreten Gegenstand verweist. Noch weniger scheint es garantiert zu sein, dass die Summe  $\zeta = \sum x(Fx)$  selbst über die Eigenschaft  $F$  verfügt. Eine Summe von Teelöffeln ist, um ein Beispiel von Bertrand Russell aufzugreifen, selbst kein Teelöffel. Es gibt aber eine Klasse von Dingen, für die es plausibel erscheint, die Axiome der klassischen extensionalen Mereologie und die Kumulativität bzw. Distributivität ihrer Eigenschaften anzunehmen. Dabei handelt es sich um Massenterme (Simons 1987, Kap. 4.6). Massen haben die Eigenschaft, dass mereologische Summen von Massen selbst wiederum Massen sind; Beispiele hierfür wären etwa Wasser und andere Stoffe. Für Massenterme gilt zudem, dass jeder echte Teil von Wasser über dieselben Eigenschaften wie Wasser verfügt – zumindest idealisiert, wenn man lediglich makroskopische Objekte im thermischen Gleichgewicht betrachtet. Zudem gilt eine Eigenschaft von Wasser, etwa dessen Transparenz, auch für die mereologische Summe aus dem Wasser in meinem Glas und dem Wasser in der Leitung.

Ein wichtiges Beispiel für solche Massenterme findet sich in der Metaphysik nach Descartes (Schütt 1990, S. 271ff.): Körper werden hier aufgefasst als Teile einer Masse, deren Essenz es ist, ausgedehnt zu sein (*res extensa*). In anderen Worten, die ausgedehnte Substanz stellt die mereologische Summe seiner Portionierungen dar, von der wir in der KEM wegen M24 ausgehen dürfen, dass sie wohldefiniert ist und die kleinste obere Schranke aller Portionierungen darstellt. Ein, gemäß Descartes, ontologisch primärer Stoff würde somit über die mereologisch reichste Struktur aufweisen. Dies scheint die Auffassung, wonach es sich bei den kartesischen Substanzen um sehr einfache Gegenstände handelt, in Frage zu stellen und impliziert vielmehr, dass es sich bei ihnen um eine Abstraktion handelt – eine Abstraktion, die uns vielleicht gar nicht mehr als solche erkenntlich ist. In Hinblick auf die These, dass Objekte nichts anderes sind als Muster oder Regelmäßigkeiten von Beziehungen, könnten Substanzen im kartesischen Sinn als diejenigen Muster gedeutet werden, welche eine (mereologisch) sehr starke Struktur aufweisen. Dabei würde dann, anders als in der traditionellen Metaphysik, von einer Substanz nicht mehr als etwas Fundamentales und allen Dingen Zugrundeliegendes gesprochen werden, auf das letztlich unsere ontologischen Untersuchungen zurückführen müssten, sondern auf etwas Abstrahiertes oder Hinzugedachtes.

Die Gegenstände, mit denen wir es im Alltag oder im Labor zu tun haben, können hingegen nur in den wenigsten Fällen mit solchen Massentermen identifiziert werden. Daher scheint es angemessen zu sein, die Adäquatheit der unterschiedlich mächtigen Axiomensysteme bezogen auf diese Gegenstände, insbesondere auf jene, welche durch naturwissenschaftliche Theorien beschrieben werden, zu untersuchen. Die Schwierigkeit besteht darin, dass die höheren Axiomensysteme keine intuitive Entsprechung mehr haben. Selbst M3 und M5, die mereologischen Supplemente, haben je nach Autor einen unterschiedlichen Status, der zwischen analytisch notwendiger (Simons 1987), kontingent wahrer (Bigelow 2010) und potentiell gar nicht erfüllter (Casati & Varzi 1999) Eigenschaft der Teilseinsrelation changiert.

## Mereologie und Boolesche Algebra

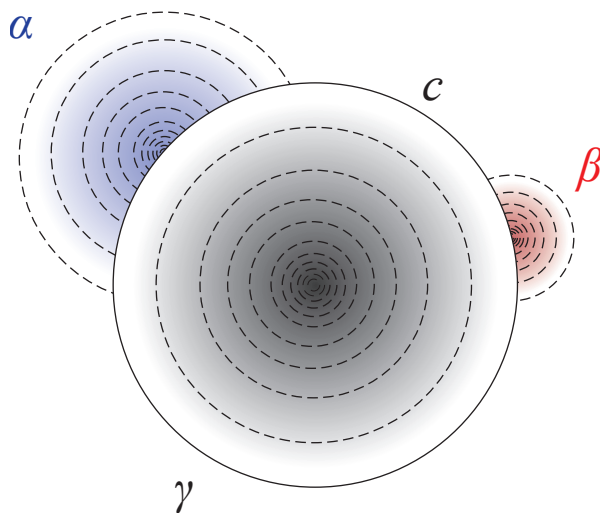
Um zu zeigen, wie das formale System der (klassischen) Mereologie eine konkrete Interpretation erfahren kann, soll kurz diskutiert werden, wie mit Hilfe der klassischen Mereologie die Euklidische Geometrie und in einem zweiten Schritt die Boolesche Algebra rekonstruiert werden können.

Um zu zeigen, wie eine punktfreie Euklidische Geometrie axiomatisiert werden könnte, erarbeitete Tarski (1956) ein System, in dem lediglich vom Begriff der (beliebig kleinen, aber) ausgedehnten Kugel ausgegangen und mit den Axiomen der KEM

verbunden wird. Punkte, Linien, Flächen und Abstände sind nicht länger die grundlegenden Begriffe der Geometrie (daher „punktfrei“), sondern lassen sich wie folgt rekonstruieren:

**Def. 1:** Ein *Punkt*  $\beta$  ist die *Klasse von Kugeln*, die *konzentrisch* zu einer gegebenen Kugel  $b$  sind.

**Def. 2:** Zwei Punkte  $\alpha$  und  $\beta$  sind äquidistant zu einem Punkt  $\gamma$ , d.h.  $d(\alpha, \gamma) = d(\beta, \gamma)$ , falls eine Kugel  $c \in \gamma$  existiert, so dass kein Element von  $\alpha$  oder  $\beta$  Teil von  $c$  oder zu  $c$  disjunkt ist. Intuitiv gesprochen, „liegen“ die Punkte  $\alpha$  und  $\beta$  auf der Kugel  $c$  (siehe Abb. 4.1).



**Abbildung 4.1:** Punkte  $\alpha, \beta, \gamma$  als Klassen von konzentrischen Kugeln (angedeutet durch gestrichelte Kreise). Die Punkte  $\alpha$  und  $\beta$  sind äquidistant zu  $\gamma$ , falls sie auf der Kugel  $c \in \gamma$  (durchgezogen) „liegen“.

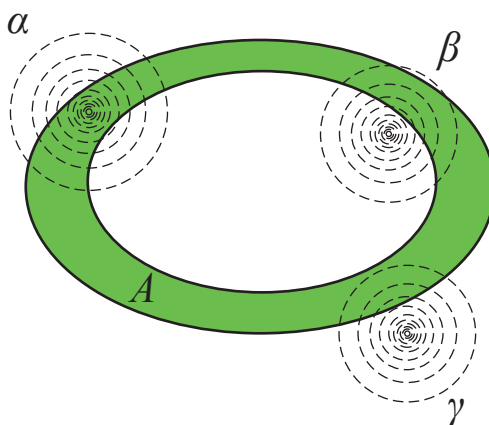
Dabei ist der Begriff der Konzentrität aus der Mereologie ableitbar, wobei eine Kugel  $a$  Teil einer Kugel  $b$  ist, wenn  $a$  von  $b$  eingeschlossen wird;  $a$  ist zudem konzentrisch zu  $b$ , wenn entweder  $a = b$ , oder alle Kugeln  $x, y$ , die  $a$  „äußerlich berühren“ und deren Mittelpunkte auf derselben Linie liegen wie der von  $a$ , die Kugel  $b$  „innerlich berühren“ [8]. Alle Beziehungen sind zudem symmetrisch, d.h.  $d(\alpha, \beta) = d(\beta, \alpha)$  und  $a$  konzentrisch zu  $b$  impliziert, dass  $b$  konzentrisch zu  $a$  ist. Die so gewonnenen Begriffe des Punktes und der Distanz genügen allen Axiomen der Euklidischen Geometrie [9].

Im selben Aufsatz gibt Tarski auch folgende zwei Definitionen, welche die Verbindung zur Booleschen Algebra ermöglicht:

Def. 3: Ein jeder *Körper* im Raum ist eine bestimmte *Summe von Kugeln*,  $A = \sum_I a_I$ .

Die Summe von Kugeln entspricht dabei der mereologischen Summendefinition aus Gl. (4.13), deren Eindeutigkeit wegen M24 garantiert ist. Zur Erinnerung: Die mereologische Summe einer Klasse von Gegenständen bezeichnet (nach Tarski) denjenigen Gegenstand, der alle Gegenstände der Klasse zum Teil hat und dessen Teile mit (mindestens) einem dieser Gegenstände überlappen.

Def. 4: Ein Punkt  $\alpha$  liegt im Inneren von  $A$ ,  $\alpha \in \text{int}(A)$ , g.d.w. eine Kugel  $x$  existiert, die gleichzeitig Element von  $\alpha$  und Teil des Gegenstandes  $A$  ist (vgl. Abb. 4.2).



**Abbildung 4.2:** Gegeben sei der Körper  $A$ . Von den drei abgebildeten Punkten (Klassen)  $\alpha, \beta, \gamma$  liegt nur der erste im Inneren von  $A$ , da nur für  $\alpha$  (mindestens) eine Kugel existiert, die sowohl Element von  $\alpha$  als auch Teil von  $A$  ist.

Zur Verbindung von Mereologie mit der Booleschen Algebra benutzt Tarski zwei weitere Postulate. Erstens, soll die Klasse der Punkte im Inneren eines Gegenstandes  $A$  einer regulären offenen (nicht-leeren) Menge  $\mathbf{A}$  entsprechen und umgekehrt:

$$\{\chi \mid \chi \in \text{int}(A)\} \cong \mathbf{A}. \quad (4.18)$$

Zweitens, entspreche die Teilsrelation zwischen zwei Gegenständen der Inklusionsrelation zwischen solchen Mengen:

$$\forall \chi (\chi \in \text{int}(A) \rightarrow \chi \in \text{int}(B)) \Leftrightarrow A < B \cong \mathbf{A} \subset \mathbf{B}. \quad (4.19)$$

Die Summe der Körper,  $A + B$ , entspreche per Definition dem kleinsten Gegenstand, der sowohl  $A$  als auch  $B$  zum Teil hat. Wegen Gl. (4.18) entspricht  $A + B$  selbst einer regulären offenen Menge  $\mathbf{S}$ ; wegen Gl. (4.19) gilt, dass  $\mathbf{A} \subset \mathbf{S} \wedge \mathbf{B} \subset \mathbf{S}$ , wobei  $\mathbf{S}$  minimal ist. Dies entspricht in moderner Schreibweise (Loeb 2014, S. 267f.):

$$\mathbf{A} + \mathbf{B} \stackrel{\text{def}}{=} \mathbf{S} = (\overline{\mathbf{A}} \cup \overline{\mathbf{B}})^0, \quad (4.20)$$

wobei  $\overline{\mathbf{M}}$  die kleinste abgeschlossene Menge bezeichnet, die  $\mathbf{M}$  enthält;  $\mathbf{M}^0$  steht für das Innere der Menge  $\mathbf{M}$ .

Es entspreche z.B. der Inhalt einer Kugel mit dem Radius  $r_i$  der (offenen) Menge  $K_i$  aller Punkte  $\{\vec{x} \in \mathbb{R}^2 \mid |x| < r_i\}$ . Bildet man die mereologische Summe aus zwei konzentrischen Kugeln,  $K_1 + K_2$  (mit  $r_2 > r_1$ ), entspricht i) deren Inhalt der (offenen und regulären) Menge aller Punkte  $\{\vec{x} \in \mathbb{R}^2 \mid |x| < r_2\}$  und ii) ist sie das (minimale) Objekt, welches sowohl  $K_1$  als auch  $K_2$  zum Teil hat.

Warum entspricht die Summe dann aber nicht einfach der Vereinigung von Mengen, also  $\mathbf{S} = \mathbf{A} \cup \mathbf{B}$ ? An einem Beispiel soll dies deutlich werden: Betrachte den 1-dimensionalen Fall, wo  $\mathbf{A} = (0, 1)$  und  $\mathbf{B} = (1, 2)$  jeweils Intervalle auf der reellen Zahlenachse darstellen. Beide sind (nichtleere) offene und reguläre Mengen  $\mathbf{X}$ , d.h. sie sind identisch zum Inhalt ihres Abschlusses (die Reihenfolge spielt eine Rolle),

$$\mathbf{X} = \overline{\mathbf{X}}^0, \quad (4.21)$$

da

$$\overline{(i, j)}^0 = [i, j]^0 = (i, j). \quad (4.22)$$

Die Vereinigung  $\mathbf{A} \cup \mathbf{B} = (0, 1) \cup (1, 2)$  ist zwar offen, erfüllt das Kriterium aus Gl. (4.21) jedoch *nicht*. Im Gegensatz dazu gilt Gl. (4.21) aber für die Menge  $(\overline{\mathbf{A} \cup \mathbf{B}})^0$ :

$$(\overline{(0, 1) \cup (1, 2)})^0 = ([0, 1] + [1, 2])^0 \quad (4.23)$$

$$= [0, 2]^0 \quad (4.24)$$

$$= (0, 2), \quad (4.25)$$

welche wiederum offen und regulär ist. Das so beschriebene Kalkül auf den offenen (regulären) Mengen bildet dann ein Boolesche Algebra, allerdings ohne Nullelement. Die enge Verwandtschaft zwischen klassischer extensionaler Mereologie und Boolescher Algebra wurde schon von Tarski (1935, S. 191; Anm. 5) selbst wie folgt beschrieben:

Die formalen Unterschiede zwischen dem erweiterten System der Booleschen Algebra und der Mereologie lassen sich bis auf einen Punkt reduzieren: auf Grund der Mereologie existiert nichts dem leeren Elemente Entsprechendes. Auch diese Differenz lässt sich jedoch durch eine gewisse Interpretation des Symbols „0“ beseitigen.

Die Ergebnisse von Tarski könnten nun auch wie folgt interpretiert werden: Die Euklidische Geometrie stellt in keiner Weise ein Abbild oder eine Repräsentation unabhängig existierender (bzw. idealisierter) Gegenstände oder gar „des Raumes“



dar, sondern ist die Konsequenz einer Verweisstruktur, die nur ausgedehnte Körper als Referenten besitzt und den Axiomen der KEM genügt. Es handelt sich also um eine Art Toy-Modell für die Erzeugung eines Symbolsystems. (Analoges gilt für die Boolesche Algebra, welche weniger den „Gesetzen des Denkens“ oder einer „natürlichen Logik“ entspricht, sondern aus dem oben genannten Verfahren unter Hinzufügung gewisser Postulate gewonnen werden kann.)

### 4.3 Erweiterte Mereologien

#### Atomismus

Unter Zuhilfenahme weiterer Kriterien lassen sich zusammen mit den formalen Definitionen aus dem vorigen Unterabschnitt die Eigenschaften der Teil-Ganzes-Beziehungen näher spezifizieren.

Dies soll nun anhand dreier Beispiele geschehen. In der mereologischen Rekonstruktion der Geometrie gingen wir vom Begriff der *ausgedehnten* Kugel aus. Intuitiv lässt sich daraus schließen, dass jeder Körper im Raum einen (ausgedehnten, echten) Teil besitzt; und da wir die Teileinsrelation für transitiv halten (M2), besitzt er sogar unendlich viele Teile. Dies muss aber nicht für alle mereologischen Systeme der Fall sein. Man spricht von einer *atomistischen Mereologie*, wenn gefordert wird, dass die Teileinsrelation letztlich an einer Stelle abbricht, nämlich im Bereich der *Atome*. Die Frage nach der Atomizität ist zwar nicht auf sie eingeschränkt, bezieht sich aber meist auf die materielle Konstitution von Objekten. Neben den Axiomen der minimalen Mereologie (M1 - M3) wird hier zusätzlich gefordert, dass es Objekte („Atome“) gibt, die keinen echten Teil besitzen:

$$at\ x \Leftrightarrow \neg \exists z(z \ll x). \quad (4.26)$$

Eine unmittelbare Konsequenz ist, dass Atome, die überlappen identisch sein müssen. Aus der Atomizität von  $x$  und  $y$  folgt mit der Definition des Überlapps, dass nur der triviale Fall  $z = x = y$  wahr sein kann:

$$at\ x \wedge at\ y \wedge \exists z(z < x \wedge z < y) \rightarrow z = x = y. \quad (4.27)$$

Die (naturphilosophische) Auffassung, dass alles in Atome zerfällt, lässt sich durch

$$AM1: \forall y \exists x(at\ x \wedge x < y)$$

ausdrücken: Wenn ein beliebiges Objekt  $y$  das Atom  $x$  zum Teil hat, so ist  $x$  entweder identisch zu  $y$  (dann wäre  $y$  selbst ein Atom) oder  $x$  ist ein echter Teil von  $y$ . Aus M3

folgt, dass dann ein  $z$ , das disjunkt zu  $x$  ist, existiert. Für dieses  $z$  gilt jedoch selbst wieder AM1, womit die Teilung solange fortgesetzt werden kann, bis  $y$  gänzlich in eine Menge aus Atomen zerfallen ist.

Das Gegenteil, nämlich die Auffassung, dass alles (bis ins Unendliche) teilbar ist, kann durch

$$\text{AM2: } \forall y \exists x (x \ll y)$$

ausgedrückt werden.

Besonders ausdrucksstark wird AM1, wenn zusätzlich das starke Supplement M5 gelten soll. Dieses hat zur Konsequenz, dass Objekte aus denselben Teilen identisch sind. Somit folgt, dass Objekte, die aus denselben Atomen gebildet werden, identisch sind. Diese Eigenschaft wurde von Nelson Goodman als „Hyperextensionalität“ bezeichnet, was ein Kriterium für den Nominalismus darstelle: „A system is nominalistic [...], if no two of its entities are generated from exactly the same atoms.“ (Goodman 1958, S. 65.).

Interessant sind insbesondere die kompositionalen Aspekte der atomistischen Mereologie. Unter Berücksichtigung von M14 (bzw. M24 für unendliche Systeme) gilt für atomistische Mereologien nämlich nicht nur AM1, dass alles in Atome zerfällt, sondern auch die komplementäre Aussage, dass jedes Objekt identisch zur Summe seiner Atome ist.

$$y = \sum x (x \ll y \wedge \text{at } x) \quad (4.28)$$

Beweis: Sei  $y = (x_1 + x_2 + \dots + x_n)$  ein Objekt, das aus  $n$  Atomen  $\{x_i\}$  gebildet ist, und sei zudem  $v_k$  ein beliebiger Teil von  $y$ ,

$$v_k < y \rightarrow v_k \circ y \Leftrightarrow v \circ (x_1 + x_2 + \dots + x_n). \quad (4.29)$$

Es lässt sich nun zeigen, dass  $v_k$  entweder selbst eines der Atome  $x_i$  ist oder ein Teilesumme der Atome von  $y$  darstellt. Wegen der Definition der mereologischen Summe aus Gl. (4.12) gilt:

$$v_k \circ (x_1 + x_2 + \dots + x_n) \Leftrightarrow \bigvee_i (v_k \circ x_i) \quad (4.30)$$

d.h.  $v_k$  überlappt mit mindestens einem der  $x_i$ . Daraus folgt, dass mindestens ein  $z$  existiert, für das gilt

$$z < v_k \wedge z < x_i \quad (4.31)$$

- Sei  $v_k$  atomar. Dann wäre obige Gleichung nur erfüllt, falls  $z = v_k = x_i$  und analog für alle weiteren Teile  $v_j$  von  $y$ . Jeder Teil  $v$  von  $y$  entspricht also genau

einem der Atome  $x_i$ . Weil die mereologische Summe unter Permutation  $\Pi(j)$  invariant ist (vgl. die rechte Seite von Gl. (4.12)), gilt:

$$y = (x_1 + x_2 + \cdots + x_n) = (x_{\Pi(1)} + x_{\Pi(2)} + \cdots + x_{\Pi(n)}). \quad (4.32)$$

Somit ist  $y$  eindeutig als Summe seiner Atome definiert.

- Für den Fall das  $v_k$  kein Atom ist, zerfällt  $v_k$  nach **AM1** aber selbst vollständig in Atome. Sei  $v'_k$  ein atomarer Teil von  $v_k$ . Wegen der Transitivität der Teilseinsrelation gilt:

$$v'_k < v_k \wedge v_k < y \rightarrow v'_k < y. \quad (4.33)$$

Obige Argumentation wäre dann auch für  $v'_k$  gültig;  $v_k$  entspricht dann selbst einer Summe aus Atomen und  $y$  lässt sich als Summe solcher Summen schreiben, wobei die Reihenfolge der Gruppierung keinen Unterschied macht:

$$y = x_1 + (x_2 + \cdots + x_n) = (x_1 + x_2) + \cdots + x_n, \quad (4.34)$$

usw.

Dies ist aber ganz erstaunlich, bedeutet es nichts anderes, als dass ein Objekt durch dessen Atome eindeutig festgelegt ist. Dies erscheint ganz natürlich, wenn wir an die materielle Konstitution von Gegenständen denken, wie sie etwa im klassischen Atomismus des 19. Jahrhundert ihren Ausdruck fand. Eine mereologische Analyse muss jedoch nicht auf materielle Konstitution von Gegenständen beschränkt sein. Rufen wir uns kurz die Interpretation in Erinnerung, wonach, möglichst allgemein gesprochen, die Mereologie von Verweisen handelt, die sich auf beliebige (konkrete) Gegenstände beziehen.

Eine Veranschaulichung liefert das Urteil „Kopf und Stiel sind die Teile des Besens“, wenn es als Kurzform für „dasjenige, worauf ich mit dem Wort ‚Besen‘ verweise, hat zwei Gegenstände zum Teil, auf welche ich mit den Wörtern ‚Kopf‘ und ‚Stiel‘ verweise“ aufgefasst wird. Nun könnte man sich fragen, ob denn die Verweise „Kopf“ und „Stiel“ den Begriff „Besen“ auf eindeutige Weise festlegen? Offensichtlich scheint es auch andere Zerlegungen zu geben, die auf dasselbe Objekt führen – etwa „der Besen ist die Summe aller Holz- und Strohmodule“.

Eine (extensionale) atomistische Mereologie stellt nun sicher, dass es keine alternative Bestimmung desselben Objekts bezüglich seiner Teile gibt, die zu neuen Bedeutungen führen könnte, so etwa wenn die Objekte „Besen = Kopf + Stiel“ und „Besen = Summe aller Holz- und Strohmodule“ auf zwei unterschiedliche Gegenstände verweisen würden. Atomistische Mereologien stehen also in einem Spannungsverhältnis zu pluralistischen Auffassungen, die von verschiedenen „Arten des Gegebenseins“ oder „Aspekten“ desselben Gegenstandes sprechen.

## Funktionalität

Am Beispiel der *Funktionalistischen Mereologie* soll nun untersucht werden, wann eine Erweiterung der mereologischen Konzepte selbst mit den Mitteln der Mereologie erfolgen kann und wann nicht. Dazu wird eine neue Relation, die des funktionalen Teilseins, eingeführt und gefragt, wie sich diese rekonstruieren ließe. Funktionalistische Mereologie impliziert, dass die Redeweise „ist Teil von“ nicht mehr nur durch die primitive Relation  $<$  ausgedrückt werden kann, sondern weiter qualifiziert werden muss:  $x$  ist genau dann ein Teil von  $y$ , wenn  $x$  und  $y$  in einem bestimmten funktionalen Verhältnis zueinander stehen. Das Herz ist ein Teil des Organismus, genau deswegen, weil es eine wichtige Funktion erfüllt (nämlich den Blutkreislauf zu ermöglichen), die zum Erhalt des Organismus beiträgt:

$$\text{FM1: } x <_{\text{F}} y \Leftrightarrow x \ll y \wedge f(x, y) \text{ [10]}$$

Das Herz nimmt eine bestimmte Funktion in Bezug auf den Organismus ein. Dabei spielt die Tatsache, dass das Herz beim Menschen an einer bestimmten Stelle zu finden ist („hinter dem Brustbein“), keine Rolle abgesehen davon, dass die Lokalisierung hinter dem Brustbein eine (wahrscheinlich kontingente) Bedingung für die Realisierung der Funktion „pumpt Blut durch den Körper“ darstellt (weil z.B. sonst nicht genügend Platz für so ein Organ wäre). Hier drückt sich das Motiv der (immer nur relativen) Unabhängigkeit der Funktion vom Substrat aus [11]. Wir können uns prinzipiell vorstellen, dass die Funktion  $f$  auch anders realisiert werden könnte, als durch ein inneres Organ, das hinter dem Brustbein sitzt und aussieht wie ein Herz (etwa durch ein künstliches Pumpsystem außerhalb des Körpers).

Welche Konsequenzen hat dies für die Transitivität der modifizierten Teilseinsrelation? Das Herz ist Teil des Organismus, welcher selbst Teil eines größeren Ökosystems ist. Daraus folgt aber nicht, dass das Herz selbst Teil des Ökosystems ist, da dessen Funktion nur relativ zum Erhalt eines bestimmten Einzelorganismus relevant ist, aber nicht direkt zum Erhalt des Ökosystems selbst beiträgt. Man könnte sich etwa vorstellen, dass ein äquivalentes Ökosystem von einer Menge an herzlosen Robotern realisiert würde, welche das exakt selbe Verhalten an den Tag legen. Andere Beispiele für das Vorkommen intransitiver Teilseinsrelationen finden sich beispielsweise bei Nicholas Rescher (1955, S. 10): Teile einer Organisation (Personen) sind nicht immer auch Teile der entsprechenden Dachgesellschaft und Mitochondrien sind Bestandteile von Zellen, aber nicht (im gleichen Sinne) Bestandteile des Gesamtorganismus.

Etwas formaler: Sei  $x <_{\text{F}} y$  und  $y <_{\text{F}} z$ , dann gilt wegen M2 zwar, dass  $x < z$  aber aus  $f(x, y) \wedge f(y, z)$  folgt nicht automatisch, dass  $f(x, z)$  und somit nicht  $x <_{\text{F}} z$ . In anderen Worten: Die Transitivitätseigenschaft ist für die Teilseinsrelation FM1

nicht mehr (axiomatisch) festgelegt, sondern hängt von der Struktur von  $f(x, y)$  ab. Die entscheidende Frage ist nun, ob dies durch rein mereologische Konzepte dargestellt werden kann, oder ob die Teileinsrelation mit nicht-mereologischen Konzepten angereichert werden muss. Die Auffassung, dass eine Zelle zwar als Teil des entsprechenden Gewebes bezeichnet werden kann, im selben Sinne aber *nicht* als Teil des Körpers, ließe dies vermuten.

Für eine jede Bewusstseinstheorie von besonderem Interesse ist der Fall, wo wir davon sprechen, dass Bewusstseinsinhalte zwar Teile eines mentalen Zustands sind und dass solche Zustände im Gehirn realisiert sind, die Bewusstseinsinhalte aber nicht Teile des Gehirns sind (oder in diesem gefunden werden können). Darauf wies etwa schon Fred Dretske (1995, S. 35) hin: Geschichten finden sich zwar in Büchern, aber was in diesen Geschichten passiert, das passiert nicht in den Büchern.

Es ist wahrscheinlich, dass sich die Relation zwischen Bewusstseinsinhalten und deren materiellen Korrelaten (falls es sie gibt) durch rein (mereo-)logische Beziehungen nicht darstellen lassen. Folgt man Autoren wie Fred Dretske (1995) oder Ruth Millikan (1989), so muss insbesondere der Verweis auf die Funktionalität mentaler Zustände beachtet werden und, in Hinblick auf eine Naturalisierung des Geistes, auf die natürlichen Entwicklung von Organismen zurückgeführt werden.

Ein weiterer Punkt betrifft die Existenz mereologischer Summen. Da Teile eines Ganzen immer auch mit ihnen überlappen, gilt mit der Definition aus FM1, dass sie „funktional überlappen“

$$x <_F y \rightarrow x \circ y \wedge f(x, y) \Leftrightarrow x \circ_F y. \quad (4.35)$$

Die Intuition, dass nur solche Objekte eine mereologische Summe bilden, die in einem solchen Funktionalverhältnis zueinander stehen, kann dann wie folgt ausgedrückt werden:

$$\text{FM2: } x \circ_F y \rightarrow \exists!(x + y).$$

Falls zwei Objekte funktional überlappen, existiert die mereologische Summe aus ihnen. Wiederum stellt sich die Frage, was eine geeignete Definition der Funktion  $f$  sei und wiederum liegt es nahe, dass diese über rein (mereo-)logische Prinzipien hinausführen wird. Die Auffassung, dass über die Zulässigkeit mereologischer Summen mit nicht-mereologischen Mitteln entschieden werden muss, wird auch von Simons (2006, S. 597) geteilt, der sogar dafür argumentiert, dass es kein rein *formales* (apriorisches) Kriterium gibt, das uns sagt, ob und unter welchen Bedingungen mereologische Summation erfolgen kann.

## Topologie

Eine Möglichkeit, das mereologische Kalkül des letzten Abschnittes um nicht-mereologische Konzepte zu erweitern, ist es, topologische Beziehungen zwischen Objekten einzuführen. Die Einführung von mereologischen und topologischen Konzepten fand sich dabei z.B. schon bei Whitehead; andere Vorläufer fanden sich in der Phänomenologie nach Husserl und der Gestaltpsychologie nach Kurt Lewin (siehe Smith 1994). Es scheint also nahezuliegen, Mereologie in Verbindung zur Topologie zu setzen. Eine *Mereotopologie*, die als Synthese von Mereologie und Topologie zu verstehen ist, wurde in den letzten Jahren z.B. von Smith (1996) oder von Casati & Varzi (1999) vorgeschlagen. Deren Motivation folgt aus der Beobachtung, dass räumliche Repräsentationen, also Darstellungen der räumlichen Verhältnisse zwischen Objekten (bzw. deren Teilen) weder durch rein mereologische noch durch rein topologische Konzepte vollständig zu beschreiben sind: Räumliche Beziehungen sind nicht notwendigerweise Teil-Ganzes-Beziehungen und Beziehungen zwischen räumlichen Objekten sind nicht notwendigerweise Beziehungen im Raum. Eine zentrale Rolle spielt hier der Begriff der (topologischen) Begrenzung.

Methodisch soll so vorgegangen werden, dass mereologische Verhältnisse nicht auf topologische zurückgeführt werden (oder umgekehrt). Zusammengenommen, so die Vermutung, ergeben sie eine reichere Struktur, die über eine reine mereologische (teilseinsbezogene) bzw. topologische (räumliche) hinausgeht. Dies soll hier kurz an einem Beispiel illustriert werden. Es sei die Relation der Berührung (Engl. *connection*) wie folgt mereologisch definiert:

$$C_m(x, y) \Leftrightarrow \exists z [x \circ z \wedge y \circ z \wedge \forall w (w < z \rightarrow w \circ x \vee w \circ y)] \quad (4.36)$$

Zwei räumliche Objekte („Gebiete“)  $x, y$  berühren sich genau dann, wenn mindestens ein gemeinsames Überlappendes  $z$  existiert, dessen Teile mit  $x$  oder  $y$  überlappen (vgl. auch PNK, Kap. VIII). Anders gesagt, es existiert kein  $v$ , das zwar ein Teil von  $z$ , aber disjunkt sowohl zu  $x$  als auch zu  $y$  wäre:

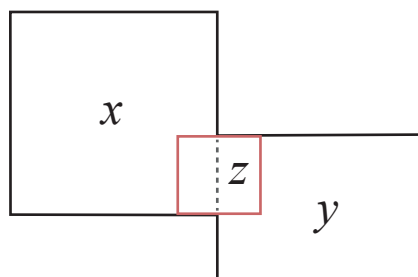
$$\forall w (w < z \rightarrow w \circ x \vee w \circ y) \Leftrightarrow \neg \exists v (v < z \wedge v \wr x \wedge v \wr y). \quad (4.37)$$

Ist dies eine gute Definition räumlicher (topologischer) Berührung? Tatsächlich kann man sich vorstellen, dass für alle Gegenstände, die sich berühren, immer mindestens ein solches  $z$  zu finden ist (Abb. 4.3).

Folgt, ausgehend von Gl. (4.36), dass jedes Objekt sich selbst berührt?

$$C_m(x, x) \Leftrightarrow \exists z [z \circ x \wedge x \circ z \wedge \neg \exists v (v < z \wedge (v \wr x \wedge v \wr x))] \quad (4.38)$$

$$\Leftrightarrow \exists z [z \circ x \wedge \neg \exists v (v < z \wedge v \wr x)] \quad (4.39)$$



**Abbildung 4.3:** Graphische Veranschaulichung von  $C_m(x, y)$ : Die beiden Hälften  $x$  und  $y$  berühren genau dann, wenn es ein gemeinsames Überlappendes  $z$  gibt, all dessen Teile mit  $x$  oder  $y$  überlappen.

Falls  $at\ x$  gilt, ist die Sache klar: Der einzige Teil, mit dem  $x$  überlappt ist  $x$  selber, daher gilt, dass  $z = x$  und wegen  $at\ x$  auch dass  $\forall w \{w < x \rightarrow w = x\}$ . Für  $x$  gilt aber auch, dass es nicht disjunkt zu sich selbst ist und somit  $C_m(x, x)$ .

Wie sieht es für nicht atomare  $x$  aus? Sei  $z$  ein beliebiger Teil von  $x$ . Dann gilt, dass  $z$  (trivialerweise) mit  $x$  überlappt. Da die Teilseinsrelation transitiv ist, gilt, dass jeder Teil von  $z$  selbst wiederum Teil von  $x$  ist und daher ebenso mit  $x$  überlappt. Es gibt also keinen Teil von  $z$ , der disjunkt zu  $x$  wäre; somit gilt  $C_m(x, x)$  immer. In beiden Fällen lässt sich die Selbstberührung also auf grundlegende Eigenschaften der Teilseinsrelation (Reflexivität und Transitivität) zurückführen.

Dies scheint aber des Guten zu viel zu sein, wie die linke Seite von Abb. 4.4 veranschaulicht. Das Objekt  $x$  erfüllt zwar  $C_m(x, x)$ , ist (im topologischen Sinn) dennoch nicht zusammenhängend. Es gilt, dass solche Objekte unserer mereologischen Definition nach selbstberührend sind, dennoch in eine Klasse von Teilen zerlegt werden können, die einerseits das Objekt vollständig ausmachen, sich aber andererseits nicht gegenseitig berühren. Es handelt sich hierbei also um ein „zerstreutes“ Objekt, das (mereologisch) mit sich selbst überlappt, aber (räumlich) unzusammenhängend ist. Für mereologische Summen  $\zeta = z_1 + z_2$  gilt:

$$\zeta \circ z_1 \wedge \zeta \circ z_2 \wedge \forall w (w < \zeta \rightarrow w \circ \zeta). \quad (4.40)$$

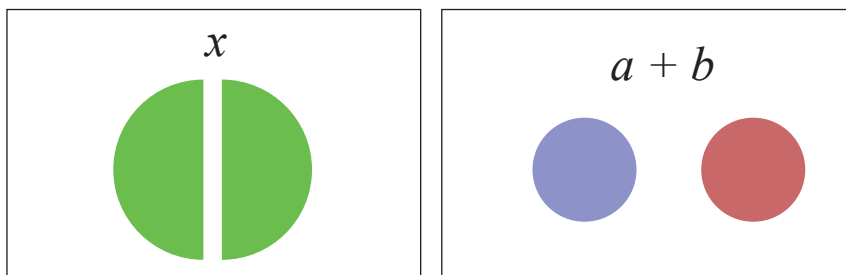
Und mit der Definition der mereologische Summe aus Gl. (4.12),  $\zeta = z_1 + z_2 \Leftrightarrow \forall w (w \circ \zeta \Leftrightarrow w \circ z_1 \vee w \circ z_2)$ , findet man schließlich

$$\zeta \circ z_1 \wedge \zeta \circ z_2 \wedge \forall w (w < \zeta \rightarrow w \circ \zeta) \quad (4.41)$$

$$\Leftrightarrow \zeta \circ z_1 \wedge \zeta \circ z_2 \wedge \forall w (w < \zeta \rightarrow w \circ z_1 \vee w \circ z_2) \quad (4.42)$$

$$\Leftrightarrow C_m(z_1, z_2), \quad (4.43)$$

d.h. Objekte, die mereologische Summen bilden, berühren im Sinne von Gl. (4.36) in ihren Summen. In einige Fällen von mereologischen Summen, scheinen sich die



**Abbildung 4.4:** Zwei Objekte, die zwar mit sich selbst überlappen („mereologisch berühren“), aber nicht der Intuition entsprechen, dass es sich dabei um räumlich kontinuierliche Objekte handelt. Links ist ein Objekt  $x$  dargestellt, welches unzusammenhängend ist und trotzdem  $C_m(x, x)$  erfüllt. Rechts ist die mereologische Summe,  $\zeta = a + b$ , zweier räumlich getrennter Objekte dargestellt, für die aber Gl. (4.39) trotzdem erfüllt ist.

Teile aber nicht zu berühren (siehe Abb. 4.4, rechts). Beschränkt man mereologische Summen nur auf Gegenstände, die sich auch berühren, würde dies vermieden werden.

Anders gesagt, scheint Gl. (4.36) (höchstens) eine notwendige Bedingung topologischer Berührung zu sein, aber keine hinreichende. Wenn wir jedoch annehmen, dass alle zulässigen Objekte auch im topologischen Sinne selbstberührend sind, also Fälle wie in Abb. 4.4 (links) ausschließen, entspricht die mereologische Berührung aus Gl. (4.36) der topologischen Berührung  $C(x, y)$ . Letztere Relation soll die folgenden Axiome erfüllen:

T1:  $C(x, x)$ : (Topologische) Berührung ist reflexiv,

T2:  $C(x, y) \Leftrightarrow C(y, x)$ : (Topologische) Berührung ist symmetrisch,

Die topologischen Axiome lassen sich nun mit den mereologischen Axiomen verbinden und formen eine *Mereotopologie*:

MT1:  $C_m(x, y) \Leftrightarrow C(x, y)$ : Mereologische Berührung impliziert topologische und *vice versa*. Unzusammenhängende Objekte wie aus Abb. 4.4 (links) sind ausgeschlossen.

Hat man pathologische Fälle axiomatisch ausgeschlossen, erscheint es sinnvoll, ein mereotopologisches Summenkriterium einzuführen: Nur Objekte, die berühren, können summiert werden (der Rumpf und die Glieder lassen sich summieren, die Glieder allein nicht):



$$\text{MT2: } C(x, y) \Leftrightarrow \exists z = x + y.$$

Dies würde dann auch Fälle wie in Abb. 4.4 (rechts) ausschließen.

Ein weiteres Axiom besagt, dass ein Teil immer „topologisch eingeschlossen“ sei von seinem Ganzen (nach Clarke 1981):

$$\text{MT3: } x < y \rightarrow \forall z (C(x, z) \rightarrow C(z, y)).$$

Falls  $x$  Teil von  $y$  ist, dann berührt jedes  $z$ , das  $x$  berührt, auch  $y$ . Da nach T1 jedes  $x$  sich selbst berührt, folgt mit  $z = x$  aus MT3 auch, dass

$$x < y \rightarrow C(x, y). \quad (4.44)$$

Die umgekehrte Richtung ist offensichtlich falsch, denn während Berührung symmetrisch ist, ist es (echtes) Teilsein nicht. Mit Gl. (4.44) folgt, dass auch Überlapp Berührung impliziert [12]:

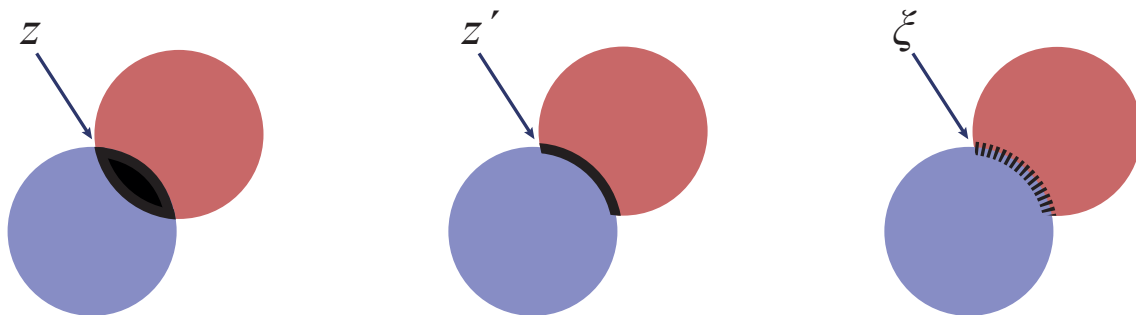
$$x \circ y \rightarrow C(x, y). \quad (4.45)$$

Interessante Fälle sind nun diejenigen, für welche zwar  $C(x, y)$  aber nicht  $x \circ y$  gilt. Man kann dann davon sprechen, dass  $x$  und  $y$  von einem  $\xi$  begrenzt werden:

$$\Xi(x, y) \Leftrightarrow C(x, y) \wedge \neg(x \circ y), \quad (4.46)$$

was dem Vorhandensein eines Objekts  $\xi$  mit folgenden Eigenschaften entspricht (vgl. Abb. 4.5):

1.  $x$  und  $y$  berühren sich „in  $\xi$ “ (mit  $\xi = \sum z_i$  für alle  $z_i$ , welche die rechte Seite von Gl. (4.36) erfüllen):
  - (a) Trivialerweise soll  $\xi$  ungleich der Summe  $x+y$  (oder eines anderen übergeordneten Ganzen) sein
  - (b) Es muss gelten, dass  $\xi \circ x \wedge \xi \circ y$ .
  - (c) Zudem müssen alle Teile von  $\xi$  ebenfalls mit  $x$  oder  $y$  überlappen.
2. Ausgeschlossen ist aber, dass  $\xi < x \wedge \xi < y$ , denn sonst würde ja gelten, dass  $x \circ y$ . Dies soll aber gerade nicht der Fall sein. Die Grenze  $\xi$  ist also nicht selbst Teil der Gebiete  $x$  und  $y$  (im Gegensatz zu  $z'$  in der Mitte von Abb. 4.5).
3.  $\xi$  ist nicht atomar, sonst würde gelten, dass aus  $\xi \circ x \wedge \xi \circ y$  folgen würde, dass  $\xi < x \wedge \xi < y$  (Atome überlappen genau dann mit einem Objekt, wenn sie ein Teil davon sind).



