UNIVERSITÄT BERN

Von freien Personen und unfreien Zombies

Philosophische Überlegungen zum Verhältnis von Erleben eigener Urheberschaft und Willensfreiheit

Masterarbeit in allgemeiner Philosophie unter Leitung von Prof. Dr. Claus Beisbart

Thomas König Matrikel #: 85-917-151 thomas.koenig@students.unibe.ch

Dieser Text hat ca. 27100 Wörter (ohne Referenzen, Inhaltsverzeichnis und redundantem Text in den Argumentationskarten).

Ich danke Claus Beisbart und Vera Hoffmann-Kolss für die anregende wohlwollend-kritisch-konstruktive Betreuung der Arbeit und Lea König für das sorgfältige Gegenlesen.

Inhalt

| 1 | Einleitung und Motivierung der Fragestellung | | | | | |
|---|--|--|-----------|--|--|--|
| 2 | Laie | Laienintuitionen zu Willensfreiheit | | | | |
| | 2.1 | Warum Laienmeinungen? | 6 | | | |
| | 2.2 | Empirische Datenbasis | 8 | | | |
| | 2.3 Laienmeinungen zur Willensfreiheit | | | | | |
| | 2.4 | 9 | | | | |
| 3 | Epistemik von Selbsterleben 13 | | | | | |
| | 3.1 | Das Phänomen der Unmittelbarkeit von Selbsterleben | 14 | | | |
| | 3.2 | Erklärungen für die Unmittelbarkeit von Selbsterleben | 14 | | | |
| | | 3.2.1 Philosophische Gründe | 15 | | | |
| | | 3.2.2 Informationstheoretische Gründe | 17 | | | |
| | | 3.2.2.1 Informationsintegration und psychophysiologische | | | | |
| | | Intransparenz | 18 | | | |
| | | 3.2.2.2 Informationsintegration und Meinheit | 19 | | | |
| | | 3.2.2.3 Informationsintegration und Privatheit | 20 | | | |
| | 3.3 | Subjektives Erleben, Repräsentation im Gehirn und Leiblichkeit | 21 | | | |
| | 3.4 | Metzingers These der transparenten Selbstrepräsentation 24 | | | | |
| | 3.5 | Fehler im Selbsterleben | 25 | | | |
| | 3.6 | Zwischenfazit | 27 | | | |
| | | | | | | |
| 4 | Ges | talt von Selbsterleben und eigener Urheberschaft | 29 | | | |
| | 4.1 | Mein-Sein von Selbsterleben | 30 | | | |
| | 4.2 | Selbst Ursache sein und basale Handlungen | 32 | | | |
| | | 4.2.1 Was sind basale Handlungen? | 32 | | | |
| | | 4.2.2 Psychophysische Token-Token Identitäten | 33 | | | |
| | | 4.2.3 Mein-Sein, Epistemik und Materialisierung basaler Handlun- | | | | |
| | | gen | 35 | | | |
| | 4.3 | Alternative Handlungsmöglichkeiten | 37 | | | |
| | 4.4 | Eigene Urheberschaft und Gültigkeitsbereichs des Selbst-Modells | 39 | | | |
| | | 4.4.1 Pathologisches Erleben von Urheberschaft | 46 | | | |
| | 4.5 | Problematik der kausalen Relevanz von erlebter eigener Urheberschaft | 46 | | | |

| 5 | Erleben von Urheberschaft im Dilemma zwischen Freiheit und | | | | | |
|----|--|---------------|---|----------|--|--|
| | Det | ermin | ismus | 50 | | |
| | 5.1 | Deter | minismus, Vorhersehbarkeit und Erleben von Urheberschaft | 52 | | |
| | | 5.1.1 | Physikalischer Determinismus | 52 | | |
| | | 5.1.2 | Definition von präreflektivem Erleben von Urheberschaft | 54 | | |
| | | 5.1.3 | Vorhersehbarkeit der Folgen eigener Willensakte dank Ge- | | | |
| | | | setzmässigkeiten | 55 | | |
| | 5.2 | Physis | sche Erklärbarkeit des Erlebens von Urheberschaft | 56 | | |
| | | 5.2.1 | Wirksamkeit dessen, was als eigene Ursächlichkeit erlebt wird | 56 | | |
| | | 5.2.2 | Vorhersehbarkeit dessen, was als selbst verursacht erlebt wird | 57 | | |
| | | 5.2.3 | Erleben von Urheberschaft und physische Gesetzmässigkei- | | | |
| | | | ten | 58 | | |
| | | 5.2.4 | Inkompatibilismus von physikalischen Begründungen von Er- | | | |
| | | | leben eigener Urheberschaft und Willensfreiheit | 62 | | |
| | | 5.2.5 | Erleben von Urheberschaft ohne Willensfreiheit | 66 | | |
| | 5.3 | | rbarkeit des Erlebens von Urheberschaft durch ein Selbst-Modell | | | |
| | | 5.3.1 | Kritik des physikalistischen Zugangs | 67 | | |
| | | 5.3.2 | Vorhersehbarkeit durch ein Selbst-Modell | 69 | | |
| | | 5.3.3 | Multiple Realisierbarkeit repräsentationaler Zustände und | | | |
| | | | Determinismus | 72 | | |
| | | 5.3.4 | Folgen der multiplen Realisierbarkeit von eigener Urheber- | | | |
| | ٠. | | schaft | 77 | | |
| | 5.4 | | ale Relevanz von repräsentationalen und physischen Zuständen | 78 | | |
| | | 5.4.1 | Kausalerklärungen und realisierungsunabhängige Abhängig- | | | |
| | | | keitsverhältnisse | 79 | | |
| | | 5.4.2 | Verursachung als Eigenschaft zweiter Ordnung | 81 | | |
| | | 5.4.3 | Willensfreie Personen und willenlose Zombies | 86 | | |
| | 5.5 | _ | venierende Kausalerklärungen und Laienüberzeugungen zu Wil- | 0.7 | | |
| | | | eiheit | 87 | | |
| | | 5.5.1 | Wahlmöglichkeiten | 87 | | |
| | | 5.5.2 | Willensfreiheit | 88 | | |
| | | 5.5.3 | Selbstbestimmung | 88 | | |
| | | 5.5.4 $5.5.5$ | Kontrolle | 89 89 | | |
| | 5.6 | | | 09 | | |
| | 5.6 | | schaften einer durch Erleben von Urheberschaft erklärbaren nsfreiheit | 90 | | |
| | | willer | Istrement | 90 | | |
| 6 | Faz | it | | 92 | | |
| Li | Literatur | | | | | |

KAPITEL 1

Einleitung und Motivierung der Fragestellung

"Als ein vernünftiges, mithin zur intelligibelen Welt gehöriges Wesen kann der Mensch die Kausalität seines eigenen Willens niemals anders als unter der Idee der Freiheit denken."

Immanuel Kant

"Ich", und wohl auch "Du", haben die starke Intuition, dass eine Person, also dieses "Ich", eine hinreichende Ursache eines physischen Ereignisses sein kann, und für dieses Ereignis als etwas das "Ich" selbst gewählt und gewollt habe verantwortlich bin. Zudem erleben wir eigenes Wollen typischerweise in überwältigender Weise nicht als verursacht durch dafür hinreichende physische Ereignisse. "Wollen" scheint also etwas zu sein, das sich der Person zuverlässig als Ursache, nicht aber als Folge von physischen Ereignissen manifestiert. Um diese Qualität unseres Wollens zu benennen, verwenden wir den Begriff der Willensfreiheit. Indem Willensfreiheit ich-fremde physische Determinanten bestimmter Handlungen ausschliesst, scheint sie mich zu sogenannten Willkürhandlungen als bewusst und willentlich wählbares, und jederzeit kontrollierbares Tun zu ermächtigen, und mich damit zur autonomen Person zu machen. Entsprechend unterscheiden wir Willkürhandlungen von unwillkürlichen Bewegungen wie etwa Reflexen: Als eindeutige Folge etwa von physiologischen Ereignissen empfinden wir Reflexbewegungen nicht als gewollt, sondern als unwillkürlich und ausserhalb dessen, was uns zur autonomen und verantwortlichen Person macht. Intuitiv scheint es uns nicht nur plausibel, dass diese Unterscheidung zwischen willkürlichem und unwillkürlichem Tun möglich ist, sondern auch, dass sie die Basis bildet für weitere wichtige Begriffe wie persönliche Verantwortung und personale Autonomie. Wenn, wie die Intuition besagt, diese Unterscheidung möglich sein sollte, setzt das die Möglichkeit von Willensfreiheit voraus.

Umgekehrt scheint es, dass es aus naturwissenschaftlicher Perspektive gar keinen Raum gibt für etwas wie Willensfreiheit: Die Naturwissenschaften gehen üblicherweise davon aus, dass sich im Prinzip jedes physische Ereignis hinreichend erklären lässt durch das Auftreten einer bestimmten physisch definierbaren Sachlage und der allgemeinen Gültigkeit von Naturgesetzen. Dieser Standpunkt wird als physikalischer Determinismus bezeichnet. Wenn ein physikalischer Determinismus wahr wäre, so wird weiter argumentiert, muss jedes Ereignis, das ein physisches Ereignis erklärt, ebenfalls ein physisches Ereignis sein. Auch unser Wollen, das ja ein physisches Ereignis erklären soll, muss folglich selbst ein physisches Ereignis sein und wäre deshalb umgekehrt ebenfalls durch physische Zustände vollständig bestimmt. Das widerspräche aber der obigen Intuition, das Wollen frei ist.¹

Unsere üblichen Intuitionen zu Willensfreiheit widersprechen damit möglicherweise naturwissenschaftlichen Annahmen eines deterministischen Verlaufs des physischen Weltgeschehens. Eine der Möglichkeiten, mit diesem Konflikt umzugehen, besteht darin, zu schliessen, dass sich freier Wille und naturwissenschaftliche Erklärungen gegenseitig ausschliessen. Diese Position wird als Inkompatibilismus bezeichnet. Inkompatibilismus kann entweder, als sogenannter Liberalismus, Determinismus ausschliessen: Das würde bedeuten, dass es Ursachen geben kann, die keine physischen Erklärungen haben, etwa in Form eines immateriellen Geistes, der willentlich etwas Physisches verursachen kann. Inkompatibilismus kann, als sogenannt harter Inkompatibilismus auch bedeuten, dass unser Wille grundsätzlich nicht frei sein kann und Erleben von Willensfreiheit deshalb eine Folge von Täuschungen sein muss. Beide dieser inkompatibilistischen Alternativen sind insofern unbefriedigend, als sie uns der Möglichkeit berauben, Dinge zu erklären, die relevant sind: Die liberalistische Alternative ist problematisch, weil sie naturwissenschaftliche Erklärungen für physische Ereignisse als grundsätzlich unvollständig darstellt, aber wie es scheint keine sinnvoll überprüfbaren Alternativen anbietet. Umgekehrt ist auch das Fazit problematisch, dass dem, was sich uns subjektiv als Wollen manifestiert, Qualitäten wie Freiheit, Gewählt-sein und Selbstbestimmtsein fehlen, weil es dann auch für eine ganze Reihe von anderen Begriffen wie Handlungen oder Verantwortung keine sinnvollen Erklärungen mehr gäbe, wir uns aber in praktischen Kontexten stark auf solche Begriffe abstützen. Die letzte Möglichkeit, den möglichen Konflikt zwischen Determinismus und Willensfreiheit aufzulösen besteht darin, die Aussage, dass sich freier Wille und Determinismus gegenseitig ausschliessen, anzugreifen. Diese Position, die als Kompatibilismus bezeichnet wird, ist aber mit der Schwierigkeit konfrontiert, zwei gewichtige, intuitiv erst einmal sehr gegensätzliche Konzepte "aneinander vorbeizubringen".