**Abbildung 4.5:** Drei Situationen, in denen sich zwei Flächen berühren. Links: Das Gebiet  $z$  überlappt beide Flächen. Mitte: Das Gebiet  $z$  kann beliebig klein gemacht werden (Grenzfall:  $z'$ ). Rechts: Rote und blaue Fläche haben keine gemeinsamen Teile, sondern sind nur durch eine Begrenzung  $\xi$  voneinander getrennt. Wie kann  $\xi$  mereotopologisch charakterisiert werden? Was ist der Unterschied zwischen der Abbildung in der Mitte und der ganz rechts?

Aus der ersten Bedingung folgt, dass

$$\xi \circ x \Leftrightarrow \exists z (z < \xi \wedge z < x), \quad (4.47)$$

d.h. dass  $\xi$  und  $x$  über mindestens einen gemeinsamen Teil verfügen; analog gilt dies für  $y$ . Das heißt aber, dass je mindestens ein (echter [13]) Teil von  $\xi$  entweder Teil von  $x$  oder Teil von  $y$  ist, nicht aber von beiden (sonst wäre wiederum  $x \circ y$  erfüllt):

$$\exists \xi \rightarrow \exists \xi_x (\xi_x \ll \xi \wedge \xi_x \ll x) \wedge \exists \xi_y (\xi_y \ll \xi \wedge \xi_y \ll y) \quad (4.48)$$

Für  $\xi_x$  ( $\xi_y$ ) können nun eine Reihe weiterer Eigenschaften abgeleitet werden:

- Wegen der Transitivität der Teilseinsrelation gilt, dass jeder Teil von  $\xi_x$  ( $\xi_y$ ) selbst Teil von  $x$  ( $y$ ) ist.
- $\xi_x$  und  $\xi_y$  sind disjunkt, andernfalls gäbe es einen gemeinsamen Teil und  $x \circ y$  würde folgen.
- $\xi_x$  und  $\xi_y$  berühren sich, genau dann wenn

$$\exists z [z \circ \xi_x \wedge z \circ \xi_y \wedge \forall w (w < z \rightarrow (w \circ \xi_x \vee w \circ \xi_y))]. \quad (4.49)$$

Sei nun  $z = \xi$ , dann folgt wegen  $\xi_x \ll \xi$ , dass  $\xi \circ \xi_x$  und analog  $\xi \circ \xi_y$ . Falls nun alle Teile von  $\xi$  mindestens mit  $\xi_x$  oder  $\xi_y$  überlappen, berühren sich  $\xi_x$  und

$\xi_y$ , obwohl sie disjunkt sind. Die Teilstruktur von  $x$  und  $y$  spiegelt sich in der Grenze  $\xi$  wider, welche selbst aus mindestens zwei Gebieten besteht, die sich berühren aber nicht überlappen, d.h. die begrenzt sind.

- Wegen MT2 folgt dann, dass die mereologische Summe  $\xi_x + \xi_y$  existiert. Da wir eben die Existenz von Teilen von  $\xi$  ausgeschlossen haben, die nicht mit  $\xi_x$  oder  $\xi_y$  überlappen, gilt,

$$\forall \xi_v (\xi_v \ll \xi \rightarrow \xi_v \circ \xi_x \vee \xi_v \circ \xi_y). \quad (4.50)$$

Aus  $\xi_v \ll \xi$  folgt  $\xi_v \circ \xi$  und die Summe  $\xi_x + \xi_y$  kann mit der Grenze  $\xi$  identifiziert werden:

$$\forall z (z \circ \xi \Leftrightarrow z \circ \xi_x \vee z \circ \xi_y \Leftrightarrow z \circ (\xi_x + \xi_y)) \quad (4.51)$$

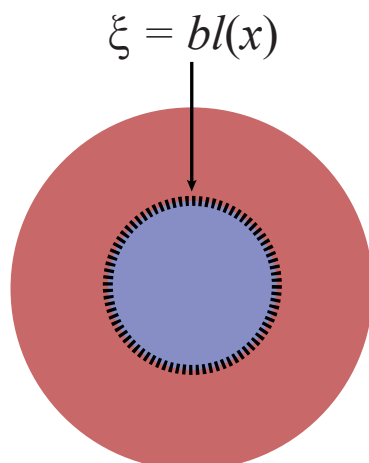
$$\Rightarrow \xi = \xi_x + \xi_y. \quad (4.52)$$

(Vgl. die Definition der mereologischen Summe aus Gl. (4.12); alle weiteren  $z$ , die in Frage kommen sind  $x, y, \xi, \xi_x, \xi_y$  für welche obige Bedingung trivialerweise erfüllt ist.)

Im Extremfall wird ein Gebiet  $x$  vollständig begrenzt (vgl. Abb. 4.6).  $\xi$  ist dann eine „Hülle“ (Engl. *blanket*),  $\xi = bl(x)$ , die  $x$  vollständig von der Außenwelt „abschirmt“ (obwohl  $x$  von allen Seiten berührt wird), d.h. es gibt keinen Teil von  $x$ , der gleichzeitig Teil seiner Umgebung,  $x^c$ , wäre. Dass es solche Gebiete gibt und dass diese wohldefiniert sind, muss zusätzlich angenommen oder empirisch begründet werden; eine wichtige Anwendung wird in nächsten Abschnitt noch vorgestellt werden. (Beachte auch, dass durch MT2 jeweils nur finite Summen definiert sind; der Fall unendlicher Summen müsste separat behandelt werden.)

Natürlich stellt sich die Frage, ob wir die Existenz von Objekten, die sich berühren, aber nicht überlappen, überhaupt zulassen sollten. Es gibt aber keinen Grund, dies von vornherein auszuschließen. Casati & Varzi (1999) sprechen davon, dass Begrenzungen (ebenso wie Löcher oder Schatten) räumliche Objekte sind, die selbst keinen Raum einnehmen. Eine ganz analoge Situation scheint für die bewusste Wahrnehmung zu gelten: Zu fragen, *wo* diese im Raum lokalisiert sei, wird vielen als unsinnig erscheinen – zumindest gemäß einer kartesischen Intuition. (Ebenso unsinnig wird es erscheinen, davon auszugehen, dass es kein Bewusstsein gäbe, nur weil wir es nicht als räumliches Objekt fassen können.)

Dennoch spüren wir Schmerzen im Fuß, haben Schmetterlinge im Bauch oder Lösen einfache Aufgaben im Kopf. Sind dies lediglich metaphorische Redeweisen? Doch



**Abbildung 4.6:** Das Gebiet  $x$  (blau) wird von einer Hülle  $\xi = bl(x)$  von der Außenwelt (rot) abgegrenzt, obwohl sie von dieser vollständig berührt wird.

kann ich nicht den visuellen Objekten eine Lage im Raum zuordnen? Höre ich nicht, dass die Töne aus einer ganz bestimmten Richtung kommen? Spüre ich nicht, dass ich von hinten angestarrt werde? Psychologisch betrachtet, projizieren wir Empfindungen in einen (nicht notwendigerweise physikalischen) Raum. Physiologisch betrachtet, hängt dies sehr eng mit der empirischen Topologie unseres Nervensystems zusammen (wir haben einen „Duft in der Nase“ und Berührung spüren wir dann, wenn wir vom entsprechenden Ort aus eine neuronale Verbindung ins Gehirn haben). Wahrnehmung ist raumhaft, ohne deswegen im Raum zu existieren.

Etwas Ähnliches lässt sich für die Zeitwahrnehmung vermuten und auch hier kann eine formale Beschreibung aus einer Erweiterung der Mereologie gewonnen werden (so etwa in der „Chronologie“ von Lejewski 1982).

## Anmerkungen zu Abschnitt 4

[1] Per Definition sind die Gegenstände, von denen die Mereologie handelt, als konkret anzunehmen, weshalb der Zusatz „konkret“ nicht immer explizit erwähnt werden wird.

[2] aus dem *Abstract*; Übersetzung RP.

[3] Genauer: Die Bedeutung der mereologischen Begriffe soll einzig und allein aus den axiomatisch festgelegten Operationen folgen, die im Rahmen eines bestimmten Kalküls zulässig sind (vgl. etwa Smith & Mulligan 1983, S. 74).

[4] Andere Notationen sind ebenfalls gebräuchlich. Beispielsweise wird in der Mereologie Leśniewskis die Teilseinsrelation als Funktion  $pt$  geschrieben. Die Aussage  $x$  ist Teil von  $y$ , im Kalkül nach Leonard & Goodman als „ $x < y$ “ geschrieben, schreibt sich nach Leśniewski dann zu „ $x \in pt(y)$ “. Es ist eine Frage der Konvention, welcher Schreibweise man dabei den Vorzug geben will (Simons 1987, S. 22f).

[5] Die (schwächere) Annahme, dass  $z$  lediglich verschieden von  $x$  sein müsste, scheint auf den ersten Blick zu genügen. Betrachtet man allerdings den Fall  $z \ll x \ll y$ , so sieht man, dass zwar  $z \ll y \wedge z \neq x$  erfüllt ist, aber nicht  $z \wr x$  gilt; vergleiche eine unendliche Serie konzentrischer Kreise ( $y (x (z (. . .)))$ ).

[6] Zeige, dass

$$z = (x + y) \rightarrow x < z \wedge y < z, \quad (4.53)$$

nicht aber umgekehrt: Bereits intuitiv scheint klar zu sein, dass  $x$  ein Teil von  $x + y$  ist. Andernfalls wäre  $\neg(x < (x + y))$  wahr. Dann würde aber M5 besagen, dass

$$\neg(x < (x + y)) \rightarrow \exists u (u < x \wedge u \wr (x + y)). \quad (4.54)$$

Somit würde ein  $u$  existieren, für das gilt:

$$u < x \rightarrow u \circ x, \quad (4.55)$$

und gleichzeitig:

$$u \wr (x + y) \Leftrightarrow \neg(u \circ (x + y)). \quad (4.56)$$

Aus der Existenz der binären Summe folgt aber, dass jedes Objekt, welches mit  $x$  überlappt, auch mit  $x + y$  überlappen muss. Es kann daher kein  $u$  geben, für welches sowohl (4.55) als auch (4.56) gilt. Analog gilt dies für  $y$ .

Es folgt somit:

$$z = x + y \rightarrow x < (x + y) \quad (4.57)$$

$$z = x + y \rightarrow y < (x + y) \quad (4.58)$$

$$\Rightarrow x < z \wedge y < z. \quad (4.59)$$

Für die Widerlegung der umgekehrten Richtung genügt es, ein Gegenbeispiel zu konstruieren. Sei also  $x < z' \wedge y < z'$ , d.h.  $z'$  würde in Folge sowohl  $x$  als auch  $y$  überlappen. Es existiere zudem ein  $\xi < z'$  mit  $\xi \wr x \wedge \xi \wr y$ , dann gilt, wegen

$$\xi < z' \rightarrow \xi \circ z', \quad (4.60)$$

dass  $\xi$  zwar mit  $z'$ , aber weder mit  $x$  noch  $y$  überlappt. Würde man also annehmen,  $z'$  sei identisch zur Summe  $x + y$ , widerspräche dies der Definition in Gl. (4.12), wonach alle Objekte, die mit dieser überlappen, auch mit  $x$  oder  $y$  überlappen müssen. Eine einfache Veranschaulichung, geht von ein Auto ( $z'$ ) aus. Dieses hat mehrere Teile, u.a. vier Räder ( $x, y$ ), und einen Motor ( $\xi$ ). Daraus folgt, dass es mit all diesen Teilen auch überlappt. Da der Motor aber nicht mit den Rädern überlappt, ist das Auto ( $z'$ ) nicht identisch zur mereologischen Summe der vier Räder ( $x + y$ , unter der Annahme diese würde existieren). Aus der Existenz einer mereologischen Summe  $z = (x + y)$  folgt zudem, dass diese die „kleinste obere Schranke“ darstellt, für die gilt, dass

$$\forall w (x < w \wedge y < w \Leftrightarrow z < w). \quad (4.61)$$

Das heißt, es gibt kein Objekt, welches sowohl  $x$  als auch  $y$  zum Teil hat, das nicht auch die Summe  $z = (x + y)$  zum Teil hätte.

Beweis: Es sei  $x < w$  und  $y < w$ , aber ohne, dass  $w$  auch  $z$  zum Teil hat, also  $\neg(z < w)$ .

Mit M5 folgt dann, dass

$$\neg(z < w) \rightarrow \exists v (v < z \wedge v \wr w). \quad (4.62)$$

Aus  $v < z$  folgt, dass  $v \circ z$  und somit

$$v \circ z \Leftrightarrow v \circ x \vee v \circ y \quad (\text{Definition der Summe } z). \quad (4.63)$$

Da aber sowohl  $x$  als auch  $y$  Teile von  $w$  sind, folgt daraus

$$v \circ w = \neg(v \wr w), \quad (4.64)$$

was im Widerspruch zur rechten Seite von Gl. (4.62) steht, und analog für die umgekehrte Richtung.

[7] Von Tarski (1956) stammt eine alternative Definition der Fusion, nämlich als dasjenige  $\zeta$ , für welches gilt:

$$\forall x (Fx \rightarrow x < \zeta) \wedge \forall y (y < \zeta \rightarrow \exists x (Fx \wedge y \circ x)). \quad (4.65)$$

Die Definition in Gl. (4.65) ist vielleicht etwas zugänglicher als die aus Gl. (4.13): Alle Objekte, welche  $F$  erfüllen, sind Teil der Summe  $\zeta$  und alle Teile von  $\zeta$  überlappen mit mindestens einem Objekt, welches  $F$  erfüllt. Die Definitionen sind nicht äquivalent, wobei Definition (4.65) etwas stärker ist. Für den Fall, dass M1-M3, M5 gilt, bilden Sie jedoch ein äquivalentes Axiomensystem (Hovda 2009, S. 56).

[8] Die formalen *mereologischen* Definitionen finden sich in (Tarski 1956, S. 26).

[9] Dies findet sich als „erstes Postulat“ in (Tarski 1956) und gründet sich auf einem Resultat des Italienischen Mathematikers Mario Pieri von 1908, wonach eine minimale Axiomatisierung der Euklidischen Geometrie durch die Klasse der Punkte zusammen mit der Äquidistanzrelation gegeben ist (siehe Marchisotto & Smith 2007).

[10] Hier stößt die mengentheoretische Veranschaulichung, die noch für die primitiven Relationen aus Kapitel 4.1 gemacht werden konnte, an ihre Grenzen. Dabei muss beachtet werden, dass es sich bei den „Funktionen“  $f(x, y)$  nicht um eine wohldefinierte Zuordnungsvorschrift zwischen Elementen einer Grundmenge und Zielmenge handelt, sondern allgemeine um Relationen zwischen Verweisen, die durch  $x$  und  $y$  ausgedrückt werden.

[11] Die Frage nach der Unabhängigkeit der Funktion vom Substrat ist letztlich eine empirische Frage; vielleicht gilt sie im Einzelfall, vielleicht aber auch nicht. Wichtig ist, dass es nicht auf der Ebene einer axiomatischen (im Sinne einer selbstevidenten) Festlegung entschieden werden kann.

[12] Beweis:

$$x \circ y \Leftrightarrow \exists z (z < x \wedge z < y). \quad (4.67)$$

Wegen Gl. (4.44) folgt nun einerseits, dass  $C(z, x)$  gilt. Andererseits gilt wegen MT3, dass  $\forall v (C(z, v) \rightarrow C(v, y))$ . Zusammengenommen folgt daraus  $C(x, y)$ . Alternativ folgt dies auch mit MT1 aus  $x \circ y \rightarrow C_m(x, y) \rightarrow C(x, y)$ .

[13] Angenommen,  $\xi$  sei selbst Teil von  $x$  ( $y$ ), dann wäre jeder Teil von  $\xi$  auch ein Teil von  $x$  ( $y$ ), insbesondere derjenige Teil, der Teil von  $y$  ( $x$ ) ist. Dann würde jedoch wieder gelten, dass  $x \circ y$ , also dass  $x$  und  $y$  einen gemeinsamen Teil haben.





## 5 Im Prozess des Teilens

*By tracing the way we represent such severance, we can [...] begin to see how the familiar laws of our own experience follow inexorably from the act of severance.*  
(George Spencer Brown: Laws of Form)

### 5.1 Projektive Mereologie

Während im letzten Abschnitt die KEM in ihren Grundzügen behandelt wurde, soll hier eine neue Darstellung des mereologischen Kalküls vorgestellt werden, welche verspricht, Mereologie wieder enger an die empirischen Wissenschaften zu führen. Dass etwa die Psychologie im 19. Jahrhundert einige Phänomene mithilfe der Mereologie klassifizierte (und somit formalisierte, siehe Smith & Mulligan 1982), erscheint manchen vielleicht nur noch als Kuriosum einer Frühphase einer exakten Wissenschaft.

Doch langsam finden mereologische Überlegungen wieder Eingang in die Kognitionswissenschaft (Eschenbach *et al.* 1994) und Sprachwissenschaft (Moltmann 1999). Neuerdings ist auch im Rahmen der „angewandten Ontologie“ (Munn & Smith 2008) außerhalb eines rein philosophischen Kontextes wieder vermehrt die Rede von Mereologie.

Dadurch erhalten auch die Fragen, nach dem „Wesen der Teileinsrelation“, „ob der Atomismus falsch ist“ oder ob ein Kalkül, „das nicht überlappende Begrenzungen kennt“, überhaupt zulässig ist, einen neuen Status: Solange keine internen Inkonsistenzen auftreten, besteht gar keine Veranlassung nicht-klassische (oder erweiterte) Mereologien von vornherein zu verwerfen. Wenn wir zudem die Auffassung vertreten, dass es bei mereologischen Analysen primär darum geht, die Strukturiertheit von Verweissystemen zu untersuchen, aus denen sich unterschiedliche (und potentiell unbegrenzt viele) symbolische Repräsentationen ableiten lassen, dann sind wir zu einer liberalen Haltung gegenüber der Zulässigkeit verschiedener mereologischer Systeme aufgefordert. Dass durch Tarski die KEM und die Boolesche Algebra als strukturgleiche Aussagenverbände rekonstruiert werden konnten (vgl. das Ende von Abschnitt 4.2), bestärkt einen vielleicht sogar in einer pluralistischen und konstruktivistischen Auffassung.

Im Folgenden soll zuerst die Unterscheidung zwischen kompositionalen und zerlegenden („projektiven“) Ansätzen vorgestellt werden. Die Frage, was eigentlich ein wohldefiniertes mereologisches Objekt ausmacht, wird einen zentralen Platz einnehmen. Eine mögliche Anwendung ergibt sich mit Blick auf das Thema Emergenz, wobei Emergenz in erster Linie als mereologisches Phänomen verstanden wird, das,

heuristisch gesprochen, in Situationen auftritt, in denen sich das Verhältnis zwischen Teilen und dem Systemganzen stark ändert oder gar umkehrt. Dass in einem spezifischen Sinn von einem abgegrenzten Objekt gesprochen werden kann, wird zu einer notwendigen Voraussetzung für die Emergenz neuer Eigenschaften.

### Komposition und Zerlegung

Wir wollen nun explizit die Kritik von Casati & Varzi (1999) aufgreifen, wonach das Wesen oder die Natur von Teilen und Ganzheiten in der klassischen Mereologie ungeklärt bliebe. Hierzu seien kurz drei Überlegungen vorgestellt werden, die im Anschluss noch weiter explizieren werden:

1. Objekte können als Resultate von Prozessen aufgefasst werden. Spezifischer für die Mereologie soll gelten, dass Teile als „abgeleitete“ und Ganzheiten als „primäre“ Objekte *relativ zu einer Operation der Projektion* verstanden werden können.
2. Für einige dieser Objekte gilt, dass sie *wohldefiniert* sind, d.h. dass sie strukturellen und anderen Kriterien genügen: Sie werden zu *Zeichen*, die auf konkrete Gegenstände verweisen. Wohldefinierte Ganzheiten, die aus mehreren Teilen bestehen, können als „Summen“ bezeichnet werden; wohldefinierte Objekte als „selbstständige“.
3. Teile als „abtrennbare“ Resultate zu betrachten, entspricht einer Idealisierung, die z.B. übersieht das Prozesse einer Einbettung bedürfen, in denen sie sich vollziehen. Eine solche Einbettung kann als der Kontext bezeichnet werden, bezüglich dessen Objekte *emergent* sind.

Die Frage nach der Natur mereologischer Objekte führt auf zwei mögliche Ansätze. Zum einen der kompositionale Ansatz, der von der Existenz wohldefinierter Objekte ausgeht und nach den angemessenen Bedingungen ihrer Summierung fragt: Seien  $x$  und  $y$  zwei Objekte. Unter welchen Bedingungen kann gesagt werden, dass die mereologische Summe  $\zeta = x + y$  existiert? In M9 fanden wir die Ansicht ausgedrückt, dass mereologische Summen immer dann existieren, wenn die Teile  $x$  und  $y$  überlappen; M14 postulierte die Existenz solcher Summen sogar dann, wenn dies nicht der Fall ist. Dies ist eine (oftmals auch kritisierte) Annahme, die in klassischen mereologischen Systemen häufig gemacht wird. Bei Simons (1987, Kap. 3) findet sich die Idee, die Zulässigkeit der mereologischen Summenbildung von einem Integrationskriterium abhängig zu machen: Mereologische Summenbildung ist nur dann erlaubt, wenn

die Objekte zusätzlich in einer bestimmten (nicht-mereologischen) Relationen zueinander stehen. Im letzten Abschnitt haben wir etwa funktionale Kriterien kennen gelernt (vgl. FM2). Ein anderes wichtiges Kriterium, das uns gleich noch einmal beschäftigen wird, knüpft die Zulässigkeit mereologischer Summation an das Vorhandensein einer Berührungsrelation (vgl. MT2), die eine schwächere Beziehung als mereologischer Überlapp darstellt (Überlapp impliziert Berührung aber nicht umgekehrt).

Diese Ideen können als Versuche betrachtet werden, zulässige Ganzheiten näher zu charakterisieren: Ganzheiten sind mereologische Summen, deren Teile ein bestimmtes Integrationskriterium,  $I$ , erfüllen,

$$I(x, y) \Leftrightarrow \exists \zeta = x + y, \quad (5.1)$$

unter der Annahme, dass die Objekte  $x$  und  $y$  gegeben sind.

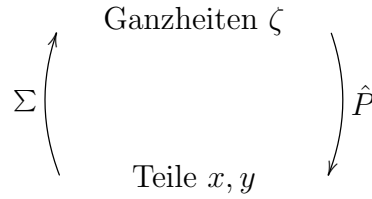
Im Gegensatz dazu steht die Vorstellung, dass Objekte erst in ganz bestimmten Kontexten als Teile einer (als primär genommenen) Ganzheit  $\zeta$  in Erscheinung treten; sie müssen als Teile erst „erzeugt“ werden. Formal ausdrücken lässt sich dies durch die Aussage, dass  $x$  genau dann ein Teil von  $\zeta$  ist, wenn es sich aus einer „Projektion“ von  $\zeta$  ergibt [1]:

$$\forall \zeta x (x < \zeta \Leftrightarrow \exists \hat{P} (x = \hat{P}\zeta)). \quad (5.2)$$

$\hat{P}$  ist dabei ein höherrangiges Objekt, welches – abgesehen von der Erzeugungsrelation in Gl. (5.2) – nicht zu  $x$  oder  $\zeta$ , möglicherweise aber zu anderen Operatoren in einem mathematischen oder mereologischen Verhältnis stehen kann.

Wichtig ist, dass es keine eindeutige Zuordnung zwischen den Objekten  $x$  und  $\zeta$  sowie dem Projektor  $\hat{P}$  gibt. Ein Projektor kann ohne nähere Einschränkung mehreren Objekten zugeordnet werden. Wenn wir in der Welt Teile vorfinden, impliziert dies jedoch, dass diese sich notwendig auf ein Ganzes beziehen und als Teile nicht isoliert von diesem vorkommen. Wir sprechen dann davon, dass Teile aus einem Ganzen „erzeugt“ worden sind, wobei dies in erster Linie auf eine konzeptuelle und weniger auf eine physikalisch-konstitutive Relation weist. Das Verhältnis von  $\zeta$ ,  $\hat{P}$  und  $x$  entspricht, metaphorisch gesprochen, dem eines Steinblocks, aus dem mehrere Objekte herausgeschlagen werden können.

Die Teil-Ganzes-Beziehung wird dann durch die Struktur der Projektoren-Algebra näher festgelegt. Insbesondere sollen Aussagen über Teil-Ganzes-Beziehungen zwischen Objekten als Aussagen über die Erzeugungsrelationen bzw. die entsprechenden Projektoren geschrieben werden.



**Abbildung 5.1:** Kompositionale und zerlegende Ansätze im Vergleich. Pfeile bezeichnen jeweils den Übergang von primären zu abgeleiteten Objekten.

Der Formalismus impliziert eine Darstellung mereologischer Beziehungen durch gerichtete Graphen. Die mereologische Struktur entspricht dabei einem Netzwerk, das aus Knoten (den mereologischen Objekten) und gerichteten Verbindungen (den Projektoren) zwischen diesen besteht. Gegeben die Ganzheit  $\zeta$ , ist die Existenz eines Teiles, gleichbedeutend mit der Existenz eines erzeugenden Projektors, was letztlich der Aussage in aus Gl.(5.2) entspricht.

Analog lassen sich auch andere mereologischen Operationen als Kombinationen von solchen Projektoren darstellen. Explizit sei dies hier für den mereologischen Überlapp angeführt:

$$x \circ y \Leftrightarrow \exists z (z < x \wedge z < y) \Leftrightarrow \exists \hat{P}_j \hat{P}_k (\hat{P}_j x = \hat{P}_k y), \quad (5.3)$$

respektive

$$x \wr y \Leftrightarrow \neg x \circ y \Leftrightarrow \neg \exists z (z < x \wedge z < y) \Leftrightarrow \forall \hat{P}_j \hat{P}_k (\hat{P}_j x \neq \hat{P}_k y). \quad (5.4)$$

Durch die Darstellung der „projektiven Mereologie“ sollen bestimmte Eigenschaften der Teil-Ganzes-Beziehung explizit als Folge der (Un)möglichkeit der Zerlegung von Ganzheiten gedeutet werden. Dazu seien komplexe Ausdrücke wenn möglich in „objekt-freier“ Form darzustellen, was bedeutet, dass sie nicht mehr auf ein bestimmtes Objekt  $x$ , mit diesen und jenen Eigenschaften Bezug nehmen, sondern für alle Zerlegungen generischer Objekte zutreffen.

Fall lediglich ausgedrückt werden soll, dass ein Projektor  $\hat{P}$  existiert, der ein Objekt  $\zeta$  in (ansonsten unbestimmte) Teile  $z$  erzeugt, soll folgende Kurzschreibweise eingeführt werden [2]:

$$\exists \hat{P} \Leftrightarrow_{\text{Df.}} \exists \hat{P} (z = \hat{P}\zeta). \quad (5.5)$$

Im nächsten Schritt lassen sich die fundamentalen mereologischen Axiome untersuchen. Ein erstes Axiom besagt, dass Projektoren idempotent sind:

$$\forall \hat{P} (\hat{P}\hat{P}\zeta = \hat{P}\zeta) \Leftrightarrow \hat{P}^n = \hat{P}. \quad (5.6)$$

Die **Idempotenz** ist dadurch motiviert, dass mehrmalige Anwendung desselben Projektors auf dasselbe Objekt führen soll.

Zusätzlich sei Projektion **assoziativ** bezüglich Multiplikation und **distributiv** bezüglich der Summation [3], d.h., dass

$$\hat{P}_1(\hat{P}_2\hat{P}_3) = (\hat{P}_1\hat{P}_2)\hat{P}_3 \quad \text{bzw.} \quad (\hat{P}_1 + \hat{P}_2)\zeta = \hat{P}_1\zeta + \hat{P}_2\zeta. \quad (5.7)$$

Insbesondere gilt [4], dass

$$\hat{P}\zeta = x \rightarrow \hat{P}x = x. \quad (5.8)$$

Gl. (5.8) ist dabei äquivalent zum Reflexivitätsaxiom  $x < x$ .

Eine wichtige Eigenschaft der Teilseinsrelation ist deren Transitivität (vgl. Axiom M1). Ein erster Versuch, diese Eigenschaft in den Projektorformalismus zu übersetzen, wäre:

$$(x = \hat{P}_1y) \wedge (y = \hat{P}_2\zeta) \rightarrow x = \hat{P}_{12}\zeta \quad \text{mit} \quad \hat{P}_{12} = \hat{P}_1\hat{P}_2 \quad (5.9)$$

Nun gilt aber, dass das Produkt aus zwei Projektoren nur dann selbst ein Projektor ist, falls die zwei Projektoren „kommutieren“, d.h. falls

$$\left[ \hat{P}_1, \hat{P}_2 \right] = 0 \Leftrightarrow (\hat{P}_1\hat{P}_2\zeta = \hat{P}_2\hat{P}_1\zeta), \quad (5.10)$$

da nur so die Idempotenzeigenschaft von  $\hat{P}_{12}$  garantiert ist [5].

Dies ist eine natürliche Einschränkung der Teil-Ganzes-Beziehung: Intuitiv sind zum Beispiel materielle Teile von funktionalen Teilen nicht notwendigerweise identisch zu den funktionalen Teilen von materiellen Teilen etc. Nicht-Kommutativität von Projektoren ist also eine mögliche formale Darstellungsweise der heuristischen Aussage, dass „verschiedene Arten von (nicht-transitiven) Teilseinsrelationen existieren“.

Es folgt schließlich die objekt-freie Darstellung des **Transitivitätsaxioms** zu:

$$\exists \hat{P}_1\hat{P}_2 \left( \left[ \hat{P}_1, \hat{P}_2 \right] = 0 \right) \rightarrow \exists \hat{P}_{12} (\hat{P}_{12} = \hat{P}_1\hat{P}_2). \quad (5.11)$$

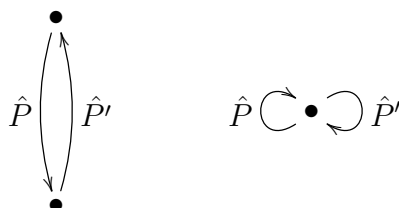
Eine zweite fundamentale Eigenschaft der (echten) Teil-Ganzes-Beziehung war deren **Asymmetrie** (M2). Intuitiv bedeutet dies im Projektorformalismus, dass es unmöglich ist, einen Teil auf sein Ganzes „rückzuprojizieren“:

$$(x \neq y \wedge x = \hat{P}y) \rightarrow \neg(y = \hat{P}'x) \quad (5.12)$$

Auch dies soll im nächsten Schritt objekt-frei formuliert werden. Es sei zuerst davon ausgegangen, dass  $x$  und  $y$  zwei verschiedene Objekte bezeichnen. Wären sowohl  $x = \hat{P}y$  als auch  $y = \hat{P}'x$  wahr, so würde folgen, dass  $x = \hat{P}\hat{P}'x$ . Dies gilt es, durch das Asymmetrieaxiom auszuschließen, welches sich objekt-frei wie folgt darstellen lässt:

$$\exists \hat{P} \rightarrow \neg \exists \hat{P}' (\hat{P}\hat{P}' = 1 \wedge \hat{P} \neq \hat{P}'). \quad (5.13)$$

Der Ausdruck in Gl. (5.13) unterscheidet sich von der ursprünglichen Formulierung von M2 dadurch, dass er sich auf alle möglichen Projektionen eines generischen Objektes bezieht, während in der KEM Asymmetrie nur für den Fall der echten Teilseinsrelation ( $\ll$ ) gilt, nicht aber für die uneingeschränkte Teil-Ganzes-Beziehung, in der möglicherweise  $x = y$  gilt. Anstelle die Unterscheidung zwischen den Operationen  $<$  und  $\ll$  auf die Nicht-Identität der betreffenden Objekte zurückzuführen, wurde daher die Klausel  $\hat{P} \neq \hat{P}'$  eingeführt, die sich nur auf die entsprechenden Projektoren bezieht. Dabei sind zwei mögliche Szenarien denkbar, die sich wie folgt als Graphen veranschaulichen lassen:



Dabei entspricht  $\bullet$  einem Objekt, während die gerichteten Pfeile den Projektoren (Teil-Ganzes-Beziehungen) entsprechen. Die linke Situation wird durch Gl. (5.13) ausgeschlossen, die rechte ist jedoch wegen  $\hat{P} = \hat{P}'$  zulässig (und entspricht dann dem Fall  $x = y$ ).

## 5.2 Objektivität

Anders als in den kompositionalen Ansätzen, in denen wir davon ausgehen können, dass Objekte  $x$  und  $y$  vorliegen, die sich als Summen zu neuen Ganzheiten kombinieren lassen, muss in der projektiven Mereologie zuerst geklärt werden, was überhaupt unter einem „guten“ Objekt zu verstehen sei.

Mereologische Objektivität zeichne sich dabei durch folgende drei Charakteristika aus:

1. Objekte sind mit sich selbst und nur mit sich selbst *identisch*.

Die wohl bekanntesten Prinzipien der Identität gehen auf Leibniz zurück. Einerseits das Prinzip der Identität des Ununterscheidbaren (*principium identitatis indiscernibilium*) und andererseits dasjenige der Ununterscheidbarkeit des Identischen:

$$pii: \forall A (A(x) \Leftrightarrow A(y)) \rightarrow x = y$$

$$pui: x = y \rightarrow \forall A (A(x) \Leftrightarrow A(y))$$

Wenn für alle möglichen Eigenschaften gilt, dass sie für  $x$  genau dann zutreffen, wenn sie auch für  $y$  zutreffen, so sind  $x$  und  $y$  identisch (bzw. umgekehrt). Oft wird angenommen, dass  $p_{ii} \Leftrightarrow p_{ui}$  gilt, was vielleicht intuitiv der Fall, aber logisch nicht zwingend ist [6].

Dies soll nun auf die Darstellungswiese der projektiven Mereologie übertragen werden, wobei anstelle der generischen Eigenschaft  $A$  lediglich von Teil-Ganzes-Beziehungen als möglichen Eigenschaften ausgegangen werden soll. Das  $p_{ii}$  folgt zu:

$$\forall \hat{P} (\hat{P}x = \hat{P}y) \rightarrow x = y, \quad (5.14)$$

und das  $p_{ui}$  zu:

$$x = y \rightarrow \forall \hat{P} (\hat{P}x = \hat{P}y). \quad (5.15)$$

Wir wollen dies nun mit Blick auf die Idee betrachten, wonach Objekte die Resultate von Prozessen sind. Es sei dazu angenommen, dass  $x$  aus der Anwendung eines Projektors  $\hat{P}_x$  gewonnen werden kann, d.h. selbst Teil einer übergeordneten Ganzheit  $\zeta$  ist (und analog für  $y$ ):

$$x = \hat{P}_x \zeta, \quad y = \hat{P}_y \zeta. \quad (5.16)$$

Wir wollen nun zeigen, wie die Identität von Objekten mit der Beziehung zwischen ihren „Erzeugenden“,  $\hat{P}_x$  und  $\hat{P}_y$ , zusammenhängt. Einsetzen in  $p_{ui}$  führt auf:

$$x = y \rightarrow \forall \hat{P}_i (\hat{P}_i \hat{P}_x \zeta = \hat{P}_i \hat{P}_y \zeta), \quad (5.17)$$

falls also  $x$  und  $y$  identisch sind, dann kann kein zweiter Projektor  $\hat{P}_i$  auf die Projektionen  $\hat{P}_x$  und  $\hat{P}_y$  der Ganzheit  $\zeta$  angewendet werden, der uns einen Unterschied erkennen lässt.

Dieses Resultat lässt sich nun im Rahmen der Projektorenalgebra weiter untersuchen. Es sei etwa der Fall betrachtet, wo  $\hat{P}_i = \hat{P}_x$ . Dann folgt aus der rechten Seite von Gl. (5.17) wegen der Idempotenz des Projektionsoperators:

$$\hat{P}_x \hat{P}_x \zeta = \hat{P}_x^2 \zeta = x = \hat{P}_x \hat{P}_y \zeta, \quad (5.18)$$

und analog für  $\hat{P}_i = \hat{P}_y$ :

$$\hat{P}_y \hat{P}_x \zeta = y. \quad (5.19)$$

Aus einer angenommenen Gleichheit von  $x$  und  $y$  folgt dann aber auch, dass

$$x = y \rightarrow (\hat{P}_x \hat{P}_y \zeta = \hat{P}_y \hat{P}_x \zeta). \quad (5.20)$$

Gl. (5.20) lässt sich unter Verwendung des Kommutators aus Gl. (5.10) abkürzen zu:

$$x = y \rightarrow [\hat{P}_x, \hat{P}_y] = 0. \quad (5.21)$$

Eine interessante Konsequenz lässt sich aus der Kontraposition ziehen:

$$\left[ \hat{P}_x, \hat{P}_y \right] \neq 0 \rightarrow x \neq y, \quad (5.22)$$

d.h. Objekte, deren erzeugende Projektoren nicht kommutieren, sind verschieden. Insbesondere lassen sich Aussagen zur Verschiedenheit von Objekten mithilfe von Gl. (5.22) rekonstruieren.

2. Das zweite Kriterium für mereologische Objektivität ist *strukturell*, d.h. Objekte sind immer nur in ganz bestimmten Teil-Ganzes-Beziehung eingebettet. Gerade im Formalismus der projektiven Mereologie mit objekt-freien Aussagen, liegt es nun nahe, solche strukturellen Kriterien näher zu untersuchen. In der KEM werden strukturelle Einschränkungen zum Beispiel durch mereologische Supplemente ausgedrückt. Als schwächstes, quasi-analytisches Kriterium haben wir in Abschnitt 4.2 das schwache Supplement (M3) kennen gelernt:

$$x \ll y \rightarrow \exists z (z \ll y \wedge z \wr x), \quad (5.23)$$

wenn  $x$  ein echter Teil von  $y$  ist, dann existiert auch ein zweiter echter Teil  $z$ , der disjunkt zu  $x$  ist. Dies lässt sich nun wie folgt in den Projektorformalismus übertragen (vgl. Gl. 5.4):

$$\exists \hat{P} (x = \hat{P}y) \rightarrow \exists z (z = \hat{P}'y \wedge \forall \hat{P}_j \hat{P}_k (\hat{P}_j z \neq \hat{P}_k x)). \quad (5.24)$$

(Dabei ist zu beachten, dass  $x$  und  $y$  in der ursprünglichen Formulierung des schwachen Supplements in M3 in einer echten Teil-Ganzes-Beziehung stehen, also verschieden sind, was in Gl. (5.24) jedoch nicht näher spezifiziert wird. Dies könnte aber leicht mithilfe des hinreichenden Kriteriums aus Gl. (5.22) berücksichtigt werden)

Einsetzen von  $x = \hat{P}y$  und  $z = \hat{P}'y$  in die rechte Seite von Gl. (5.24) transformiert die Gleichung in eine Aussage, über die Zulässigkeit von Projektionen eines generisches Objektes  $y$ :

$$\exists \hat{P} (x = \hat{P}y) \rightarrow \exists \hat{P}' \forall \hat{P}_j \hat{P}_k (z = \hat{P}'y \wedge (\hat{P}_j \hat{P}'y \neq \hat{P}_k \hat{P}y)). \quad (5.25)$$

(Aus  $\forall \hat{P}_j \hat{P}_k (\hat{P}_j \hat{P}'y \neq \hat{P}_k \hat{P}y)$  folgt mit  $\hat{P}_j = \hat{P}$  und  $\hat{P}_k = \hat{P}'$  zudem, dass  $\left[ \hat{P}, \hat{P}' \right] \neq 0$ , was automatisch die Nichtidentität von  $x$  und  $z$  impliziert.) Mit der Kurzschreibweise aus Gl. (5.5) führt dies letztlich auf folgende objekt-freie Form:

$$\exists \hat{P} \rightarrow \exists \hat{P}' \forall \hat{P}_j \hat{P}_k (\hat{P}_j \hat{P}' \neq \hat{P}_k \hat{P}). \quad (5.26)$$

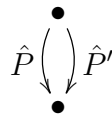


Das schwache Supplement wird somit zur Aussage über die mögliche Struktur eines Netzwerkes von Projektoren: Gegeben ein Projektor  $\hat{P}$ , dann existiert mindestens ein weiterer Projektor  $\hat{P}'$ , sodass alle darauf folgenden Projektionen zu unterschiedlichen Resultaten ( $\hat{P}_j \hat{P}' \neq \hat{P}_k \hat{P}$ ) führen. Auch hier erweist sich die Kontraposition als nützlich [7]:

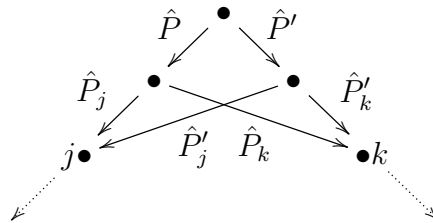
$$\exists \hat{P}_j \hat{P}_k \forall \hat{P}' (\hat{P}_j \hat{P}' = \hat{P}_k \hat{P}) \rightarrow \neg \exists \hat{P}, \tag{5.27}$$

d.h. wenn (mindestens) ein Paar,  $\hat{P}_j, \hat{P}_k$ , existiert, sodass für alle Projektoren,  $\hat{P}'$ , gilt, dass  $\hat{P}_j \hat{P}' = \hat{P}_k \hat{P}$ , dann kann  $\hat{P}$  keinen echten Teil erzeugen.

Nun lässt sich anhand der graphischen Darstellung als Netzwerk leicht einsehen, dass durch Gl. (5.26) eine Reihe von Strukturen ausgeschlossen sind, etwa das triviale Modell,



oder

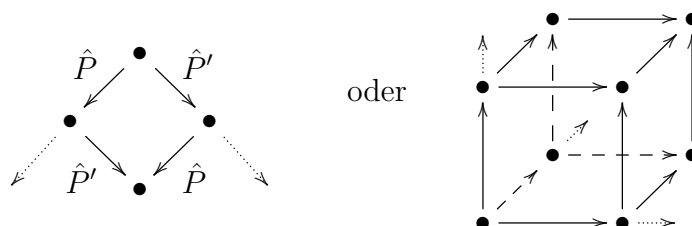


und „unendliche Ketten“, in denen jeder Knoten eines Netzwerks mithilfe verschiedener Kombination von Projektoren, ausgehend von unterschiedlichen Knoten, erreicht werden kann.

Betrachten wir noch einmal den Spezialfall  $\hat{P}_j = \hat{P}$  und  $\hat{P}_k = \hat{P}'$ . Gl. (5.27) lässt sich nun vereinfachen zu:

$$\forall \hat{P}' (\hat{P} \hat{P}' = \hat{P}' \hat{P}) \Leftrightarrow \forall P' ([\hat{P}, \hat{P}'] = 0) \rightarrow \neg \exists \hat{P}. \tag{5.28}$$

Falls also alle Paare von Projektoren eines Objekts kommutieren, erzeugen sie keine (echten) Teile, oder umgekehrt: Für jeden Projektor eines Objektes, existiert (mindestens) ein anderer nicht-kommutierender Projektor; einige Modelle, die auf diese Weise ausgeschlossen sind:



Im Prinzip können nun für alle weiteren Axiome, die in Abschnitt 4.2 vorgestellt wurden, objekt-freie Ausdrücke bestimmt werden, so etwa für das starke Supplement (M5): Falls  $x$  kein Teil von  $y$  ist, dann existiert ein zu  $y$  disjunktes  $z$ , welches Teil von  $x$  ist,

$$\neg(x < y) \rightarrow \exists z (z < x \wedge z \wr y). \quad (5.29)$$

Unter Hinzuziehen der Erzeugenden  $\hat{P}_x$  und  $\hat{P}_y$  lässt sich dies analog wie zuvor ausdrücken:

$$\exists \hat{P}_x \hat{P}_y \neg \exists \hat{P} (\hat{P}_x = \hat{P} \hat{P}_y) \rightarrow \exists \hat{P}' \forall \hat{P}_j \hat{P}_k (\hat{P}_j \hat{P}' \hat{P}_x \neq \hat{P}_k \hat{P}_y). \quad (5.30)$$

Wir wollen diesen Ausdruck mit den Ersetzungen  $\hat{P}_x = \hat{P}_1$  und  $\hat{P}_y = \hat{P}_2$  etwas umschreiben und davon ausgehen, dass die Projektoren  $\hat{P}_1$  und  $\hat{P}_2$  existieren und zwei Objekte aus einer Ganzheit erzeugen, sodass gilt:

$$\forall \hat{P} (\hat{P}_1 \neq \hat{P} \hat{P}_2) \rightarrow \exists \hat{P}' \forall \hat{P}_j \hat{P}_k (\hat{P}_j \hat{P}' \hat{P}_1 \neq \hat{P}_k \hat{P}_2). \quad (5.31)$$

Betrachten wir kurz den Spezialfall  $\hat{P}_j = \hat{P}'$ . Dies ist gerechtfertigt, da aus der rechten Seite von Gl. (5.31) wegen des Allquantors, folgt, dass

$$\forall \hat{P} (\hat{P}_1 \neq \hat{P} \hat{P}_2) \rightarrow \exists \hat{P}' \forall \hat{P}_j \hat{P}_k (\hat{P}_j \hat{P}' \hat{P}_1 \neq \hat{P}_k \hat{P}_2) \quad (5.32)$$

$$\rightarrow \exists \hat{P}' \forall \hat{P}_k (\hat{P}' \hat{P}' \hat{P}_1 \neq \hat{P}_k \hat{P}_2). \quad (5.33)$$

Die Kontraposition von Gl. (5.33) führt mit der Idempotenz,  $\hat{P}' \hat{P}' = \hat{P}'$ , auf:

$$\exists \hat{P}_k \forall \hat{P}' (\hat{P}' \hat{P}_1 = \hat{P}_k \hat{P}_2) \rightarrow \exists \hat{P} (\hat{P}_1 = \hat{P} \hat{P}_2), \quad (5.34)$$

d.h., in Worten, falls jede Projektion des ersten Objekts durch einer Projektion des zweiten Objekts dargestellt werden kann, dann lässt sich die Beziehung  $\hat{P}_1 = \hat{P} \hat{P}_2$  finden. Gilt nun auch der umgekehrte Fall, wonach jede Projektion aus dem zweiten Objekt durch mindestens eine Projektionen aus dem ersten erhalten werden kann, folgt in analoger Weise  $\hat{P}_2 = \hat{Q} \hat{P}_1$ . Wegen der Asymmetrie, muss jedoch gesamthaft gelten, dass

$$\hat{P} = \hat{Q} = 1, \quad (5.35)$$

und somit:

$$\exists \hat{P}_k \hat{P}_l \forall \hat{P}' ((\hat{P}' \hat{P}_1 = \hat{P}_k \hat{P}_2) \wedge (\hat{P}' \hat{P}_2 = \hat{P}_l \hat{P}_1)) \rightarrow \hat{P}_1 = \hat{P}_2. \quad (5.36)$$

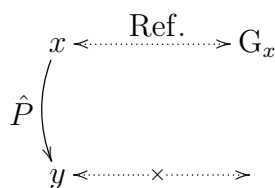
Dies entspricht der Extensionalitätsforderung auf Ebene von Projektoren.

3. Das dritte und letzte Kriterium für mereologische Objektivität bezieht sich auf die Interpretation der Mereologie als Verweissystem und betrifft das *Vorhandensein eines Verweises* auf einen (konkreten) Gegenstand.

Manchmal wird in diesem Zusammenhang von einer prinzipiellen Diesheit oder *haecceitas* gesprochen, die referenzierte Gegenstände (und entsprechend die referierenden Objekte) als ausweisbare Individuen kennzeichnet. Gerade wenn „das Subjekt“ als der ausgezeichnete Bezugspunkt des Bewusstseinsaktes verobjektiviert wird, scheint die Redeweise einer prinzipiellen Diesheit zumindest implizit mitzuschwingen.

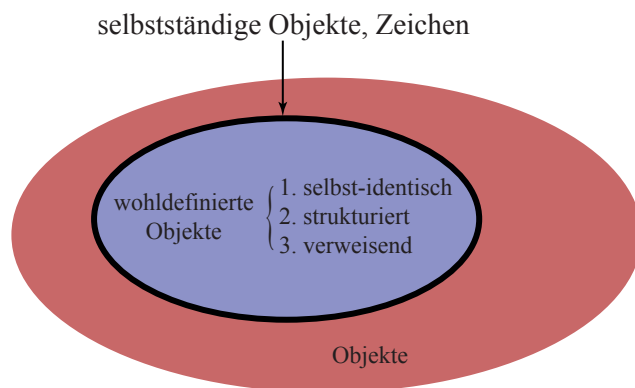
Eine andere wichtige Unterscheidung geht auf Husserl (LU, Bd. II: §3.1) zurück, der auch von „selbstständigen und unselbstständigen Teilen“ spricht (siehe auch Simons 1987, Kap. 8 und 9): Während die Existenz selbstständiger Teile nicht abhängig von der Existenz anderer Objekte ist, gilt dies gerade nicht für unselbstständige Teile, die in ihrer Existenz auf die Ganzheit angewiesen sind, in der sie vorkommen. Umgelegt auf das vorgestellte Kalkül soll dies bedeuten, dass unselbstständige Teile auf keinen konkreten Gegenstand verweisen, aber nur innerhalb eines (übergeordneten) Verweissystems (d.h., als Teile eines Verweises) existieren, was noch weiter operationalisiert werden kann. So wird z.B. später in Abschnitt 6.2 noch gezeigt werden, dass die Unterscheidung zwischen selbstständigen Objekten und unselbstständigen Teilen eine Reihe an neuartigen Überlegungen bei der Beschreibung von emergenten Phänomenen innerhalb von Wahrnehmungsprozessen, wo (selbstständige) „Zeichen“ ihren (unselbstständigen) Teilen gegenüberstehen, ermöglicht.

In Abb. 5.2 ist das grundlegende Schema dargestellt: Während das selbstständige Objekt  $x$  auf einen konkreten Gegenstand verweist, ist dies nicht der Fall für den unselbstständigen Teil  $y$ . Insbesondere muss nicht gelten, dass  $y$  auf einen Gegenstand  $G_y$  verweist, der in einer Beziehung zum Gegenstand  $G_x$  steht, die isomorph zu  $\hat{P}$  wäre. Motiviert ist dies durch ein Beispiel, das uns in Abschnitt 6.2 wieder begegnen wird: Es bezeichne  $x$  ein wahrgenommenes Objekt und  $y$  dessen Farbe;  $y$  wird im Allgemeinen *nicht* auf einen (unabhängigen) Gegenstand verweisen (etwa auf „die Farbe“); oder es bezeichne  $x$  ein wahrgenommenen Ton und  $y$  dessen Lautstärke;  $y$  wird im Allgemeinen *nicht* auf einen (unabhängigen) Gegenstand namens „die Lautstärke“ verweisen. Anders gesagt, sind die wahrgenommenen Qualitäten von Objekten solche unselbstständigen Teile.



**Abbildung 5.2:** Verweisende Objekte  $x$  und unselbstständige Teile  $y = \hat{P}x$ , die selbst auf keinen Gegenstand verweisen.

Zusammengefasst: *Wohldefinierte* Objekte („Zeichen“) sind *selbst-identisch*, genügen *strukturellen* (mereologischen) Anforderungen und *verweisen* auf Gegenstände (Abb. 5.3). Der Hauptunterschied zwischen einem kompositionalen und zerlegenden Ansatz ist darin auszumachen, dass im ersteren von solchen Objekten ausgegangen wird, während zerlegende Darstellungen daran interessiert sind, wie Objekte (als Resultate von Operationen) überhaupt erst erzeugt werden.



**Abbildung 5.3:** Wohldefinierte mereologische Objekte („Zeichen“) als Unterklasse aller Objekte.

Zuletzt sei an dieser Stelle auch noch etwas zur Existenz mereologischer Summen gesagt, deren Behandlung kompositionale und zerlegende Ansätze verbindet:

Erstens gelte, dass mereologische Summen, so sie existieren, wohldefiniert sind und daher auf (konkrete) Gegenstände verweisen und in Strukturen entsprechend den Summationsprinzipien (z.B. M9, FM2, M14 oder MT2) eingebunden sind. Dies macht den Hauptunterschied zwischen mereologischen Summen und bloßen Klassen oder Mengen von Objekten aus. Zweitens gelte, da Summen  $\sigma$  als kompositionale Objekte aus der Menge ihrer Teile, z.B.  $\{x, y, z\}$ , gebildet werden, dass deren Teile entweder selbst wohl-definierte Objekte darstellen (Abb. 5.4, links) oder unselbstständige Teile relativ zu einer Ganzheit  $\zeta$  bezeichnen. Ist jedoch letzteres der Fall, so existiert eine

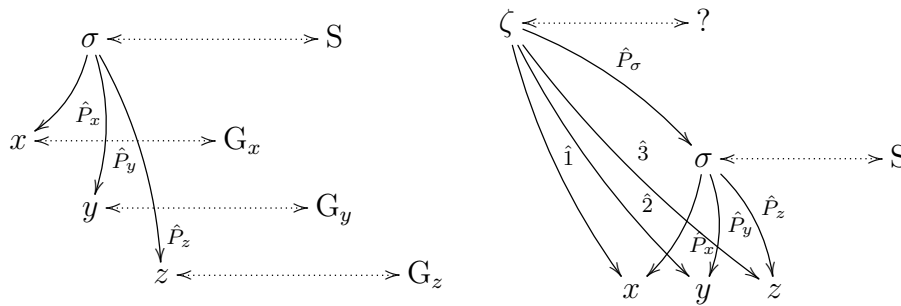
Projektion  $\hat{P}_\sigma$ , welche die Summe direkt als Teil dieser Ganzheit erzeugt [8] (Abb. 5.4, rechts), wobei dann wegen der Transitivität (Gl. (5.11)) zu gelten hat, dass

$$[\hat{P}_\sigma, \hat{P}_j] = 0 \quad \forall j \in \{x, y, z\} \tag{5.37}$$

$$\text{mit } j = \hat{P}_j \sigma \text{ und } \sigma = \sum j. \tag{5.38}$$

(Im einfachsten Fall entspricht die Ganzheit  $\zeta$  der Summe  $\sigma$ , die „Projektion“  $\hat{P}_\sigma$  ist dann einfach die Identität. Dies gilt jedoch nicht mehr, wenn weitere  $\hat{P}'$  zu  $\zeta$  existieren, mit  $\forall \hat{P}_j \hat{P}_k (\hat{P}_j \hat{P}' \neq \hat{P}_k \hat{P}_\sigma)$ .) Die Summe  $\sigma$  ist bezogen auf  $\zeta$  ein selbstständiger Teil, bezogen auf die Teile  $x, y, z$  eine wohldefinierte Ganzheit.

Ein so spezifiziertes System kann dann in Analogie zur klassischen Mereologie als „projektive extensionale Mereologie“ (PEM) bezeichnet werden.



**Abbildung 5.4:** Mereologische Summe  $\sigma = x+y+z$  verweist auf Gegenstand S. Links:  $\sigma$  als Summe selbstständiger Objekte, die ebenfalls auf Gegenstände verweisen; rechts:  $\sigma$  als Summe unselbstständiger Teile,  $\sigma$  kann dann immer als Teil einer übergeordneten Ganzheit  $\zeta$  geschrieben werden.

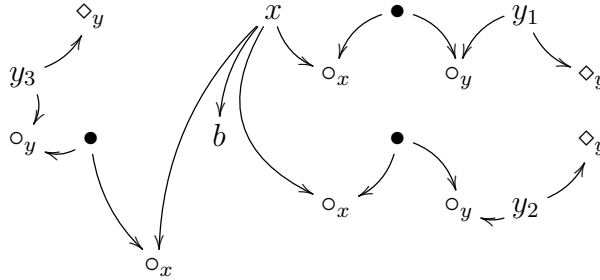
### 5.3 Emergenz und Topologie

Der in Abschnitt 2.3 eingeführte Begriff der Emergenz soll nun noch einmal betrachtet werden. Mit dem Slogan „Das Ganze ist mehr als die Summe seiner Teile“ wird ein intuitives Verständnis von Emergenz ausgedrückt. Eine philosophische Reflexion darüber bedingt jedoch, dass man sich über den Begriff des Teils im Klaren ist. Dies wurde z.B. schon von Hempel & Oppenheim (1948, S. 149) erkannt, welche einer starken Form des Emergentismus kritisch eingestellt waren:

Before we can significantly ask whether a characteristic  $W$  of an object  $w$  is emergent, we shall [...] state the intended meaning of the term “part of”. This can be done by defining a specific relation  $\ll$  and stipulating

that those and only those objects which stand in  $\ll$  to  $w$  count as parts or constituents of  $w$ .

Genauso sehr muss man sich jedoch im Klaren werden, wie wir überhaupt Objekte als solche erkennen können. Zur Explikation dieser Frage kann auf die mereotopologischen Axiome aus Abschnitt 4.3 und die PEM zurückgegriffen werden. Dabei wird sich zeigen, dass Emergenz wesentlich als kontextuelles Phänomen zu verstehen ist. In Abb. 5.5 ist ein generisches System gezeigt, in dem die Objekte  $x$  und  $\{y_i\}$  berühren, aber nicht überlappen (vgl. die Diskussion zu Gl. (4.36)). Die  $\{\bullet\}$  bezeichnen dabei die „Grenzen“ zwischen je zwei Objekten  $x$  und  $y_i$ , also diejenigen Objekte, welche mit ihnen überlappen. Zudem bezeichnen  $\{\circ_i\}$  die Teile der Grenzen  $\{\bullet\}$  gemäß Gl. (4.48) aus Abschnitt 4.3.



**Abbildung 5.5:** Ein Netzwerk mereologischer Beziehungen. Projektion (Teilsein) wird dargestellt als gerichteter Pfeil. Das Objekt  $x$  wird durch die Grenzobjekte  $\{\bullet\}$  von den Objekten  $\{y_i\}$  getrennt; verschiedene Summen können auf diesem Netzwerk postuliert werden (Details siehe Text).

Unter Berücksichtigung der Extensionalität können ausgehend von diesen Objekten eine Reihe von Summen postuliert werden, etwa die „Ganzheit“  $\zeta$ ,

$$\zeta = x + \{\bullet\}_+ + y_1 + y_2 + y_3 = b + \{\circ_x\}_+ + \{\circ_y\}_+ + \{\diamond_y\}_+, \quad (5.39)$$

bestehend aus dem Objekt  $x$ :

$$x = b + \{\circ_x\}_+, \quad (5.40)$$

dessen „Hülle“  $bl(x)$ ,

$$bl(x) = \{\bullet\}_+ = \{\circ_x\}_+ + \{\circ_y\}_+, \quad (5.41)$$

und dessen „Umgebung“  $x^c$ ,

$$x^c = y_1 + y_2 + y_3 = \{\circ_y\}_+ + \{\diamond_y\}_+. \quad (5.42)$$

(Dabei steht  $\{\triangleleft\}_+$  für die „Zwischensumme“, der nicht näher indizierten Objekte  $\{\triangleleft\}$ .) Schließlich kann der „Abschluss“  $cl(x)$ ,

$$cl(x) = x + \{\bullet\}_+ = b + \{\circ_x\}_+ + \{\circ_y\}_+, \quad (5.43)$$

definiert werden, wobei jeweils angenommen wird, dass ein mereologisches Objekt identisch zur Summe aller seiner Teile ist. Ferner gilt auch, dass

$$\zeta = x + \underbrace{bl(x)}_{cl(x)} + x^c, \quad (5.44)$$

also insbesondere, dass der Abschluss gleich der Summe eines Objekts mit seiner Hülle ist:

$$cl(x) = x + bl(x). \quad (5.45)$$

Bei genauerer Betrachtung zeigt sich jedoch folgender Unterschied zwischen den eben postulierten Summen: Gemäß MT2 seien nur diejenigen Objekte (mereologisch) summierbar, die sich untereinander berühren. Auch wenn nicht alle Objekte paarweise berühren, wie etwa  $y_1$  und  $y_2$ , so erlaubt MT2 dennoch eine Summierung über eine „Zwischensumme“:  $y_1$  berührt  $x$  und kann somit zu  $y_1 + x$  summiert werden; nun berühren jedoch auch  $y_2$  und  $y_1 + x$  und können somit summiert werden, usw.

Auf diese Weise lässt sich einsehen, dass die Summen  $\zeta$  und  $cl(x)$  laut MT2 existieren. Gleiches gilt jedoch nicht für die Objekte  $bl(x)$  und  $x^c$  [9].

Daraus folgt aber auch, dass die Objekte  $bl(x)$  und  $x^c$  *keine* selbstständigen Teile von  $\zeta$  im Sinne der PEM darstellen. Für den selbstständigen Teil  $cl(x)$  gilt hingegen, dass:

$$\exists \hat{P}_{cl} (\hat{P}_{cl}\zeta = cl(x)), \quad (5.46)$$

wobei  $\hat{P}_{cl}$  die Kuratowski-Axiome erfüllt, d.h. dass  $\hat{P}_{cl}$  eine Topologie auf  $\zeta$  induziert und, bildlich gesprochen, die Ganzheit  $\zeta$  so formt, dass aus  $\zeta$  ein abgegrenztes Objekt  $x$  entsteht. Die Übersetzung der Kuratowski-Axiomen in den Formalismus der PEM laute dabei wie folgt:

0.  $0 \in cl(x) \Leftrightarrow 0 < \hat{P}_{cl}\zeta$ . Dies kann dadurch garantiert werden, dass per Definition ein „Leerverweis“ als Null-Element immer Teil eines jeden mereotopologischen Objektes ist.
1.  $x \in cl(x) \Leftrightarrow \hat{P}_x\zeta < \hat{P}_{cl}\zeta$ ; vgl. Gl. (5.43) und (5.46).
2.  $cl(cl(x)) = cl(x) \Leftrightarrow \hat{P}_{cl}^2\zeta = \hat{P}_{cl}\zeta$ . Wegen der Idempotenz führt wiederholtes Projizieren zu keiner Änderung in den projizierten Objekten [10].

3.  $cl(x_1 + x_2) = cl(x_1) + cl(x_2) \Leftrightarrow \hat{P}_{cl}(\zeta_1 + \zeta_2) = \hat{P}_{cl}\zeta_1 + \hat{P}_{cl}\zeta_2$ . Es sei nun  $\zeta_1 + \zeta_2$  als Summe zweier Projektionen einer übergeordneten Ganzheit,  $\zeta_0$ , darstellbar:

$$\zeta_1 + \zeta_2 = \hat{P}_1\zeta_0 + \hat{P}_2\zeta_0. \quad (5.47)$$

Einsetzen in obige Bedingung:

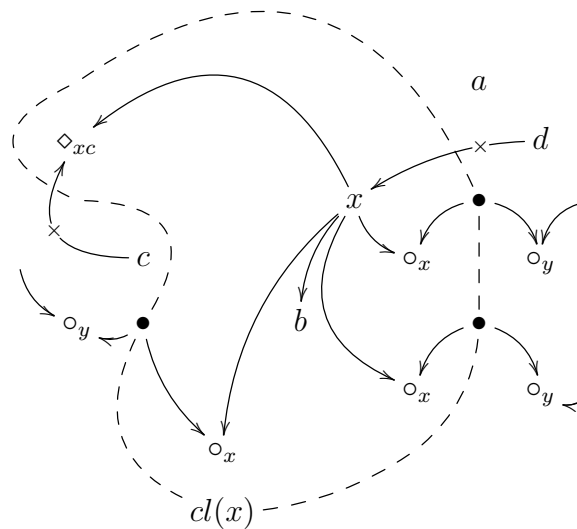
$$\hat{P}_{cl}(\hat{P}_1\zeta_0 + \hat{P}_2\zeta_0) \stackrel{?}{=} \hat{P}_{cl}\hat{P}_1\zeta_0 + \hat{P}_{cl}\hat{P}_2\zeta_0, \quad (5.48)$$

was dann der Fall wäre, wenn die Zerlegung  $\{\hat{P}_1, \hat{P}_2\}$  die Ganzheit  $\zeta_0$  in nicht-berührende Teile zerlegt (was gleichbedeutend dazu ist, dass  $\zeta_1 + \zeta_2$  kein wohl-definiertes Objekt der PEM bezeichnen kann).

Der letzte Punkt wird z.B. dadurch garantiert, dass es keine Zerlegung einer Ganzheit gibt, welche diese in überlappende Teile zerlegt, von denen sich einer *außerhalb* der Hülle des anderen befindet (vgl. Abb. 5.6) oder in anderen Worten: dadurch, dass  $x$  vollständig durch die Hülle  $bl(x)$  von seiner „Außenwelt“ abgeschirmt wird. Zerlegungen, die den oben genannten Bedingungen genügen, implizieren eine neuartige topologische Struktur, was gemäß der Theorie kontextueller Emergenz Grundvoraussetzung zur Beobachtung emergenter Phänomene ist. In diesem Sinne sind daher Änderungen in der mereo(topo-)logischen Struktur Bedingungen für das Hervortreten von Objekten [11]. Den Objekten entsprechen dann aber, metaphysisch gesprochen, keine „Gegenstände für-sich“, sondern sie sind immer abhängig von der Struktur des (mereologischen) Netzwerkes, in welches sie in ihrer Funktion als (potentielle) Verweise auf konkrete Gegenstände eingebunden sind.

Das Objekt  $w$ , von welchem im Zitat von Hempel & Oppenheim weiter oben die Rede war, ist wesentlich durch dessen strukturelle Abgeschlossenheit gegenüber seiner „Außenwelt“, also durch seine Position innerhalb eines Systems von Verweisen, charakterisiert. Neben der von Hempel & Oppenheim geforderten Notwendigkeit, die Bedeutung von „Teil-von- $w$ -zu-sein“ zu explizieren, muss daher ebenfalls der Kontext, der überhaupt auf das Objekt  $w$  führt, näher spezifiziert werden. In einem ersten Schritt können wir mithilfe der PEM versuchen, die Zerlegungsweisen, die einer solchen Emergenz von Objekten zugrundeliegen, formal darzustellen.





**Abbildung 5.6:** Topologie und projektive Mereologie. Teilsein wird wiederum dargestellt als gerichteter Pfeil; die Objekte  $\{•\}$  grenzen  $x$  von seiner Umgebung ab. Die Objekte  $a, b, c$  und  $d$  können als „Test-Objekte“ betrachtet werden (Details siehe Text).

Eine Veranschaulichung, die bereits in Abschnitt 3.1 diskutiert wurde, liefern zelluläre Systeme. Den topologischen Kontext bildet dort die Partitionierung des Phasenraums in informationell getrennte Bereiche (also solche, die durch eine „Markov-Hülle“ von ihrer Außenwelt abgeschirmt wird), was in ergoidschen Systemen auf ein (emergentes) Verhalten führt (vgl. Friston 2013).

Nun impliziert aber sowohl die formale Darstellung der PEM als auch die konkrete Veranschaulichung mithilfe zellulärer Systeme, dass es sich bei den als emergent bezeichneten Objekten weniger um die *relata* statischer Strukturen handelt, sondern wesentlich um die Ergebnisse von (koordinierten) Prozessen: Einmal ist dies abstrakt dargestellt durch Projektionen, das andere mal konkret anhand der physikalisch-chemischen Reaktionen, die in biologischen Organismen stattfinden und zur Realisierung der Markov-Abschirmung beitragen. Dann sind jene Prozesse aber immer auch als eingebunden in (natürliche) Aktivitäten oder, anders gesagt, als *kontextuell* zu betrachten.

## Anmerkungen zu Abschnitt 5

[1] Die Wahl des Projektformalismus ist primär dessen Nützlichkeit geschuldet. Man könnte diese Beziehung zu einer allgemeinen semiotischen Relation verallgemeinern (siehe etwa Burgin 2010, Kap. 2).

[2] Formal korrekt wäre es, an dieser Stelle eines Existenzprädikat für Operatoren zu definieren:

$$\exists! \hat{P} \Leftrightarrow_{\text{Df.}} \exists \hat{Q} (\hat{Q} = \hat{P}), \quad (5.49)$$

mit

$$\forall \hat{P} (\exists! \hat{P} \Leftrightarrow \exists \zeta z (z = \hat{P}\zeta)). \quad (5.50)$$

Dies besagt, dass ein Projektor genau dann existiert, wenn es aus einem Ganzen einen Teil erzeugt. Die Formeln müssen dann entsprechend angepasst werden.

[3] Beachte, dass daraus nicht zwangsweise folgt, dass

$$\hat{P}(x_1 + x_2) = \hat{P}x_1 + \hat{P}x_2, \quad (5.51)$$

als Distributivität bezüglich der Summation der Teile.

[4]

$$x = \hat{P}\zeta = \hat{P}^2\zeta = \hat{P}(\hat{P}\zeta) = \hat{P}x. \quad (5.52)$$

[5] Wann sind Produkte der Form  $\hat{P}_1\hat{P}_2$  idempotent? Genau dann, wenn  $[\hat{P}_1, \hat{P}_2] = 0$ :

$$(\hat{P}_1\hat{P}_2)^2\zeta = \hat{P}_1\hat{P}_2\hat{P}_1\hat{P}_2\zeta = \hat{P}_1(\underbrace{\hat{P}_1\hat{P}_2 - [\hat{P}_1, \hat{P}_2]}_0)\hat{P}_2\zeta = \hat{P}_1^2\hat{P}_2^2\zeta = \hat{P}_1\hat{P}_2\zeta. \quad (5.53)$$

[6] Mit Bezug auf die Quantenmechanik wird z.B. manchmal davon gesprochen, das Prinzip der Identität des Ununterscheidbaren sei verletzt: Paulis Ausschlussprinzip z.B. besagt, dass der Zustand von  $n$  Elektronen durch eine antisymmetrische Wellenfunktion  $\Psi$  darzustellen ist, beispielsweise in der Ortsbasis:  $\{(x_1, x_2, \dots, x_n)\}$ . Dies bedeutet, dass beim Vertauschen der Positionen zweier Teilchen die Gesamtwellenfunktion des Systems bis auf einen negativen Vorfaktor unverändert bleibt (beliebige Vertauschungen mehrerer Teilchen können als Produkte von Zwei-Teilchen-Vertauschungen geschrieben werden, daher gilt für diese das Folgende in analoger Weise):

$$\Psi(x_1, \dots, x_j, x_k, \dots, x_n) = -\Psi(x_1, \dots, x_k, x_j, \dots, x_n). \quad (5.54)$$

Da aber die messbaren Eigenschaften als Wahrscheinlichkeiten vom Betragsquadrat von  $\Psi$  abhängen, kann keine Vertauschung zweier Teilchen einen beobachtbaren Unterschied ausmachen. Falls, wie im Fall oben, alle Teilchen dieselben „intrinsischen“ Eigenschaften

haben (z.B. Masse, Ladung, Spin) und eine Vertauschung keinen beobachtbaren Unterschied ausmacht, lässt sich davon sprechen, dass zwei Teilchen derselben Sorte *ununterscheidbar* sind. Gemäß *pii* müssten die Teilchen nun als identisch angesehen werden.

Eine mögliche Deutung behauptet, dass die Quantenmechanik ein Beispiel für Situationen liefere, in denen das *pii* verletzt sei. Schließlich liege hier ein Fall vor, wo mehreren Teilchen in alle (messbaren) Eigenschaften übereinstimmen (vgl. French & Redhead 1988). Zu sagen, dass es sich hier um einen Zustand aus mehreren Teilchen handelt, setzt aber selbst wiederum eine Art *haecceitas* als Eigenschaft der Dinge oder der Substanz voraus, deren Annahme aber vielleicht nur eine Folge einer „*fallacy of the misplaced concreteness*“ ist (Whitehead, PR; für die metaphysischen Implikationen vgl. (Seibt 1996); für ein Diskussion in Hinblick auf die Quantenmechanik siehe (Esfeld 2001, S. 260ff.)). Eine alternative Deutung würde besagen, dass gar nicht mehr von einzelnen „Teilchen“ gesprochen werden könne, sondern nur noch von (symmetrisierten) Mehr-Teilchen-Zuständen (oder Anregungen von Quantenfeldern), wobei dann der Fall der Verschränkung noch genauer zu behandeln wäre. Am *pii* könnte dann als kognitive (oder transzendente) Norm festgehalten werden.

Das oben Ausgeführte gilt nicht nur für den Fall der Elektronen sondern kann gemäß des Spin-Statistik-Theorems der relativistischen Quantenmechanik (Fierz 1939) auf beliebige Elementarteilchen erweitert werden.

[7] Unter Verwendung der Identität:

$$\neg\exists P\forall Q (f(P) \neq f(Q)) \Leftrightarrow \exists Q\forall P (f(P) = f(Q)). \quad (5.55)$$

[8] Dass die Summe direkt „ohne Umweg über die Teile“ aus der Ganzheit mittels Projektion erzeugt wird, kann z.B. mit Verweis auf die Extensionalität gezeigt werden: Angenommen, es bezeichnen die  $\{x_i\}$  die unselbstständigen Teile der Summe  $\sigma$ , so dass  $\sigma$  über keine weiteren Teile verfügt. Aus deren Unselbstständigkeit folgt, dass  $x_i = \hat{P}_{x_i}\zeta$  wobei  $\zeta$  die entsprechende Ganzheit zu den  $x_i$  darstellt. Nun gilt gemäß der Extensionalität, dass  $\zeta = \sigma$  genau dann, wenn  $\zeta$  und  $\sigma$  aus denselben Teilen bestehen. Andernfalls muss  $\zeta$  über weitere Teile disjunkt zu  $\sigma$  verfügen. Da aber alle  $x_i$  (per Annahme) auch Teile von  $\sigma$  sind, gilt immer  $\sigma = \hat{P}_\sigma\zeta$ . (Ohne Verweis auf die Extensionalität wären Ganzheiten  $\zeta$  vorstellbar, die nicht in einer Ganzes-Teil-Beziehung zu  $\sigma$  stehen.)

[9] Alternativ ließe sich dies auch anhand der allgemeinen Definitionen der Umgebung und der Hülle einsehen: Die Umgebung von  $x$  ist wie folgt definiert:

$$x^c = \sum z (C(x, z) \wedge x \wr z), \quad (5.56)$$

also als Summe derjenigen Objekte, die  $x$  berühren, aber nicht mit  $x$  überlappen. Es gibt jedoch möglicherweise  $z$ , welche zwar die rechte Seite von Gl. (5.56) erfüllen, aber gar kein anderes  $z'$  berühren. Dann können  $z$  und  $z'$  gemäß MT2 nicht summiert werden und das Objekt  $x^c$  existiert nicht. Analog gilt dies auch für die Hülle  $bl(x)$ , welche als Summe aller begrenzenden Objekte von  $x$  definiert werden kann.

[10] Dass dies eben nicht für die Objekte  $bl(x)$  und  $x^c$  gilt, lässt sich wie folgt einsehen: Angenommen, ein solcher Projektor würde existieren. Dann ließe sich jedoch zeigen, dass dieser die Idempotenzeigenschaft nicht erfüllen würde, denn:

$$bl(bl(x)) = 0 \neq bl(x) \rightarrow \hat{P}_{bl}^2 \zeta \neq \hat{P}_{bl} \zeta. \quad (5.57)$$

und

$$(x^c)^c = x \neq x^c \rightarrow \hat{P}_c^2 \zeta \neq \hat{P}_c \zeta. \quad (5.58)$$

Da die Idempotenzeigenschaft aber eine fundamentale Eigenschaft in der PEM darstellt, wäre dies im Widerspruch zur ursprünglichen Annahme.

[11] Selbstverständlich gilt dies nicht für alle möglichen Topologien, sondern nur für solche, die i) in Zusammenhang zu weiteren, notwendigen Bedingungen stehen (in Analogie zur statistisch-mechanischen Beschreibung mikroskopischer Systeme) und ii) bestimmten Robustheitskriterien entsprechen, wobei darüber im Formalismus der projektiven Mereologie bisher noch nichts gesagt wurde. Prinzipiell ist es jedoch denkbar, dass solche Robustheitskriterien auch auf der Ebene mereologischer Beschreibung implementiert werden können.

## 6 Wahrnehmung im Kontext

*I believe that 'consciousness' [...] is the name of a nonentity.* (William James: Does 'Consciousness' Exist?)

### 6.1 Ding und Prozess

#### Das personale Selbst als Muster von Beziehungen

Dieser Abschnitt untersucht die Frage, wie sich für substantiell gehaltene Dinge, d.h. Gegenstände, die wir für greifbare, selbstständige und persistente Entitäten halten, als (relativ) stabile Muster oder invariante Strukturen einer darunter liegenden bezeichnenden Aktivität verstehen lassen.

Ähnlich ließe sich auch Fechners „phänomenalistische Konzeption von Substanz“ (Heidelberger 2002, S. 50) deuten, wonach der Substanzbegriff keinen Grundbegriff bilde, der für dasjenige steht, was selbstständig existieren würde und Träger von Eigenschaften wäre, sondern, dass „Substanz“ nichts als ein (nützlicher) Hilfsbegriff sei, der sich letztlich in den gesetzmäßigen Verbindungen von Einzelercheinungen auflösen lässt [1]. Fechner entwickelte diesen Gedanken wesentlich im Rahmen des von ihm mitbegründeten Feldes der Psychophysik, also des Studiums der Beziehung von psychischen und physikalischen Phänomenen und fand eine umfassende (direkte oder indirekte) Rezipierung – so etwa im Konstitutionsprinzip bei Rudolf Carnap oder in der analytischen Philosophie des Geistes von Herbert Feigl, aber auch bei den zeitgenössischen Vertretern einer nicht-reduktiven Psychologie (z.B. Velmans 2009, Price & Barrell 2012).

Folgt man nun der Fechner'schen Ansicht, ist es suggestiv, auch das Bewusstsein verobjektivierend als ganz spezielle Ordnung solcher Einzelereignisse zu betrachten, wenn es denn schon nicht als substantiell aufgefasst wird. Wir werden im Folgenden noch einen Schritt weiter gehen und *Bewusstsein* nicht so sehr mit einem Muster identifizieren, als vielmehr nur noch mit der *Form* des Prozesses der Mustererkennung. Es bezeichnet dann eine „Nicht-Entität“ im Sinne von William James.

Doch zuerst wollen wir den Gedanken, wonach sich bestimmte psychische Dinge letztlich in Strukturen auflösen lassen, noch etwas näher betrachten. Dabei könnte zum Beispiel vermutet werden, das „personale Selbst“ sei nichts als ein ausgezeichnetes Muster von Beziehungen.

Von Thomas Metzinger (2003a) wurde der Vorschlag gemacht, das Selbst als (episodisch – also etwa während des Wachzustandes – aktive) Repräsentation derjenigen

Struktur zu deuten, welche die Repräsentationen von Erfahrungsgegenständen, also etwa unsere Wahrnehmungen der Umwelt und des Körpers, aber auch Prozesse der Entscheidungsfindung, in einer systematischen Weise verbindet. Eine solche *Meta*-repräsentation sei nun genau das, was im Gehirn vor sich gehe, weswegen auch davon gesprochen werden könne, das Selbst würde vom Gehirn „erzeugt“. Auf diese Weise könnte auch versucht werden, zu verstehen, wie ein komplexer Organismus nach Zielen und Bedürfnissen „handelt“ oder Gründe zur nachträglichen Plausibilisierung von solchen Handlungen hervorbringt.

Diese Reduktion des Selbst scheint vielen höchst anti-intuitiv zu sein, was aber für sich genommen noch kein Grund ist, der gegen diese Theorie spricht. Was aus der Sicht einer substanzkritischen Auffassung bemängelt werden könnte, ist eher deren fehlende Radikalität: Nicht nur das Selbst sei erzeugt, auch die Objekte, die wir als physische bezeichnen – Sterne, Tische oder Zellen –, seien immer nur die Resultate von Prozessen oder relativ zu diesen *emergent*.

Um seine Theorie des „Nicht-Selbst“ zu plausibilisieren, liefert Metzinger (2011) einige Gründe, die dafür sprechen sollen, das Selbst nicht für einen objektiv existierenden Gegenstand zu halten, womit argumentativ Raum für seine Theorie geschaffen wäre. Metzingers Argumente gegen die (objektive) Existenz eines Selbst können jedoch ohne allzu große Abänderungen auch als Argumente gegen die (objektive) Existenz neuronaler Systeme, des Gehirns oder gar physikalischer Gegenstände *per se* übernommen werden [2]. Die eigentlichen Gründe, die letztlich darüber entscheiden, ob wir von „dem“ Selbst oder doch lieber von einem bestimmten biologischen Objekt („dem“ Gehirn) sprechen, sind jedoch gar nicht in einer solchen Argumentation zu finden. Ein Skeptiker könnte daher einwenden, dass keine logische Notwendigkeit bestehe, eine bestimmte Redeweise zu übernehmen; man wird höchstens zu ihr überredet oder gewöhnt sich an diese. Während nun aber Metzinger an der dinglichen Existenz eines personalen Selbst zweifelt, scheint er dennoch die gängigste Form des Physikalismus zu übernehmen.

Ganz anders Galen Strawson (2009), der in seiner „minimalen Theorie des Selbst“ noch einen Schritt weiter geht. Zwar sei es richtig, dass die Vorstellung eines gegenständlichen personalen Selbst zurückzuweisen sei, allerdings gilt, dass zu jeder Erfahrung ein irreduzibler (a-personaler) subjektiver Kern gehöre, der weder einer komplexen empirischen Struktur entspreche – etwa einem „Selbstmodell“ im Sinne Metzingers –, noch über eine Form von Ich-Bewusstsein verfügen würde. Ganz im Gegensatz zur Vorstellung, dass sich psychische Dinge *restlos* auflösen lassen, folgert Strawson, dass eine solches Minimalsubjekt („*thin subject*“; Strawson 2009, S. 323) notwendigerweise zu jedem Erfahrungsprozess hinzugedacht werden müsse. Ein solches Subjekt bezeichnet ein einheitliches Ding (ein „*sesmet*“; Strawson 2009, S.

204), was gleichzeitig zentral für das Verständnis des Strawson'schen Panpsychismus ist, der davon ausgeht, dass jegliche Form von Materie (auch) über psychische Eigenschaften verfügt. Wohingegen Metzinger das personale Selbst als Ergebnis rein physikalischer Beziehungen betrachtet, gilt für Strawson, dass sich das personale Selbst ebenfalls als Resultat verstehen lässt, allerdings als Resultat eines Prozesses, der sich auf ein oder mehrere Minimalsubjekte gründet.

Beide Autoren berufen sich stellenweise auf die Buddhistische Lehre [3] der Nicht-Existenz eines individuellen, dauerhaften Selbst oder einer Seele (Lehre des „anātman“), wobei dort die Einsicht in die Nicht-Existenz des Selbst einen wesentlichen Schritt auf dem Weg zur Erleuchtung darstellt. Dies steht im Gegensatz zur Hinduistischen Advaita Vedanta, die gerade die Einsicht in die All-Gegenwart des Selbst (Einheit von „ātman“ und „brahman“) als großes spirituelles Ziel ansieht. Beide Lehren zielen auf die Erlösung des Menschen, unterscheiden sich jedoch grundlegend in der Art und Weise, wie diese Form der Erlösung zu erlangen sei. Ein Versuch, zwischen beiden Auffassungen zu vermitteln, besteht nun darin zu argumentieren, dass hier wesentlich zwischen verschiedenen Aspekten, Modi oder Kontexten unterschieden werden müsse (ähnlich der „Syādvāda“- Lehre des Jainismus) – gewissermaßen sind beide Lehren korrekt, je nachdem unter welchem Aspekt wir sie betrachten.

Solchen Versöhnungsstrategien haftet natürlich immer der Makel einer gewissen vereinnahmenden Beliebigkeit an. Die Moral, die sich dennoch daraus ziehen lässt, ist, dass vielleicht nur ein Mangel an Unterscheidungsfähigkeit der Grund dafür ist, dass in den Beschreibungen psychischer Phänomene die Tendenz vorherrscht, entweder alles vorschnell zu substantivieren oder, umgekehrt, alles auf die Dynamik materieller Teilchen reduzieren zu wollen – im ersten Fall weil man glaubt, diese Phänomene selbst entsprächen wirklich existierenden Gegenständen, im zweiten Fall, weil man glaubt, hinter diesen „konstruierten“ (von wem und woraus?) Phänomenen stünden nur real existierende materielle Teilchen.

Die Frage danach, welchen Gegenständen wir Realität zuschreiben und welche als nur scheinbare oder abstrahierte Existenzen betrachtet werden müssen, führt auf den Begriff der Emergenz zurück, wenn wir diesen als Ausdruck des Übergangs zwischen Seinsbereichen betrachten.