Zum Verhältnis von Willensfreiheit und Determinismus ist in der Philosophie schon viel gesagt worden (z.B. Fischer u.a. 2009). Ich werde mich in der folgen-

 $^{^1{\}rm Siehe}$ zum Beispiel Kim (2005) für einen prominenten Vertreter dieses Standpunkts.

den Arbeit darum bemühen, einen eigenständigen Beitrag zu dieser Debatte zu leisten, indem ich davon ausgehe, dass es für unsere alltäglichen Überzeugungen zu Willensfreiheit eine allgemeine und philosophisch interessante Erklärung gibt. Insbesondere interessiert es mich, wie sich unser unmittelbares und alltägliches Erleben, die Fähigkeit zu haben, gewisse Ereignisse in der Welt "selbst gewählt" und "freien Willens" verursachen zu können erklären lässt. Dieses Erleben wird in der Literatur als präreflektives Erleben von Urheberschaft (Prereflective Sense of Agency, Gallagher (2000)) bezeichnet. Darauf aufbauend will ich in einem zweiten Schritt klären, in welchem Verhältnis die Grundlagen dieses Erlebens von eigener Ursächlichkeit zu einem physischen Determinismus stehen. Ich erhoffe mir durch dieses Vorgehen fundiertere Aussagen darüber, inwiefern durch Erleben von Urheberschaft begründete Überzeugungen zu Willensfreiheit mit Determinismus kompatibel sind, wo ein physischer Determinismus solcherart begründete Überzeugungen zu Willensfreiheit einschränken, und ob und zu welchem Grad solche Einschränkungen für unsere typische Selbstwahrnehmung als willensfreie Personen problematisch sind.

Vom Aufbau her werde ich im Kapitel 2 erarbeiten, was die typischen Intuitionen zu Willensfreiheit und Determinismus beinhalten. Dazu stütze ich mich auf Literatur aus der experimentellen Philosophie, die dazu einiges Wissen bereitstellt. Im Kapitel 3 wird es darum gehen, die epistemischen Grundlagen von Selbsterleben zu klären, wobei ich mich dazu sowohl auf philosophische wie auch auf neurobiologische Thesen beziehen werde, die auch mit empirischen Befunden belegt werden. Ausgehend von dieser Grundlage eigenen Selbsterlebens untersuche ich im Kapitel 4, auf welchen Grundlagen präreflektives Erleben eigener Urheberschaft beruht und stelle eine enge Verbindung zwischen Erleben von Urheberschaft und Erleben von Wollen her. Diese Untersuchung wird uns zur Frage führen, wie sich das, was wir als eigene und willentliche Urheberschaft erleben, zu einem physischen Determinismus verhält. Das Kapitel 5 wird sich dieser Frage dann systematisch annehmen, indem es eine Reihe von Prämissen, die sich aus den vorhergehenden Kapiteln ergeben haben, aussagenlogisch verknüpft und in Bezug setzt zu bestehenden philosophischen Überlegungen zur Frage, was es ausmacht, dass wir sinnvollerweise von "Ursächlichkeit" sprechen. Ausgehend von dem Ansatz, dass sich subjektive Überzeugungen zu Willensfreiheit aus dem Erleben eigener Urheberschaft erklären, werde ich zum Abschluss dieser Arbeit kurz auf die typischen Laienaussagen zu Willensfreiheit zurückkommen, um ausgehend von den gemachten Uberlegungen mögliche Begründungen für solche Aussagen anzubieten. Und damit zur Sache!

Laienintuitionen zu Willensfreiheit

2.1 Warum Laienmeinungen?

"Willensfreiheit" ist kein einheitlicher Begriff. In der philosophischen Forschungsgemeinschaft existieren verschiedene Verständnisse von Willensfreiheit, die unterschiedliche Ansprüche an das stellen, was erfüllt sein muss, um Willensfreiheit anzunehmen (Keil 2018). Gleichzeitig scheint es so, dass wir auch ohne philosophische Vorbildung "schon wissen", was Willensfreiheit bedeutet. Dieses Kapitel wird sich darum bemühen, anhand von Literatur aus der experimentellen Philosophie diese Intuitionen zu Willensfreiheit abzuholen und zur Grundlage der weiteren Untersuchungen zu machen.² Dafür gibt es mehrere Gründe:

Erstens findet die Debatte um Willensfreiheit nicht im luftleeren Raum statt: Das Vorhandensein einer geeigneten Form von Willensfreiheit scheint Voraussetzung für lebenspraktische Fähigkeiten wie etwa zu kohärenter Selbstbestimmung und Verantwortung zu sein. Diesen Fähigkeiten geben wir Wert (Kane 2007, S. 6). Die Frage, welches Verständnis von Willensfreiheit denn wünschenswert sei (Dennett 1984), stellt sich also in einem Kontext, der normativ aufgeladen ist und der sich nicht einfach von allgemein üblichen Begründungen für moralische Urteile in unserem praktischen Alltag lösen lässt. Mit scheint es deshalb bei der gegebenen Fragestellung sinnvoll, erst einmal von einem Verständnis von Willensfreiheit auszugehen, das sich am Alltagsverständnis orientiert: Dieses Alltagsverständnis von Willensfreiheit sollte nämlich (vielleicht mehr als einige akademischere Varianten) von einer normativ motivierten Notwendigkeit geprägt sein, das intuitiv angenommene Vorhandensein unserer alltäglichen Fähigkeit zu Selbstbestimmung und Verantwortung zu begründen. Es macht darum Sinn, vorzugsweise unser Alltagsverständnis von Willensfreiheit auf seine Rechtfertigungsgründe zu hinterfragen.