### **Die Emergenz des Bewusstseins?**

Es soll hierzu noch einmal die Diskussion rund um die Emergenz des Bewusstseins aufgegriffen werden, die bereits in der Einleitung im Rahmen des Panpsychismus von Interesse war (vgl. Abschnitt 1.2):

*Natura non facit saltum* – die Natur macht keine Sprünge; so drückt Strawson

(2016, These 7) seine Überzeugung aus, dass Bewusstsein nicht aus der unbewussten Materie hervorgehen könne, und ganz analog ließe sich argumentieren, dass räumlich Konkretes nicht aus dem unausgedehnten Abstrakten, Intrinsisches nicht aus dem Relationalen und Asymmetrisches nicht aus dem Symmetrischen hervorgehen könne. Reduziert man die Frage nach der Möglichkeit von Emergenz auf die Frage, ob eine bestimmte Art von Gegenständen aus anderen, grundsätzlich verschiedenen Gegenständen entstehen könne, landet man letztlich wieder bei metaphysischen Überzeugungen, welche als Gründe dafür herangezogen werden, wie die jeweilige Antwort ausfallen wird. Dieser Letztverweis auf metaphysische Überzeugungen scheint jedoch immer dann höchst unbefriedigend zu sein, wenn es diese Überzeugungen selbst sind, die in Frage gestellt werden.

Wenn wir statt diesen generellen Betrachtungen jedoch zum eigentlichen Inhalt des Diskurses – das Verhältnis von Bewusstsein und Materie – zurückkehren, lässt sich auf systematischer Ebene folgendes feststellen:

1. Bewusstsein wird (potentiell) als Objekt von Theorien verstanden, analog zu Elektronen, Genen oder Wirbelstürmen.
2. Viele Argumente in der Philosophie des Geistes gehen von einer Substanzontologie aus, die, auch wenn dies manchmal explizit zurückgewiesen wird, die Unterscheidung zwischen den *Trägern (things)* und ihren Eigenschaften (*properties*) nicht überwinden kann: Eigenschaften werden als Attribute von Substanzen verstanden und Eigenschaftsbeziehungen, sind letztlich Beziehungen zwischen Dingen.
3. Selbst wenn eine dezidiert anti-emergentistische Position wie der Strawson'sche Panpsychismus vertreten wird, schleicht sich der Emergenzdiskurs durch die Hintertür wieder ein. Aus diesem Blickwinkel erscheint es dann so, als wäre der Panpsychismus keine überlegene philosophische Position, sondern nur eine solche, die letztlich neuen „metaphysischen Ballast“ auflädt, aber ansonsten gar nicht besser dastünde als die physikalistischen Positionen, zu deren Elimination sie angetreten ist (Stephan 2016).

Anscheinend ist das Entstehen von Neuartigem ein Phänomen, das eng mit Geist und Bewusstsein verknüpft ist: Wie ließe sich sonst verstehen, dass etwas so Andersartiges wie Bewusstsein aus der Materie entstünde? Dies wird häufig so verstanden, dass irgendwo in der fiktionalen Kette Materie – Körper – Gehirn – Geist – Bewusstsein ein „Sprung“ auszumachen wäre und nun eine Theorie antreten müsste, um diesen Sprung verständlich zu machen oder gar zu erklären.



Der in dieser Arbeit eingeschlagene Weg versucht hingegen Emergenz als strukturelle Eigenschaft natürlicher (geschlossener) Prozesse zu verstehen. In den vorigen Abschnitten wurde an mehrere Stellen bereits auf die Kontextualität verwiesen: Dass wir überhaupt (Verbindungen zwischen) Regularitäten erkennen können, ist der Setzung eines Kontextes geschuldet. Mathematisch kann dies mit Verweis auf die mereologische Struktur eine Theorie (oder „der Partitionierung eines Phasenraumes“) sogar exakt dargestellt werden. Ein drohender Subjektivismus wird zudem dadurch entschärft, dass emergente Beschreibungen gewisse Stabilitätsbedingungen erfüllen müssen.

Das Motiv der Kontextualität ist uns neben dieser allgemeinen Betrachtungsweise an mehreren Stellen begegnet, etwa beim Übergang von syntaktischer zu pragmatischer Information (Abschnitt 2.3) oder bei der Herausbildung von Zeichen in mereotopologischen Netzwerken (Abschnitt 5.3). Der Begriff der Emergenz kann zudem nur dann auf den Bereich wissenschaftlicher Erkenntnis umgelegt werden, wenn er als Verbindungsweise verschiedener Beschreibungen (also „epistemologisch“) verstanden wird, insbesondere wenn es uns im Folgenden um eine Beschreibung der Erscheinungsweise von Objekten selbst gehen wird.

Dies sei an dieser Stelle in Anlehnung an die „Subjekt-Objekt Terminologie“ ausgedrückt, wie sie typischerweise von den Nachfolgern Kants und insbesondere etwa von Arthur Schopenhauer verwendet wurde. Diese Redeweise ist mit Vorsicht zu genießen, erscheint aber für die folgende Illustration geeignet – zumindest wenn wir ihr eher den Stellenwert einer Metapher zukommen lassen, der wir nicht in allen Punkten folgen müssen.

Wahrgenommene Eigenschaften seien demnach weniger als Attribute von Substanzen, sondern vielmehr als Beziehungen zwischen zwei „Polstellen“ der Wahrnehmung zu verstehen: als Relationen zwischen einem subjektiven und einem objektiven Pol, welche beide als notwendig hinzuzudenkende Bezugspunkte der Vorstellungswelt, der „Welt der Objekte“, angesehen werden müssen.

Eine geometrisches Bild soll diese Redeweise veranschaulichen: Eine Kugel wird aufgespannt durch zwei Pole. Die Position eines beliebigen Punktes auf der Oberfläche dieser Kugel kann unter Angabe zweier Koordinaten angegeben werden, die sich auf die *relative* Lage des Punktes zu diesen beiden Polen beziehen, wie es etwa in der geographischen Darstellung mithilfe von Längen- und Breitengraden üblich ist. Metaphorisch gesprochen „zerfallen“ kategorische Eigenschaften in Relationen zwischen den beiden Polen, so wie sich Positionen auf einer Kugel anhand ihrer relativen Lage darstellen lassen. Auf diese Weise ausgedrückt, entspricht das Entstehen einer neuartigen Eigenschaften dem Hervortreten einer Beziehung, was selbst wiederum zum Gegenstand einer höherrangigen Beschreibung werden kann.

Die Frage, wie dieses Hervortreten zu verstehen sei, erscheint als frei von mysteriös anmutenden Referenzen auf Vital- und Konfigurationskräfte oder kausal wirksame *Qualia*; sie verwandelt sich in die Frage nach der Erklärbarkeit bzw. Verständlichmachung eines relationalen Phänomens relativ zum gegenwärtigen Zustand eines Systems und zur Notwendigkeit der Explikation eines Kontextes.

Dies sagt uns freilich noch gar nichts über die metaphysische Frage, nach der „intrinsic Natur“ der Pole oder wie dies aus der Perspektive der „Nullten Person“ (Harman 2009) zu verstehen sei. Und erst wenn wir uns an dieser Stelle in metaphysische Spekulationen begeben und vielleicht gar zum Schluss kommen, dass es hier eine „Lücke zu überwinden gäbe“, wird die Sache mysteriös.

Doch klammern wir diese Frage vorläufig aus. Es geht uns also weniger um einen Übergang zwischen Seinsweisen, sondern vielmehr um das, was während des Vollzugs einer „bipolaren“ Aktivität vor sich geht. Dies wollen wir auf die Frage nach dem Bewusstsein umlegen: Sofern etwas *über* den subjektiven Pol dieses Prozesses *als Objekt* gesagt werden soll, bleibt unklar, wie aus einer Ansammlung von Gegenständen ein subjektives Bewusstsein entstünde, da hier das *explanandum* im *explanans* bereits vorausgesetzt würde: Dass wir überhaupt Gegenstände *als Objekte* erkennen können, setzt ein Bewusstsein bereits voraus – und auch ein Verweis auf den Begriff der Emergenz kann uns hier nicht weiterhelfen. In diesem Sinne ist Strawson also zuzustimmen, dass der Begriff der Emergenz, wie er gewöhnlich verstanden wird und an Beispielen aus den Naturwissenschaften veranschaulicht werden kann, nicht viel taugt für eine *Erklärung* des Bewusstseins.

Dennoch können uns diese Überlegungen dazu bringen, dass wir das *hard problem* überdenken: Warum existieren überhaupt Aktivitäten in der Natur, welche zur Emergenz von Objekten führen, und können wir dies (wissenschaftlich) beschreiben? Der erste Teil ist dabei eine Frage der spekulativen Metaphysik und auf der gleichen Stufe anzusiedeln wie die Fragen: „Warum ist überhaupt Seiendes und nicht vielmehr nichts?“, „Was ist das Wesen der Materie?“ oder „Woher kommt das Universum und warum gab es einen Urknall?“. Der zweite Teil hingegen ist *vielleicht* tatsächlich eine Frage, die auch wissenschaftliche Erkenntnisse produzieren könnte. Unter diesem Gesichtspunkt soll nun noch einmal auf die Wahrnehmung zurückgekommen werden.

Die beiden Annahmen, welche die folgenden Abschnitte leiten, lauten, dass erstens, Objekte (und die zu Objekten gemachten Subjekte) keine „ewigen Dinge“ sind, sondern vielmehr die (emergenten) Glieder einer Kette von Zeichenprozessen darstellen; und dass zweitens, Mereologie von den Beziehungen zwischen (bezeichnenden) Objekten, ihren Teilen und den Gegenständen, auf die sie verweisen, handelt.

Sollten diese Vermutungen wissenschaftlich Bestand haben, so würden ausgehend

von ihnen folgende Fragen beantwortet werden oder zumindest eine neue Darstellung erfahren:

1. Inwieweit sind solche Objekte „Teile“ des Bewusstseins? Oder Allgemeiner: Was sind die Konstituenten von Bewusstsein und wie sind dessen Inhalte strukturiert? Kann die Beziehung zwischen physischen Gegenständen (der Tasse und ihren physikalischen Eigenschaften), ihrer Wahrnehmung und den entsprechenden Wahrnehmungsinhalten (die „vorgestellte Tasse“) als Ganzes-Teil-Beziehung gefasst werden? Welche Rolle spielt dabei das phänomenale Bewusstsein (oder „die *Qualia*“)?
2. Die Struktur von Bewusstsein scheint wesentlich einheitlich zu sein. Diese Einheit ist eine strukturelle Eigenschaft, die das Bewusstsein gegenüber den Gegenständen der physischen Welt auszeichnet. Kant spricht in der *KrV* davon, dass eine (nicht-empirische) Einheit von Bewusstsein die Bedingung der Möglichkeit von Erfahrung überhaupt darstellt („transzendente Apperzeption“ [4]). Sonach würde die Vielfalt der physikalischen Erscheinungen, wenn wir sie als Objekte wahrnehmen oder vorstellen, notwendig auf eine Einheit im Bewusstsein verweisen. Doch was ist dabei mit Einheit und Vielheit genau gemeint?

## 6.2 Aspekte der Wahrnehmung

Vor dem Hintergrund der PEM wollen wir nun noch einmal das Beispiel der Wahrnehmung paradigmatisch als Entwicklungsstufe des bewussten Denkens diskutieren [5]. Dabei gehen wir von einer Situation aus, die ähnlich des beschriebenen Falls in Abschnitt 3.1 ist. Dort haben wir die Vorstellung kennengelernt, wonach Wahrnehmung als Konstruktion betrachtet werden könne, in der aus der Anregung der Sinnesorgane Wahrnehmungsobjekte als Bestandteile eines Modells generiert werden, das als Grundlage für Manipulation und Interaktion mit der jeweiligen Umwelt dient. Der dabei verwendete mathematische Formalismus verdeutlichte, dass dies nichts mit einer vermeintlichen „Essenz“ von Wahrnehmung – z.B. mit der Abbildung einer physikalischen Wirklichkeit – zu tun hat, sondern eine Konsequenz des Drangs nach Selbsterhaltung ist (oder im Jargon der Kybernetiker gesprochen: eine Konsequenz von „Regulierung“; wir werden gleich dafür argumentieren, dass hinter diesen Metaphern nichts anderes steckt als eine bestimmte Art der zeichenartigen Bezugnahme). Eine anschließende Kritik an diesem Bild lautete (Abschnitt 3.2), dass dieses aber doch gar nichts über das Bewusstsein aussagen würde: Weder spielt das Bewusstsein

irgendeine funktionale Rolle noch wird irgendwie ersichtlich, wie sich nun bewusste Wahrnehmungen von gerechneten Inferenzen unterscheiden würden. Worauf schließen wir also in der Wahrnehmung genau und was hat dies mit Bewusstsein zu tun?

Eine Beobachtung, die der Phänomenologie Husserls entstammt, besagt, dass wir nicht den Gegenstand selbst wahrnehmen, sondern dass uns immer nur ein bestimmter Aspekt dieses Gegenstandes erscheint. Aber auch die moderne empirische Wahrnehmungsforschung geht heute davon aus, dass zwischen wahrgenommenen Objektteilen, also etwa den vermeintlichen Projektionen physikalischer Gegenstände auf die Netzhaut, und den daraus konstruierten visuellen Objekten selbst, also den Objekten, die wir bewusst wahrnehmen, zu unterscheiden ist (Singh & Hoffman 1997). Dies wollen wir nun wie folgt darstellen: Es stelle  $x$  ein Wahrnehmungsobjekt dar. Dann gilt für den eigentlich präsentierten Aspekt  $y$  in unserer Notation, dass

$$y = \hat{P}x. \quad (6.1)$$

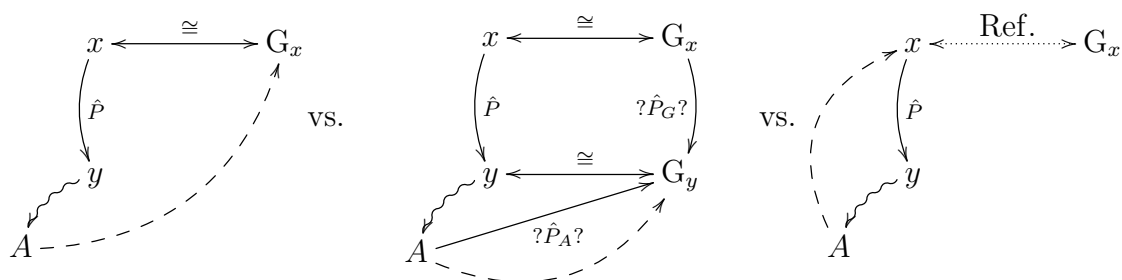
Gl. (6.1) soll deutlich machen, dass  $x$  und  $y$  in einer Teil-Ganzes-Beziehung stehen, und dass von  $y$  nur insofern als Objekt gesprochen werden kann, als dass es in einer prozessualen Relation,  $\hat{P}$ , zu  $x$  steht.

Das Problem, dass wir anscheinend nur Teile eines Objektes wahrnehmen, aber dennoch meinen, wir hätten den Gegenstand  $G$  wahrgenommen, wurde in Anlehnung an Husserl z.B. von Reinhardt Grossmann (1973, Kap. 19) diskutiert, der dafür argumentiert, dass primär nur die Objekte selbst als eigentliche Inhalte des Bewusstseins gegeben sind. Dies gilt gerade auch dann, wenn wir gar nicht alle seine Teile gesehen haben, was ja für visuelle Wahrnehmung allgemein nicht der Fall ist. Objekte sind nach Grossmann nicht identisch zu der Klasse ihrer Teile, sondern wesentlich Strukturen auf diesen Teilen: Das Objekt  $x = \sum_i y_i$  bezeichnet etwas anderes als die Klasse  $K_x = \{y_i | y_i \ll x\}$ , nämlich eine strukturierte Ganzheit. Dies lässt die Möglichkeit offen, Objekte (via Strukturen) zu erkennen, ohne dabei die Klasse aller ihm zugehörigen Teile wahrzunehmen.

Führen wir uns noch einmal die Definition aus Abschnitt 5.2 vor Augen, wonach *wohldefinierte* mereologische Objekte auf (konkrete) Gegenstände verweisen und zusätzlich gewisse strukturelle Bedingungen erfüllen, im Gegensatz zu *unselbstständigen* Teilen oder Klassen solcher Teile. Dies stellt einen möglichen Unterschied zwischen der Klasse  $K_x$  und dem Objekt  $x$  dar. Während also  $x$ , sofern es wohldefiniert ist, notwendig auf einen Gegenstand  $G_x$  verweist, gilt dies nicht für die Klasse  $K_x$ . Allerdings scheint dies eher ein Taschenspielertrick zu sein; Definitionen lösen in der Regel keine Probleme.

In vielen neueren Diskussionen wird den Teilen oder Aspekten  $y$  ein qualitativer Charakter zugesprochen: Der visuellen Wahrnehmung einer Frucht entspricht der (phänomenale) grüne Fleck, den ich sehe. Dies scheint nun aber schwer mit dem Bild in Einklang zu bringen zu sein, wonach die Früchte „da draußen“ Dinge sind, die selbst gar keine phänomenalen Qualitäten besitzen. Woher kommen diese Qualitäten dann? Entstehen sie im Bewusstsein? Werden sie der Wahrnehmung „hinzuaddiert“? Sind sie einfach nur die Eigenschaften der Gegenstände, wie sie vom Gehirn eines Lebewesens repräsentiert werden (vgl. Dretske 1995, Kap. 3)?

Gehen wir davon aus, dass i) Objekte auf Gegenstände verweisen und ii) wir in der Wahrnehmung von Teilen oder Aspektgestalten (unbewusst) auf die Anwesenheit von Dingen (d.h. auf Gegenstände oder andere Objekte) schließen, lässt sich die Situation graphisch wie in Abb. 6.1 (links) darstellen, etwa wenn wir einen grünen runden Fleck wahrnehmen und dabei implizit darauf schließen, eine Frucht vor uns zu haben.



**Abbildung 6.1:** Wahrnehmung als Konstruktion von Dingen durch ein Subjekt  $A$ ; drei Modelle. Kontinuierliche, gerichtete Pfeile entsprechen dabei den Teil-Ganzes-Beziehung ( $\hat{P}$ ) zwischen Wahrnehmungsobjekten  $x$  und den eigentlich präsentierten ( $\rightsquigarrow$ ) Aspekten oder Teilobjekten  $y$ . Gestrichelte Pfeile bezeichnen immer nur reflexiv einsehbare Schlüsse („Konstruktionen“); die  $G_\alpha$  bezeichnen Gegenstände, die entweder als Abbild ( $\cong$ ) oder als Referent der entsprechenden Objekte  $\alpha$  betrachtet werden (Details siehe Text).

Auf den Einwand, dass wir ja gar nie ein Objekt selbst, sondern immer nur Aspekte oder Teile des Objekts wahrnehmen, wird manchmal zugegeben, dass wir letztlich die Existenz des Gegenstandes „konstruieren“ würden, obwohl wir ihn selbst gar nicht wahrnehmen. Und da dieser Konstruktion selbst kein wahrgenommener Gegenstand (Aspekt) entspricht, sind solche Konstruktionen transparent (vgl. auch die Diskussion in Abschnitt 3.2).

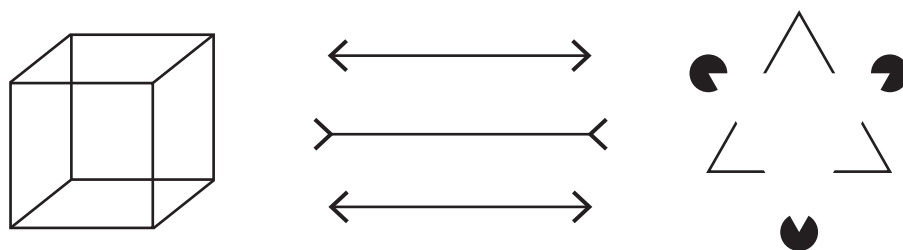
Der Begriff der Konstruktion ist jedoch ein problematisch und muss näher spezifiziert werden. Im visuellen Fall sind viele geneigt, von Wahrnehmungsobjekten als „vom Gehirn konstruierten“ Repräsentationen von Gegenständen zu sprechen, die den realen (physikalischen und nicht-konstruierten) Gegenständen gegenüberstehen.

Wenn wir jedoch andere Sinnesmodalitäten betrachten – etwa das Hören, Riechen oder Tasten –, stellen wir fest, dass es keinen Unterschied zwischen dem Gegenstand selbst und seiner phänomenalen Darstellung zu geben scheint: Wir verfügen über keine „Repräsentation“ eines Tons auf „unserem Schirm“, die sich vom Ton „da draußen“ unterscheiden würde: Der „phänomenale Ton“ bezeichnet denselben Gegenstand, den der physikalische bezeichnet.

Diese berechtigte Kritik am Repräsentationalismus führt allerdings schnell auf einen direkten Realismus, der davon ausgeht, dass wir die Dinge so wahrnehmen, wie sie sind, ohne den Umweg über Inferenzen oder Repräsentationen zu gehen. Während die Fokussierung auf den visuellen Fall zur Annahme einer „verdoppelten Wirklichkeit“ – Abbild vs. abgebildeter Gegenstand – verleitet, führt die Alternative oftmals auf die Annahme, wir würden stattdessen die Gegenstände (oder ihre Eigenschaften) selbst wahrnehmen. Analog könnte dies auch über die Wahrnehmung von „inneren“ Zuständen gesagt werden, wohingegen in einer repräsentationalistischen Theorie z.B. „Schmerz“ typischerweise mit einer Registrierung von Gewebeschaden, also mit einer Repräsentation, gleichgesetzt wird (so etwa bei Dretske 1995, Kap. 4).

Offensichtlich werden die Konstruktionsleistungen schließlich dann, wenn den wahrgenommenen Objekten gar kein physikalischer Gegenstand entspricht. Aus den Untersuchungen der visuellen Wahrnehmung sind unzählige Fälle bekannt, die uns bewusst machen, dass die Objekte, die wir „sehen“, nirgendwo außerhalb unserer Wahrnehmung existieren oder über Eigenschaften verfügen, die sie „objektiv“ gar nicht haben. Bekannte Beispiele sind etwa der Necker-Würfel, bei dem abwechselnd zwischen verschiedenen dreidimensionalen Perspektiven hin- und hergewechselt wird, die Müller-Lyer-Illusion, die drei gleich lange Linien zeigt, welche jedoch unterschiedlich lang erscheinen, oder das Kaniza-Dreieck, bei dem es zur Entstehung eines Objekts (weißes Dreieck) in der Wahrnehmung kommt (Abb. 6.2).

Aber auch auf einer abstrakteren, begrifflichen Ebene sind wir mit derartigen Konstruktionsleistungen konfrontiert, z.B. wenn wir davon sprechen, wir würden „die Welt“ oder „die Natur“ wahrnehmen, und dabei fälschlicherweise meinen, uns auf einen real existierenden Gegenstand zu beziehen. Obwohl es zulässig sein mag, von diesen Dingen als konstruierte (fiktive) Objekte zu sprechen, gibt es gar keine diesen entsprechenden (konkreten) Gegenstände, die hier in ein Verhältnis gebracht werden könnten. Neben diesen beiden Beispielen gibt es noch eine Reihe anderer Objekte, deren Referenz fraglich ist: Kobolde, Einhörner, Bewusstseine (als Objekte in „der Natur“) oder Dinge wie Zahlen und die idealisierten Objekte der Wissenschaft.



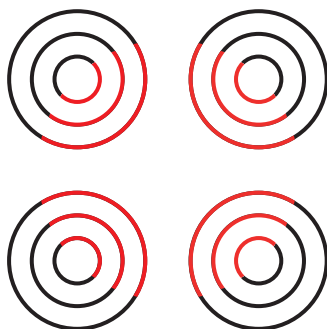
**Abbildung 6.2:** Konstruktionen der Wahrnehmung. Der Neckar-Würfel (links) veranschaulicht die Bistabilität der Gestaltwahrnehmung, die Müller-Lyer-Illusion (Mitte) zeigt drei (nur scheinbar) unterschiedlich lange Linien und das Kaniza-Dreieck (rechts), verdeutlicht die Entstehung eines visuellen Objekts in der Wahrnehmung.

Natürlich bedeutet dies nicht, dass es keine Unterschiede zwischen diesen Objekten gibt. Von einigen sagen wird, dass sie nur Einbildungen oder Hirngespinnste sind, andere scheinen Gegenstand zeitloser wissenschaftlicher Theorien zu sein. Wiederum andere sind ungreifbar, aber ihre objektive Existenz scheint vielen gewiss zu sein... Im Allgemeinen ist es nicht einfach, darüber Auskunft zu geben, welchen Objekten nun ein Gegenstand entspricht oder nicht, da wir ja nicht einfach aus unserem Denken heraus treten können.

In Abb. 6.1 (mitte) wird die Aspekthaftigkeit der Wahrnehmung etwas anders interpretiert. Aus der Reflexion über die Tatsache, dass wir immer nur Teilobjekten oder Aspekten  $y$  von Objekten begegnen, wird gefolgert, dass diese selbst auf Gegenstände  $G_y$  verweisen würden, welche die eigentlich wahrgenommenen Gegenstände darstellten. Der qualitative Charakter der Wahrnehmung von  $y$  – „das Grün“ der Frucht – könnte dann als echte (Teil-)Eigenschaft des Gegenstandes, auf den das Objekt „Frucht“ verweist, verstanden werden,  $G_y = \hat{P}_G G_x$ . Farben wären somit Eigenschaften der Gegenstände. Es ist jedoch nicht trivial, plausibel zu machen, wie dies im Rahmen des Physikalismus verstanden werden könnte (vgl. die Diskussion zum Farbrealismus bei Byrne & Hilbert 2003).

Alternativ kann auch behauptet werden, dass der Farbeindruck zum Subjekt gehören würde – Farben wären somit Eigenschaften (oder Teile) des („Zustandes“ des) Bewusstseins,  $G_y = \hat{P}_A A$ . Beide Ansichten sind problematisch: Im ersten Fall, falls wir davon ausgehen, dass physikalische Gegenstände eigentlich gar keine Farben besitzen, sondern nur physikalisch primäre Eigenschaften wie Ausdehnung, Masse, Ladung etc. Im zweiten Fall, falls wir die Existenz eines substantiellen Bewusstseins ablehnen, dem irgendwelche Eigenschaften zugeschrieben werden könnten. Die Wahrnehmung der Frucht mag die Qualität eines Farbeindruckes besitzen, aber we-

der haben physikalische Gegenstände Farben noch existiert ein gefärbtes etwas, das in unseren Köpfen herumspukt. Besonders eindrücklich lässt sich dies veranschaulichen, wenn wir optische Täuschungen betrachten, bei der es zur Wahrnehmung eines gefärbten Objekts kommt, obwohl „objektiv“ ein solches Objekt gar nirgends existiert (Abb. 6.3; nach van Tuijl 1975). Die Mechanismen, welche dieser Täuschung zugrunde liegen, sind bis heute nicht restlos geklärt.



**Abbildung 6.3:** Entstehung eines gefärbten Objekts. Alle Pixel, die nicht auf einer der Konturen zu liegen kommen, sind weiß. Eine Färbung von Teilen der Konturen in den Ecken führt jedoch zum Entstehen einer glühenden roten Scheibe in der Mitte des Bildes.

Ein gedanklicher Schritt zur Auflösung dieses Dilemmas wird in der Darstellungsweise der projektiven Mereologie ermöglicht: Einem Objektteil muss schlechterdings nicht, wie zuvor angenommen, ein Gegenstand mit diesen und jenen (realen) Eigenschaften entsprechen, falls solche Objekte in erster Linie nur als *relata* einer (projektiven) Beziehung aufgefasst werden. Erst wenn bestimmte Objektteile als *wohldefinierte* Objekte in der Wahrnehmung auftreten, referenzieren diese überhaupt auf einen Gegenstand.

Weder konstruieren wir Gegenstände „in der Welt“, noch „bilden“ wir sie ab. Wir schließen auf Objekte, die auf Gegenstände verweisen aber immer nur Teile unserer Erfahrung sind. In Abb. 6.1 (rechts) ist der Wahrnehmungsakt als konstruktives Schließen dargestellt. „Hinter“ den (selbstständigen) Objekten  $x$ , die uns nur in ihren (unselbstständigen) Teilen  $y$  gegeben sind, stehen bestimmte Gegenstände  $G_x$ , von denen wir zwar annehmen, dass sie existieren und die Referenten der Objekte  $x$  darstellen, die wir also solche allerdings nie wahrnehmen. Wahrnehmung bewegt sich, in einer Metapher gesprochen, immer im Raum der Objekte und nicht im Raum der Gegenstände.

Dies kann also keines Falls so verstanden werden, dass wir bestimmte Gegenstände in



der Wahrnehmung quasi-mystisch „erschauen“ oder von ihnen unmittelbare Kenntnis („*acquaintance*“; Russell 1914, Goff 2015) erlangen; und da wir agnostisch bezüglich der Frage bleiben müssen, was nun genau hinter den Objekten steht, ist diese Interpretation der Wahrnehmung auch nicht identisch zum Repräsentationalismus: Wir „schließen“ von (präsentierten) Teilen auf (intentionale) Objekte, nicht jedoch auf die Eigenschaften von Gegenständen, die zu diesen Objekten in einem Modell- oder Abbildverhältnis stehen würden. Gleichzeitig unterscheidet sich diese Auffassung von einer typischen Lesart des Idealismus. Nur weil wir auf *x als wohldefiniertes Objekt* erst in der Wahrnehmung schließen, folgt nicht automatisch, dass es keine Gegenstände außerhalb unserer Wahrnehmung gäbe (daraus lässt sich aber wiederum keine naiv-realistische Position ableiten).

Allerdings bleibt die Frage unbeantwortet, wie wir mit der Tatsache umzugehen haben, dass *y* phänomenale Qualitäten zu besitzen scheint, insbesondere da wir ja meinen, die Objekte *x* wahrzunehmen, obwohl uns nur deren Teile *y* präsent sind. Wenn wir die phänomenalen Qualitäten von *y* nicht wie in Abb. 6.1 (mitte) gegenständlich (z.B. als *Qualia*) denken, dann müssen wir etwas zur Rolle des Subjekts *A* sagen, das manchmal als der „permanente Bezugspunkt“ aller bewusster Aktivität aufgefasst wird und von dem gesagt wird, ihm würden diese oder jene Eigenschaft „erscheinen“.

Mit Vorgriff auf die folgende Diskussion wollen wir dem die Ansicht gegenüberstellen, dass sich die Phänomenalität der Wahrnehmung aus deren Einbindung in einen Verweisungszusammenhang ergibt, in welchen die Objekte eingebunden sind. Subjekte kommen dabei weder als Träger noch als Betrachter qualitativer Eigenschaften vor. Für ein besseres Verständnis muss aber zuerst die sog. „Einheit des Bewusstseins“ genauer untersucht werden.

### 6.3 Aktivität und Kontext

#### Die Einheiten des Bewusstseins

Gemäß der Vorstellung, wonach die Naturwissenschaften eine Seinsstufe nach der anderen ergründen würden, ließe sich vermuten, dass schließlich auch die „Phänomene des Bewusstseins“ als Objekte einer Wissenschaft in Zukunft erschlossen und analysiert werden können.

Dem scheint aber die Auffassung zu widersprechen, dass Bewusstsein ja überhaupt erst Bedingung der Möglichkeit solcher Analysen darstellt. In anderen Worten: Beschreibt die Wissenschaft nicht mehr die Struktur der Welt an sich, sondern lediglich deren Struktur, wie sie vom Bewusstsein abstrahiert wird, so begegnen wir einer

grundsätzlichen Form der Selbstbezüglichkeit, wenn wir nach der Struktur des Bewusstseins fragen und es als Teil derselben Welt verstehen wollen.

Besonders akut wird dies nun, wenn die sog. „Einheit des Bewusstseins“ als mereologische Beziehung zwischen getrennten biologischen Prozessen oder einzelnen Zuständen des Erlebens rekonstruiert werden soll (Bayne 2010, Wiese & Metzinger 2012). Eine *transzendente* Einheit des Bewusstseins müsse nämlich, so die Kantische Idee, einer (mereologischen) Beschreibung von Erfahrungsgegenständen *voraus* gehen; wir können diese Einheit nicht direkt erfahren, sondern höchstens als Bedingung von Erfahrung reflexiv erkennen.

Dennoch lässt sich hoffen, dass wir zumindest einige Aussagen zur *empirischen* Struktur des Bewusstseins machen können. Ein Beispiel dafür liefert die von Bayne & Chalmers (2003) formulierte „Subsumptionsthese“. Demnach ist jeder phänomenale Zustand  $\varphi$  eines Organismus (korrespondierend z.B. zu einer bestimmten Einzelwahrnehmung) in einen globalen Zustand  $\psi$  eingebettet, der als „das Bewusstsein“ zum jeweiligen Zeitpunkt bezeichnet werden könnte. Die Relation, die dieser Einbettung zugrunde liegt, kann als mereologische verstanden werden:

$$\psi = \sum_i \varphi_i, \quad (6.2)$$

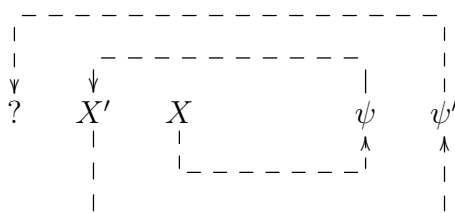
wobei  $\sum$  die mereologische Summe über die einzelnen Wahrnehmungszustände bezeichnet (alternativ könnte die Summe auch über weitere phänomenale Zustände laufen), d.h. das Bewusstsein  $\psi$  ist identisch zur mereologischen Summe der  $\varphi_i$ . Wie in den vorherigen Abschnitten argumentiert wurde, ist eine unqualifizierte Summenbildung allerdings umstritten [6]. Es sollten weitere Bedingungen, die eine Summenbildung erlauben, von den Zuständen  $\varphi_i$  erfüllt werden, was gleichzeitig die Intuition ausdrücken soll, „das Ganze sei mehr als die Summe seiner Teile“. Kandidaten dafür wären Berührung oder Funktionalität der jeweiligen Zustände  $\varphi_i$  (in einer Phänomenologischen Theorie) oder in Bezug auf die Gegenstände, auf welche jene verweisen (in einer repräsentationalistischen Theorie). Man könnte aber auch umgekehrt folgern, dass der Unterschied zwischen bewussten und nicht-bewussten Gegenständen, etwa zwischen den mentalen Zuständen von Menschen und funktionalen Zuständen eines Computers, gerade darin bestünde, dass die Summe aus Gl. 6.2 für erstere (und nur für diese) immer und uneingeschränkt existiert. Dies wäre ein ganz grundsätzliches Argument, dass auf die „Essenz“ des Bewusstseins abzielen würde. Von phänomenalen Zuständen  $\varphi_j$  könnte dann vermutet werden, dass diese immer nur als Teile eines umfassenden Objekts  $\psi$  existieren:

$$\varphi_j \rightarrow \exists \psi \left( \psi = \sum_i \varphi_i \wedge \forall \varphi_i (\varphi_i = \hat{P}_i \psi) \right), \quad (6.3)$$

wobei  $\varphi_j$  einen beliebigen phänomenalen Zustand und  $\{\varphi_i\}$  die Klasse aller derzeit aktualisierter phänomenaler Zustände bezeichne.

Nun würde so eine Beschreibung vielleicht sogar einen deskriptiven Fortschritt gegenüber den üblichen Versuchen zur Beschreibung des Bewusstseins als empirisch zugängliches Objekt darstellen. Allerdings wird dabei die transzendente Problematik, welche die Frage nach der Einheit des Bewusstseins überhaupt erst interessant macht, außen vorgelassen.

Glauben wir nämlich, dass die Einheit des Bewusstseins die Einheit der wahrgenommenen oder vorgestellten Gegenstände garantiert, landen wir schnell in einen infiniten Regress, falls Bewusstsein wiederum gegenständlich aufgefasst wird (Abb. 6.4): Dass wir einen Gegenstand als Einheit (d.h., als Referent eines selbstständigen Objekts) vorstellen können, wird auf die Einheit des Bewusstseins zurückgeführt; wenn wir aber nun die Einheit des Bewusstseins (wiederum als Referent eines Objekts verstanden) auf die Einheit eines bestimmten (z.B. physikalischen) Gegenstandes zurückführen – so etwa die Vorstellung, dass die Einheit des Bewusstseins in der temporalen Einheit (Synchronisation) neuronaler Aktivität begründet sei (Engel & Singer 2001) –, landen wir automatisch bei der Frage, was denn nun wieder die Einheit dieses Gegenstandes begründet usw. Wenn wir also, ähnlich wie Kant, daran glauben, die Einheit der wahrgenommenen Objekte folge der „synthetischen Einheit“ einer geistigen Aktivität – in anderen Worten: wenn wir daran glauben, dass es vielleicht eine Wirklichkeit, aber sicherlich keine *objektive* Wirklichkeit außerhalb des Bewusstseins gibt –, dann begegnen wir Schwierigkeiten, wenn wir versuchen, von „dem“ Bewusstsein mit diesen und jenen (objektiven) Eigenschaften zu sprechen.



**Abbildung 6.4:** Zirkuläre Struktur der Vergegenständlichung. Bewusstsein als einheitlicher Gegenstand  $\psi$  wird auf die Einheit (physikalischer) Gegenstände  $X$  zurückgeführt, deren Einheit wiederum durch Rückführung garantiert werden muss.

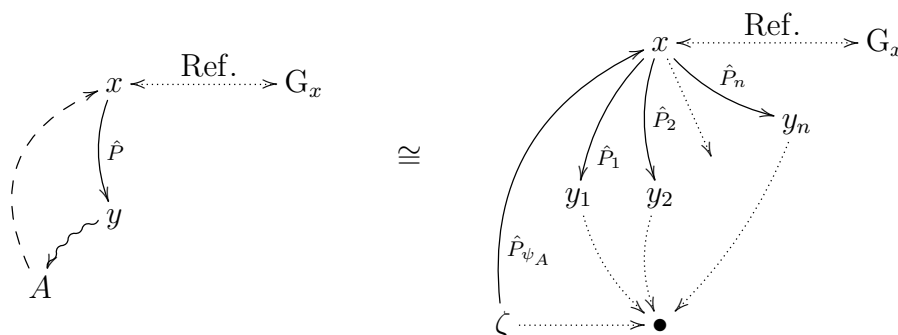
Dies soll nun in den Kontext des letzten Teilabschnittes zur Aspektgestaltung der Wahrnehmung gestellt werden. In den Darstellungen dieses Abschnittes wurde bisher eine große Vereinfachung getroffen: Es wurde stillschweigend davon ausgegangen, dass, ausgehend von den Teilen  $y_i$ , die in einem Subjekt  $A$  „zusammenkommen“, auf

ein entsprechendes Objekt  $x$  vom Subjekt  $A$  geschlossen wird – auch wenn diese Ansicht empirisch motiviert war (vgl. Abschnitt 3.1).

Die Frage nach der Einheit des Bewusstseins kann als Frage danach verstanden werden, wie dieser Prozess des Schließens überhaupt möglich ist. Während es in Abschnitt 6.2 noch um die Beziehungen zwischen den (intentionalen) Objekten, ihren Teilen und den Gegenständen, auf die sie verweisen, ging, scheint hier nach der Darstellung des Vollzuges eines „subjektiv“ genannten Prozesses gefragt zu werden. Nun droht dies aber, genau im zuvor erwähnten Regress zu resultieren, wenn wir die Frage nach der Einheit des Bewusstseins mit Verweis auf „das“ Subjekt (oder „das“ Bewusstsein), „in dem“ sich irgendetwas vollzieht, beantworten wollen.

Aber selbst wenn wir über das Subjekt nicht vielmehr sagen können, als dass es den Pol eines Prozesses bildet, können wir dennoch versuchen, den Vollzug dieses Prozesses verständlich werden zu lassen, d.h. zu versuchen, Bewusstsein als die Form eines solchen Prozesses zu verstehen.

Dies ist in Abb. 6.5 veranschaulicht, wobei wir uns in der Analyse wiederum auf einen einfachen Wahrnehmungsprozess beschränken wollen.



**Abbildung 6.5:** Übergang von einer kompakten Darstellung (links) zu einer ausführlicheren Darstellung (rechts) der einheitlichen Struktur der Wahrnehmung. Die Rede von einem gegenständlichen Subjekt, „in dem Sinnesdaten zusammenkommen“ ( $\rightsquigarrow$ ) und aus denen schließlich auf die wahrgenommenen Objekte geschlossen wird ( $-\rightarrow$ ), soll durch die Vorstellung ersetzt werden, dass Objektwahrnehmung ein synthetischer Prozess ist, der aber immer auch auf einen irreduziblen Pol, einen „blinden Fleck“ ( $\bullet$ ) verweist (Details siehe Text).

In einer ausführlicheren Darstellung, zerfallen Objekte  $x$  in Objekt-Teile  $y_i$ , welche in ihrer Einzelwahrnehmung so in eine Aktivität eingebunden sind, dass deren Summe  $\zeta = \sum y_i$  in der Wahrnehmung des intentionalen Objektes  $x$  resultiert. Die Subsumptionsthese von Bayne & Chalmers (2003) kann nun als Forderung nach Selbstkonsistenz dieses Vorgangs verstanden werden: In der Wahrnehmung wird erst das Objekt  $x$  aus einer Ganzheit geformt, deren Elemente selbst aus den Projektio-

nen eben dieses Objektes herrühren. Wir bezeichnen die Form einer Aktivität,  $\psi_A$ , die eine solche zirkuläre Struktur aufweist, als Bewusstseinsaktivität (oder kolloquial einfach nur als *Bewusstsein von A*) [7].

Es gilt nun, dass

$$y_j = \hat{P}_j x \quad (6.4)$$

$$x = \hat{P}_{\psi_A} \zeta, \text{ mit } \zeta = \sum y_i \quad (6.5)$$

$$\Rightarrow x = \hat{P}_{\psi_A} (\hat{P}_1 x + \hat{P}_2 x + \dots + \hat{P}_n x). \quad (6.6)$$

(Dabei wurden der Einfachheit wegen die Details der mereotopologischen Emergenz von Objekten aus Abschnitt 5.1 nicht berücksichtigt. Im Allgemeinen werden die wahrgenommenen Teile in eine mannigfaltige Struktur eingebunden sein: Wir erfahren ja niemals ein vollständig isoliertes Objekt  $x$ , sondern nehmen es vor dem Hintergrund einer sinnlichen Vielheit wahr, aus deren Ganzheit das Objekt  $x$ , gemäß der mereotopologischen Ansicht, emergiert, sofern es über eine „Hülle“ verfügt, die  $x$  von einer „Umgebung“,  $x^c$ , trennt.)