Zweitens scheint es mir auch vom Verfahren her sinnvoll, mich bei der Frage, welches Verständnis von Willensfreiheit denn möglicherweise in einem interessan-

²Experimentelle Philosophie ist eine empirische Forschungsrichtung, die mit experimentellen Mitteln herauszufinden versucht, wie philosophische Fragen im Allgemeinen beurteilt werden.

ten Verhältnis zu Erleben von Urheberschaft steht, am praktischen Alltagsverständnis von Willensfreiheit zu orientieren, weil unser Erleben von Urheberschaft ebenfalls eine sehr alltägliche und praktische Erfahrung ist. Wenn eigenes Erleben von Urheberschaft ein intuitiver Grund für die Annahme von Willensfreiheit ist (Nichols 2011), und wenn dieses Erleben allgemein gültige Strukturen hat, dann sollte das auch zu relativ einheitlichen und unkontroversen Intuitionen zu Willensfreiheit führen. Diese Intuitionen lassen sich dann mit den Mitteln der experimentellen Philosophie als subjektive Urteile untersuchen (Sytsma und Livengood 2015). Ich erwarte mir also für die folgende Untersuchung von Willensfreiheit, dass das, was wir zu Selbsterleben und Erleben von Urheberschaft feststellen werden, in der Lage ist, Laienurteile zu Willensfreiheit, wie sie durch die experimentelle Philosophie erfasst werden, zu begründen.

Die experimentelle Philosophie und die Sozialpsychologie haben in den letzten zwanzig Jahren eine solide Basis an Befunden zu Laienurteilen über Willensfreiheit erarbeitet. Wichtig und kongruent mit der hier gegebenen Fragestellung ist, dass es den Autor*innen dieser Literatur bei ihrer systematischen Untersuchung von Laienüberzeugungen zu Willensfreiheit nicht um ein schwer fassbares und abstraktes Konzept ging, sondern um etwas, was einen recht konkreten und alltäglichen Bezug zu bewussten Entscheidungsmöglichkeiten und Handlungsoptionen hat (Monroe, Dillon und Malle 2014; Feldman, Baumeister und Wong 2014), wie sie auch das Erleben von Urheberschaft prägen. Gleichzeitig waren die Autor*innen dieser Studien häufig umsichtig genug, auch die Intuitionen zu anderen in dieser Fragestellung relevanten Begriffen wie Determinismus, Dualismus und Reduktionismus zu erfassen. Ich will diese Befunde im Folgenden kurz zusammenfassen:

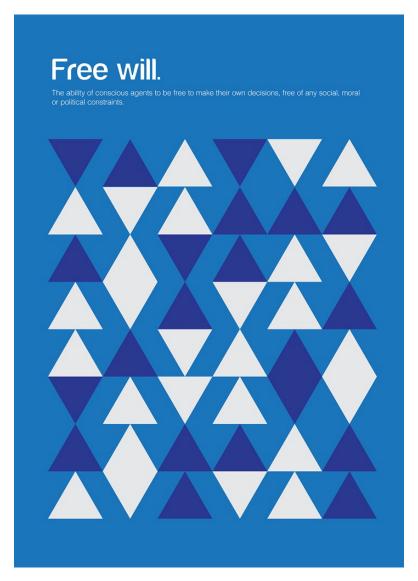


Abb. 2.1: Free Will, von Genis Carreras, (Carreras 2013)

2.2 Empirische Datenbasis

In einer Studie mit insgesamt über 1000 Versuchspersonen haben Nadelhoffer u.a. (2014) versucht, einen Fragebogen ("Free will inventory") zu entwickeln und zu validieren, der den Grad an individuellem Glauben an die drei Konstrukte "Willensfreiheit", "Determinismus" und "Dualismus/Antireduktionismus" quantifiziert. Über eine Reihe von Unterstudien und einer abschliessenden konfirmatorischen Faktorenanalyse konnten für jedes dieser Konstrukte ein Satz von Aussagen iden-

tifiziert werden, der dieses Konstrukt unabhängig von den jeweils anderen Konstrukten erfasst (Nadelhoffer u.a. 2014, S. 34-35). Mit diesem Instrument wurde in einer weiteren, gross angelegten und kulturübergreifenden Studie mit insgesamt 1800 Versuchspersonen aus den USA und Singapur quantifiziert, wie stark die Versuchspersonen diesen jeweiligen Aussagen zustimmten, und was sich daraus in Bezug zu den überwiegenden Laienmeinungen zu diesen drei Konstrukten ergab (Wisniewski, Deutschländer und Haynes 2019). ich werde im folgenden Unterkapitel für das Konstrukt der Willensfreiheit die eingedeutschten Aussagen aus dem Free will inventory mit den aggregierten Zustimmungsraten aus der Arbeit von Wisniewski, Deutschländer und Haynes (2019) darstellen.

2.3 Laienmeinungen zur Willensfreiheit

Für den Faktor "Willensfreiheit" fanden Nadelhoffer u. a. (2014) die folgenden fünf Aussagen, die das, was "Willensfreiheit" in der Laienmeinung bedeutet, gut erfassen:

- 1. Personen können jederzeit anders handeln. ("People always have the ability to do otherwise.")
- 2. Personen haben immer Willensfreiheit. ("People always have free will.")
- 3. Das Leben einer Person ist völlig selbstbestimmt. ("How people's lives unfold is completely up to them.")
- 4. Letztlich haben Personen die vollständige Kontrolle über ihre Entscheidungen und Handlungen. ("People ultimately have complete control over their decisions and their actions.")
- 5. Personen haben freien Willen, selbst wenn ihre Wahlmöglichkeiten durch äussere Umstände vollständig eingeschränkt sind. ("People have free will even when their choices are completely limited by external circumstances.")

Empirisch fand sich, dass kulturübergreifend jeweils über 80% der Versuchspersonen diesen Aussagen zu Willensfreiheit überwiegend zustimmten (Wisniewski, Deutschländer und Haynes 2019, Abb. 2.2). Gut übereinstimmend damit fand eine Studie von Nahmias, Shepard und Reuter (2014), dass über 80% der Versuchspersonen der Meinung war, dass eine Person frei entscheiden könne, auch wenn die getroffene Entscheidung wissenschaftlich vorhersehbar wäre.

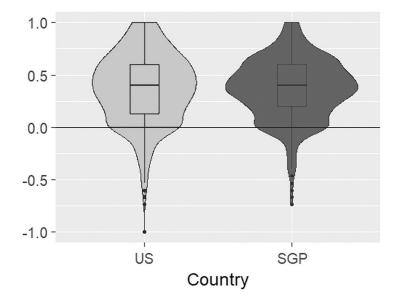


Abb. 2.2: Laienüberzeugungen zu Willensfreiheit. Die Violin-und Box-plots zeigen die Verteilung von Glauben an Willensfreiheit in den USA (US, links/hellgrau) und Singapur (SGP, rechts/dunkelgrau). Werte über Null bedeuten Glaube an Willensfreiheit, Werte unter Null bedeuten, dass ein solcher Glaube abgelehnt wird, Werte von Null bedeuten Indifferenz. Aus Wisniewski, Deutschländer und Haynes (2019).

2.4 Prämissen aus der experimentellen Philosophie

Nachdem wir nun besser wissen, welche Aussagen Laien mit dem Begriff der Willensfreiheit verbinden und dass sie mehrheitlich davon überzeugt sind, dass eine solchermassen verstandene Willensfreiheit besteht, können wir dieses Wissen für die These dieser Arbeit verwenden. Diese These ist, dass Erleben von Urheberschaft diese Laienmeinungen begründen kann. Ich werde darum aus den von Nadelhoffer u.a. (2014) identifizierten fünf Aussagen zum Laienverständnis eine Reihe von fünf Prämissen formulieren, die diesen Zusammenhang beinhalten und die sich am Ende dieser Arbeit überprüfen lassen sollten. Diese Prämissen sehen folgendermassen aus:

Prämisse 2.4.1: Präreflektives Erleben von Urheberschaft beinhaltet Wahlfreiheit

Die Form des subjektiven Zugangs zu eigener Urheberschaft kann erklären, dass das Subjekt überzeugt ist, die eigenen Handlungen wählen zu können.

Prämisse 2.4.2: Präreflektives Erleben von Urheberschaft beinhaltet Willensfreiheit

Die Form des subjektiven Zugangs zu eigener Urheberschaft kann erklären, dass das Subjekt überzeugt ist, dass der eigene Wille frei sei.

Prämisse 2.4.3: Präreflektives Erleben von Urheberschaft beinhaltet Selbstbestimmung

Die Form des subjektiven Zugangs zu eigener Urheberschaft kann erklären, dass das Subjekt überzeugt ist, über sich selbst bestimmen zu können.

Prämisse 2.4.4: Präreflektives Erleben von Urheberschaft beinhaltet Entscheidungs- und Handlungskontrolle

Die Form des subjektiven Zugangs zu eigener Urheberschaft kann erklären, dass das Subjekt überzeugt ist, die eigenen Entscheidungen und Handlungen kontrollieren zu können.

Prämisse 2.4.5: Präreflektives Erleben von Urheberschaft beinhaltet äusserliche Unanstastbarkeit

Die Form des subjektiven Zugangs zu eigener Urheberschaft kann erklären, dass das Subjekt überzeugt ist, selbst dann einen freien Willen zu haben, wenn seine Wahlmöglichkeiten durch äussere Umstände vollständig eingeschränkt sind.

Damit wissen wir für meine Fragestellung erst einmal genug darüber, was es zu erklären gilt, wenn Laienmeinungen zu Willensfreiheit erklärt werden sollen. In den folgenden beiden Kapiteln 3 und 4 wird es darum gehen, eine Erklärung für diese Laienmeinungen durch "Erleben eigener Urheberschaft" zu liefern, die ich dann im Kapitel 5 auf ihre Schlüssigkeit prüfen werde.

KAPITEL 3

Epistemik von Selbsterleben

Die zentrale These dieser Arbeit lautet, dass Überzeugungen zu Willensfreiheit, wie sie sich in Laienmeinungen manifestieren, im subjektiven Erleben von Urheberschaft eine Erklärung finden. Erleben von Urheberschaft ist sinngemäss jede Erfahrung, die besagt, ein Ereignis "selbst" zu verursachen oder verursacht zu haben. Zu dieser Erfahrung gehört notwendigerweise ein Subjekt, dem die Erfahrung bewusst ist, etwas "selbst" verursacht zu haben (Searle 2000, S. 561)³. Wenn es also zu Erleben von Urheberschaft kommt, dann gibt es eine Klasse von Ereignissen (typischerweise, wie wir später sehen werden, willentliche Akte), die sich dadurch auszeichnet, dass es ein Subjekt gibt, das die Aussage, "selbst" dieses Ereignis verursacht zu haben, als wahr empfindet. Für das Subjekt begründet sich der Wahrheitsanspruch dieser Aussage erst einmal damit, dass es Urheberschaft für dieses Ereignis phänomenal erlebt und dass das Ereignis deshalb für das Subjekt Teil eines ihm eigenen willentlichen Akts ist.

Sprachlich äussert die Person dies in einer bestimmten Form, indem sie das Pronomen "ich" verwendet und mit einem Verb verbindet, das das Ereignis bezeichnet: "Ich schreibe". Die Begründung einer solchen Aussage durch das subjektive Erleben fehlt hier typischerweise, wohl weil sie so selbstverständlich ist. Wir dürfen das aber ergänzen: Mit einer Begründung durch das Erleben von Urheberschaft würde der obige Satz dann vielleicht so aussehen: "Ich weiss, dass ich schreibe, weil ich ein Erleben habe, das ich als selbst schreibend empfinde." Bevor ich mich jedoch mit der Handlungskomponente weiter beschäftigen werde, will ich die Art und Weise, wie sich dieses "Selbst" epistemisch verhält, weiter klären. Dann werde ich darauf zurückkommen, wie sich die Epistemik von Ereignissen darstellt, von denen jemand behauptet, selbst Urheber*in gewesen zu sein.