Gl. (6.6) lässt sich beliebig iterieren:

$$x = \hat{P}_{\psi_A} (\hat{P}_1 x + \hat{P}_2 x + \dots + \hat{P}_n x) \quad (6.7)$$

$$= \hat{P}_{\psi_A} (\sum \hat{P}_i) x = \hat{P}_{\psi_A} (\sum \hat{P}_i) \hat{P}_{\psi_A} (\sum \hat{P}_i) x \quad (6.8)$$

$$= \dots = [\hat{P}_{\psi_A} (\sum \hat{P}_i)]^n x, \quad (6.9)$$

was äquivalent zur Aussage ist, dass  $x$  einen „Eigenwert“ zur Operation  $\hat{P}_{\psi_A} (\sum \hat{P}_i)$  darstellt. Dies soll durch die Schreibweise  $[x]_{\psi_A}$  ausgedrückt werden:

$$[x]_{\psi_A} \Leftrightarrow x = \lim_{n \rightarrow \infty} [\hat{P}_{\psi_A} (\sum \hat{P}_i)]^n x. \quad (6.10)$$

Ganz ähnliche Konzepte finden sich bei den Diskussionen zur „Kybernetik zweiter Ordnung“ von Wahrnehmungs-Koordinationsprozessen (von Foerster 1981) oder, allgemeiner, zur Struktur kreisförmiger und selbstbezüglicher Prozesse (Kauffman 2003). Objekte können dann als emergent bezüglich einer solchen Struktur aufgefasst werden.

Es soll nun noch einmal verdeutlicht werden, warum der Wahrnehmungsprozess wesentlich als mereologischer darzustellen war: Dort wo davon gesprochen wird, dass die (empirische) Einheit des Bewusstseins „erklärt“ werden könne, wird gewöhnlich davon ausgegangen, dass der Inhalt des Bewusstseins einer einheitlichen Szenerie gleicht, die das Resultat einer neuronal vermittelten Zusammenfassung einer Vielzahl von Teil-Objekten ist („*phenomenal binding*“; Revonsuo 1999). Dabei kommt den Inhalten des Bewusstseins aber die erklärungsbedürftige Eigenschaft zu, einheitliche Objekte darzustellen, die zwar in diskrete Teile zerlegt werden können, aber

selbst keine körnige Struktur aufweisen – anders materielle Gegenstände, von denen meist angenommen wird, sie wären in einer fundamentalen Weise gekörnt, wobei deren Teile als entsprechende Zusammenfassungen diskreter Atome oder Teilchen verstanden werden können (vgl. auch die Diskussion zur „Hyperextensionalität“ aus Abschnitt 4.3).

Sowohl in der Phänomenologischen Analyse nach Husserl, der von der Wahrnehmung von Aspekten ausgeht, als auch in einer neurobiologischen Darstellung, die eine *einheitliche* „phänomenale Realität“ (Metzinger 1995a, S. 33) auf Basis von Anregungen *separater* Sinnesmodalitäten verständlich machen will, nehmen wir die Objekte des Bewusstseins immer über ihre Teile wahr.

Dadurch wird die eine Seite eines Wechselspiels dargestellt, nämlich die, welche von Teilen zu Ganzheiten fortschreitet. Würden wir tatsächlich die Objekte  $x$  direkt und ohne Umweg über die unselbstständigen Aspekte  $y_i$  wahrnehmen, so bliebe die Einheit der wahrgenommenen Objekte eine unmittelbare und nicht, wie in Gl. (6.10) ausgedrückt, eine (im *Bewusstsein*) vermittelte. **Fiat lux!**

Dies muss jedoch um die dazu komplementäre Darstellung ergänzt werden, die von einer Ganzheit ausgeht, aus der sich Teile herausbilden. Im Bewusstsein werden erst die Objekte untereinander und deren Teile als Projektionen unterschieden (also: sichtbar gemacht). In diesem Sinne werden sie erst im Bewusstsein *realisiert*. Synthetische Einheit und Unterscheidung sind dabei als je-relative und gegenläufige Teilprozesse zu verstehen, wodurch noch einmal die These der Koemergenz von Teilen und Ganzheiten sowie von Objekten und deren Erscheinungsweisen verdeutlicht wird.

Wir haben also bewusste Wahrnehmung als kreisförmigen Prozess dargestellt, der Objekte als stabile Resultate hervorbringt. Doch inwiefern müssen wir hier von einer „Aktivität“ sprechen und was bedeutet es zu sagen, „das Subjekt“  $A$  sei nichts als der „Pol“ einer solchen Aktivität? Der Aktivitätscharakter lässt sich dadurch einsehen, dass die Gesamtheit der selbstständigen Objekte und deren Teile zwar ein Bild all dessen liefern, was sich an einen solchen Prozess überhaupt („objektiv“) feststellen lässt, allerdings eines, das nicht vollständig ist, um diesen als sich vollziehend zu verstehen, ein Bild also, dem es notwendigerweise an „Tiefenschärfe“ fehlt. Wahrnehmung, so die Annahme, entspricht jedoch einem solchen, sich (kreisförmig) vollziehenden Prozess und eine Zerlegung in dessen „objektive“ Glieder würde unvollständig bleiben.

Betrachten wir dazu noch einmal die Kugelmetapher eines Erfahrungsprozesses, der durch zwei Pole aufgespannt wird. Ein einführendes Beispiel in die mathematische Topologie handelt davon, dass zwar immer zweidimensionale, flache Karten erzeugt

werden können, die lokal einen Abschnitt der Kugeloberfläche abbilden, dass es aber keine zusammenhängende, nicht-redundante Karte der gesamten Oberfläche gibt. In Analogie gilt: Erfahrung lässt sich zwar mithilfe einer Objektbeschreibung „lokal verobjektivieren“, aber es gibt *keine globale Verobjektivierung* von Erfahrung.

Diese Überlegungen lassen sich nun auf unsere Darstellung übertragen. Die Gesamtheit  $\zeta$ , die aus der Klasse der Objektteile  $y_i$  besteht und aus der das selbstständige Objekt  $x$  hervorgeht, verweist immer auch auf einen „blinden Fleck“ außerhalb ihrer selbst. Dieser hinzuzudenkende Bezugspunkt kann als „subjektiver Pol“ des Wahrnehmungsprozesses bezeichnet werden, der aber grundsätzlich gar keinen Gegenstand „in der Welt“ oder eben ein Objekt bezeichnet: Es gibt kein „Subjekt“, das als Objekt innerhalb der Wahrnehmung auftauchen könnte, da sich der Inhalt jeder Wahrnehmung bereits vollständig in den verweisenden Objekten  $x$  und ihren Objektteilen  $y_i$  erschöpft.

Wir nehmen keine Subjekte wahr, so wie wir Bäume im Wald wahrnehmen. Wenn wir genauer hinsehen sehen wir vielleicht Zweige und Blätter, aber kein Subjekt. Das Beiwort „subjektiv“ soll anzeigen, dass etwas auf ganz bestimmte Weise zu den Objekten korreliert. Subjekte, um in einer anderen Metapher zu sprechen, entsprechen den Perspektiven auf die Objekte, nicht einem Objekt selbst. *Bewusstsein* kann als die Form derjenigen Aktivität gedeutet werden, die überhaupt verweisende Objekte relativ zu einem nicht-sichtbaren Bezugspunkt hervorbringt, aber es kann nicht davon gesprochen werden, dass hier Bewusstsein „im“ Subjekt zu finden wäre oder dass Bewusstsein selbst ein Objekt unter vielen darstellen würde.

### **Wie es ist, eine Fledermaus zu sein**

Nach der Diskussion der einheitlichen Struktur des Bewusstseins, soll nun etwas zu dessen Phänomenalität gesagt werden. Dies betrifft das Rätsel der Qualitäten. Wie wir in der Einleitung bereits diskutiert hatten, führt die Annahme von elementaren *Qualia* oftmals zu einer Aufspaltung in ein An-sich-Sein der Natur und deren Wahrnehmung, etwas das von Whitehead als zu vermeidende „Bifurkation“ bezeichnet wurde (CN, Kap. 2). Um nicht selbst in ganz ähnliche Schwierigkeiten zu geraten, soll eine schwächere, nicht-substantialistische Auffassung von Phänomenalität vorgestellt werden, welche diese primär als Struktureigenschaft von Wahrnehmungsprozessen ausweist.

Die Darstellungsweise mithilfe der *projektiven Merelogie* hebt den Aspekt der Seinsweise der wahrgenommenen Teile hervor, der immer nur relativ zu einem Wahrnehmungskontext ist. Dieser Wahrnehmungskontext ist, wie im letzten Abschnitt argumentiert wurde, ein synthetisierender. Die Phänomenalität der Wahrnehmung, die

meist nur formal unbefriedigend (Dennett 1988) durch Verweis auf eine un(an)greifbare „what-it-is-like-ness“ beschrieben wird, entspricht dabei der Art und Weise, wie Gegenstände als Objekte in Erscheinung treten.

Phänomenalität, wie wir sie hier verstehen wollen, ist keine unstrukturierte und intrinsische Eigenschaft der Dinge, sondern wesentlich eine strukturierte und relationale „Gesamteigenschaft“ eines Prozesses der (die synthetisierte Klasse der) Objekte hervorbringt. Solche Prozesse sind jedoch, erstens, situiert – sie passieren nicht einfach, sondern unterliegen biologische, sozialen oder kulturellen Einschränkungen. Zudem haben wir, zweitens, gesehen, dass die Objekte, die als Resultate dieser Prozesse betrachtet werden können, eine ganz eigene mereologische Struktur besitzen. Drittens, verweisen diese Objekte auf externe (reale) Gegenstände [8].

- Ein erstes Beispiel liefert die Wahrnehmung von Farben. Der Kristall ist nicht „blau an-sich“, sondern er erscheint blau, sofern er als Objekt wahrgenommen wird, also sofern er den (mereotopologisch emergenten) Inhalt eines Wahrnehmungsprozesses bezeichnet. Daraus folgt aber nicht, dass einem Gegenstand von „dem Bewusstsein“ irgendetwas hinzugefügt würde. Ganz ähnlich lässt sich das Beispiel der glühenden roten Scheibe aus Abb. 6.3 erklären. Es gibt ja unabhängig vom Kontext, in dem Wahrnehmung stattfindet, gar keine Scheibe, die überhaupt repräsentiert werden könnte und der dann eine Farbe hinzugefügt würde.

Aus diesem Grund soll hier ein wesentlich holistisches Bild vorgeschlagen werden: Erstens leitet sich die Phänomenalität nicht aus den „Einzeleigenschaften“ eines bestimmten Objektes oder einer Klasse von diesen ab, sondern hat mit deren Einbindung in einen verweisenden Kontext zu tun; und zweitens impliziert die Eigenschaft, dass einzelne Objektteile selbst als verweisende Objekte erscheinen können – d.h., dass immer auch den Objektteilen eine ihnen eigene Phänomenalität zugesprochen werden kann, wenn diese als verweisende Objekte in Erscheinung treten, also etwa wenn wir unsere Aufmerksamkeit auf diese richten und sie selbst zu einer Ganzheit werden –, dass Phänomenalität *keine kompositionale Struktur* aufweist, sondern nur, dass *ausgehend* von einer (phänomenalen) Ganzheit (phänomenale) Teile erzeugt (oder im Formalismus der PEM: „projiziert“) werden können – vgl. auch die Diskussion zur „Differenzierung des Erlebens“ bei Prentner (2016, S. 119ff.) für ein Modell, das phänomenales Bewusstsein anhand dreier Modalitäten oder Momente charakterisiert, die sich im Rahmen ihrer Situierung in biologischen, psychologischen und sozialen Kontexten (oder alternativ gesprochen: in ihren materiellen, funktionalen und systemischen Bedingtheiten) entwickeln.



- Als zweites Beispiel lässt sich der stufenweise Verstehensprozess, wie wir ihn etwa beim Vollzug eines Beweises oder der sinnstiftenden Verkettung bedeutender Einzelereignisse erleben, anführen. Oftmals wird davon gesprochen, ein Ergebnis „noch nicht richtig verstanden“ zu haben, obwohl man jeden der einzelnen Schritte nachvollziehen kann und deren Korrektheit einsieht. Das richtige Verständnis und die damit einhergehende phänomenale Veränderung treten erst dann ein, wenn man das Ergebnis „in seiner Gesamtheit“ erkannt hat, wenn man es also – um dies in der hier eingeführten Terminologie auszudrücken – als selbstständiges Objekt wahrgenommen hat. Dies wirkt aber wiederum zurück auf die Bedeutung der einzelnen Schritte, da diese nun als Teile einer Ganzheit erkannt wurden. Die Phänomenalität der Ganzheit ist *rückgekoppelt* an die Erscheinungsweise der Teile, aus deren Gesamtheit sie hervorgeht. Dies ist aber keine *magische* Eigenschaft, die allein menschliches Bewusstsein als „echt“ intentional ausweisen würde.

Ob es sich bei beiden Beispielen um mehr als nur eine Veranschaulichung eines Gedanken handelt, kann unter mehreren Gesichtspunkten untersucht werden, z.B.

phänomenologisch: Was ist die entsprechende (Teil-Ganzes-)Struktur, die in einem *konkreten Wahrnehmungsakt* realisiert wird?,

evolutionär: Wie entstehen *Typen solcher Strukturen* im Zuge der Evolution eines Netzwerkes von Gegenständen?,

korrelativ: In welcher Konstellation treten verschiedene Objekte innerhalb des *konkreten* Falls verschiedener Bezugnahmen auf? Oder: Wie *korreliert* dabei eine Gruppe von Objekten zu einer anderen?,

psycho-physisch: Lassen sich *Verallgemeinerungen* treffen? Lässt sich gar eine Art „Brückenprinzip“ [9] formulieren?

Dies sind *keine reduktiven* Fragestellungen, was verdeutlichen soll, dass es sich nicht um eine „Erklärung des phänomenalen Bewusstseins“ o.ä. handelt, wie es die meisten Vertreter einer naturalistischen Theorie des Bewusstseins vor Augen haben. Genauso wenig aber handelt es sich dabei um eine (philosophische) Theorie, die uns darüber aufklärt, „wie es für jemanden ist, eine Erfahrung zu machen“. Erfahrungen sind vielmehr als irreduzibel und unhintergebar angenommen; gleichwohl sollen sie nicht als außer-natürlich betrachtet werden.

Eine besondere Schwierigkeit scheint dabei darin zu bestehen, dass die qualitativen Inhalte des Bewusstseins nicht experimentell zugänglich zu machen sind. Selbst wenn

wir die Schilderungen von Personen betrachten, die uns erzählen, wie es sich anfühlt, eine bestimmte Sinneswahrnehmung zu durchlaufen – also etwa einen als hell empfundenen Punkt in kurzer Distanz zu sehen oder zwei als harmonisch empfundene Töne zu hören, die zeitlich versetzt erklingen – lässt sich nicht mit Bestimmtheit sagen, ob es sich dabei um die empfundene Qualität der Wahrnehmung selbst, um die Art und Weise des Berichtens von Wahrnehmungserfahrung oder um die Vergegenwärtigung der Erinnerung an eine solche handelt. Wie in Abschnitt 3.3 bereits ausgeführt, sind Erste-Personendaten nicht etwas rein Subjektives, und die Idee, Subjekte seien lediglich die Pole eines Wahrnehmungsprozesses, sprechen dagegen, *überhaupt irgendwo* Subjektives „in der Welt“ zu entdecken.

Allerdings bedeutet dies nicht, dass wir nicht versuchen könnten, Bewusstsein als Aktivität wissenschaftlich zu beschreiben. Wahrnehmung zeichnet sich z.B. auch durch dessen Strukturiertheit aus, welche etwa mithilfe psychophysischer Experimente untersucht werden können.

Während das Sehen zum Beispiel die räumliche Struktur des Wahrnehmungsaktes verdeutlicht, also die Tatsache, dass wir visuelle Objekte lokalisieren (oder „in den Raum werfen“), so verdeutlicht das Hören dessen zeitliche Struktur: Aus einer Abfolge wahrgenommener Töne entsteht die Wahrnehmung einer Melodie. Anders als bei der räumlichen Struktur der wahrgenommenen Objekte ist diese zeitliche Struktur jedoch unweit schwerer auf eine „tatsächliche“ zeitliche Abfolge externer Gegenstände reduzierbar und stellt paradigmatisch ein Strukturelement des Wahrnehmungsaktes dar, das erst mithilfe geeignet formalisierter Analogieschlüsse mit außerpsychischen Daten verglichen werden kann (für Details siehe z.B. Sieroka 2015). Wie lässt sich diese Strukturiertheit näher verstehen? Anders als in den Ausführungen zur Aspektgestaltung der Wahrnehmung, bei denen es darum ging, dass uns in der Wahrnehmung von Objekten nur dessen Teile präsent sind (= Strukturiertheit der Objekte), gilt hier, dass Bewusstsein, welches wir mit der Form eines geschlossenen Prozesses identifiziert haben, selbst wiederum als differenzierbare Ganzheit anzusehen ist (= Strukturiertheit einer Aktivität). Ganz analog zum Fall der Objekt-Teil-Beziehung, entspricht dies keiner offensichtlichen (und unmittelbaren) Einsicht, sondern zeigt sich erst in einer reflexiven Betrachtung.

Folgt man etwa Husserls Phänomenlogischer Analyse des Zeitbewusstseins (PIZ), so lässt sich der gesamte Bewusstseinsakt als zusammengesetzt aus den Momenten der „Protention“, der „Retention“ und der „Urimpression“ betrachten. Keiner dieser Momente stellt für sich genommen ein (intentionales) Objekt dar oder verweist auf ein solches. Erst in ihrer Gesamtheit konstituieren sie den Wahrnehmungsakt, dessen Objekt dann ein zeitliches ist mit all seinen Qualitäten.

Ähnlich wie bei den visuellen Objekten können wir zum Beispiel Töne als im Raum

stehende lokalisieren, mit etwas Übung können wir dies sogar zur Orientierung nutzen und erschaffen dadurch einen „gehörten Raum“. Dabei werden einige einem so gebildeten Raumbegriff vielleicht einen anderen Stellenwert als dem visuellen zuweisen. Doch letztlich unterscheidet sich dieser von jenem nur dadurch, dass wir dem einen eine äußere Realität zusprechen, die an den visuellen Objekten „haftet“, während der andere „bloß vorgestellt“ oder „nur gedacht“ wird. Realitätszuschreibungen folgen jedoch einer Eigengesetzlichkeit, die nicht in den Gegenständen selbst zu finden ist. Sobald wir gewillt sind, zuzugeben, dass es sich in beiden Fällen „nur“ um einen im Bewusstsein konstruierten Raumbegriff handelt, werden wir auf die Frage, wie es sich wohl anfühle, wie eine Fledermaus räumliche Objekte mittels Echo zu orten, antworten: „Genauso so wie es sich anfühlt, sie zu sehen – nur anders“.

## Anmerkungen zu Abschnitt 6

[1] In diesem Kontext lässt sich auch folgende Bemerkung von Leonard & Goodman (1940, S. 55) verstehen:

It becomes clear that the practice of supposing that things are what the x's and y's of Principia mathematica denominate and that qualities are necessarily to be interpreted as logical predicates thereof, rather than vice versa, is purely a matter of habit. The dispute between nominalist and realist as to what actual entities are individuals and what are classes is recognized as devolving upon matters of interpretative convenience rather than upon metaphysical necessity.

Explizit spielt dies auf den Streit zwischen Realisten und Nominalisten an. Dem zugrunde, so ließe sich vermuten, liegt jedoch der Verdacht, wonach Metaphysik weniger mit einer nicht-empirischen und nur der Vernunft zugänglichen Erkenntnis von Notwendigkeit zu tun hätte, als vielmehr mit der Interpretation formal darstellbarer Systeme.

[2] Und zwar wie folgt:

- *(Onto-)logischer Antirealismus*: Es gibt keine *rein logischen* Gründe, welche die Annahme eines physikalistischen Realismus erzwingen würden. Ob wir an der Realität des Selbst oder des Gehirns zweifeln, ist rein logisch betrachtet letztlich Geschmackssache. Es müssen daher noch weitere Kriterien berücksichtigt werden.
- *Semantischer Antirealismus*: Eine begriffliche Analyse der Systemtheorie – also dem Fundament von Metzingers Theorie – steht noch aus, d.h. es ist gar nicht klar, ob sie überhaupt auf echte Gegenstände verweist, oder eher eine rein epistemische Funktion hat, z.B. ein Ordnungsprinzip gewisser wissenschaftlicher Aussagen darstellt. Systemische Beschreibungen wie diejenigen der systemischen Kognitionswissenschaften sind vielleicht sogar nur Platzhalter für mechanistische Erklärungen der molekularen Neurowissenschaften (vgl. den „erbarmungslosen Reduktionismus“ von Bickle 2006), welche selbst wiederum nur bestimmte Ausdrucksweisen innerhalb eines Sprachspiels darstellen könnten und unter Umständen gar nicht auf etwas Reales verweisen würden.
- *Methodologischer Antirealismus*: Aussagen über das Gehirn können, der physikalistischen Mehrheitsmeinung zufolge, letztlich auf Aussagen über physikalische Objekte (z.B. Elementarteilchen) reduziert werden. Dem Gehirn kommt daher, außer vielleicht aus pragmatischen Gründen, keine wirklich fundamentale Rolle als *explanans* zu. Die Annahme seiner objektiven Existenz entspricht vielmehr einer gewissen „explanatorischen Haltung“ (Dennett 1987), die wir einnehmen müssen, da wir nicht in

der Lage sind, zufriedenstellende Erklärungen auf der Ebene von Elementarteilchen zu produzieren. Zudem sind die als letzte Träger jeder wissenschaftlichen Erklärung postulierten Teilchen kaum als Gegenstände im herkömmlichen Sinne zu verstehen (die Redeweise von „Teilchen“ ist irreführend). Außerdem ist es fraglich, ob sich der Erfolg, den elementaristische Konzeptionen in den physikalischen Wissenschaften zweifelsohne gehabt haben, auf andere Bereiche übertragen lässt.

- *Epistemologischer Antirealismus*: Erkenntnis ist immer vermittelt und begrifflich situiert, so auch jede Erkenntnis über Gehirne oder andere physikalische Gegenstände. Keine wissenschaftliche Erkenntnis kann selbst darüber entscheiden, ob hinter ihr ein realer Gegenstand steht und welcher Art dieser ist: Vom epistemologischen Standpunkt aus betrachtet sind die Objekte der Physik ebenso real wie die „Götter Homers“ (Quine 1951) – oder eben das Selbst.

[3] Dies ist natürlich nicht der Platz für eine umfassende Abhandlung über klassische Indische Philosophie. Für einen Überblick siehe etwa (Lorenz 1998). Zu betonen sei etwa, dass es keine einheitliche Hinduistische oder Buddhistische Lehre gibt, sondern eine Vielzahl rivalisierender Schulen. Noch dazu gilt es, sprachliche Besonderheiten zu berücksichtigen, die ganz unterschiedliche Interpretationen implizieren. Bei der Überlegung indischer Philosophen, expliziter als in vielen Philosophien des Westens, spielt zudem Ethik eine zentrale *metaphysische* Rolle und es ist unmöglich, hier exakt zwischen Religion und Philosophie zu unterscheiden.

[4] „Allein die Verbindung (*conjunctio*), eines Mannigfaltigen überhaupt, kann niemals durch Sinne in uns kommen, und kann also auch nicht in der reinen Form der sinnlichen Anschauung zugleich mit enthalten sein, [denn diese beruhe vielmehr auf einer] Verstandeshandlung, die wir mit der allgemeinen Benennung *Synthesis* belegen würden, um dadurch zugleich bemerklich zu machen, daß wir uns nichts, als im Objekte verbunden, vorstellen können, ohne es vorher selbst verbunden zu haben“, KrV, B 130.

[5] Man könnte den Eindruck gewinnen, dass es sich dabei um eine modifizierte Form von Brentanos These handelt, wonach bewusste geistige Zustände charakterisiert sind durch deren *Intentionalität* (siehe auch Crane 2001a). Viele moderne Philosophen, wie etwa Chalmers (1996), würden dem widersprechen und stattdessen *Phänomenalität* als wesentliches Merkmal des bewussten Geistes betrachten. Wieder andere setzen auf den Begriff der *Subjektivität* (Gallagher & Zahavi 2008). Letztlich lassen sich diese definitorischen Streitigkeiten in dem hier vorgestellten Modell auflösen, ohne dabei eine der metaphysischen Hintergrundannahmen – Materialismus, Dualismus, Idealismus – zum Sieger zu erklären: „Erfahrendes Subjekt“ und „intentionales Objekt“ werden dabei durch die Rede von zwei Polen eines natürlichen Prozesses ersetzt, der wesentlich durch seine Phänomenalität gekennzeichnet ist. Die reiche Struktur des menschlichen Bewusstseins, welches den Ausgangspunkt Phänomenologischer Analysen darstellt, ist davon abgeleitet.

[6] Außer natürlich wenn wir meinen, dass es sich bei der Einheit des Bewusstseins um eine Eigenschaft der *res cogitans* handelt. Dann scheint eine allgemeine Summationsbeziehung angemessen, welche die rechten „Pakete“ des Mentalen herausgreift. Immerhin ließe sich dann noch folgern, dass zumindest gelten müsste, dass

$$\forall\phi(\phi \circ \varphi_i \Leftrightarrow \phi \circ \psi), \quad (6.11)$$

d.h., dass die Definition der mereologischen Summe erfüllt wäre, wonach jeder Zustand, der mit (mindestens) einem der phänomenalen Zustände überlappt, auch mit  $\psi$  überlappen würde. Alle Zustände  $\phi$  entstammen jedoch – dies ist ja die dualistische Annahme – der *res cogitans* und obiges Kriterium könnte maximal als Abgrenzungskriterium dienen, das zwei Träger von Bewusstsein innerhalb der denken Substanz unterscheidet.

[7] Natürlich entspricht unsere Darstellung wiederum einer Verobjektivierung dieses Prozesses. Genauer müsste es daher heißen, dass  $\psi_A$  nicht Objekt derselben Stufe oder desselben logischen Typus ist wie  $x$  oder  $y_i$ . Dennoch kann  $\psi_A$  selbst mereologisch analysiert werden; so entspricht etwa die Bezeichnung als „Moment“ ganz der Husserl’schen Terminologie, wonach „Momente“ die „ontologisch abhängigen“ Teile einer Gesamtheit (in unserem Fall des gesamten Wahrnehmungsaktes) bezeichnen, vgl. (LU, Bd. II: §3) und Simons (1987, Kap. 8).

[8] Gewisse Ähnlichkeit scheinen hier zum „starken Intentionalismus“ von Crane (2001a, S. 77ff.) zu bestehen, der die Phänomenalität des Bewusstseins aus dem Inhalt und der Art des Gegebenseins intentionaler Objekte abgeleitet sieht. Allerdings folgt aus dieser Lesart eine potentielle Schwierigkeit, auf die Bayne (2010, S. 70f.) hinweist: Auch wenn sich der Inhalt des Bewusstseins (als einheitliches Objekt) aus mehreren intentionalen Objekten zusammensetzt, so ist dennoch nicht klar, dass jener selbst wiederum ein intentionales Objekt mit derselben doppelten Struktur Inhalt/Art des Gegebenseins ist. Dies allerdings würde die Bedingung dafür sein, dass sich Phänomenalität als abgeleitete Eigenschaft verstehen lässt.

[9] Dies ist angelehnt an einen Vorschlag von Chalmers (2004), bezieht sich jedoch nicht wie dort auf die mögliche Verbindung von Erste- und Dritte-Personenbeschreibungen (vgl. dazu die Ausführungen in Abschnitt 3.3), sondern auf die Korrelation von Objekten.

## 7 Résumé: Bewusstsein und Natur

*Die Naturwissenschaft beschreibt und erklärt die Natur nicht einfach, so wie sie „an sich“ ist. Sie ist vielmehr ein Teil des Wechselspiels zwischen der Natur und uns selbst.* (Werner Heisenberg: Physik und Philosophie)

### 7.1 Eine Emergenztheorie des Bewusstseins

Auf die einleitende Frage, ob überhaupt in einer wissenschaftlichen Weise vom Bewusstsein gesprochen werden könne, wurde im Verlaufe dieser Arbeit grob zwischen zwei Ansätzen unterschieden, die beide die realistische Ansicht teilen, wonach Bewusstsein als wirkliches Naturphänomen verstanden werden müsse:

**Erstens, Theorien von Bewusstsein als Abbildungen** oder Modelle externer Gegenstände. Dabei spielt es nur auf den ersten Blick eine Rolle, ob hier von den Zuständen eines einzelnen Objekts, einer ganzen Klasse, einer abstrakten Struktur oder einer beobachtbaren Dynamik die Rede ist. Solche Objekte (Klassen, Strukturen, Dynamiken) bezeichnen Dinge außerhalb eines Erfahrungsprozesses.

Die entsprechenden Formulierungen stellen alle, mit mehr oder weniger Raffinesse, gegenständliche Auffassungen dar, die Denken und Bewusstsein als identisch zu oder abgeleitet von einer physikalischen Struktur (Physikalismus) oder einer Funktion auf solchen Strukturen (Funktionalismus [1]) denken.

Charakteristisch für derartige Theorien ist zudem, dass unsere Intuitionen bezüglich des Bewusstseins und die Rolle der Phänomenologie in den mathematischen Beschreibungen der Theorie aufgehen, sei es, weil es zu einem Identitätsschluss gekommen ist, der bestimmte theoretische Entitäten mit vermeintlich intuitiv erfassten Eigenschaften des Bewusstseins identifiziert, oder sei es, weil eine vollständige Theorie in der Lage ist, jegliche Intuitionen über das Bewusstsein zu ersetzen oder als Annäherung an das von der Theorie spezifizierte Verhalten auszuweisen. Je nachdem, wie hier methodisch vorgegangen wird, können diese Ansätze als *top-down*, von der Phänomenologie des Bewusstseins ausgehende, oder als *bottom-up*, von der Darstellung materieller Phänomene ausgehende, Verfahren unterschieden werden (Tab. 7.1, links).

Nun ist die Literatur bekanntlich voll von Problemen, die mit solchen Theorien einher gehen. Exemplarisch wurden etwa die Inversionen von Bewusstseinsinhalten, die Beschreibung von (bewussten) Handlungen, die Rolle des Bewusstseins in funktionalistischen Theorien und das Rätsel der (bewussten) mentalen

Verursachung kurz diskutiert. Das Ziel dieser Arbeit lag jedoch nicht darin, für die Gültigkeit des einen oder anderen Arguments oder gegen eine ganz bestimmte Position zu argumentieren, sondern die Grundproblematik einer Voraussetzung zu beleuchten, die sich darin äußert, dass so vom Bewusstsein gesprochen werden könne, wie von anderen Gegenständen „da draußen“ auch, seien das nun Elektronen, Zahnräder, Wirbelstürme oder Gehirnzellen. Doch wie ließe sich eine Theorie von Bewusstsein anders verstehen, als dass wir mit ihrer Hilfe das Bewusstsein als einen (natürlichen) Gegenstand „erschauen“ ( $\theta\epsilon\omega\rho\epsilon\acute{\iota}\nu$ )?

**Zweitens, Theorien von *Bewusstsein* als Form** einer Aktivität oder eines Prozesses. Bewusstsein, dieser Ansicht zufolge, ist kein Gegenstand, der in der Welt vorkommen würde, und auch – etwas abstrakter gesprochen – kein Objekt wie es Zahlen, Geschichten und Funktionen sind, die vielleicht als Gegenstände ebenfalls nicht in der Welt vorkommen, aber diese im weitesten Sinne strukturieren.

Nun scheint es eine Reihe von Prozessen zu geben, auf welche dies ebenfalls nicht zutrifft und die am ehesten als relationale charakterisiert werden können – beispielhaft die Wahrnehmung, für die gilt, dass deren Inhalte weder auf die „äußeren“ Gegenstände noch auf die „inneren“ Arten der Darstellung dieser Gegenstände reduziert werden kann.

Es ist allerdings ein weiter Weg, der von diesen sehr allgemeinen Betrachtungen hin zu einer Theorie von Bewusstsein führt. Ein erster Schritt besteht in einer formalen Darstellung relationaler Phänomene, in der die Eigenschaften von Objekten als Ergebnisse von wechselseitigen Beziehungen rekonstruiert werden.

Paradigmatisch hierfür wurde in dieser Arbeit die Teilseinsrelation untersucht, die vielleicht als Ausdruck der elementarsten relationalen Eigenschaft überhaupt betrachtet werden kann. Dabei wurde zuerst eine mögliche axiomatische Rekonstruktion der klassischen Formulierung von Teilsein (KEM) und deren Erweiterung (durch Atome, Funktionen, Topologien) besprochen. Anschließend wurde eine Interpretation von Mereologie nahegelegt, die diese als formale Darstellung der Art und Weise auffasst, wie Zeichensysteme entstehen oder verwendet werden, anstelle einer stark metaphysischen Lesart, die in der Mereologie ihren formalontologischen Zweig sieht, der zum Beispiel auf die Darstellung der Struktur eines Ding-an-sichs oder von Substanzen abzielt.

Etwas expliziter gemacht wurde dies im Modell der „Projektiven Mereologie“, in dem Teilsein als bedingte (oder kontextuelle) Eigenschaft angesehen wird: Keinem Objekt kommt „für sich“ die Eigenschaft zu, ein Teil oder ein Ganzes



zu sein, sondern immer nur relativ zur Existenz eines Projektionsoperators, der ein Objekt als Teil eines anderen „erzeugt“.

Ausgehend davon wurde etwa die Unterscheidung eingeführt zwischen „wohl-definierten Objekten“, die auf konkrete Gegenstände verweisen, und zwischen „unselbstständigen Teilen“, die auf keine solchen Gegenstände außerhalb des Systems der Teil-Ganzes-Beziehungen, in denen jene Teilobjekte stehen, verweisen. Der Projektorformalismus ermöglicht eine relativ einfache Darstellung der Teil-Ganzes-Beziehungen als gerichteten Graphen und, in Verbindung mit mereotopologischen Überlegungen, eine Darstellung von Emergenzphänomenen. Über Bewusstsein, so die Überlegung, soll mithilfe dieser theoretischen Werkzeuge auf neuartige Weise formal gesprochen werden können. Dabei geht es jedoch weniger um einer Darstellung des Bewusstseins als solchem, sondern darum, diejenigen Prozesse etwas näher zu beleuchten, die wir für gewöhnlich mit Bewusstsein in Verbindung bringen.

Mit diesen unterschiedlichen Theoriekonzeptionen sind zwei unterschiedliche Positionen zur Emergenz verbunden:

1. Emergenz als die Verbindung von „Seinsweisen“. Dabei wird davon ausgegangen, dass Emergenz denjenigen Prozess bezeichnet, der verschiedene Arten oder Seinsweisen von Gegenständen hervortreten lässt oder ineinander überführt. Dabei lassen sich schwache und starke Ausprägungen unterscheiden, die entweder die sehr liberale Haltung vertreten, wonach die emergenten Eigenschaft einfach Systemeigenschaften bezeichnen, die zwar nicht bereits in den Systemteilen realisiert sind, aber dennoch in einer naturgesetzlichen Beziehung zu diesen stehen, oder die radikale Auffassung, wonach emergente Eigenschaften als prinzipiell irreduzible betrachtet werden müssen. Wenn nun Bewusstsein als solche emergente Systemeigenschaft betrachtet würde, stehen wir jedoch vor dem Dilemma, das die radikale Auffassung *zu radikal* erscheint, um von irgendwelchem Nutzen für die Wissenschaft zu sein und die liberale Haltung *zu schwach* ist, um typische Streitfragen, wie etwa das *hard problem* zu adressieren [2].

2. Emergenz als das Entstehen von Objekten. Dabei wird der Begriff der Emergenz nicht als ontologischer aufgefasst, sondern steht für die Herausbildung neuartiger Objektbeschreibungen innerhalb eines Systems von Verweisen. Ein solcher Emergenzbegriff wird bei genauerer Analyse wesentlich als kontextueller, struktur- oder musterrelativer zu verstehen sein.

Dies soll des Weiteren nicht so verstanden werden, dass Emergenz einen Vorgang bezeichnen würde, der sich losgelöst und „nur im Verstand“ vollzieht, was letzt-

lich wiederum einem dualistischen Bild entspräche – dem einer physikalischen Welt, die ihrer (emergenten) Erscheinung im Geiste eines Beobachters gegenübersteht. Emergenz bezeichnet viel mehr den Vorgang der Entstehung von Wahrnehmungsobjekten innerhalb von (natürlichen) Prozessen, die sich auf ihre Umwelt beziehen. Eine Emergenztheorie versucht, dies zu explizieren und wo möglich zu formalisieren.

Theorie von <i>Bewusstsein</i> als (emergenter) Gegenstand	Emergenztheorie des <i>Bewusstseins</i> als Form eines Prozesses
→ <i>top-down</i> : Bewusstsein wird auf theoretische Entitäten oder Elemente von Mengen <i>abgebildet</i> .	1. <i>semiotisch</i> : Objekte als Zeichen für Gegenstände.
oder	2. <i>mereologisch</i> : Struktur von Objekten.
→ <i>bottom-up</i> : Intuitionen über Bewusstsein werden durch eine Theorie der neuronalen Dynamik <i>ersetzt</i> .	3. <i>phänomenologisch</i> : Erscheinungs- weise von Objekten.
Emergenz: Verbindung zwischen Gegenstands- bereichen oder Seinsweisen	Entstehung von Objekten, relativ zu einem Kontext

**Tabelle 7.1:** Eine kurze Übersicht verschiedener Ansätze zu einer Theorie von Bewusstsein. Die derzeit gängigen Theorien entsprechen gegenständlichen Auffassungen. Ihre Schwierigkeiten wurden in den Abschnitten 2 und 3 behandelt. Eine alternative, prozessuale Auffassung von Bewusstsein wurde in dieser Arbeit vorgestellt, wobei auf die semiotischen, mereologischen und phänomenologischen Aspekte näher eingegangen wurde (Abschnitte 3-6). In beiden Auffassungen wird manchmal vom Begriff der Emergenz gesprochen.

Zusammen mit den Überlegungen zur Mereologie ergibt sich eine **Emergenztheorie des Bewusstseins**, welche die Inhalte des Bewusstseins, als wohldefinierte Objekte betrachtet, die zu ihren Teilen in einer Teil-Ganzes-Beziehung stehen und gleichzeitig auf Gegenstände außerhalb der Erfahrung verweisen. Emergenz bezieht sich nicht auf das Bewusstsein selbst, welches nicht als Gegenstand „in der Natur“ sondern als Form desjenigen Prozesses betrachtet wird, innerhalb dessen Objekte hervortreten. Dieser lässt sich anhand dreier Merkmale skizzieren (Tab. 7.1, rechts): Zum einen handelt es sich dabei um den Verweischarakter der Objekte, was sowohl durch empirische Untersuchungen als auch durch grundlegende Überlegungen zur Rolle des

Bewusstseins (Abschnitt 3) nahegelegt wurde. Daneben sind diese Objekte auf eigentümliche Weise strukturiert, wobei hier primär auf die mögliche mereologische Charakterisierung dieser Objekte eingegangen wurde (Abschnitt 5 und 6). Schließlich gilt es – für viele zeitgenössische Denker *das* Merkmal des Bewusstseins –, ihre qualitative Erscheinungsweise zu begreifen (Abschnitt 6). Etwas expliziter wurde dies anhand dreier formaler Darstellungen der Wahrnehmungsphänomenologie ausgeführt:

- Die Aspektgestaltung des Wahrnehmungsaktes
- Die Einheit der Wahrnehmungsobjekte
- Die Phänomenalität („what-it-is-likeness“) der Wahrnehmung

Dabei wurde ein Modell vorgestellt, das sich grob an der Analyse der Wahrnehmung als „Abschattungsprozess“ orientiert und diese als synthetische Leistung identifiziert, deren Ergebnisse phänomenal als Ganzheiten charakterisiert werden können, die sich aus physikalisch beschreibbaren Teilen zusammensetzen. Die Qualitäten der Objekte ergeben sich dabei aus ihrer Lage innerhalb des Verweissystems, in welches sie eingebettet sind (und sind in diesem Sinne *nicht intrinsisch*).

Im Anschluss soll eine solche Emergenztheorie des Bewusstseins noch in den Kontext einiger zeitgenössischer naturalistischer Projekte gestellt werden.

## 7.2 Brücken zur Wissenschaft

### Physik und naturalisierte Phänomenologie

Die Idee, Phänomenologie weniger als rein innerphilosophische Tradition zu begreifen und diese stattdessen eng an naturwissenschaftliche Verfahren heranzuführen, kann als „naturalisierte Phänomenologie“ bezeichnet werden. Das bekannteste Beispiel ist wohl die Neurophänomenologie Francisco Varelas. Beispielhaft wurden dabei Husserls Analyse des Zeitbewusstseins aus den PIZ und eine Beschreibung der neurobiologischen Grundlage temporaler Reizverarbeitung als sich gegenseitig bedingende („*mutually constraining*“) Darstellungen betrachtet (Varela 1999). Neben Varela gibt es auch weitere Vertreter einer naturalisierten Phänomenologie, die statt der neurophysiologischen Ebene auf systemisch-kognitiver (van Gelder 1999) oder überhaupt zellulärer (Thompson 2007) Ebene ansetzen. Auch die Idee, unser bewusst wahrgenommenes affektives Erleben würde aus einer systemischen Verstärkung einer bereits zellulär realisierten Empfindung resultieren (Damasio & Carvalho 2013),

könnte zumindest als phänomenologisch-inspirierte aufgefasst werden.

Für das Projekt einer naturalisierten Phänomenologie, gleich auf welcher Ebene nun angesetzt wird, ist die Formulierung geeigneter mathematischer oder anderer formaler Modelle zentral, da diese das Verbindungsglied zwischen der Phänomenologischen Analyse und der Darstellung physikalisch-biologischer Prozesse bildet. Dass dabei der Theorie dynamischer Systeme eine exklusive Rolle zukommen würde, die quasi „in der Natur der Sache“ läge, ist eine Vorannahme, die auf einer vermuteten Verwandtschaft zwischen Systemtheorie und dem Aktivitätscharakter des Bewusstseins gegründet ist. Rein pragmatisch betrachtet, entspricht dies einer vorschnellen Einengung auf ein bestimmtes mathematisches Modell, die übersieht, dass solche Modelle ganz generell als ontologisch-neutrale Werkzeuge zur Beschreibung psychischer und physischer Phänomene betrachtet werden können (vgl. dazu insbesondere Sieroka 2015, Kap. 9.3).

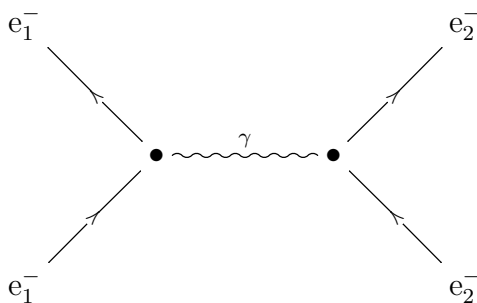
Nun gewinnt man bei allzu schnellem Hinsehen leicht den Eindruck, dass Phänomenologie lediglich als deskriptive Vorstufe zu neuro- oder kognitionswissenschaftlichen Experimenten zu betrachten wäre, in deren Design mehr oder weniger detailreiche Erkenntnisse einer erst-personalen Betrachtung des Bewusstseins einfließen würden. Aus der Perspektive einer Emergenztheorie des Bewusstseins liegt es auf der Hand, dass ein solches Verständnis zurückzuweisen ist, da es ja gerade darum gehen sollte, wie denn überhaupt Objekte als Bewusstseinsinhalte auftreten, oder im Husserl'schen Jargon gesprochen, wie es denn überhaupt zur Invarianz gewisser Inhalte der bewussten Wahrnehmung kommt. Aber genau dies wird in der Theorie dynamischer Systeme genauso wenig manifest wie in anderen mathematischen Modellen, die ihren Fokus auf zelluläre oder neurobiologische Physiologie legen.

Dies ist ein Punkt, der bereits von Varela (1998) selbst impliziert wird, wenn er darauf hinweist, dass bewusste Erfahrung („*experience*“) nicht Effekt oder Begleiterscheinung eines ganz bestimmten, objektiv beschreibbaren physikalischen Vorganges darstellt, sondern den *Ausgangspunkt* einer jeden (letztlich auch wissenschaftlichen) Tätigkeit markiert (vgl. dazu auch Bitbol 2002 und Zahavi 2013): Es geht weniger darum, wie bestimmte Objekte des Bewusstseins in ihrer vermeintlich unmittelbaren Qualität zu beschreiben wären, als vielmehr darum, wie diese Objekte überhaupt entstehen können.

Dagegen ließe sich vielleicht einwenden, dass es sich beim oben Gesagten höchstens um interessante Spekulationen handle, die von philosophischem Interesse sind, aber wenig mit den experimentellen und quantitativen Studien zu tun haben, welche die naturwissenschaftliche Verfahrensweise auszeichnen würde. Anstelle von mysteriösen philosophischen Ausführungen wären Erklärungen zweckmäßiger, die diese vermeintlich „außer-natürlichen“ und unklaren Prozesse der Verobjektivierung auf

ausgewählte Naturprozesse, wie sie durch Physik, Chemie und Biologie beschrieben werden, zurückführen. Und letztlich entspräche die Frage nach den Objekten, vielleicht nur der Frage nach der rechten Auflösung: Wenn wir nur genau genug hinsähen, würden wir schon „die Objekte“ ausmachen können und damit auf die „eigentlich“ existierenden physikalischen Gegenstände schließen können.

Doch dies bleibt eine eher zweifelhafte Heuristik, die von der physikalischen Situation nur bedingt gedeckt wird. Richten wir unseren Blick ins „Innere der Natur“, beginnen sich die ursprünglich gegebenen Objekte aufzulösen. Im Nachhinein lassen sich die uns so vertraut erscheinenden Objekte als hochstufige Verdichtungen fundamentalerer Prozesse interpretieren. Doch diese Prozesse selbst sind, gemessen an der alltäglichen Erfahrung, höchst sonderbar, was wie folgt illustriert werden kann: Mithilfe der sog. Feynman-Diagramme können die elementaren Prozesse der Wechselwirkung bildhaft dargestellt und letztlich quantitativ bestimmt werden (Feynman 1949). Dabei werden Interaktionen (beispielsweise zwischen zwei Elektronen) als Austausch sog. „virtueller Teilchen“ beschrieben, die nur noch wenig gemein haben mit Objekten im üblichen (alltäglichen) Sinn. Allgemein gesprochen, werden in der Quantenfeldtheorie Teilchen als Anregungszustände von Feldern verstanden, deren Interaktion wiederum über virtuelle Teilchen dargestellt werden kann [3].



**Abbildung 7.1:** Feynman-Diagramm für einen Streuprozess zwischen zwei Elektronen  $e_j^-$ , der in erster Ordnung als Austausch virtueller Photonen ( $\gamma$ ) dargestellt wird. (Die Zeit verläuft nach oben.)

Aber auch wenn wir einfache chemische Reaktionen betrachten und ausgehend von ihnen zu immer komplexer werdenden Systemen fortschreiten – Reaktionsnetzwerke, Zellen, ganze Organismen –, stoßen wir auf zahlreiche Beispiele für die Entstehung qualitativ neuartiger Objekte, welche sich im Rahmen der Dynamik des Systems ausbilden. (Und dazu braucht es auch keine zusätzliche Zutat, die kausal ins Naturgeschehen eingreifen würde.)

Die primäre Herausforderung für das Projekt einer naturalisierten Phänomenologie besteht also darin, einen Formalismus zu finden, der versucht, einen solchen Ob-

jektivierungsprozess darzustellen, ohne, dass es so aussähe, als würde er nur die Dynamik bereits vorhandener „für-sich existierender“ Teile beschreiben (etwas, das auch als Kritik an einer ontologisch-naiven Lesart der Theorie dynamischer Systeme verstanden werden kann).

## Gehirn und Quantenbiologie

Die Notwendigkeit, quantenphysikalische Modelle zur Erklärung von Bewusstsein hinzuzuziehen, wird oft bezweifelt, nicht nur, weil jene neurobiologisch relevante Quanteneffekte postulieren, sondern auch, weil sie sich dabei oft an oder sogar jenseits der Grenzen der etablierten Quantentheorie positionieren. Die Faszination von quantenphysikalischen Modellen ist dennoch ungebrochen und wird vermutlich noch lange anhalten. Letztlich gibt es wenig Gründe, die dagegen sprechen, dass derartige Projekte nicht als Alternativprojekte zum neurobiologischen Dogma der Bewusstseinsforschung verfolgt werden sollten.

Der aktuell wahrscheinlich meist-diskutierte Ansatz ist die „Orch OR“-Theorie von Stuart Hameroff und Roger Penrose (2014), die, erstens, davon ausgeht, dass die Reduktion des quantenmechanischen Wellenpakets (der „Kollaps der Wellenfunktion“) einem objektiven Ereignis in der Natur entspricht („objektive Reduktion“), welches, zweitens, mit Momenten einer vor-bewussten, psychischen Aktivität einhergeht, die schließlich, drittens, von biologischen Strukturen, den Mikrotubuli, so „orchestriert“ werden, dass daraus eine Form von Bewusstsein hervorgeht. Alle drei Punkte sind umstritten, der erste weil er eine realistische Auffassung der Quantenmechanik nahelegt, die auf eine bestimmte (unbestätigte) Theorie der Quantengravitation verweist, der zweite, weil er eine Form des Panpsychismus postuliert und der dritte, weil er die Existenz von quantenmechanischen Kohärenzen im Gehirn postuliert. Jeder dieser Punkte für sich wäre Gegenstand heftiger Kontroversen, aber zusammengenommen ergibt sich eine explosive Mischung, welche für viele die Grenzen der Wissenschaftlichkeit sprengt.

Aus der Sicht einer Emergenztheorie des Bewusstseins sind insbesondere die beiden ersten Annahmen problematisch, welche die realistische und letztlich gegenständliche Auffassung nahelegen, wonach (Vor-)Bewusstsein mit bestimmten objektiven Vorgängen „in der Welt“ zu identifiziert sei. Es ist jedoch interessant, aus der Sicht von Penrose (1994) die Hauptmotivation für die Notwendigkeit einer solchen Annahme kurz zu betrachten: die Existenz nicht-algorithmisch beschreibbarer Vorgänge in der Natur. Penroses vielleicht bekanntestes Beispiel stützt sich dabei auf eine Interpretation des Unvollständigkeitssatzes von Kurt Gödel, eine Interpretation, die aber umstritten ist (siehe etwa Feferman 1995). Ein etwas anschaulicheres Beispiel, dass

zumindest der Idee nach dem Penrose'schen Gedankengang entspricht, haben wir in Abschnitt 6.3 kennen gelernt, als wir die Phänomenologie eines Verstehensprozesses untersucht haben. Dabei ging es darum, dass die Bedeutung eines mathematischen Beweises manchmal auch dann nicht erkannt wird, wenn die einzelnen Schritte innerhalb einer Beweiskette vollständig nachvollzogen worden sind: Das letzte „Aha-Moment“ scheint nicht das Produkt eines regelhaften Verfahrens zu sein. Anders als bei Penrose wurde dieses jedoch mit der Emergenz eines Objektes im Bewusstsein verknüpft und nicht mit der Existenz von nicht-algorithmischen aber objektiven Quantenprozessen.

Die vielleicht meiste Aufmerksamkeit bekam allerdings die dritte Annahme, wonach kohärente Superpositionen von Mikrotubuli zu einer Art Verstärkung psychischer Aktivität und letztlich zum Entstehen von Bewusstsein führen würde, was Patricia Churchland (1998, S. 121) zum berüchtigten Verdikt verleitet hat, wonach die Postulierung von „Feenstaub in den Synapsen ungefähr genauso brauchbar zur Erklärung von Bewusstseinsphänomenen wäre wie Kohärenz von Mikrotubuli“. Auf der weniger brachialen Ebene gab es Schätzungen von Tegmark (2000) und Erweiterungen von Hagan *et al.* (2002), die sich mit der Frage nach den Zeitskalen der Dekohärenz solcher Prozesse beschäftigt haben, allerdings zu gegensätzlichen Schlüssen gekommen sind. Die Mehrheit der Physiker und Biologen stehen der Möglichkeit solcher Prozesse – selbst angesichts neuerer Erkenntnisse zur Aufrechterhaltung von Kohärenz und zur „Rekohärenz“ in verrauschten, thermischen Umgebungen (Briegel & Popescu 2014) – eher skeptisch gegenüber (was aber letztlich eine soziologische Beobachtung darstellt).

Dass Quanteneffekte im Gehirn dennoch eine Rolle spielen könnten, wurde unlängst auch vom renommierten Physiker Matthew Fisher (2015) vorgeschlagen, der anders als Hameroff & Penrose jedoch von der räumlich getrennten Verschränkung zwischen den Kernspins zweier Phosphationen ausgeht, die durch deren molekulare Umgebung isoliert werden können. Dies hätte zur Folge, dass die Verschränkung über genügend lange Zeit stabil bliebe und dem Gehirn als Art computationale Resource zur Verfügung stünde. Motivation lieferten experimentelle Untersuchungen, die Isotopeneffekte bei der Verabreichung von Antidepressiva beobachteten (Sechzer *et al.* 1986), was gemäß jeder klassischen Theorie neurochemischer Wirksamkeit ausgeschlossen wäre. Bei Fishers Theorie handelt es sich allerdings „nur“ um die Vorstellung, im Gehirn würden Quantenprozesse nutzbar gemacht werden, ähnlich wie in der Photosynthese, wo Quantenprozesse eine Rolle beim Energietransport spielen sollen (Sarovar *et al.* 2010), oder bei der Orientierung von Vögeln entlang des Erdmagnetfelds (Ritz *et al.* 2000). Im Gegensatz zur Theorie von Hameroff & Penrose gibt es keine Identifizierung zwischen Bewusstsein (oder einer Vorform des

Bewusstseins) mit einer „objektiven Reduktion“ von Quantenzuständen o.ä., was Fisher zum nicht-revisionären Vertreter einer als „Quantenbiologie“ (Flemming & Scholes 2014) bezeichneten Forschungsrichtung macht, die Quanteneffekte in biologischen Systemen nachweisen will.

Allerdings ist fraglich, ob die Möglichkeit, das Gehirn in Teilen als Quantencomputer zu beschreiben, uns mehr über das Bewusstsein verrät, als klassische computationale Theorien des Gehirns es tun, da auch ein derartiges Modell uns nicht klar machen könnte, wie oder warum es zu Bewusstseinsinhalten in Form verweisender Objekte kommt. Zwar würden quantenbiologische Theorien neuartige neuronale Operationen spezifizieren, aber sie wären in Bezug auf das Bewusstsein mit denselben Schwierigkeiten konfrontiert wie die bereits besprochenen funktionalistischen und computationalistischen Theorien auch.

In eine ganz andere Richtung geht ein Zweig der modernen Kognitionspsychologie, welcher versucht, formale Analogien zwischen quantenmechanischen und kognitiven Phänomenen nutzbar zu machen (Wang *et al.* 2013). Dabei geht es, anders als in den zuvor genannten Theorien, in erster Linie um strukturelle Eigenschaften kognitiver Prozesse wie etwa die Nicht-Kommutativität von Projektoren oder anderen Operatoren und weniger um die Existenz eines quantenphysikalischen Mechanismus im Gehirn.

Dies ist eng verbunden mit aktuellen Entwicklungen im Bereich des *artificial life* und des maschinellen Lernens, die suggerieren, dass ein exklusiver Fokus auf das Gehirn als „Substrat der Psyche“ unter Umständen dazu führt, dass wichtige theoretische Einsichten übersehen werden. Ein interessantes Beispiel, das dies deutlich macht, ist das Modell der „projektiven Simulation“ (Briegel & De Las Cuevas 2012), demgemäß das Verhalten eines Akteurs auf der Basis nicht-deterministischer Bewegungen innerhalb eines virtuellen Raumes („*memory space*“) bestimmt wird. Das Modell ist im doppelten Sinne projektiv: einerseits, weil es, metaphorisch gesprochen, dem Akteur ermöglicht, sich in vorstellbare Situationen hineinzusetzen [4], andererseits, nun wörtlicher genommen und in Übereinstimmung mit den Überlegungen aus Abschnitt 5, weil der Übergang zwischen zwei Episoden im „Gedächtnis“ des Akteurs als Projektion innerhalb eines virtuellen Raumes verstanden wird.

Eine jede Form von Situiertheit oder Reflexion im Verhalten des Akteurs wird dabei nicht durch symbolische Prozesse beschrieben, wie wir sie üblicherweise mit der Datenverarbeitung eines Computers in Verbindung bringen, sondern folgt aus den strukturellen Eigenschaften des Akteurs und der Rückkopplung an seine Umgebung. Etwas spekulativer sprechen Briegel & Müller (2015) sogar davon, dass hier die Geschichtlichkeit von Erfahrung und Identität sichtbar würde und letztlich



sogar bedeutsame Konsequenzen für das philosophische Problem der Willensfreiheit [5] gezogen werden können.

### Kognitionswissenschaften und theoretische Psychologie

Wie auch immer man zu solchen Spekulationen stehen mag, scheint in den Kognitionswissenschaften einer bestimmten Auffassung vermehrt Aufmerksamkeit zuzukommen, welche „Geist“ als (physikalisch-biologische) Aktivität versteht und nicht mehr mit der Instantiierung einer geeigneten Funktion gleichsetzt, die durch explizite input-output-Kopplungen definierbar wäre.

Dieser Umstand ist umso interessanter, als dass sie als kleinster gemeinsamer Nenner einiger sehr verschiedener Ansätze betrachtet werden kann, welche durchaus entgegengesetzte Grundannahmen über das Verhältnis zwischen Körper und Psyche treffen. Als Beispiele lassen sich hier die Theorie des prädiktiven Geistes (Hohwy 2013), der erweiterten und verkörperten Kognition (Varela *et al.* 1991, Clark & Chalmers 1998, Gallagher 2005) oder der symbolischen Wahrnehmung (Hoffman *et al.* 2015) aufzählen [6]. Daneben existiert eine Unzahl an Arbeiten, die das Für und Wider dieser Theorien abwägen (z.B. Hickok 2015, Michel *et al.* 2016) oder die Querverbindungen herstellen (z.B. Clark 2013, Gallagher & Allen 2016).

Eine strikte Kategorisierung wird dadurch erschwert, dass viele klassische philosophische Positionen von den oben genannten Vertretern teilweise entgegengesetzt bewertet werden (Internalismus – Externalismus, Repräsentationalismus – Direkter Realismus, Objektivismus – Konstruktivismus, etc.). Aus diesem Grund scheint es umso angebrachter zu sein, sich auf ein ganz grundsätzliches strukturelles Merkmal dieser Ansätze zu beschränken, welches anschließend in Verbindung zur hier vorgestellten Emergenztheorie des Bewusstseins gebracht werden soll.

Dabei kann vom Begriff eines „Organismus“ ausgegangen werden, der einen (z.B. statistisch-entropisch definierten) relativ stabilen Bereich beschreibt, welcher von seiner Außenwelt *informationell entkoppelt* ist, was bedeutet, dass jegliche Form von Wissen, die ein Organismus über seine Umgebung in Erfahrung bringen könnte, immer bereits einer Filterung, Modifizierung, oder artspezifischer Verarbeitung durch seine „Sinnesorgane“ ausgesetzt ist. Gleichzeitig sind Organismen jedoch *aktiv an ihre Außenwelt gekoppelt*, d.h. sie können auf ihre Umgebung einwirken. Organismen und deren Umgebung entwickeln sich parallel. Einerseits existieren also keine expliziten mentalen Repräsentation der Umgebung „im Geist“ des Organismus, andererseits ist die Möglichkeit zur aktiven Teilhabe am natürlichen Geschehen durch spezifisch dafür vorgesehene „Motororgane“ garantiert. Exemplarisch wurde dies etwa bereits in Abschnitt 3.1 im Formalismus der „Markov-Hüllen“

(Pearl 1988, Friston 2013) dargestellt, die gleich einer Membran das Innere einer Zelle von seiner Außenwelt abschirmen und gleichzeitig dafür sorgen, dass es eine zirkuläre, kausale Struktur zwischen den Zuständen der Umgebung und denen der Zelle gibt (vgl. auch Abb 3.1).

Eine noch radikalere Form nimmt dies in der symbolischen Wahrnehmungstheorie von Hoffman *et al.* (2015) an, die statt von „Organismen“ von der zellulären Struktur eines (künstlichen oder natürlichen) „bewussten Akteurs“ (Hoffman & Prakash 2014) spricht und – ähnlich zur Phänomenologischen Reduktion – nicht mehr davon ausgeht, dass zwischen den Zuständen des Akteurs und den Zuständen der Außenwelt ein abbildendes oder kausales Verhältnis existiert. Als Konsequenz können dann auch nicht mehr die Zustände des Akteurs und seiner Außenwelt mit denen der Motor- und Sinnesorgane respektive dessen *objektiver* Umgebung identifiziert werden. Diese ergeben sich als *wahrgenommene* aus der Interaktion mehrerer solcher Akteure. Die Inhalte unsere Wahrnehmungen sind nicht mehr Abbilder oder Schlüsse auf die objektiven (= wahren) Eigenschaften externer Gegenstände, sondern Produkte der Evolution eines solchen Netzwerkes.

Nun scheinen diese Ansätze, gerade *keine* Theorie des Bewusstseins im herkömmlichen Sinn zu formulieren, und insbesondere ist nicht ersichtlich, wie sich aus ihnen eine derartige Theorie ableiten ließe. Dies unterscheidet sie daher auch von vielen zeitgenössischen Ansätzen aus der Bewusstseinsforschung, die, wie im Verlaufe der Arbeit dargelegt wurde, als Versuche einer gegenständlichen Bewusstseinstheorie betrachtet werden müssen. Aus der Sicht einer Emergenztheorie des Bewusstseins ist dies jedoch begrüßenswert, insofern, als dass jene Ansätze genau die Frage, wie das Bewusstseins als Gegenstand dieser Welt charakterisiert werden könnte, durch die Frage ersetzen, wie Objekte als Proxys (oder „Zeichen“) für sinnvolle Bezugnahmen fungieren können. Allen Ansätzen gemein ist es, dass sie plausibel machen müssen, wie es überhaupt zu (zielgerichtetem oder biologisch nützlichem) Verhalten kommen kann, wenn es keine expliziten mentalen Repräsentationen von objektiven physikalischen Gegenständen gibt, die als Grundlage von Handlungen und Entscheidungen dienen könnten.

### 7.3 Metaphysik, Natur und Panpsychismus

Es sollen nun abschließend die Emergenztheorie des Bewusstseins in den Kontext einer realistischen Naturmetaphysik gestellt werden. Um dabei Paradoxien der Selbstanwendung zu vermeiden, soll dies im Rahmen eines kontextuellen Aspektmonismus geschehen.

## Netzwerke und Aktivitäten

Wir scheinen im Rahmen unserer bisherigen Darstellung unnötig viele Existenzannahmen getroffen zu haben: Es existiert demnach ein Reich der Gegenstände, das mittels verweisenden (emergenten) Objekten erschlossen wird, welche als Teile den Subjekten erscheinen. Das Ergebnis wäre ein semiotisch angereicherter Dualismus, ein eher wenig zufriedenstellendes Resultat. Der Physikalismus versucht nun, dies dadurch zu umgehen, dass er sowohl die Gegenstände als auch die Subjekte und alle Zeichenobjekte als Klassen und Konfigurationen einer einzigen Art von Dingen – den „physikalischen“ – ausweist. Anti-physikalistische Argumente bestreiten diese Möglichkeit, etwa indem sie dafür argumentieren, dass es neben physikalischen Wahrheiten noch andere gibt, dass rein physikalische Erklärungen eine Erklärungslücke hinterlassen oder dass Welten möglich sind, in denen Wesen existieren, die physikalisch gleich beschaffen sind wie wir, jedoch über kein Bewusstsein verfügen.

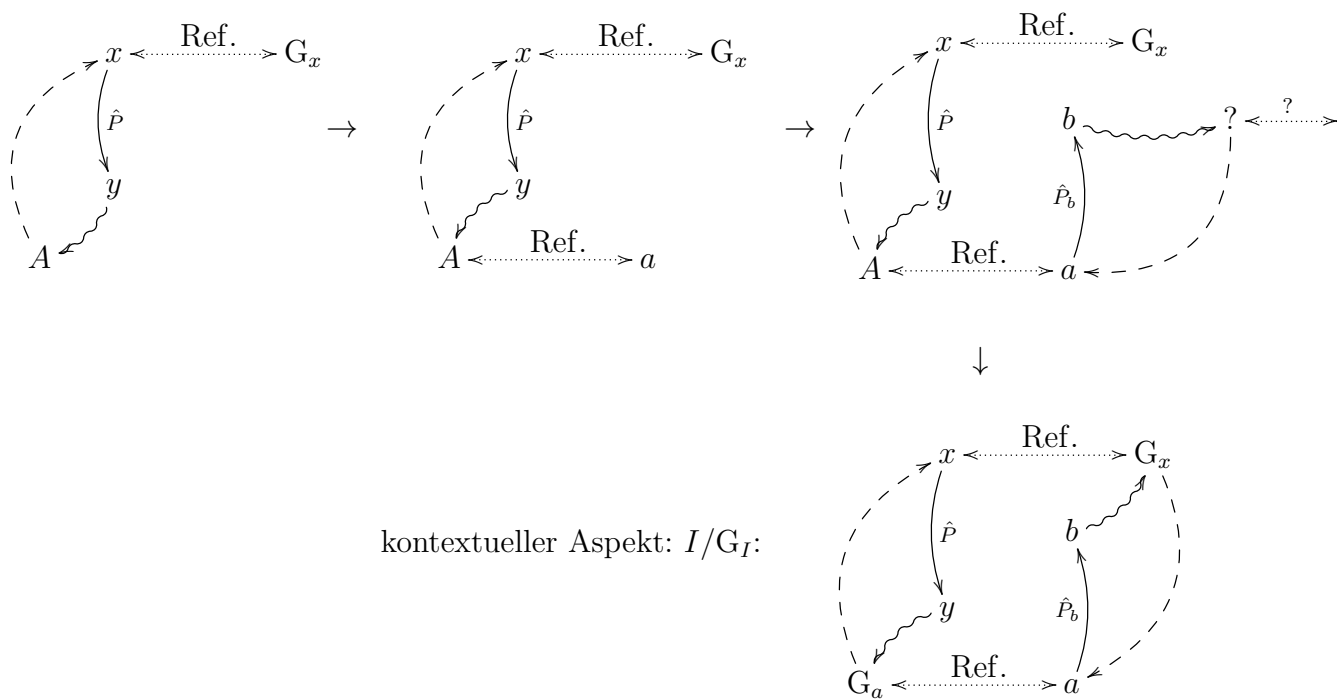
Doch anstelle eine solch starke metaphysische Annahme zu rechtfertigen oder zurückzuweisen, kann zuerst einmal versucht werden, die Komplexität der dargestellten Situationen zu verringern, *ohne* dabei überhaupt eine Reduktionsannahme zu treffen. Ein erster Schritt in diese Richtung wurde in der Diskussion der Einheit des Bewusstseins bereits genommen, indem das substantielle Subjekt in einen Prozess aufgelöst wurde, dem zwar ein „Subjektpol“ zugrunde liegt, dem aber für sich genommen keine substantialistische Interpretation zukommt. Dies entspricht einer Auffassung, die kategorische Eigenschaften als immer nur kontextuelle ausweist. Nun sei zusätzlich noch angenommen, dass diese „Subjekte“ selbst nichts anderes als die Referenten eines Zeichens darstellen: Wann immer wir uns auf sie beziehen wollen, geschieht dies in verobjektivierender Weise, d.h. aber, dass wir uns *zeichenartig* auf sie beziehen und gar nicht (sprachlich oder symbolisch) darüber Auskunft geben können, wie Subjekte nun „eigentlich“ oder „intrinsisch“ beschaffen wären.

Wenn aber Subjekte immer nur als via Zeichen hypostasierte Gegenstände aufzufassen sind und wir annehmen, dass die Existenz eines Zeichens immer bereits auf ein oder mehrere andere Subjekte verweist, dann scheinen wir in einen Regress zu landen, der nie zu einem Abschluss zu kommt.

Eine Möglichkeit, wie sich eine solche scheinbar endlose Verkettung semiotischer Beziehungen als abgeschlossene interpretieren ließe, beruht auf der Annahme einer kontextuellen Aspekthaftigkeit: Einerseits lassen sich Subjekte selbst als Pole eines Zeichenprozesses auffassen, die selbst zum Zeichen werden, andererseits lassen sich Subjekte als diejenigen Gegenstände denken, die sich mithilfe von Zeichen auf andere Gegenstände beziehen. Oder anders ausgedrückt: In der Wahrnehmung (aus der

„Perspektive der ersten Person“) erscheinen uns andere Subjekte als zeichenartige Objekte; in einer Beschreibung (aus der „Perspektive der dritten Person“) erscheinen Subjekte als aufeinander bezogene Gegenstände (Abb. 7.2).

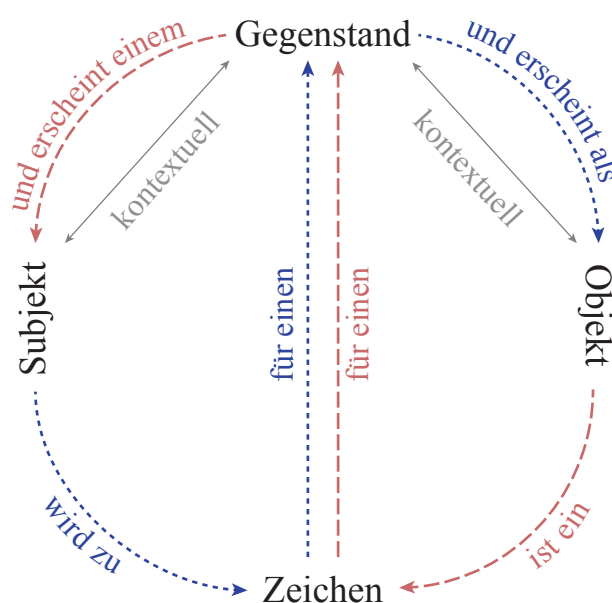
Die Redeweise von „Subjekten“ und „Gegenständen“ soll nicht darüber hinwegtäuschen, dass es sich bei diesen Ausdrücken letztlich nur um die Darstellungen der *relata* einer kontextuellen Beziehung handelt; sofern von „Gegenständen“ gesprochen wird, meint dies lediglich, dass hier ein Subjekt, weil es über ein Zeichen (d.h. über ein verweisendes Objekt) vermittelt wahrgenommen wird, aus der Sicht eines anderen Subjektes als Referent eben dieses Zeichens und somit als gegenständlich hypostasiert wird. Es gibt kein „absolutes Subjekt“ und keine „letzten Gegenstände“, denen unabhängig von ihren Beziehungen irgendwelche Eigenschaften zukommen würden.



**Abbildung 7.2:** Projektive Mereologie und Wahrnehmung als gegenseitiger Symbolaustausch: Der Austausch von verweisenden Objekten (Zeichen) zwischen Gegenständen kann als gegenseitige Wahrnehmung zweier Subjekte verstanden werden kann (= Annahme kontextueller Aspekte).

Nun, so ließe sich einwenden, bliebe aber dennoch ein letzter Dualismus zurück, nämlich einer, der zwischen Zeichen (= wohldefinierten Objekten) und zeichenverwendenden Gegenständen (oder kontextuell: Subjekten) unterscheiden würde – und der letztlich nichts anderes ist als die Unterscheidung zwischen einer nun zur Zeichenkette umgedeuteten Materie und eines zeichenverwendenden und vergegenständlichten Geistes.

Gerade eine graphische Darstellung, die ja immer bestimmte Punkte mit Gegenständen  $G$  oder mit bezeichnenden Objekten  $x$  identifiziert, würde dies vermuten lassen. Auch hier ließe sich erwidern, dass es *gar nicht um eine absolute als vielmehr um eine kontextuelle Bestimmung* gehen kann: Was in der Abbildung einen Gegenstand darstellt, ist selbst wiederum nur Objekt einer (übergeordneten) Bezugnahme; und was als Objekt dargestellt wird, ist selbst wiederum nur als Resultat einer (zugrundeliegenden) Tätigkeit zu verstehen. Ob wir etwas Subjekt (Gegenstand) oder Objekt (Zeichen) nennen, hängt letztlich von der jeweiligen Position ab, die wir ihm in einer Kette semiotischer Prozesse, einer „Erfahrungsreihe“, zusprechen (Abb. 7.3).



**Abbildung 7.3:** Objekte und Subjekte als Zeichen für Gegenstände. Die Objekte unseres Bewusstseins sind Zeichen, die für externe Gegenstände stehen, wie sie einem Subjekt erscheinen. Doch wenn wir von „dem Subjekt“ sprechen, benennen wir damit wiederum nur ein Zeichen, das für andere externe Gegenstände steht – das Subjekt wird „verobjektiviert“. Einem nie endenden Kreislauf lässt sich dadurch entkommen, dass wir annehmen, dass Gegenstände je nach Kontext als Objekte oder Subjekte innerhalb eines geschlossenen zeichenaustauschenden Netzwerkes zu verstehen sind.

Das hier gezeichnete Bild ist ähnlich der radikal-empiristischen Herangehensweise von William James (2006): Sobald wir ein „Das“ wahrnehmen, ist dieses bereits als vermittelte Konstruktion auszuweisen. „Für-sich“ gibt es weder Subjekte noch (selbstständige) Objekte, sondern nur die nicht-differenzierte „reine Erfahrung“, die in verschiedenen Kontexten als Zeichen für einen Gegenstand auftritt. Eine Möglichkeit, eine solche „reine Erfahrung“ darzustellen, besteht darin, diese

als Netzwerk zeichenaustauschender Gegenstände zu beschreiben, wobei die Knotenpunkte je nach kontextueller Bestimmung als Zeichen oder Bezeichnendes aufzufassen sind. Die „reine Erfahrung“ muss jedoch zwangsläufig eine angenommene bleiben, da wir diese gar nicht anders als über Objektivierungen, also eben nicht in ihrer „Absolutheit“ erfassen können. Unsere ontologischen Bestimmungen sind relative. Durch die Darstellung als geschlossenes Netzwerk wird ersichtlich, dass es wenig Sinn macht, nach dem Anfangs- oder Endpunkt, nach einem letzten Wesen oder einem unbewegten Beweger zu fragen.

### Gemeinsamkeiten und Unterschiede zur Kosmologie Whiteheads

Der Gebrauch von Mereologie bei der Darstellung der Erscheinungsweise von Objekten in Verbindung mit einer prozessualen Metaphysik legt einen Vergleich mit Whiteheads Methode der „*extensive abstraction*“ nahe, die den Übergang von der Wahrnehmung ausgedehnter Ereignisse (ihrer „*apprehension*“) zur bewussten Objektwahrnehmung beschreibt und paradigmatisch in den Kapiteln VI und VII der PNK darlegt wird.

Tatsächlich bestehen einige Ähnlichkeiten zwischen der Emergenztheorie des Bewusstseins und dem System Whiteheads. Schematisch lässt sich letzteres in zwei Schritten darstellen: Erstens, lässt sich die Wahrnehmung von Ereignissen anhand von Momenten („*constants of externality*“) charakterisieren, die sich als Definiertheit und Position („*here-present*“ und „*now-present*“) zweier aufeinander verweisender Ereignisse – ein wahrnehmendes („*percipient*“) und ein wahrgenommenes („*apprehended*“) Ereignis – innerhalb einer Dauer, als deren (mereotopologische) Extension und Einbettung sowie als Kontinuität zum gesamten Prozess der Natur („*community with nature*“) bezeichnen lassen (Tab. 7.2, linke Spalte). Wahrnehmung ist eine Relation *innerhalb* der Natur, die zwischen zwei Ereignissen besteht, und nicht ein Vorgang, der einen Wahrnehmenden außerhalb der Natur mit dieser verbindet.

Wichtig ist, dass es hier *nicht* um das bewusste Wahrnehmen oder die Apperzeption eines Objekts – Objekt im weitesten Sinne, hier zu verstehen als singulärer und relativ permanenter Bezugspunkt des bewussten Denkens – sondern um eine vorbewusste Wahrnehmung von Ereignissen geht, wobei Ereignisse zwar ontologisch fundamentale und mereologisch strukturierte, aber nicht-diskrete und vergängliche Entitäten darstellen.

Aus diesen Ereignissen lassen sich jedoch, zweitens, die Objekte des Bewusstseins als zugehörige „*recognita*“ denken, die aus einer Abstraktion hervorgehen. Analog der Einteilung in wahrnehmendes und wahrgenommenes Ereignis ergeben sich so das wahrnehmende oder perzipierende Objekt („*percipient object*“) und die wahr-

genommenen Eigenschaften und Objekte („*sense-objects*“ bzw. aus ihnen bestehend die „*perceptual objects*“, die unter Umständen von einem „*perceptual judgement*“ begleitet werden); vgl. auch Abb. 7.4, in welcher die Begriffe „Stuhl“ und „Fuß“ Wahrnehmungsobjekte darstellen, die aus einem wahrgenommenen Ereignis hervorgehen, wohingegen das perzipierende Objekt als „objektiviertes Subjekt“ [7] des Wahrnehmungsprozesses bezeichnet werden kann.

Hinzu kommen die „Situationen“ und „Konditionierungen“, die (Teile derjenigen) Ereignisse sind, die mit den wahrgenommenen Ereignissen in einer Relation stehen: Objekte sind eingebettet in die Situation, innerhalb derer sie als selbst-identische Permanenzen auftreten, Konditionierungen, etwa naturgesetzliche Ursachen, die außerhalb der Dauer der zugrundeliegenden *apprehension* liegen, bestimmen den „Charakter“ eines solchen Objekts [8].

Whiteheads Methode der *extensive abstraction* ist nun eine Darstellung des Übergangs von den ausgedehnten Ereignissen hin zu den diskreten Objekten. Dabei „entstehen“ die diskreten Objekte ähnlich wie Punkte als Grenzwerte einer Folge immer kleiner werdender konzentrischer Kreise „entstehen“. Dieses Bild entspricht ganz der am Ende von Abschnitt 4.2 diskutierten Rekonstruktion der Geometrie durch Alfred Tarski, der den Begriff des Punktes vom Begriff der ausgedehnten Kugel ableitet. Betrachten wir nun die Objektivitätskriterien der Emergenztheorie des Bewusstseins finden wir die korrespondierende Ansicht, dass sich Objekte als wohldefinierte, d.h. selbst-identische, strukturierte und verweisende Ergebnisse eines (bewussten) Wahrnehmungsprozesses ergeben. Ganz wie bei Whitehead gehen Objekte aus einem als natürlich bezeichneten Prozess hervor und bilden nicht dessen Grundlage. Anders als bei Whitehead beschreiben die mereotopologischen Beziehungen jedoch *nicht* die Seinsweise einer zugrundeliegenden Wirklichkeit (die Extensionalität der Ereignisse), sondern charakterisieren die Struktur von Bewusstseinsakten, woraus sich auch die Bezeichnung als „Emergenztheorie“ ableitet (zur Erinnerung: Es geht nicht um die Emergenz des Objekts „Bewusstsein“, sondern um die Emergenz von Objekten im, oder besser noch: während des Bewusstseins, zum Beispiel gemäß der Forderung, dass Objekte das Innere mereotopologisch abgeschlossener Ganzheiten darstellen; vgl. Abschnitt 5.1).

In Tab. 7.2 (rechte Spalte) sind überblicksartig die zentralen Charakteristika einer Emergenztheorie des Bewusstseins gelistet. Auffällig ist dabei, dass nur der letzte Punkt tatsächlich eine metaphysische Behauptung darstellt, der in den PNK die Postulierung einer globalen Zusammengehörigkeit von Ereignissen (*community of nature*) entspricht: Die Gegenstände ergeben sich als bestimmte Konfigurationen netzwerkartiger Aktivität und begründen im Rahmen einer dual-Aspekt-Annahme einen Monismus im Sinne von William James.

Methode der <i>extensive abstraction</i>		Emergenztheorie des Bewusstseins
<i>Apprehension</i>	<i>Consciousness</i>	
1. Definiertheit		Subjekt/Objekt als Resultate eines Wahrnehmungsprozesses
2. Position	Objekt/ob. Subjekt	Wohldefinierte Objekte als selbst-identische Entitäten
3. Referenz	↔ Situationen	Referenz eines w.d. Objekts auf einen konkreten Gegenstand
4. Extension und Einbettung	↔ Konditionierungen	Mereologisches Verhältnis zwischen einem Objekt, seinen Teilen und einer zugehörigen Ganzheit
5. Kontinuität		Natur als Netzwerk von Aktivitäten

**Tabelle 7.2:** Whiteheads Methode der *extensive abstraction* im Vergleich zur mereotopologischen Emergenz von Objekten. Wie entstehen (↔) Objekte des Bewusstseins aus der Wahrnehmung von Ereignissen? – Wie entstehen Objekte innerhalb von Netzwerken zeichenverwendender Aktivität? *Mereologie* bezieht sich dabei im einen Fall auf die *ontologische Struktur von Ereignissen* der vor-bewussten Wahrnehmung (Whitehead); im anderen Fall auf die *Darstellung von emergenten Phänomenen* innerhalb der Wahrnehmung (Emergenztheorie des Bewusstseins).



Drei mögliche Einwände gegen das System Whiteheads drängen sich auf, deren scheinbare Natürlichkeit Whitehead vermutlich auf einen tiefsitzenden Glauben an eine „Bifurkation der Natur“ (CN, Kap. 2) zurückführen würde. Eine kurze Diskussion dieser Einwände lässt noch einmal die Gemeinsamkeiten und Unterschiede zwischen Whiteheads System und einer Emergenztheorie des Bewusstseins stärker hervortreten.

**Erstens:** *Wer* abstrahiert? Wodurch oder *in welchem Medium* wird abstrahiert?

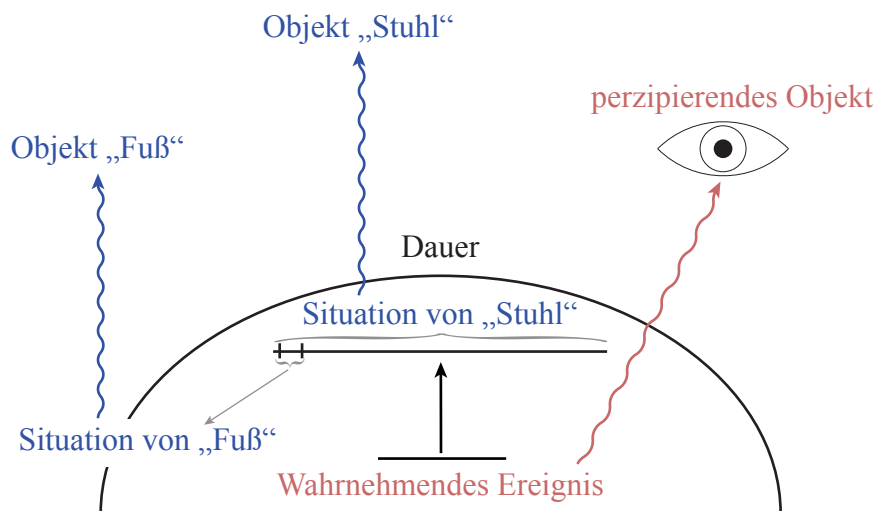
Selbstverständlich würde Whitehead den Vorwurf, dass in seiner Methode der *extensive abstraction* immer ein angenommener Dualismus mitschwingt, scharf zurückweisen, dennoch bleibt die Beantwortung obiger Fragen äußerst schwierig, ohne dabei auf scheinbar dualistisches Vokabular zurückzugreifen. Seine Theorie der *prehension*, die er in PR als nicht-kognitive Vorstufe zur *apprehension* entwickelt, kann als Versuch einer solchen Antwort interpretiert werden, die Whitehead den Ruf eines „Panexperientialisten“ (Griffin 1977) eingebracht hat. Letztlich wird Natur als „Prozeß der Erfahrung von Erfahrung“ (Hampe 1990, S. 115) verstanden, wodurch die Vorstellung, dass jede psychische Relation auf ein Außerhalb der Natur verweist, zugunsten der Ansicht, wonach jeder Form von Relationalität in der Natur auch immer ein psychisches Moment zukommt, zurückgewiesen wird.

**Zweitens:** Was ist der Ausgangsstoff (das „Material“), aus dem sich die bewusst wahrgenommenen Objekte ergeben? Diese Frage ist entscheidend, da wir nur hier – im eigenen Bewusstsein – tatsächlich („empirisch“) mit psychischen Vorgängen konfrontiert sind. Eine jede Extrapolation – sei es die Rede vom „Geist in der Maschine“ oder vom „Bewusstsein in der Natur“ – ist letztlich Spekulation, die über unsere eigene Erfahrung hinausführt.