 $^{^3}$ Ich denke nicht, dass nicht bewusstseinsfähige Ereignisse für das gegebene Projekt von Interesse sind.

3.1 Das Phänomen der Unmittelbarkeit von Selbsterleben

Die Besonderheiten der Epistemik des eigenen Selbst sind in der Philosophie seit langem bekannt. Ludwig Wittgenstein hat zum Beispiel in seinem "Blue Book" festgestellt, dass es in Sätzen, die ein subjektives Erleben bezeichnen, für das Subjekt unmissverständlich ist, was mit dem Pronomen "ich" gemeint ist (Wittgenstein 1958, S. 67). Wittgenstein illustriert das am Beispiel von Zahnschmerzen. Wenn jemand behauptet "Ich habe Zahnschmerzen", kann man, wie Wittgenstein feststellt, sinnvollerweise fragen: "Bist du sicher, dass es Zahnschmerzen sind?", weil sich die Person darin täuschen kann, dass es Zahnschmerzen und nicht andere Schmerzen sind, die sie verspürt. Es ist aber für Wittgenstein aus logischen, und nicht aus empirischen Gründen sinnlos zu fragen: "Bist du sicher, dass es du bist, die die Zahnschmerzen verspürt?", weil es unmöglich ist, dass sich die angesprochene Person in dem täuscht, was sie mit "ich" bezeichnet. Es kann deshalb a-priori niemand anders sein als sie selbst, die diese Schmerzen erlebt. Deshalb ist die zweite Frage absurd: Sie impliziert die Möglichkeit, gegen dieses A-priori verstossen zu können. "Ich" muss damit etwas bezeichnen, das der Person in einer unmittelbaren Form gegeben ist und das ihr ohne weitere deduktive Schritte gewiss ist.

3.2 Erklärungen für die Unmittelbarkeit von Selbsterleben

Es lohnt sich ist an dieser Stelle, der Frage nachzugehen, warum dieses "Ich" der sich so äussernden Person nicht anders als unmittelbar gegeben sein kann. Ich werde im Folgenden in einem ersten Teil kurz zwei philosophisch motivierte und theoretische Gründe angeben, warum wir auf grosse Probleme stossen, wenn wir von einem herleitenden, also mittelbaren Zugang zum "Ich" ausgehen, und dass darum ein unmittelbarer Zugang die plausiblere Option ist. Im zweiten, ausführlicheren Teil, werde ich zu zeigen versuchen, dass sich das gleiche Fazit auch ergibt, wenn wir uns auf naturalistische Thesen von (Selbst-) Bewusstsein stützen. Ich werde dafür einerseits auf zwei etablierte Theorien zurückgreifen, nämlich erstens auf die sogenannte Informations-Integration-Theorie (IIT) von Bewusstsein von Guilo Tononi (Tononi 2004) aus den Neurowissenschaften, und zweitens auf Thomas Metzingers Selbst-Modell-Theorie der Subjektivität (Metzinger 2003; Metzinger 2007). Diese Theorien sind für die gegebene Fragestellung interessant, weil sie Aussagen an den Schnittstellen zwischen biologischen Systemen, Funktionen und Erlebensweisen machen. Zusätzlich werde ich auch kurz auf einige klassische Befunde aus den Verhaltens- und Neurowissenschaften eingehen, die die eher abstrakt hergeleiteten Argumente empirisch hintermauern. Es wird sich zeigen, dass sich aus diesen mit naturwissenschaftlichen Ansichten kompatiblen und empirisch gestützten Theorien eine Reihe von Eigenschaften von Selbsterlebens herleiten lassen, die mit unseren vortheoretischen Intuitionen im Einklang sind und die für die weitere Argumentation von Bedeutung sein werden.

Ich möchte jedoch explizit festhalten, dass es mir unangebracht scheint (und es auch nicht stattfinden wird), aus den hier dargestellten Theorien auf eine vollständige und abschliessende Erklärung von (Selbst-)Bewusstsein zurückgreifen zu wollen. Um Antworten auf die Frage zu finden, ob und in welcher Form sich aus der Untersuchung des Erlebens von Urheberschaft relevante Aussagen zur Debatte um Willensfreiheit machen lassen, soll und wird es hier reichen, einen Satz von Eigenschaften von Selbsterleben zu identifizieren, der einerseits für das Erleben von Urheberschaft notwendig ist, und der gleichzeitig hinreichende Bedingungen für eine der möglichen Verstehensweisen von Willensfreiheit darstellt. Was einen solchen Satz von Eigenschaften ausmachen wird, ist erstens, dass diese Eigenschaften einen Bezug zu materiell realisierbaren raum-zeitlichen Folgebeziehungen eines "Selbst" haben, zweitens, dass sie besondere relationale Eigenschaften zu diesem "Selbst" haben, und drittens, dass sie zumindest Teilaspekte der phänomenalen Eigenschaften von Selbsterleben erklären können.

3.2.1 Philosophische Gründe für Unmittelbarkeit von Selbsterleben

Philosophisch betrachtet lässt sich eine Antwort darauf finden, warum es ein unmittelbares Selbsterleben gibt, indem wir auf Wittgensteins (Wittgenstein 1958, S.107) Feststellung zurückkommen, dass beim "Ich-Selbst" Bezeichnendes ("Ich als Subjekt") und Bezeichnetes ("Ich als Objekt") identisch sind. Wenn wir davon ausgehen wollten, dass die Gegebenheit von "ich" nicht unmittelbar, sondern erst durch deduktiv herleitendes Verständnis zustande käme, ergäben sich eine Reihe von Widersprüchen und Problemen, die entstehen, weil es bei deduktiven Begründungen eine Instanz braucht, die die Wahrheit oder Falschheit einer Aussage feststellen kann. Im gegebenen Fall ist das aber aus den folgenden Gründen nicht möglich:

• In einer deduktiven Begründung müssen der zu beurteilende Sachverhalt (ich habe Zahnschmerzen) und die Instanz, die begründen soll, dass ich es bin, der Zahnschmerzen hat, notwendigerweise voneinander unabhängig sein. Es muss bei einer deduktiven Begründung nämlich grundsätzlich möglich sein, dass die urteilende Instanz zum Schluss kommt, dass der zu beurteilende Sachverhalt nicht gegeben ist. Wenn urteilende Instanz und zu beurteilender Sachverhalt gleichzeitig identisch sein müssen, ist diese Unabhängigkeit nicht mehr gegeben. So ist es zum Beispiel gar nicht möglich, dass die urteilende

Instanz feststellen kann, dass der zu beurteilende Sachverhalt nicht wahr ist, weil das Urteilen selber vom Wahrsein dieses Sachverhalts abhängig ist. Eine deduktive Begründung, dass ich es bin, der Zahnschmerzen erlebt, ist deshalb als logischer Akt auf einer Metaebene gar nicht möglich.

• Wenn es neben dem Subjekt eine unabhängig urteilende Instanz gäbe, die feststellen müsste, dass ich es bin, der Zahnschmerzen hat, wäre dieser urteilende Akt ein Akt dieser Instanz. Die Instanz müsste also zusätzlich Aussagen machen können wie: "Ich stelle fest, dass ich es bin, der Zahnschmerzen hat." Damit kommt es zu einem infiniten Regress, weil sich daraus die Frage ergeben würde: "Bist sicher, dass es du bist, der feststellt, dass es du bist, der die Zahnschmerzen verspürt?" Weil sich diese Gegenfrage beliebig fortsetzen lässt, wird es unmöglich, überhaupt je zu einem abschliessbaren Urteil zu kommen.⁴

Um solche Arten von Problemen zu vermeiden, muss folglich das, was "Ich" bezeichnet, in einer unmittelbaren Form immer bereits gegeben sein. In dem Satz "Ich weiss, dass ich schreibe, weil ich ein Erleben habe, in dem ich mich als "schreibend" empfinde", ist es kein deduktiver Akt, sondern eine a-priori gegebene Tatsache, dass das Objekt "Ich", das als Träger der Eigenschaft, "schreibend zu sein" bezeichnet wird, identisch ist mit dem Subjekt "Ich", das aktiv diesen Träger als "schreibend" bezeichnet. "Ich-Selbst" bildet damit eine Klammer dem Subjekt "Ich", das sich durch bewusstseinsfähige und intentionale mentale Zustände wie "Erleben von ...", "Überzeugt sein dass..." auszeichnet, und dem Objekt "Ich", das die für materielle Gegenstände üblichen physischen Eigenschaften wie einen Ort, eine Ausdehnung, oder ein Gewicht hat. Das ist insofern interessant, als damit eine Form von Identität zwischen bewusstseinsfähigen und intentionalen Zuständen auf der einen Seite und physischen Entitäten auf der anderen Seite zu bestehen scheint. Diese Feststellung wird im weiteren Text noch eine wichtige Rolle spielen und uns zum Begriff des "Selbst-Modells" führen. Gleichzeitig ergibt sich aus der unmittelbaren Gegebenheit dessen, was als "Ich" bezeichnet wird, eine klare Abgrenzung zu allem "Nicht-Ich": "Nicht-Ich" ist alles, was nicht in dieser unmittelbaren Form gegeben ist.

⁴Die Problematik von Selbstprädikationen stellt sich nicht nur in diesem Kontext. Platon hat sich zum Beispiel im "Parmenides" im Rahmen der Frage nach der Ontologie seiner Ideenlehre mit dem Problem auseinandergesetzt, ob die Idee einer Eigenschaft selbst diese Eigenschaft hat (Zum Beispiel: Ist die Idee des Schönen selbst schön?) (Platon 1987). Für Platons Parmenides führt diese Frage in unlösbare Probleme und ist deshalb nicht zu beantworten. Eines dieser Probleme ist, dass es auch hier zu einem infiniten Regress kommt: Wenn die Idee I die Eigenschaft F hat, durch deren Teilhabe die Dinge $\{D_1; D_2; ...; D_n\}$ die Eigenschaft F haben, gibt es eine Menge $\{I; D_1; D_2; ...; D_n\}$, die sich durch das Vorhandensein von F' in allen ihren Elementen definiert. Dieses Vorhandensein von F' ist aber nach der Ideenlehre genau dadurch gegeben, dass es eine Idee I' von I' gibt, die wiederum Teil der Menge I'; I

3.2.2 Informationstheoretische Gründe für die Unmittelbarkeit von Bewusstsein

Interessanterweise lässt sich das gleiche Fazit auch aus Theorien des Bewusstseins herleiten, die eher aus den Neurowissenschaften inspiriert sind: Eine der gegenwärtig einflussreichsten Thesen zu den funktionalen Korrelaten von Bewusstsein sagt, dass Bewusstsein mit der Fähigkeit eines Systems korrespondiert, Information zu integrieren (Information Integration Theory, IIT, Tononi (2004)). Im Gehirn wird, so die These, diese Integration ermöglicht, indem grosse Teile des neuronalen Gewebes, und insbesondere der Neokortex, eine extrem ausgeprägte Netzwerkstruktur haben, sodass jedes Element dieses Netzwerks mit sehr vielen anderen Elementen dieses Netzwerks in Abhängigkeitsbeziehungen steht. Der Zustand eines Elements, und damit die Information, die es trägt, ist deshalb abhängig von den Zuständen sehr vieler anderer Elemente, und folglich auch von den Informationen, die diese anderen Elemente tragen. Information wird damit über all die Elemente integriert, wobei, so die These, die Menge an integrierter Information mit der "Menge" an Bewusstsein zusammenhängt.