Wie in der vorigen Antwort angedeutet wird, lassen sich die Objekte des Bewusstseins als Abstraktionen eines psychischen (oder wie man in der zeitgenössischen analytischen Philosophie zu sagen pflegt: „proto-bewussten“) Prozesses denken. Zusätzlich liefert uns nun Whitehead aber eine formalontologische Charakterisierung dieser Prozesse mithilfe mereologischer Überlegungen – etwas das als formale Naturmetaphysik oder als Teile einer „Kosmologie“, wie sie Whitehead in PR bezeichnen wird, gelten kann.

Dass der mereologische Kalkül bei Whitehead primär von den Strukturen der Ereignisse handelt und dass sich die Teileinsrelation nur in abgeleiteter Weise auf die Objekte des Bewusstseins bezieht, wird in Kapitel VII der PNK am Beispiel eines Stuhls illustriert: Während wir vielleicht in abkürzender Rede sagen würden, dass der Fuß einen Teil des Stuhls bildet, müssten wir kor-

rekterweise damit meinen, dass zwischen dem Ereignis (also der „Situation“), aus dem via *extensive abstraction* das Objekt „Stuhl“ hervorgeht, und dem Ereignis (der „Situation“), aus dem das Objekt „Fuß“ hervorgeht, eine Teil-Ganzes-Beziehung existiert. Zwischen „dem Stuhl“ und „dem Fuß“ besteht nur indirekt eine Teil-Ganzes-Beziehung, denn Objekte selbst sind nichts als Abstraktionen der eigentlich wirklichen Ereignisse. Mereologie ist Formal*ontologie* (Abb. 7.4).



**Abbildung 7.4:** Mereologische Strukturiertheit von Ereignissen. Wellenlinien führen auf die Objekte des Bewusstseins (Methode der *extensive abstraction*). Urteile über Teil-Ganzes-Beziehungen zwischen ihnen („der Fuß ist Teil des Stuhls“) sind nur abgeleitet von den „eigentlichen“ Teil-Ganzes-Beziehungen auf der Ebene von Ereignissen (Situationen).

Anders in einer Emergenztheorie des Bewusstseins: Die Abschnitte 4 und 5 können dahingehend interpretiert werden, dass Mereologie nicht die fundamentale Seinsweise der Welt beleuchten würde, sondern vielmehr eine Darstellung derjenigen Relationen ist, die letztlich mit der Entstehung von Objekten einhergehen. Dies wird etwa durch die Ansicht bestärkt, wonach es gar keine nicht-kontextuelle Einteilung in Ganzheiten und Teile gibt, die unabhängig von konkreten theoretischen und praktischen Verfahren existiert. Spezifisch dient das System der PEM zur Beschreibung des Vorganges – eben jenes Vorganges, der für gewöhnlich als „Bewusstwerdung“ bezeichnet wird –, der sich während der Herausbildung verweisender Objekte in semiotischen Netzwerken vollzieht, nicht aber zur Charakterisierung eines *metaphysischen Bildes*.

**Drittens:** Wie lässt sich die Auffassung begründen, dass es sich bei der mereologisch beschriebenen Struktur der Ereignisse nicht selbst „bloß“ um eine Abstraktion handelt? Objekte sind gemäß Whitehead Abstraktionen von Ereignissen, aber woher kommt es, dass in „objektiver“ Weise über Ereignisse gesprochen werden kann? Sind nicht auch die Beschreibungen, die dabei gegeben werden, letztlich sozial erlernte Formen der Ausdrucksweise und müsste daher nicht eine Form der Sozialität immer bereits vorausgesetzt werden?

Aus heutiger Perspektive drängen sich solche Fragen auf, Whitehead waren sie wahrscheinlich fremd. Es muss daher letztlich Spekulation bleiben, wie Whitehead auf diesen fundamentalen Zweifel an der Möglichkeit von Metaphysik reagiert hätte. Die Emergenztheorie des Bewusstseins muss sich diesen Fragen allerdings stellen.

Dabei ließe sich entgegnen, dass Emergenz ja gerade beschreiben soll, wie ein zeichenhaftes Objekt als Fixpunkt zweier Teilprozesse zu verstehen ist: eines synthetisierenden, Einheit stiftenden und eines unterscheidenden, Vielheit erzeugenden (vgl. Abschnitt 6.2). Auch wenn an dieser Stelle keine genaueren Ausführungen gegeben werden sollen, ist es durchaus vorstellbar, dass sich dieser Gedanke befriedigend auf die Wechselwirkung zwischen individueller und kollektiver Zeichenverwendung übertragen lässt. Dies kann (und soll) nicht das Projekt einer Fundamentalmetaphysik rehabilitieren, aber zumindest die Konsistenz einer Emergenztheorie des Bewusstseins behaupten.

## Perspektive und Panpsychismus

Daran anknüpfend sei hier noch die sog. „Perspektive der zweiten Person“ angesprochen, die manchmal in der Auseinandersetzung mit dem Bewusstsein vernachlässigt wird, aber ganz zentral ist, wenn deren philosophische und ethische Konsequenzen beleuchtet werden sollen.

In einer Emergenztheorie des Bewusstseins wäre die Zweite-Personenperspektive gleichzusetzen mit einer Situation, in der Objekte (bzw. das System, welches aus ihnen gebildet wird) so aufgefasst werden, *als ob* sie als stellvertretend für ein je anderes Subjekt (als „Du“) anzusprechen wären. Dies betrifft das Kontinuitätsargument des Panpsychismus: Da wir nicht in der Lage sind, eine fixe Grenze zwischen bewussten und nicht-bewussten Lebewesen und Dingen zu ziehen, folgern Panpsychisten mithilfe eines Analogieschlusses, dass prinzipiell *allem* eine Form von Bewusstsein zuzusprechen sei (vgl. hierfür Prentner 2014, S. 104f.).

Etwas mit „Du“ anzusprechen impliziert meist, dass wir meinen, es mit einer *Person* zu tun zu haben; und dazu gehört mehr als nur die durch Analogie vermittel-

te Überzeugung, einem empfindenden Wesen gegenüberzustehen – etwas, das zum Beispiel höchst aktuell ist, wenn es um die Frage nach Personenrechten für Tiere geht –, allerdings würden viele eine Bewusstseinsfähigkeit als notwendiges Kriterium hierfür voraussetzen: Für den Fisch, der sich keiner Schmerzen bewusst werden kann, brauche es kein Recht auf ein schmerzfreies Leben. Mit der Maschine – sei sie eine biologische oder eine aus Silizium –, die nur „im Dunkeln“ vor sich hinrechnet, könne umgegangen werden, wie es uns passt.

Durch die semiotische Auflösung der Redeweise von Erster-, Zweiter- und Dritter-Personen-Perspektive, wie sie hier vorgeschlagen wurde, erhält das Kontinuitätsargument eine neue Wendung, nicht weil auf einmal erklärt würde, warum wir ein System von Objekten das eine Mal als „Es“ und das andere Mal als „Du“ ansprechen, sondern weil nun gar nicht mehr darauf abgezielt wird, ein Bewusstsein „in“ allen Objekten zu erkennen: Objekte sind prinzipiell nur Zeichen und es macht wenig Sinn zu fragen, ob Zeichen bei Bewusstsein sind.

Doch sollte dies nicht zu vorschnellen Schlüssen verführen. Zwar ist außerhalb des Prozesses, zu dem ein Ich als *relatum* gehört, d.h. außerhalb der je-eigenen leiblichen Erfahrung, Bezugnahme in einer nicht-verobjektivierten Form ja gar nicht erfahrbar; wir sehen – im wörtliche Sinne – kein Bewusstsein. Doch dies zeigt nur die Grenzen unserer Wahrnehmungsfähigkeit an, nicht aber die Grenzen der Wirklich über unsere Wahrnehmung hinaus. Wenn wir z.B. die Erfahrungsberichte von Testsubjekten mit den Bildern ihrer neuronalen Aktivität vergleichen, dann vergleichen wir nicht „etwas Subjektives“ mit „etwas Objektivem“, sondern zwei verschiedene Arten von Zeichenketten. Wir würden dennoch nicht folgern, dass es sich bei den Versuchspersonen lediglich um eine nicht weiter schützenswerte Ressource handelt, die von uns beliebig benutzt werden kann (und sei es nur zum Zwecke der Erkenntnisgewinnung).

Und selbst wenn wir nun über unser eigenes Bewusstsein Auskunft geben wollen, wenn wir etwas *über* das Bewusstsein sagen wollen – etwa wenn wir es zum *Objekt* eines introspektiven Berichtes machen – benutzen wir wiederum Zeichen, welche z.B. die Phenomenalität des Wahrnehmungsaktes benennen. Daraus entsteht zwangsläufig die Schwierigkeit auf die Frage zu antworten, „was denn nun das Bewusstsein eigentlich sei“, worauf sich viele verlegen auf eine scheinbar intuitiv gegebene Einsicht berufen – oder wie es Ned Block (1978, S. 281) ausgedrückt hat: „Wenn du noch fragen musst, wirst du’s niemals nicht verstehen.“

Weisen wir ein solches unmittelbares Wissen über Dinge, die uns als Objekte gegeben sind, als Mythos zurück, müssen wir schließen, dass alles, was wir über Bewusstsein und Wahrnehmung (als Objekt) aussagen, bereits wieder als über Zeichen vermittelt (und nicht als abbildende Beschreibung „der Sache selbst“) anzusehen ist. Die

Ausführungen dieses Abschnittes sind also selbst nur Versuche einer über Zeichen vermittelten Beschreibungsweise des so verobjektivierten Erfahrungsprozesses. Mehr kann auch die Philosophie nicht leisten.

Dem Vorwurf der Zirkularität könnte man dadurch begegnen, dass man bestreiten würde, überhaupt auf eine bestimmte metaphysische Grundlegung angewiesen zu sein – insofern, als dass man ja nur „rein deskriptiv“ vorgehe. Dies scheint ein pragmatischer Schachzug zu sein, der philosophisch jedoch nicht sonderlich befriedigend ist. Letztlich blieben immer Zweifel an der Wahrhaftigkeit einer vorgeblich rein deskriptiv agierenden Philosophie zurück. Zum Beispiel dürfte kein ausdrucksstarker Naturbegriff angenommen werden, der auf eine bestimmte metaphysische Haltung verweisen würde. Analog gilt dies auch für „die Welt“ „das Gute“ oder auch „die Erfahrung“, wenn jene als Allgemeinbegriffe gebraucht würden, die eine universelle (Nicht-)Zugehörigkeit ausdrücken.

Eine Alternative wäre es, einen realistischen Naturbegriff zu postulieren, der wesentlich durch dessen Produktivität gekennzeichnet ist: Die Natur gleiche demnach einem Netzwerk produktiver Tätigkeiten, welche wir letztlich als psychische interpretieren können, sofern wir sie unter dem Aspekt der „Innerlichkeit“ betrachten. Dies ist gerechtfertigt, da innerhalb eines Kreisganges von erzeugenden/bezeichnenden und erzeugten/bezeichneten Objekten (und zurück) kein absoluter Standpunkt eingenommen werden kann. Etwas ähnliches ließe sich für eine materialistische Interpretation sagen, die unter dem Aspekt der „Äußerlichkeit“ getroffen wird, allerdings wird heutzutage die Frage, warum es überhaupt zur materiellen (nicht-psychischen) Phänomenen kommt, als viel weniger rätselhaft empfunden. Die metaphorische Redeweise, dass „etwas unter diesem und jenen Aspekt“ betrachtet werden kann, verführt vielleicht zum fälschlichen Glauben an einen Standpunkt *außerhalb* des Naturgeschehens. Die Rede von der „Innerlichkeit“ und der „Äußerlichkeit“ soll lediglich anzeigen, dass es sich dabei um unterschiedliche Kontexte handelt, in denen uns die *relata* der Naturtätigkeit erscheinen.)

Dies würde auch nicht implizieren, dass „Steine“ oder „Tische“ irgendwie „bei Bewusstsein“ wären – ein Verdikt, das wohl auch der „Panexperientialist“ Whitehead so nicht treffen würde –, sondern, dass jene als Resultate von Prozessen zu betrachten sind, deren Form letztlich als bewusste gedeutet werden kann.

Zusammenfassend ließe sich also festhalten:

- I Die Welt, ist alles, was erscheinen kann, aber Erscheinungen benötigen einen Kontext.
- II Alle Objekte, derer wir uns bewusst sind, können als Zeichen verstanden werden.
- III Bezeichnende Objekte erscheinen einem wahrnehmenden Subjekt, das selbst zu einem Zeichen wird.
- IV Einer nie abbrechenden Verkettung von Bezeichnungen ließe sich durch Postulierung eines geschlossenen Netzwerkes Rechnung tragen, das Subjekte und Objekte als kontextuelle Aspekte betrachtet.
- V In einem solchen Netzwerk stellt die spontane Entstehung neuartiger Objekte ein wesentliches Merkmal dar. Ein solches Netzwerk ist eine Darstellung von Natur.

Wenn nun Bewusstsein, erstens, gar keinen Gegenstand „irgendwo in der Natur“ bezeichnet, d.h., wenn es sich bei „dem“ Bewusstsein selbst nur um ein Zeichen für die Form eines sich vollziehenden Prozesses handelt, und wenn ich, zweitens, anerkenne, dass sich dieser Prozess mindestens in mir selbst vollzieht; dann scheinen sich mir *ceteris paribus* nur zwei Alternativen darzubieten: ein Solipsismus, der nur von der Existenz einer einzigen wahren Bezugnahme auf die Welt ausgeht, oder ein universeller Psychismus, der in den Objekten den Ausdruck einer mannigfaltigen geistigen Tätigkeit erkennt.



## Anmerkungen zu Abschnitt 7

[1] In der Regel sind Funktionalisten auch Physikalisten. Rein logisch lassen sich Spielarten des Funktionalismus aber durchaus mit nicht-physikalistischen Ansichten vereinbaren, solange diese ein Substrat kennen, das der Realisierung der entsprechenden Funktion zugrunde liegt – so etwa ein „funktionalistischer Panpsychismus“ (Prentner 2014, S. 95f. und Abb. 1) oder der „nicht-reduktive Funktionalismus“ von Chalmers (1996).

[2] Tatsächlich ist sie sogar zu schwach, um überhaupt interessante Klassifikationen innerhalb der etablierten Wissenschaft zu ermöglichen (Stephan 2006). Auch die Monismusthese („alles ist physikalisch“; vgl. Abschnitt 2.3) macht aus „schwacher Emergenz“ zwar einen ontologischen Begriff, allerdings einen, der nicht sonderlich vielsagend ist.

[3] Virtuelle Teilchen unterscheiden sich von gewöhnlichen Teilchen dadurch, dass sie in der Regel keine asymptotisch stabilen Zustände solcher Streuprozesse bezeichnen. Philosophisch könnte man also davon sprechen, dass jenen „für-sich“ genommen keine Existenz außerhalb der Prozesse, in denen sie auftreten, zukommt. Absolut gesehen gibt es jedoch gar keine deutliche Unterscheidung zwischen virtuellen Teilchen und gewöhnlichen Teilchen. Beide haben eine endliche Lebensdauer, die jedoch für wirkliche Teilchen um ein Vielfaches höher ist. Während virtuelle Teilchen manchmal als Fluktuationen eines Feldes bezeichnet werden, spricht man bei gewöhnlichen Teilchen von relativ stabilen Anregungszuständen eines Feldes.

[4] „[This] allows the agent to detach from primary experience and to project itself into conceivable situations.“ (Briegel & De Las Cuevas 2012, S. 2).

[5] Ohne, dass es von den Autoren explizit erwähnt wird, lässt das Modell der *projective simulation*, welches zwischen Nichtdeterminiertheit und Determiniertheit vermittelt, eine Verwandtschaft zum „Zweiphasenmodell des freien Willens“ (Doyle 2010) bei William James vermuten.

[6] Angemerkt sei hier, dass es dabei primär um (philosophisch gesehen) basale Handlungen und wenig komplexe Formen der Entscheidungsfindung geht. Die soziale Dimension wird bis auf wenige Ausnahmen vernachlässigt (aber siehe z.B. Menary 2007, Friston & Frith 2015).

[7] Auf gar keinen Fall sollte der Begriff „Subjekt“ mit einem personalen, andauernden Selbst gleichgesetzt werden (vgl. hierzu etwa Hampe 1990, § 15). Viel eher gleicht Whiteheads *percipient object* dem, was in Abschnitt 6 als Reifikation des „subjektiven Pols“ eines Erfahrungsprozesses bezeichnet wurde: „[It] is a recognizable permanence [...] which is the unity of awareness whose recognition leads to the classification of a train of percipient events as the natural life associated with one consciousness.“ (PNK, § 22.4).

[8] Ein möglicher Grund, warum Whitehead die Bedeutung von „Konditionierungen“ hervorhebt, die über die eigentlich wahrgenommene Ereignisse (von denen ausgehend Objekte als Abstraktionen abgeleitet werden) hinausreichen, besteht darin, dass sich scheinbar nur auf diese Weise, die Existenz von überindividuell gültigen Naturgesetzen, welche die Objekte unseres Bewusstseins „konditionieren“, mit einer individuell-psychisch Einzelwahrnehmung vereinbaren lässt: Wenn wir mit Whitehead annehmen, dass alle Objekte nichts als Abstraktionen einer zugrundeliegenden Wahrnehmungstätigkeit (der *apprehension*) sind, dann scheint die Frage offensichtlich zu sein, wie denn nun der Status der gesetzmäßigen Verbindungen dieser Objekte zu beurteilen sei: Naturgesetze handeln von Dingen, die es (außer als Abstraktion) gar nicht gibt; doch warum folgen dann die Inhalte des Bewusstseins letztlich diesen Gesetzen? Anders gefragt, wie können Gesetze, die immer nur von subjektiven Dingen handeln, objektive Gültigkeit besitzen?







## Parergon

Dieser Abschnitt dient hauptsächlich der Klärung einiger Begrifflichkeiten, die im Verlaufe dieser Arbeit verwendet wurden. Neben den Anmerkungen am Ende der jeweiligen Kapitel soll er auch die Bezüge zu verwandten Ideen verdeutlichen und einige Motive, die sich durch die Arbeit ziehen, noch einmal explizieren. Zudem soll er die Lektüre der Arbeit erleichtern.

### Mentalistische Ausdrücke

Bezeichnung	Synonyme	Definition
Erfahrungsprozess		Psychische Aktivität
Bewusstsein	Erleben	Form eines EPs
Subjekt	Minimales Selbst	Pol eines EPs, „blinder Fleck“ der Wahrnehmung
Ich	Empirisches Ich, Ego, empirisches Selbst	Empirische Struktur, die zu einem Su. korreliert
Ich-Bewusstsein	Selbstbewusstsein (empirisch)	Bezugnahme auf ein Ich, die sich innerhalb derselben empirischen Struktur vollzieht; wird manchmal zur vollständigen „Selbsttransparenz“ mystifiziert
Personales Selbst	Person	Baut auf dem empirischen Ich auf, speist sich aber auch aus Erinnerungen, Selbstnarrationen, sozialen Interaktionen; ist wesentlich plastisch und bedingt ein Ich-Bewusstsein

**Tabelle P1:** Eine Übersicht über das verwendete mentalistische Vokabular.

## Objekte und Gegenstände

Bezeichnung	Synonyme	Definition
Welt		(Abstrakte) Zusammenfassung von Dingen
Ding	Entität	Umfasst die Klasse der Objekte und der Gegenstände
Gegenstand	Substantielles Ding, Individuum	Konkret, als unabhängig existierend angenommen
Objekt	Nicht-substantielles Ding	Teile eines EPs, als bedingt existierend angenommen
Zeichen	Selbstständiges oder wohldefiniertes Objekt	Ob., welches auf (konkreten) Gegenstand, der außerhalb des Bewusstseins liegt, verweist
Teilobjekt	Unselbstständiger Teil	Ob., welches selbst auf keinen Gegenstand verweist; Teil eines selbstständigen Objekts

**Tabelle P2:** Eine Übersicht über das verwendete nicht-mentalistiche Vokabular.

## Glossar

**Abbildung** „im strengen Sinn“ bezeichnet eine strukturerhaltende und umkehrbare (isomorphe) Darstellung der erfahrbaren Welt durch eine Beschreibung. Dabei muss implizit davon ausgegangen werden, dass es sich um eine Beziehung zwischen strukturierten oder theorieartigen Bereichen handelt (anders wäre die Redeweise von einem „Isomorphismus“ nicht intelligibel). Anstelle von strenger Abbildungsbeziehung ließe sich auch von „realistischer Repräsentation“ sprechen.

**Bewusstseinsphänomen** bezeichnet als Überbegriff all diejenigen Dinge, die wir mit denjenigen psychischen Vorgängen identifizieren, welche wir als „bewusste“ bezeichnen, so etwa die Wahrnehmung, die Entscheidungsfindung, oder die Intention etc. Der Begriff des „Phänomens“ steht dabei primär für alles, was Gegenstand einer wissenschaftlichen Beschreibung werden kann, vergleichbar den elektrischen, thermischen oder chemischen Phänomenen. Dabei suggeriert der Phänomenbegriff jedoch immer bereits, dass hier *jemandem etwas erscheint*. Er legt, anders gesagt, die Struktur einer Welt an-sich, die irgendwie in Erscheinung tritt, nahe. Die Gretchenfrage einer „Wissenschaft vom Bewusstsein“ ist

somit, ob es eine sinnvolle Rede von „Phänomenen“ überhaupt geben kann, die von dieser scheinbar korrespondierenden ontologischen Annahme losgelöst ist. Dies ist nicht zu verwechseln mit der Phänomenologie als eigenständiger philosophischer Disziplin. Für den Phänomenologen ist die Ausklammerung ontologischer Urteile sogar Teil seiner grundsätzlichen Methode, wenn er versucht, die Wahrnehmungsinhalte so zu beschreiben, wie sie erscheinen.

### Das Rätsel der ...

... **Phänomenalität** bezeichnet die Frage, wie und warum es sich „irgendwie anfühlt“ in einem mentalen (bewussten) Zustand zu sein. Eine Denkrichtung versucht, Phänomenalität als real existierende Eigenschaft von Objekten zu betrachten (die Eigenschaft der „Röte“ wäre eine Eigenschaft des wahrgenommenen Apfels, sein **Quale**). Eine andere Denkrichtung, vergegenständlicht das Bewusstsein und schreibt diesem Phänomenalität als Eigenschaft zu.

Beide Ansichten basieren jedoch auf einem substantialistischen Bild von Erfahrung, welches hier zugunsten eines kontextuellen zurückgewiesen werden soll.

... **Selbstbezüglichkeit** stellt sich dann, wenn davon ausgegangen wird, dass Bewusstsein die Voraussetzung dafür ist, dass es sich bei etwas überhaupt um ein *Symbolsystem* handelt. Oft wird dieser Aspekt überhöht dargestellt (Transzendentalphilosophie) bzw. mystifiziert (zum „reinen Selbstbewusstsein“). Naturalistisch gedacht wird Selbstbezüglichkeit zum Gegenstand einer Systemtheorie des Mentalen.

... **Einheit** ist die Frage nach der Konstitution des Bewusstseins. Wie kombinieren sich materielle Wesenheiten zu psychischen oder umgekehrt? Auf der Ebene der Erfahrung erscheint uns der Inhalt des Bewusstseins einheitlich gegeben, auf der Ebene des Gehirns scheinen dazu modulare oder zumindest raum-zeitlich getrennte physikalische Abläufe zu korrelieren. Doch wie folgt daraus die „Einheit des Bewusstseins“?

**Emergenz** bezeichnet das Auftreten systemischer Eigenschaften, die neuartig gegenüber den Eigenschaften der Teile sind. Epistemologisch kann Emergenz als Herausbildung einer singulären Beschreibung aus einer zugrundeliegenden Beschreibung verstanden werden und ist kontextuell, d.h. die (pragmatischen und theoretischen) Kriterien, nach denen eine Beschreibung gewählt wird, entstammen nicht der zugrundeliegenden Beschreibung selbst.

Spezifisch innerhalb eines EPs betrifft Emergenz die Entstehung eines selbst-

ständigen (verweisenden) Objekts zu dessen unselbstständigen (und daher nicht-verweisenden) Teilen.

**KEM** steht für das Kalkül der klassischen extensionalen Mereologie. Eine wichtige Vorstufe ist die minimale extensionale Mereologie (**MEM**), die neben den Basisaxiomen die zwei Supplemente umfasst, aber keine Aussagen zur Zulässigkeit von Summenbildung trifft. Diese stellt den Ausgangspunkt für eine Erweiterung des Systems durch Atome, Funktionen und Topologien dar.

**Natur** bezeichnet die Gesamtheit oder die Klasse aller möglichen Erfahrungen. Bestimmte Arten der Systematisierung und Formalisierung von Erfahrungszusammenhängen werden als **naturwissenschaftliche** bezeichnet. Spekulativ soll Natur als umfassendes Netzwerk spontaner Aktivitäten gedeutet werden.

**Panpsychismus** kam historisch in vielen Variationen vor, wird heute aber vor allem so verstanden, dass den Objekten der Mikrophysik ein (Proto-)Bewusstsein innewohnt („Quarks sind bei Bewusstsein“) oder dass bestimmte Systeme als (proto-)bewusst angesehen werden (etwa weil sie über genügend große „informationelle Komplexität“ verfügen), obwohl sie sich in einigen Punkten sehr von Menschen und anderen Tieren unterscheiden. Doch dies sind nur zwei, mit einigen Schwierigkeiten behaftete Ausprägungen des Panpsychismus. Diesen gegenüber steht etwa die Auffassung, dass Produktivität in natürlichen Prozessen letztlich als psychisch zu deuten ist.

**Projektive Mereologie** beschreibt das Verhältnis zwischen Ganzheiten und Teilen als Projektionsrelation. Projektive Mereologie entspricht einem dekompositionalen Ansatz, der nach den möglichen Zerlegungsweisen eines Objekts fragt, als Gegenstück zur Frage nach den Bedingungen mereologischer Summation einer Vielzahl bereits wohldefinierter Objekte. Wird formalisiert im Rahmen der projektiven extensionalen Mereologie (**PEM**).

**Quale** bezeichnet den spezifischen Erlebnisgehalt einer Erfahrung, der scheinbar über alle physikalische Beschreibung hinausreicht und dabei auf eine „intrinsische“ Eigenschaft des Bewusstseins verweist, welche dieses vor den restlichen Gegenständen der Welt auszeichnet. Der Plural (*Qualia*) wird manchmal synonym zu „Bewusstsein“ gebraucht und erweckt leicht den Eindruck einer substanzialistischen Redeweise vom Mentalen. Eine naturalistische Erklärung des Zusammenhangs zwischen *Qualia* und Zuständen des Gehirns kann als „*hard problem*“ bezeichnet werden (vgl. auch das **Rätsel der Phänomenalität**).

**Topologie** bezeichnet allgemein die Lehre räumlicher, aber nicht notwendigerweise metrischer Verhältnisse; spezifisch in der Mathematik die Theorie der Invari-

---

anten unter stetiger und umkehrbarer Verformung. Konkret werden in dieser Arbeit die Begriffe der Berührung und des Abschlusses im Rahmen erweiterter mereologischer Kalküle (**MT**) behandelt.





## Zitatnachweise und Siglen

Den Kapiteln vorangestellte Zitate entstammen folgenden Quellen:

- I Friedrich Wilhelm Joseph Schelling, *Ideen zu einer Philosophie der Natur*, 1797, Buch II: §7, 353; nach der historisch-kritischen Ausgabe der Bayerischen Akademie der Wissenschaften, Frommann-Holzboog, Stuttgart, 2001.
- II Ludwig Wittgenstein, *Vorlesungen 1930-1935*, Anm. 286; Suhrkamp, Frankfurt am Main, 1989.
- III W. Ross Ashby, Design for a brain, S. 379; in *Electrical Engineering*, **20**: 379–383, 1948.
- IV Hermann Weyl, *Riemanns Geometrische Ideen, ihre Auswirkung und ihre Verknüpfung mit der Gruppentheorie*, S. 7; Berlin, Springer, 1988.
- V George Spencer Brown, *Laws of Form*, S. v; London, George Allen and Unwin Ltd, 1969.
- VI William James, Does ‘Consciousness’ Exist?, S. 477; in *Journal of Philosophy, Psychology and Scientific Methods*, **1**: 477–491, 1904.
- VII Werner Heisenberg, *Physik und Philosophie*, S. 56; Hirzel, Stuttgart, 5. Auflage von 1990.

Durch die folgenden Siglen sei verwiesen auf:

- |     |                         |  |
|-----|-------------------------|--|
| KrV | Immanuel Kant,          | <i>Kritik der reinen Vernunft</i> (2. Auflage, 1787).                    |
| KU  | —                       | <i>Kritik der Urteilskraft</i> (1788).                                   |
| LU  | Edmund Husserl,         | <i>Logische Untersuchungen</i> (2. Band, 1901).                          |
| PIZ | —                       | <i>Zur Phänomenologie des inneren Zeitbewußtseins</i> (1928).            |
| PNK | Alfred North Whitehead, | <i>An Enquiry Concerning the Principles of Natural Knowledge</i> (1919). |
| CN  | —                       | <i>The Concept of Nature</i> (1920).                                     |
| PR  | —                       | <i>Process and Reality: An Essay in Cosmology</i> (1929).                |



## Literatur

- Anderson, P. W. (1972), 'More Is Different', *Science* **177**(4047), 393–396.
- Ashby, W. R. (1956), *An Introduction to Cybernetics*, Chapman & Hall, London.
- Ashby, W. R. (1962), Principles of the self-organizing system, in H. von Foerster & G. W. Zopf Jr. (Hg.), 'Principles of Self-Organization: Transactions of the University of Illinois Symposium', Pergamon Press, London, S. 255–278.
- Aspect, A., Grangier, P. & Roger, G. (1981), 'Experimental Tests of Realistic Local Theories via Bell's Theorem', *Physical Review Letters* **47**(7), 460–463.
- Atmanspacher, H. (2014), '20th Century Varieties of Dual-aspect Thinking', *Mind and Matter* **12**(2), 245 – 289.
- Atmanspacher, H. & beim Graben, P. (2007), 'Contextual emergence of mental states from neurodynamics', *Chaos and Complexity Letters* **2**(2-3), 151–168.
- Atmanspacher, H. & Bishop, R. C. (2007), 'Stability conditions in contextual emergence', *Chaos and Complexity Letters* **2**(2-3), 139–150.
- Atmanspacher, H. & Kronz, F. (1999), Relative Onticity, in A. Amman, H. Atmanspacher & U. Müller-Herold (Hg.), 'On Quanta, Mind and Matter. Hans Primas in Context', Springer, Berlin, S. 273–298.
- Baars, B. J. (1988), *A Cognitive Theory of Consciousness*, Cambridge University Press, New York.
- Bachelard, G. (1987), *Die Bildung des wissenschaftlichen Geistes. Beitrag zu einer Psychoanalyse der objektiven Erkenntnis*, Suhrkamp, Frankfurt am Main.
- Balduzzi, D. & Tononi, G. (2008), 'Integrated Information in Discrete Dynamical Systems: Motivation and Theoretical Framework', *PLoS Computational Biology* **4**(6), e1000091.
- Bayne, T. (2010), *The Unity of Consciousness*, Oxford University Press, Oxford.
- Bayne, T. & Chalmers, D. J. (2003), What is the Unity of Consciousness?, in A. Cleeremans (Hg.), 'The Unity of Consciousness: Binding, Integration, Dissociation', Oxford University Press, Oxford.
- Bedau, M. A. (2008), 'Is Weak Emergence Just in the Mind?', *Minds and Machines* **18**(4), 443–459.

- beim Graben, P. & Atmanspacher, H. (2009), Extending the Philosophical Significance of the Idea of Complementarity, *in* H. Atmanspacher & H. Primas (Hg.), 'Recasting Reality. Wolfgang Pauli's Philosophical Ideas and Contemporary Science', Springer, Berlin, S. 99–113.
- beim Graben, P., Barrett, A. & Atmanspacher, H. (2009), 'Stability criteria for the contextual emergence of macrostates in neural networks', *Network: Computation in Neural Systems* **20**(3), 178–196.
- Bell, J. S. (1964), 'On the Einstein-Podolsky-Rosen Paradox', *Physics* **1**(3), 195–200.
- Bennett, B., Hoffman, D. D. & Prakash, C. (1989), *Observer Mechanics. A Formal Theory of Perception*, Academic Press, Inc., San Diego, CA.
- Bickle, J. (2006), 'Reducing mind to molecular pathways: explicating the reductionism implicit in current cellular and molecular neuroscience', *Synthese* **151**(3), 411–434.
- Bickle, J. (2012), A brief history of neuroscience's actual influences on mind-brain reductionism, *in* S. Gozzano & C. S. Hill (Hg.), 'New Perspectives on Type Identity', Cambridge University Press, Cambridge, S. 88–110.
- Bigelow, J. (2010), 'Quine, mereology, and inference to the best explanation', *Logique et Analyse* **53**(212), 465.
- Bishop, R. C. & Atmanspacher, H. (2006), 'Contextual Emergence in the Description of Properties', *Foundations of Physics* **36**(12), 1753–1777.
- Bitbol, M. (2001), 'Non-representationalist theories of knowledge and quantum mechanics', *SATS* **2**, 37–61.
- Bitbol, M. (2002), 'Science as if situation mattered. Dissolving the "Hard Problem of Consciousness" from our "Being There"', *Phenomenology and the Cognitive Science* **1**, 181–224.
- Block, N. (1978), Troubles with Functionalism, *in* C. W. Savage (Hg.), 'Perception and Cognition', University of Minnesota Press, Minneapolis, MN, S. 261–325.
- Block, N. (1990), 'Inverted Earth', *Philosophical Perspectives* **4**, 53–79.
- Block, N. (1995), 'A confusion about a function of consciousness', *Behavioral and Brain Sciences* **18**, 227–247.
- Block, N. (2011), 'Perceptual consciousness overflows cognitive access', *Trends in Cognitive Sciences* **15**(12), 567–575.

- Bohm, D. (1952a), 'A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I', *Physical Review* **85**(2), 166–179.
- Bohm, D. (1952b), 'A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. II', *Physical Review* **85**(2), 180–193.
- Bohr, N. (1935), 'Can Quantum-Mechanical Description of Physical Reality be Considered Complete?', *Physical Review* **48**, 696–702.
- Boogerd, F. C., Bruggeman, F. J., Richardson, R. C., Stephan, A. & Westerhoff, H. V. (2005), 'Emergence and Its Place in Nature: A Case Study of Biochemical Networks', *Synthese* **145**(1), 131–164.
- Bourbaki, N. (1961), 'Die Architektur der Mathematik', *Physikalische Blätter* **17**. Erster Teil erschienen in Band 17, S. 161–166; zweiter Teil im selben Band, S. 212–218.
- Briegel, H. J. & De Las Cuevas, G. (2012), 'Projective simulation for artificial intelligence', *Scientific Reports* **2**(400). doi:10.1038/srep00400.
- Briegel, H. J. & Müller, T. (2015), 'A Chance for Attributable Agency', *Minds and Machines* **25**(3), 261–279.
- Briegel, H. & Popescu, S. (2014), A perspective on possible manifestations of entanglement in biological systems, in Mohseni *et al.* (2014), S. 277–310.
- Brüntrup, G. (2011), Alter Wein in neuen Schläuchen. Die Renaissance des Panpsychismus in der gegenwärtigen Philosophie des Geistes, in T. Müller & H. Watzka (Hg.), 'Ein Universum voller 'Geiststaub'? Der Panpsychismus in der aktuellen Geist-Gehirn-Debatte', mentis, Paderborn, S. 23–59.
- Brüntrup, G. & Jaskolla, L., Hg. (2016), *Panpsychism: Philosophical Essays*, Oxford University Press, Oxford.
- Burgin, M. (2010), *Theory of Information. Fundamentality, Diversity and Unification*, World Scientific, Singapur.
- Buzsáki, G. (2006), *Rhythms of the Brain*, Oxford University Press, Oxford.
- Byrne, A. & Hilbert, D. R. (2003), 'Color Realism and Color Science', *Behavioral and Brain Sciences* **26**, 1–46.
- Calosi, C. & Graziani, P., Hg. (2014), *Mereology and the Sciences*, Springer, Cham.

- Calosi, C. & Tarozi, G. (2014), Parthood and Composition in Quantum Mechanics, *in* Calosi & Graziani (2014), S. 52–84.
- Carhart-Harris, R. L. & Friston, K. (2010), ‘The default-mode, ego-functions and free-energy: a neurobiological account of Freudian ideas’, *Brain* **133**(4), 1265–1283.
- Carnap, R. (1928), *Der Logische Aufbau der Welt*, Felix Meiner Verlag, Leipzig.
- Casati, R. & Varzi, A. C., Hg. (1999), *Parts and Places. The Structure of Spatial Representation*, MIT Press, Cambridge, MA.
- Cerullo, M. A. (2015), ‘The Problem with Phi: A Critique of Integrated Information Theory’, *PLoS computational biology* **11**(9), e1004286.
- Chalmers, D. J. (1995), ‘Facing Up to the Problem of Consciousness’, *Journal of Consciousness Studies* **2**, 200–219.
- Chalmers, D. J. (1996), *The Conscious Mind. In Search of a Fundamental Theory*, Oxford University Press, Oxford.
- Chalmers, D. J. (2000), What is a Neural Correlate of Consciousness?, *in* T. Metzinger (Hg.), ‘Neural Correlates of Consciousness: Empirical and Conceptual Questions’, MIT Press, Cambridge, MA, S. 17–40.
- Chalmers, D. J. (2004), How Can We Construct a Science of Consciousness?, *in* M. Gazzaniga (Hg.), ‘The Cognitive Neurosciences III’, MIT Press, Cambridge, MA, S. 1111–1120.
- Chalmers, D. J. (2016), The Combination Problem for Panpsychism, *in* Brüntrup & Jaskolla (2016), S. 19–47.
- Churchland, P. M. (1981), ‘Eliminative Materialism and the Propositional Attitudes’, *The Journal of Philosophy* **78**(2), 67–90.
- Churchland, P. S. (1998), Brainshy: Nonneural theories of conscious experience, *in* S. Hameroff, A. W. Kaszniak & A. C. Scott (Hg.), ‘Toward a Science of Consciousness II: The Second Tucson Discussions and Debates’, MIT Press, Cambridge, MA, S. 109–126.
- Clark, A. (2013), ‘Whatever next? Predictive brains, situated agents, and the future of cognitive science’, *Behavioral and Brain Sciences* **36**(3), 181–204.
- Clark, A. & Chalmers, D. J. (1998), ‘The extended mind’, *Analysis* **58**(1), 7–19.

- Clarke, B. L. (1981), 'A Calculus of Individuals Based on Connection', *Notre Dame Journal of Formal Logic* **22**(3), 204–219.
- Conant, R. C. & Ashby, W. R. (1970), 'Every good regulator of a system must be a model of that system', *Int. J. Systems Sci.* **1**(2), 89–97.
- Crane, T. (2001a), *Elements Of Mind: An Introduction to the Philosophy of Mind*, Oxford University Press, Oxford.
- Crane, T. (2001b), The significance of emergence, in B. Loewer & G. Gillett (Hg.), 'Physicalism and its Discontents', Cambridge University Press, Cambridge, S. 207–224.
- Crane, T. & Mellor, D. H. (1990), 'There is No Question of Physicalism', *Mind* **99**(394), 185–206.
- Crick, F. (1994), *The Astonishing Hypothesis: The Scientific Search for the Soul*, Simon & Schuster, London.
- Crick, F. & Koch, C. (2003), 'A framework for consciousness', *Nature Neuroscience* **6**(2), 119–126.
- Damasio, A. & Carvalho, G. B. (2013), 'The nature of feelings: evolutionary and neurobiological origins', *Nature Reviews Neuroscience* **14**(2), 143–152.
- Davidson, D. (1990), *Handlung und Ereignis*, Suhrkamp, Frankfurt am Main.
- Dayan, P., Hinton, G. E., Neal, R. M. & Zemel, R. S. (1995), 'The Helmholtz Machine', *Neural Computation* **7**, 889–904.
- deCharms, R. C., Maeda, F., Glover, G. H., Ludlow, D., Pauly, J. M., Soneji, D., Gabrieli, J. D. E. & Mackey, S. C. (2005), 'Control over brain activation and pain learned by using real-time functional MRI', *Proceedings of the National Academy of Sciences* **102**(51), 18626–18631.
- Dehaene, S. & Naccache, L. (2001), 'Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework', *Cognition* **79**, 1–37.
- Dennett, D. C. (1987), *The Intentional Stance*, MIT Press, Cambridge, MA.
- Dennett, D. C. (1988), Quining Qualia, in 'Consciousness in Contemporary Science', Oxford University Press, Oxford, S. 42–77.
- Dennett, D. C. (1991a), *Consciousness Explained*, Little, Brown & Company, Boston.

- Dennett, D. C. (1991b), 'Real Patterns', *The Journal of Philosophy* **88**(1), 27–51.
- Dennett, D. C. (2003), 'Explaining the Magic of Consciousness', *Journal of Cultural and Evolutionary Psychology* **1**, 7–19.
- Descartes, R. (1996), *Les passions de l'âme*, Felix Meiner Verlag, Hamburg. Französisch-Deutsch (Orig. 1649).
- Descartes, R. (2009), *Meditationen. Mit sämtlichen Einwänden und Erwiderungen*, Felix Meiner Verlag, Hamburg. Original: *Meditationes de prima philosophia, in qua Dei existentia et animae immortalitas demonstratu* (1641).
- Doyle, B. (2010), 'Jamesian free will, the two-stage model of William James', *William James Studies* **5**, 25–50.
- Dretske, F. I. (1983), 'Précis of Knowledge and the Flow of Information', *Behavioral and Brain Sciences* **6**, 55–90.
- Dretske, F. I. (1995), *Naturalizing the Mind*, MIT Press, Cambridge, MA.
- du Bois-Reymond, E. (1974), *Vorträge über Philosophie und Gesellschaft*, Meiner, Hamburg.
- Dunne, J. W. (1934), *The Serial Universe*, Faber & Faber Limited, London.
- Eddington, A. (1928), *The Nature of the Physical World*, Macmillan, New York.
- Edelmann, G. & Tononi, G. (2000), *A Universe Of Consciousness: How Matter Becomes Imagination*, Basic Books, New York.
- Einstein, A., Podolsky, B. & Rosen, N. (1935), 'Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?', *Physical Review* **47**(10), 777–780.
- Engel, A. K. & Singer, W. (2001), 'Temporal binding and the neural correlates of sensory awareness', *Trends in Cognitive Sciences* **5**(1), 16–25.
- Eschenbach, C., Habel, C. & Smith, B., Hg. (1994), *Topological Foundations of Cognitive Science*, Graduiertenkolleg Kognitionswissenschaft, Hamburg.
- Esfeld, M. (2001), *Holism in Philosophy of Mind and Philosophy of Physics*, Kluwer, Dordrecht.
- Esfeld, M. (2003), 'Do relations require underlying intrinsic properties? A physical argument for a metaphysics of relations', *Metaphysica* **4**, 5–25.



- Esfeld, M. (2008), *Naturphilosophie als Metaphysik der Natur*, Suhrkamp Taschenbuch Wissenschaft, Frankfurt am Main.
- Feferman, S. (1995), 'Penrose's Goedelian Argument: A review of Roger Penrose's *Shadows of the Mind*', *Psyche* **2**(7), 21–32.
- Feigl, H. (1958), The "Mental" and the "Physical", in H. Feigl, M. Scriven & G. Maxwell (Hg.), 'Concepts, Theories, and the Mind-Body Problem', Minnesota University Press, Minnesota, S. 370–497.
- Feyerabend, P. K. (1963), 'Materialism and the Mind-Body Problem', *The Review of Metaphysics* **17**(1), 49–66.
- Feynman, R. P. (1949), 'Space-Time Approach to Quantum Electrodynamics', *Physical Review* **76**(6).
- Fierz, M. (1939), 'Über die relativistische Theorie kräftefreier Teilchen mit beliebigen Spin', *Helvetica Physica Acta* **12**, 3–37.
- Findlay, S. D. & Thagard, P. (2012), 'How parts make up wholes', *Frontiers in Physiology* **3**(455). doi: 10.3389/fphys.2012.00455.
- Fine, A. (2007), 'Bohr's Response to EPR: Criticism and Defense', *Iyyun: The Jerusalem Philosophical Quarterly* **56**, 1–26.
- Fisher, M. P. A. (2015), 'Quantum cognition: The possibility of processing with nuclear spins in the brain', *Annals of Physics* **362**, 593–602.
- Flemming, G. & Scholes, G. D. (2014), Quantum biology: introduction, in Mohseni *et al.* (2014), S. 3–24.
- French, S. & Redhead, M. (1988), 'Quantum Physics and the Identity of Indiscernibles', *The British Journal for the Philosophy of Science* **39**(2), 233–246.
- Friston, K. (2010), 'The free-energy principle: a unified brain theory?', *Nature Reviews Neuroscience* **11**(2), 127–138.
- Friston, K. (2013), 'Life as we know it', *Journal of The Royal Society Interface* **10**, 20130475. doi: 10.1098/rsif.2013.0475.
- Friston, K. & Frith, C. (2015), 'A Duet for one', *Consciousness and Cognition* **36**, 390–405.
- Friston, K., Kilner, J. & Harrison, L. (2006), 'A free energy principle for the brain', *Journal of Physiology-Paris* **100**(1-3), 70–87.

- Fuchs, C. A. & Schack, R. (2014), Quantum Measurement and the Paulian Idea, in H. Atmanspacher & C. A. Fuchs (Hg.), 'The Pauli-Jung Conjecture and its Impact Today', Imprint Academic, Exeter, S. 93–108.
- Fuster, J. M. (2005), *Cortex and Mind. Univing Cognition*, Oxford University Press, Oxford.
- Gallagher, S. (2005), *How the body shapes the mind*, University Press, Oxford.
- Gallagher, S. & Allen, M. (2016), 'Active Inference, enactivism and the hermeneutics of social cognition', *Synthese* . doi:10.1007/s11229-016-1269-8.
- Gallagher, S. & Zahavi, D. (2008), *The phenomenological mind: An introduction to the philosophy of mind and cognitive science*, Routledge, London.
- Gendlin, E. T. (1997), *Experiencing and the Creation of Meaning*, Northwestern University Press, Evanston.
- Goff, P. (2006), Experiences Don't Sum, in Strawson & Freeman (2006), S. 53–61.
- Goff, P. (2015), Real Acquaintance and Physicalism, in P. Coates & S. Coleman (Hg.), 'Phenomenal Qualities: Sense, Perception, and Consciousness', Oxford University Press, Oxford, S. 121–146.
- Goodman, N. (1958), 'On Relations That Generate', *Philosophical Studies* 9(5-6), 65–66.
- Griffin, D. R. (1977), Whitehead's Philosophy and Some General Notions of Physics and Biology, in J. B. Cobb Jr. & D. R. Griffin (Hg.), 'Mind in Nature', University Press of America, S. 122–134.
- Gröblacher, S., Paterek, T., Kaltenbaek, R., Brukner, Č., Żukowski, M., Aspelmeyer, M. & Zeilinger, A. (2007), 'An experimental test of non-local realism', *Nature* 446(7138), 871–875.
- Grossmann, R. (1973), *Ontological Reduction*, Indiana University Press, Bloomington, IN.
- Gödel, K. (1931), 'Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme', *Monatshefte für Mathematik und Physik* 38, 173–189.
- Hagan, S., Hameroff, S. R. & Tuszynski, J. A. (2002), 'Quantum computation in brain microtubules: Decoherence and biological feasibility', *Physical Review E* 65(061901).

- Haken, H. (1983), *Synergetik. Eine Einführung*, Springer, New York. 2. deutsche Auflage.
- Halmos, P. R. (1968), *Naive Mengenlehre*, Vandenhoeck & Ruprecht, Göttingen.
- Hameroff, S. (2001), 'Anesthesia: The "Other Side" of Consciousness', *Consciousness and Cognition* **10**(2), 217–229.
- Hameroff, S. & Penrose, R. (2014), 'Consciousness in the universe: A review of the 'Orch OR' theory', *Physics of Life Reviews* **11**, 39–78.
- Hampe, M. (1990), *Die Wahrnehmung der Organismen*, Vandenhoeck & Ruprecht, Göttingen.
- Hampe, M. (2007), *Eine kleine Geschichte des Naturgesetzbegriffs*, Suhrkamp, Frankfurt am Main.
- Hampe, M. (2014), *Die Lehren der Philosophie*, Suhrkamp, Frankfurt am Main.
- Harman, G. (2009), Zero-person and the psyche, in D. Skrbina (Hg.), 'Mind that Abides', John Benjamins Publishing Co., Amsterdam, S. 153–182.
- Harré, R. & Llored, J.-P. (2011), 'Mereologies as the grammars of chemical discourses', *Foundations of Chemistry* **13**(1), 63–76.
- Heidelberger, M. (2002), Wie das Leib-Seele Problem in den Logischen Empirismus kam, in M. Pauen & A. Stephan (Hg.), 'Phänomenales Bewußtsein–Rückkehr zur Identitätstheorie', mentis, Paderborn, S. 40–72.
- Hempel, C. G. & Oppenheim, P. (1948), 'Studies in the Logic of Explanation', *Philosophy of Science* **15**, 135–175.
- Hensen, B., Bernien, H., Dréau, A. E., Reiserer, A., Kalb, N., Blok, M. S., Ruitenberg, J., Vermeulen, R. F. L., Schouten, R. N., Abellán, C., Amaya, W., Pruneri, V., Mitchell, M. W., Markham, M., Twitchen, D. J., Elkouss, D., Wehner, S., Taminiau, T. H. & Hanson, R. (2015), 'Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres', *Nature* **526**(7575), 682–686.
- Hertz, H. (1894), *Prinzipien der Mechanik in neuem Zusammenhang dargestellt*, Johann Ambrosius Barth, Leipzig.
- Hickok, G., Hg. (2015), *Psychonomic Bulletin & Review*, Bd. 22, S. 1477–1576.

- Hochberg, L. R., Serruya, M. D., Friehs, G. M., Mukand, J. A., Saleh, M., Caplan, A. H., Branner, A., Chen, D., Penn, R. D. & Donoghue, J. P. (2006), 'Neuronal ensemble control of prosthetic devices by a human with tetraplegia', *Nature* **442**(7099), 164–171.
- Hoffman, D. D. (2006), 'The scrambling theorem: A simple proof of the logical possibility of spectrum inversion', *Consciousness and Cognition* **15**, 31–45.
- Hoffman, D. D. (2008), 'Conscious Realism and the Mind-Body Problem', *Mind and Matter* **6**(1), 87–121.
- Hoffman, D. D. & Prakash, C. (2014), 'Objects of consciousness', *Frontiers in Psychology* **5**(577). doi: 10.3389/fpsyg.2014.00577.
- Hoffman, D. D., Singh, M. & Prakash, C. (2015), 'The Interface Theory of Perception', *Psychonomic Bulletin & Review* **22**(6), 1480–1506.
- Hohwy, J. (2013), *The Predictive Mind*, Oxford University Press, Oxford.
- Hohwy, J. (2015a), Prediction error minimization, mental and developmental disorder, and statistical theories of consciousness, in R. J. Gennaro (Hg.), 'Disturbed Consciousness: New Essays on Psychopathology and Theories of Consciousness', MIT Press, Cambridge, MA.
- Hohwy, J. (2015b), The Neural Organs Explains the Mind, in T. Metzinger & J. Windt (Hg.), 'Open MIND', MIND Group, Frankfurt am Main. doi: 10.15502/9783958570016.
- Hovda, P. (2009), 'What is Classical Mereology?', *Journal of Philosophical Logic* **38**(1), 55–82.
- Hovda, P. (2014), Natural Mereology and Classical Mereology, in Calosi & Graziani (2014), S. 141–159.
- Husserl, E. (2009), *Logische Untersuchungen*, Felix Meiner Verlag, Hamburg.
- Husserl, E. (2013), *Zur Phänomenologie des inneren Zeitbewußtseins*, Felix Meiner Verlag, Hamburg.
- Jaeger, G. (2009), *Entanglement, Information, and the Interpretation of Quantum Mechanics*, Springer, Heidelberg.
- James, W. (2006), *Pragmatismus und radikaler Empirismus*, Suhrkamp, Frankfurt am Main.

- Juarrero, A. (1999), *Dynamics in Action. Intentional Behavior as a Complex System*, MIT Press, Cambridge, MA.
- Kant, I. (1998), *Kritik der reinen Vernunft*, Felix Meiner, Hamburg.
- Kant, I. (2006), *Kritik der Urteilskraft*, Felix Meiner, Hamburg.
- Kauffman, L. (2003), 'Eigenforms – Objects as Tokens for Eigenbehaviors', *Cybernetics & Human Knowing* **10**(3-4), 71–88.
- Kauffman, S. (1995), *At Home in the Universe: The Search for Laws of Self-Organization and Complexity*, Oxford University Press, Oxford.
- Kemmerling, A. (1996), 'Descartes über das Bewusstsein', *Studia philosophica* **55**, 85–114.
- Kleist, H. v. (1878), 'Über die allmähliche Verfertigung der Gedanken beim Reden', *Nord und Süd* **4**, 3–7.
- Koch, C. (2005), *Bewusstsein – ein neurobiologisches Rätsel*, Elsevier, München.
- Küppers, B.-O. (1990), *Information and the Origin of Life*, MIT Press, Cambridge, MA.
- Ladyman, J., Lambert, J. & Wiesner, K. (2013), 'What is a complex system?', *European Journal for Philosophy of Science* **3**(1), 33–67.
- Ladyman, J. & Ross, D. (2007), *Every Thing Must Go: Metaphysics Naturalized*, Oxford University Press, Oxford.
- Laughlin, R. B. (2007), *Abschied von der Weltformel. Die Neuerfindung der Physik*, Piper, München. Paperback edition; first edition from 2005.
- Lejewski, C. (1982), Ontology: What's Next?, in W. Leitfellner, E. Kraemer & J. Schank (Hg.), 'Langauge and Ontology. Proceedings of the 6th International Wittgenstein Symposium.', Hölder-Pichler-Tempsky, Wien, S. 173–185.
- Leonard, H. S. & Goodman, N. (1940), 'The Calculus of Individuals and its Uses', *Journal of Symbolic Logic* **5**, 45–55.
- Leśniewski, S. (1916/1992), Foundations of the General Theory of Sets. I, in S. J. Surma *et al.* (Hg.), 'Collected Works', Kluwer, Dordrecht, S. 129–173.
- Levine, J. (1983), 'Materialism and Qualia: The Explanatory Gap', *Pacific Philosophical Quarterly* **64**, 354–361.

- Llored, J.-P. & Harré, R. (2014), Developing the Mereology of Chemistry, *in* Calosi & Graziani (2014), S. 189–217.
- Locker, A. (1981), Meta-theoretical Presuppositions for Autopoiesis – Self-Reference and “Autopoiesis”, *in* M. Zeleny (Hg.), ‘Autopoiesis. A Theory of Living Organization’, North Holland, New York, S. 211–235.
- Locker, A. (1998), The Present Status of General Systems Theory, 25 Years after Ludwig von Bertalanffy’s Decease. A Critical Overview, *in* G. E. Lasker (Hg.), ‘Advances in Artificial Intelligence and Engineering Cybernetics, Vol. IV’, The International Institute for Advanced Studies in Systems Research and Cybernetics, Ontario, S. 8–16.
- Loeb, I. (2014), From Mereology to Boolean Algebra: The Role of Regular Open Sets in Alfred Tarski’s Work, *in* K. Mulliagn, K. Kijania-Placek & T. Placek (Hg.), ‘The History and Philosophy of Polish Logic. Essays in Honour of Jan Woleński’, Palgrave Macmillan, Basingstoke, S. 259–277.
- Lorenz, K. (1998), *Indische Denker*, C. H. Beck, München.
- Luisi, P. L. (2002), ‘Emergence in Chemistry: Chemistry as the embodiment of emergence’, *Foundations of Chemistry* 4, 1–18.
- Marchisotto, E. A. & Smith, J. T. (2007), *The Legacy of Mario Pieri in Geometry and Arithmetic*, Birkhäuser, Boston - Basel - Berlin.
- Marr, D. (1982), *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*, W. H. Freeman, San Francisco, CA.
- Maturana, H. R. & Varela, F. J. (1980), *Autopoiesis and Cognition. The Realization of the Living*, D. Reidel, Dordrecht.
- Mausfeld, R. (2012), Der Schein des Realen. Die empiristische Fehlkonzeption der Wahrnehmung und das Wahrnehmungsattribut ‚phänomenal real‘, *in* S. Kluck & S. Volke (Hg.), ‘Näher dran? Zur Phänomenologie des Wahrnehmens’, Karl Alber, Freiburg, S. 192–219.
- Menary, R. (2007), *Cognitive Integration*, Palgrave Macmillan, Basingstoke.
- Metzinger, T. (1995a), Einleitung: Das Problem des Bewußtseins, , S. 15–53.
- Metzinger, T. (2003a), *Being No One. The Self-Model Theory of Subjectivity*, MIT Press, Cambridge, MA.

- Metzinger, T. (2003b), 'Phenomenal transparency and cognitive self-reference', *Phenomenology and the Cognitive Sciences* **2**(4), 353–393.
- Metzinger, T. (2011), The No-Self Alternative, in S. Gallagher (Hg.), 'Oxford Handbook of the Self', Oxford University Press, Oxford, S. 279 – 296.
- Metzinger, T., Hg. (1995b), *Bewußtsein – Beiträge aus der Gegenwartsphilosophie*, mentis, Paderborn.
- Michel, J. G., Boström, K. J. & Pohl, M., Hg. (2016), *Ist der Geist im Kopf? Die These des erweiterten Geistes in Philosophie und Wissenschaft*, mentis, Münster.
- Millikan, R. G. (1989), 'Biosemantics', *Journal of Philosophy* **86**, 281–297.
- Mohseni, M., Omar, Y., Engel, G. & Plenio, M. B., Hg. (2014), *Quantum Effects in Biology*, Cambridge University Press, Cambridge.
- Moltmann, F. (1999), *Parts and Wholes in Semantics*, Oxford University Press, Oxford.
- Munn, K. & Smith, B., Hg. (2008), *Applied Ontology. An Introduction*, ontos verlag, Frankfurt am Main.
- Nagel, E. (1961), *The Structure of Science. Problems in the Logic of Explanation*, Harcourt, Brace & World, New York.
- Nagel, T. (1974), 'What is it like to be a bat?', *The Philosophical Review* **83**, 435–450.
- Nagel, T. (1996), *Letzte Fragen*, Wissenschaftliche Buchgesellschaft, Darmstadt. Erweiterte Deutsche Neuauflage.
- Oizumi, M., Albantakis, L. & Tononi, G. (2014), 'From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0', *PLoS Computational Biology* **10**(5), e1003588.
- Pattee, H. H. (2013), 'Epistemic, Evolutionary, and Physical Conditions for Biological Information', *Biosemiotics* **6**(1), 9–31.
- Pearl, J. (1988), *Probabilistic reasoning in intelligent systems: networks of plausible inference*, Morgan Kaufmann, San Fransisco, CA.
- Peirce, C. S. (1983), *Phänomen und Logik der Zeichen*, Suhrkamp, Frankfurt am Main.

- Peirce, C. S. (1991), *Naturordnung und Zeichenprozeß. Schriften über Semiotik und Naturphilosophie. Mit einem Vorwort von Ilya Prigogine. Herausgegeben und eingeleitet von Helmut Pape*, Suhrkamp, Frankfurt am Main.
- Penrose, R. (1994), *Shadows of the Mind*, Oxford University Press, Oxford.
- Perry, E. K., Ashton, H. & Young, A. H., Hg. (2002), *Neurochemistry of Consciousness: Neurotransmitters in Mind*, Advances in Consciousness Research, John Benjamins Publishing Co., Amsterdam.
- Perry, J. R. (2001), *Knowledge, Possibility, and Consciousness*, MIT Press, Cambridge, MA.
- Petitot, J., Varela, F. J., Pachoud, B. & Roy, J.-M., Hg. (1999), *Naturalizing Phenomenology. Issues in Contemporary Phenomenology and Cognitive Science*, Stanford University Press, Stanford, CA.
- Pockett, S. (2014), 'Problems with theories that equate consciousness with information or information processing', *Frontiers in Systems Neuroscience* **225**. doi: 10.3389/fnsys.2014.00225.
- Poggio, G. F. & Poggio, T. (1984), 'The Analysis of Stereopsis', *Annual Review of Neuroscience* S. 379–412.
- Popper, K. (1978), Three Worlds, in 'The Tanner Lecture on Human Values', S. 1–27.
- Prentner, A. (2004), *Bewusstseinsverändernde Pflanzen von A-Z*, Springer, Wien.
- Prentner, R. (2014), 'A Framework for Critical Materialists', *Mind and Matter* **12**(1), 93–118.
- Prentner, R. (2016), Erweiteter Geist – Erweitertes Bewusstsein?, in Michel *et al.* (2016), S. 109–131.
- Prentner, R. (2017), 'Chemistry, Context and the Objects of Thought', *Foundations of Chemistry* **19**(1), 29–41.
- Price, D. J. & Barrell, J. J. (2012), *Inner Experience and Neuroscience*, MIT Press, Cambridge, MA.
- Primas, H. (1994), Realism and Quantum Mechanics, in D. Prawitz, B. Skyrms & D. Westerståhl (Hg.), 'Logic, Methodology and Philosophy of Science IX', Elsevier Science B.V., S. 609–631.



- Primas, H. (1998), 'Emergence in exact natural science', *Acta Polytechnica Scandinavica* **91**, 83–98.
- Quine, W. V. (1951), 'Two Dogmas of Empiricism', *The Philosophical Review* **60**, 20–43.
- Rescher, N. (1955), 'Axioms for the Part Relation', *Philosophical Studies* **6**(1), 8–11.
- Revonsuo, A. (1999), 'Binding and the Phenomenal Unity of Consciousness', *Consciousness and Cognition* **8**(2), 173–185.
- Ritz, T., Adem, S. & Schulten, K. (2000), 'A Model for Photoreceptor-Based Magnetoreception in Birds', *Biophysical Journal* **78**(2), 707–718.
- Ros, A. (2005), *Materie und Geist. Eine philosophische Untersuchung*, mentis, Paderborn.
- Ros, A. (2008), 'Mentale Verursachung und mereologische Erklärung. Eine einfache Lösung für ein komplexes Problem', *Deutsche Zeitschrift für Philosophie* **56**, 176–203.
- Rovelli, C. (1996), 'Relational Quantum Mechanics', *International Journal of Theoretical Physics* **35**(8), 1637–1678.
- Russell, B. (1914), 'On the Nature of Acquaintance. Preliminary Description of Experience', *The Monist* **24**(1), 1–16.
- Russell, B. (1927), *The Analysis of Matter*, Kegan Paul, Trench & Trubner, London.
- Sarovar, M., Ishizaki, A., Fleming, G. R. & Whaley, K. B. (2010), 'Quantum entanglement in photosynthetic light-harvesting complexes', *Nature Physics* **6**(6), 462–467.
- Schaffer, J. (2010), 'Monism: The Priority of the Whole', *Philosophical Review* **119**(1), 31–76.
- Scheibe, E. (1997), *Die Reduktion physikalischer Theorien. Ein Beitrag zur Einheit der Physik. Teil I: Grundlagen und elementare Theorie*, Springer, Berlin.
- Scheibe, E. (2006), *Die Philosophie der Physiker*, C.H. Beck, München.
- Schlosshauer, M. A. (2008), *Decoherence and the Quantum-to-Classical Transition*, Springer, Heidelberg.
- Schopenhauer, A. (1972), *Die Welt als Wille und Vorstellung*, 3. Auflage, Brockhaus, Wiesbaden. Folgt der erweiterten 2. Auflage aus dem Jahr 1844.

- Schütt, H.-P. (1990), *Substanzen, Subjekte und Personen. Eine Studie zum Cartesischen Dualismus*, Manutius Verlag, Heidelberg.
- Schüz, M. (1986), *Die Einheit des Wirklichen. Carl Friedrich von Weizsäckers Denkweg*, Günther Neske, Pfullingen.
- Schwartenbeck, P., FitzGerald, T. H. B., Mathys, C., Dolan, R., Kronbichler, M. & Friston, K. (2015), 'Evidence for surprise minimization over value maximization in choice behavior', *Scientific Reports* S. 1–14.
- Seager, W. (1995), 'Consciousness, information and panpsychism', *Journal of Consciousness Studies* **2**(3), 272–288.
- Seager, W. (2006), The 'intrinsic nature' argument for panpsychism, in Strawson & Freeman (2006), S. 129–145.
- Seager, W. (2012), *Natural Fabrications. Science, Emergence and Consciousness*, Springer, Heidelberg.
- Searle, J. R. (1980), 'Minds, brains, and programs', *Behavioral and Brain Sciences* **3**, 417–457.
- Searle, J. R. (1991), *Intentionalität. Eine Abhandlung zur Philosophie des Geistes*, Suhrkamp, Frankfurt am Main.
- Searle, J. R. (1992), *The Rediscovery of the Mind*, MIT Press, Cambridge, MA.
- Sebel, P., Bowdle, T., Ghoneim, M., Rampil, I., Padilla, R., Gan, T. & Domino, K. (2004), 'The incidence of awareness during anesthesia: A multicenter united states study', *Anesthesia and Analgesia* S. 833—841.
- Sechzer, J. A., Lieberman, K. W. & Alexander, G. J. (1986), 'Aberrant parenting and delayed offspring development in rats exposed to lithium', *Biological Psychiatry* **21**, 1258–1266.
- Seibt, J. (1996), 'The Myth of Substance and the Fallacy of Misplaced Concreteness', *Acta Analytica* **15**, 119–139.
- Sellars, W. (1956), Empiricism and the philosophy of mind, in H. Feigl & M. Scriven (Hg.), 'Minnesota Studies in the Philosophy of Science, vol. I', University of Minnesota Press, Minneapolis, MN, S. 253 — 329.
- Seth, A. K. (2013), 'Interoceptive inference, emotion, and the embodied self', *Trends in Cognitive Sciences* **17**(11), 565–573.

- Sewell, G. (2002), *Quantum Mechanics and Its Emergent Macrophysics*, Princeton University Press, Princeton.
- Shannon, C. E. (1993), *Collected Papers*, IEEE Press, New York.
- Shannon, C. E. & Weaver, W. (1949), *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, IL. Reprinted in 1998, with a foreword by Richard E. Blahut and Bruce Hajek.
- Sieroka, N. (2007), ‘Hertzian Pictures of Quantum Field Theory’, *Philosophia Naturalis* **44**(1), 88–113.
- Sieroka, N. (2009), Schellingsches Natur- und Materieverständnis im und um das 20. Jahrhundert, in ‘Brüsseler Kongress der Internationalen Fichte-Gesellschaft’.
- Sieroka, N. (2015), *Leibniz, Husserl, and the Brain*, Palgrave Macmillan, Basingstoke.
- Silberstein, M. (2014), ‘Experience Unbound: Neutral Monism, Contextual Emergence and Extended Cognitive Systems’, *Mind and Matter* **12**, 289 – 340.
- Simons, P. (1987), *Parts. A Study in Ontology*, Clarendon Press, Oxford.
- Simons, P. (2003), ‘The Universe’, *Ratio* **16**(3), 236 – 250.
- Simons, P. (2006), ‘Real Wholes, Real Parts: Mereology without Algebra’, *The Journal of Philosophy* **103**(12), 597–613.
- Singh, M. & Hoffman, D. D. (1997), ‘Constructing and representing visual objects’, *Trends in Cognitive Sciences* **1**(3), 98–102.
- Skrbina, D. (2005), *Panpsychism in the West*, MIT Press, Cambridge, MA.
- Smart, J. J. C. (1959), ‘Sensations and Brain Processes’, *The Philosophical Review* **68**(2), 141–156.
- Smart, J. J. C. (2006), Ockhamist Comments on Strawson, in Strawson & Freeman (2006), S. 158–162.
- Smith, B. (1994), Introduction, in C. Eschenbach, H. C. & S. B. (Hg.), ‘Topological Foundations of Cognitive Science’, Graduiertenkolleg Kognitionswissenschaft, Hamburg.
- Smith, B. (1995), ‘Zur Kognition räumlicher Grenzen: Eine mereotopologische Untersuchung’, *Kognitionswissenschaft* **4**, 177–184.

- Smith, B. (1996), 'Mereotopology: a theory of parts and boundaries', *Data & Knowledge Engineering* **20**, 287–303.
- Smith, B. & Mulligan, K. (1982), Pieces of a Theory, in B. Smith (Hg.), 'Parts and Moments. Studies in Logic and Formal Ontology', Philosophia Verlag, München, Wien., S. 15–109.
- Smith, B. & Mulligan, K. (1983), 'Framework for Formal Ontology', *Topoi* **2**, 73–85.
- Solms, M. (2013), 'The Conscious Id', *Neuropsychanalysis* **15**(1), 5–19.
- Stephan, A. (1999), *Emergenz: Von der Unvorhersagbarkeit zur Selbstorganisation*, Dresden University Press, Dresden.
- Stephan, A. (2006), Zur Rolle des Emergenzbegriffs in der Philosophie des Geistes und in der Kognitionswissenschaft, in D. Sturma (Hg.), 'Philosophie und Neurowissenschaften', Suhrkamp, Frankfurt am Main, S. 146–166.
- Stephan, A. (2016), Emergence and Panpsychism, in Brüntrup & Jaskolla (2016), S. 334 – 348.
- Strawson, G. (2006), Realistic Monism: Why Physicalism Entails Panpsychism, in Strawson & Freeman (2006), S. 3–31.
- Strawson, G. (2008), *Real Materialism and other essays*, Oxford University Press, Oxford.
- Strawson, G. (2009), *Selves: An Essay in Revisionary Metaphysics*, Oxford University Press, Oxford.
- Strawson, G. (2016), Mind and Being: The Primacy of Panpsychism, in Brüntrup & Jaskolla (2016), S. 75–112.
- Strawson, G. & Freeman, A., Hg. (2006), *Consciousness and its place in nature. Does physicalism entail panpsychism?*, Imprint Academic, Exeter.
- Tarski, A. (1935), 'Zur Grundlegung der Booleschen Algebra. I', *Fundamenta Mathematicae* **24**, 177–198.
- Tarski, A. (1956), Foundation of the Geometry of Solids, in 'Logic, Semantics, Metamathematics', Clarendon Press, Oxford, S. 24–29. Übersetzte und überarbeitete Fassung von 'Le fondements de la géométrie des corps', erschienen 1929 in den *Annales des Société Polonais des Mathématiques*.

- Tegmark, M. (2000), 'Importance of quantum decoherence in brain processes', *Phys. Rev. E* **61**(4), 4194–4206.
- Thompson, E. (2007), *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*, Harvard University Press, Cambridge, MA.
- Thompson, E. & Varela, F. J. (2001), 'Radical embodiment: neural dynamics and consciousness', *Trends in Cognitive Sciences* **5**(10), 418–425.
- Tononi, G. (2008), 'Consciousness as Integrated Information: a Provisional Manifesto.', *Biol. Bull.* **215**, 216–242.
- Tononi, G. & Koch, C. (2015), 'Consciousness: here, there and everywhere?', *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**(1668), 20140167–20140167.
- Tsuchiya, N. & Koch, C. (2009), The relationship between consciousness and attention, in S. Laureys & G. Tononi (Hg.), 'The Neurology of Consciousness', Academic Press, New York, S. 63–78.
- van Gelder, T. (1999), Wooden iron? Husserlian Phenomenology meets Cognitive Science, in Petitot *et al.* (1999), S. 245–265.
- van Tuijl, H. F. J. M. (1975), 'A new visual illusion: Neonlike color spreading and complementary color induction between subjective contours', *Acta Psychologica* **39**(6), 441–445.
- Varela, F. J. (1975), 'A Calculus for Self-reference', *International Journal of General Systems* **2**(1), 5–24.
- Varela, F. J. (1998), A Science of Consciousness as If Experience Mattered, in S. Hammeroff, A. W. Kaszniak & A. C. Scott (Hg.), 'Toward a Science of Consciousness II: The Second Tucson Discussions and Debates', MIT Press, Cambridge, MA, S. 31–44.
- Varela, F. J. (1999), The specious present: A neurophenomenology of time consciousness, in Petitot *et al.* (1999), S. 266–314.
- Varela, F. J., Thompson, E. & Rosch, E. (1991), *The Embodied Mind: Cognitive Science and Human Experience*, MIT Press, Cambridge, MA.
- Velmans, M. (2009), *Understanding Consciousness*, Routledge, London.

- Vollenweider, F. X. (1994), Kortiko-Subkortikale Dysbalance während Ketamin- und Psilocybin- induzierten Bewusstseinszuständen, *in* A. Dittrich, A. Hofmann & L. H. (Hg.), 'Welten des Bewusstseins. Band 3: Experimentelle Psychologie, Neurobiologie und Chemie', VWB - Verlag für Wissenschaft und Bildung, Berlin, S. 25–43.
- von Bertalanffy, L. (1968), *General System Theory: Foundations, Development, Applications*, Braziller, New York.
- von Foerster, H. (1981), Objects: tokens for (eigen-) behaviors, *in* 'Observing Systems', Intersystems Publications, Seaside, CA, S. 274–285.
- Wackerbauer, R., Witt, A., Atmanspacher, H., Kurths, J. & Scheingraber, H. (1994), 'A comparative classification of complexity measures', *Chaos, Solitons & Fractals* **4**(1), 133–173.
- Wallace, D. (2012), *The Emergent Multiverse. Quantum Theory according to the Everett Interpretation*, Oxford University Presse, Oxford.
- Wang, Z., Busemeyer, J. R., Atmanspacher, H. & Pothos, E. M. (2013), 'The Potential of Using Quantum Theory to Build Models of Cognition', *Topics in Cognitive Science* **5**, 672–688.
- Whitehead, A. N. (1919), *An Enquiry Concerning the Principles of Natural Knowledge*, Cambridge University Press, Cambridge.
- Whitehead, A. N. (1920), *The Concept of Nature*, Cambridge University Press, Cambridge.
- Whitehead, A. N. (1929), *Process and Reality: An Essay in Cosmology*, Macmillan, New York.
- Wiese, W. & Metzinger, T. (2012), Desiderata for a mereotopological theory of consciousness, *in* 'Being in Time: Dynamical Models of Phenomenal Experience', John Benjamins Publishing Co., S. 185–209.
- Zahavi, D. (2013), 'Naturalized Phenomenology: A Desideratum or a Category Mistake?', *Royal Institute of Philosophy Supplements* **72**, 23–42.
- Zermelo, E. (1908), 'Untersuchungen über die Grundlagen der Mengenlehre', **65**, 261–281.

## Autorenindex

- Abellán, C. 32
- Adem, S. 180
- Albantakis, L. 60, 67
- Alexander, G. J. 180
- Allen, M. 183
- Amaya, W. 32
- Anderson, P. W. 51
- Ashby, W. R. 25, 71, 74, 95
- Aspect, A. 6
- Aspelmeyer, M. 32
- Atmanspacher, H. 27, 33, 56, 67–69, 180
- Baars, B. J. 32, 78
- Bachelard, G. 85
- Balduzzi, D. 51
- Barrell, J. J. 147
- Barrett, A. 69
- Bayne, T. 159, 160, 172
- Bedau, M. A. 56, 67
- beim Graben, P. 56, 69
- Bell, J. S. 6
- Bennett, B. 59
- Bernien, H. 32
- Bickle, J. 66, 170
- Bigelow, J. 106
- Bishop, R. C. 27, 56, 68
- Bitbol, M. 59, 177
- Block, N. 59, 80, 193
- Blok, M. S. 32
- Bohm, D. 32
- Bohr, N. 6
- Boogerd, F. C. 49
- Bourbaki, N. 60
- Bowdle, T. 34
- Branner, A. 81
- Briegel, H. 180
- Briegel, H. J. 180, 197
- Bruggeman, F. J. 49
- Brukner, Č. 32
- Brüntrup, G. 34
- Burgin, M. 51, 144
- Busemeyer, J. R. 180
- Buzsáki, G. 51
- Byrne, A. 157
- Calosi, C. 99
- Caplan, A. H. 81

- Carhart-Harris, R. L. 21, 76
- Carnap, R. 86, 147
- Carvalho, G. B. 177
- Cerullo, M. A. 60
- Chalmers, D. J. 1, 15, 20, 43, 59, 67, 83, 95, 159, 160, 171, 172, 183, 197
- Chen, D. 81
- Churchland, P. M. 45
- Churchland, P. S. 180
- Clark, A. 183
- Clarke, B. L. 119
- Conant, R. C. 71
- Crane, T. 33, 34, 49, 171, 172
- Crick, F. 1, 32
- Damasio, A. 177
- Davidson, D. 45
- Dayan, P. 90
- De Las Cuevas, G. 180, 197
- deCharms, R. C. 81
- Dehaene, S. 32
- Dennett, D. C. 4, 9, 27, 67, 78, 165, 170
- Descartes, R. 10
- Dolan, R. 76
- Domino, K. 34
- Donoghue, J. P. 81
- Doyle, B. 197
- Dréau, A. E. 32
- Dretske, F. I. 20, 67, 95, 114, 155
- du Bois-Reymond, E. 1
- Dunne, J. W. 10
- Eddington, A. 17, 86
- Edelmann, G. 1
- Einstein, A. 4
- Elkouss, D. 32
- Engel, A. K. 160
- Esfeld, M. 17, 32, 43, 144
- Feferman, S. 180
- Feigl, H. 147
- Feyerabend, P. K. 45
- Feynman, R. P. 177
- Fierz, M. 144
- Findlay, S. D. 99
- Fine, A. 32
- Fisher, M. P. A. 180
- FitzGerald, T. H. B. 76
- Fleming, G. R. 180
- Flemming, G. 180
- French, S. 144
- Friehs, G. M. 81



- Friston, K. 21, 51, 74, 76, 90, 92, 94, 143, 183, 197
- Frith, C. 197
- Fuchs, C. A. 59
- Fuster, J. M. 1
- Gabrieli, J. D. E. 81
- Gallagher, S. 26, 171, 183
- Gan, T. 34
- Gendlin, E. T. 85
- Ghoneim, M. 34
- Glover, G. H. 81
- Goff, P. 19, 157
- Goodman, N. 100, 112, 170
- Grangier, P. 6
- Griffin, D. R. 191
- Gröblacher, S. 32
- Grossmann, R. 154
- Gödel, K. 180
- Hagan, S. 180
- Haken, H. 51, 92, 93
- Halmos, P. R. 61
- Hameroff, S. 1, 21, 32, 180
- Hameroff, S. R. 180
- Hampe, M. 59, 98, 191, 197
- Hanson, R. 32
- Harman, G. 150
- Harré, R. 99
- Harrison, L. 90
- Heidelberger, M. 147
- Hempel, C. G. 139, 142
- Hensen, B. 32
- Hertz, H. 42
- Hilbert, D. R. 157
- Hinton, G. E. 90
- Hochberg, L. R. 81
- Hoffman, D. D. 37, 59, 98, 153, 183
- Hohwy, J. 76, 90, 183
- Hovda, P. 99, 124
- Husserl, E. 97, 136, 167, 172
- Ishizaki, A. 180
- Jaeger, G. 32
- James, W. 25, 147, 184, 188, 197
- Juarrero, A. 24
- Kalb, N. 32
- Kaltenbaek, R. 32
- Kant, I. 26, 150, 153, 159, 160
- Kauffman, L. 25, 163
- Kauffman, S. 51

- Kemmerling, A. 10
- Kilner, J. 90
- Kleist, H. v. 81
- Koch, C. 1, 32, 37, 80
- Kronbichler, M. 76
- Kronz, F. 56
- Küppers, B.-O. 51
- Kurths, J. 67
- Ladyman, J. 43, 56, 67, 92
- Lambert, J. 67, 92
- Laughlin, R. B. 51
- Lejewski, C. 121
- Leonard, H. S. 100, 170
- Leśniewski, S. 97, 100
- Levine, J. 1
- Lieberman, K. W. 180
- Llored, J.-P. 99
- Locker, A. 26
- Loeb, I. 109
- Lorenz, K. 171
- Ludlow, D. 81
- Luisi, P. L. 49
- Mackey, S. C. 81
- Maeda, F. 81
- Marchisotto, E. A. 124
- Markham, M. 32
- Marr, D. 81
- Mathys, C. 76
- Maturana, H. R. 25
- Mausfeld, R. 59, 81
- Mellor, D. H. 33
- Menary, R. 197
- Metzinger, T. 81, 147, 159, 163
- Millikan, R. G. 114
- Mitchell, M. W. 32
- Moltmann, F. 127
- Mukand, J. A. 81
- Müller, T. 180
- Mulligan, K. 45, 97, 123, 127
- Naccache, L. 32
- Nagel, E. 66, 68
- Nagel, T. 1, 60
- Neal, R. M. 90
- Oizumi, M. 60, 67
- Oppenheim, P. 139, 142
- Padilla, R. 34
- Paterek, T. 32
- Pattee, H. H. 51

- Pauly, J. M. 81
- Pearl, J. 74, 183
- Peirce, C. S. 83, 90, 98
- Penn, R. D. 81
- Penrose, R. 32, 180
- Perry, J. R. 81
- Pockett, S. 51
- Podolsky, B. 4
- Poggio, G. F. 81
- Poggio, T. 81
- Popescu, S. 180
- Popper, K. 14, 33
- Pothos, E. M. 180
- Prakash, C. 59, 183
- Prentner, A. 1
- Prentner, R. 26, 27, 99, 166, 193, 197
- Price, D. J. 147
- Primas, H. 10, 27, 33, 68
- Pruneri, V. 32
- Quine, W. V. 171
- Rampil, I. 34
- Redhead, M. 144
- Reiserer, A. 32
- Rescher, N. 114
- Revonsuo, A. 163
- Richardson, R. C. 49
- Ritz, T. 180
- Roger, G. 6
- Ros, A. 59, 65
- Rosch, E. 183
- Rosen, N. 4
- Ross, D. 43, 56
- Rovelli, C. 27
- Ruitenberg, J. 32
- Russell, B. 17, 102, 103, 106, 157
- Saleh, M. 81
- Sarovar, M. 180
- Schack, R. 59
- Schaffer, J. 97
- Scheibe, E. 66, 96
- Scheingraber, H. 67
- Schlosshauer, M. A. 33
- Scholes, G. D. 180
- Schopenhauer, A. 150
- Schouten, R. N. 32
- Schulten, K. 180
- Schütt, H.-P. 106
- Schüz, M. 96

- Schwartenbeck, P. 76
- Seager, W. 17, 18, 20, 51, 56
- Searle, J. R. 51, 67, 74, 98
- Sebel, P. 34
- Sechzer, J. A. 180
- Seibt, J. 144
- Sellars, W. 67
- Serruya, M. D. 81
- Seth, A. K. 76
- Sewell, G. 68
- Shannon, C. E. 51, 67
- Sieroka, N. 44, 96, 167, 177
- Silberstein, M. 24
- Simons, P. 27, 97, 99, 100, 102, 103, 105, 106, 115, 123, 128, 136, 172
- Singer, W. 160
- Singh, M. 59, 153, 183
- Skrbina, D. 15
- Smart, J. J. C. 19, 66
- Smith, B. 45, 97, 115, 123, 127
- Smith, J. T. 124
- Solms, M. 1
- Soneji, D. 81
- Stephan, A. 27, 49, 56, 67, 93, 150, 197
- Strawson, G. 15, 18, 19, 32, 147, 149, 150
- Taminiau, T. H. 32
- Tarozzi, G. 99
- Tarski, A. 97, 107, 110, 124, 188
- Tegmark, M. 180
- Thagard, P. 99
- Thompson, E. 24, 34, 177, 183
- Tononi, G. 1, 21, 32, 37, 51, 60, 67
- Tsuchiya, N. 80
- Tuszynski, J. A. 180
- Twitchen, D. J. 32
- van Gelder, T. 177
- van Tuijl, H. F. J. M. 157
- Varela, F. J. 24, 25, 177, 183
- Velmans, M. 21, 60, 78, 83, 88, 147
- Vermeulen, R. F. L. 32
- Vollenweider, F. X. 1
- von Bertalanffy, L. 24
- von Foerster, H. 163
- Wackerbauer, R. 67
- Wallace, D. 33
- Wang, Z. 180
- Weaver, W. 67

- 
- |  |                         |
|--|-------------------------|
| Wehner, S. 32  | Witt, A. 67             |
| Westerhoff, H. V. 49   |                         |
| Whaley, K. B. 180  | Zahavi, D. 26, 171, 177 |
| Whitehead, A. N. 97, 98, 105, 115, 144,<br>165, 188, 191, 192, 197 | Zeilinger, A. 32        |
| Wiese, W. 159  | Zemel, R. S. 90         |
| Wiesner, K. 67, 92   | Zermelo, E. 103         |
|  | Żukowski, M. 32         |