Es ist nicht ganz unerwartet umstritten, ob diese These bereits eine hinreichende Erklärung von Bewusstsein liefert (Chalmers 2007). Das ist aber im Kontext der gegebenen Fragestellung nicht relevant. Für den gegebenen Zweck genügt es, einige wie mir scheint unproblematische Dinge festzuhalten: Erstens macht Tononis These die Annahme plausibel, dass die Integration von Information eine notwendige Bedingung dafür ist, ein relevantes Mass an Bewusstsein entstehen zu lassen. Zweitens ist die These einerseits theoretisch kompatibel mit naturwissenschaftlichen Erklärungen, hat aber umgekehrt auch eine gewisse Erklärungskraft für eine Reihe von empirischen Beobachtungen: Empirischen Support für eine solche These findet sich zum Beispiel, wenn man die Neurophysiologie von Tiefschlaf betrachtet, wo die neuronalen Interaktionen im Kortex deutlich eingeschränkt sind, oder bei epileptischen Anfällen, bei dem Teile des Gehirns in sehr einfache, und damit informationsarme Interaktionsmuster geraten. In beiden Fällen kommt es zu grossen Einschränkungen oder dem totalen Verlust von Bewusstsein bei der betroffenen Person.

Wenn aber eine derart integrierte Information nun mit dem korrespondiert, was unser bewusstes Erleben ist, dann lässt sich eine Reihe von Eigenschaften dieses bewussten Erlebens erklären, die für die gegebene Fragestellung relevant sind. Dabei geht es erstens um Eigenschaften des Zugangs zu diesem Erleben, die ich in den folgenden Abschnitten zu psychophysiologischer Intransparenz, Meinheit, und Privatheit darstellen will. Weiter werden wir uns im Abschnitt 3.3 mit der Frage beschäftigen, wie sich dieses Erleben subjektiv manifestiert, um schliesslich im Abschnitt 3.4 damit zu beschäftigen, wie "Ich selbst" Teil dieses Erlebens bin.

3.2.2.1 Informationsintegration und psychophysiologische Intransparenz

Was meint psychophysiologische Intransparenz? Die Annahme, dass die Information, deren wir bewusst sind, notwendigerweise das Produkt vieler als gegenseitige Abhängigkeitsverhältnisse beschreibbarer Integrationsvorgänge ist, erklärt paradoxerweise, dass unmöglich werden kann, die konkreten Ursachen dieser Information zurückverfolgen zu können. Das ist so, weil bei solchen Integrationsvorgängen Information zusammengefasst und transformiert wird, und dieser Vorgang typischerweise irreversibel ist. Diese Irreversibilität ist eine einfache Konsequenz davon, dass es bei Integrationsvorgängen typischerweise mehr als eine Möglichkeit gibt, auf der Ebene, die die integrierte Information enthält, einen bestimmten Zustand zu verursachen.⁵ Eine direkte Folge davon ist, dass anhand dessen, was ein integrierter Zustand an Information enthält, nicht mehr eindeutig auf das zurückgeschlossen werden, was diesen integrierten Zustand verursacht hat: Es besteht ein inverses Problem.⁶

In der Biologie scheint es plausibel, dass die Eigenschaft eines Organismus, Information in einer Art und Weise zu integrieren, die förderlich für die eigene Fitness ist, einen Selektionsvorteil darstellt und sich damit in der Evolution systematisch entwickelt hat. Diese Eigenschaft wird dann zum Vorteil, wenn unterschiedliche Informationen, die für den Organismus jeweils in gleicher Weise relevant sind, zu einem möglichst gleichen Zustand integriert wird und damit zu Repräsentationen führen, die für den Organismus eine einheitliche und relevante Bedeutung enthalten und zu einem entsprechenden Verhalten führt. Wenn da ein Tiger ist, ist Flucht ein angepasstes Verhalten, wobei es keine Rolle spielt, auf welchen Zellen der Retina sich welche Teile dessen abbilden, das ich als "Tiger" erkennen muss. Wichtig ist, dass "Tiger" allgemein zuverlässig erkannt werden, unabhängig davon, wie sich ein bestimmter Tiger zu einem bestimmten Zeitpunkt sensorisch manifestiert.

⁵Dem Argument der multiplen Realisierbarkeit begegnen wir hier nicht zufällig, es ist ein klassischer funktionalistischer Einwand gegen Type-Type Identitätstheorien von mentalen und physischen Zuständen (Beckermann 2008, S. 137). Als stark funktionalistisch geprägte Theorie ist die multiple Realisierbarkeit eine der inhärenten Eigenschaften von Tononis Informationsintegrationstheorie. Nicht überraschend wird multiple Realisierbarkeit auch in dieser Arbeit eine wichtige Rolle spielen.

⁶Ein einfaches Beispiel für ein durch Integration verursachtes inverses Problem: Der von einer Person zurückgelegte Weg ist ein Integral aus den jeweiligen Zeiten, zu denen sich die Person mit bestimmten Geschwindigkeiten bewegt hat. Damit ist die zurückgelegte Wegstrecke zwar vollständig dadurch erklärt, dass sich die Person zu jeweils bestimmten Zeiten mit einer jeweils bestimmten Geschwindigkeit bewegt hat. Es ist aber unmöglich, anhand des Wissens der zurückgelegten Wegstrecke zu wissen, wann sich die Person mit welcher Geschwindigkeit bewegt hat, weil dieselbe Wegstrecke durch eine beliebige Anzahl unterschiedlicher Geschwindigkeitsverläufe der Person erklärbar ist.

Wenn aber die Fähigkeit eines Systems, Information zu integrieren, gemäss Tononis These mit der Bewusstseinsfähigkeit des Systems korrespondiert, ergibt sich das Fazit, dass sich Zugang zu materiellen Erklärungen solcher Repräsentationen und das Haben von Bewusstsein solcher Repräsentationen gegenseitig ausschliessen: Je integrierter ein Zustand, desto mehr ist er einerseits bewusstseinsfähig, desto weniger lässt er sich aber anhand dessen, was er bewusst repräsentiert, erklären. Für die weitere Argumentation will ich darum folgendes Fazit festhalten:

Fazit 3.2.1: Psychophysiologische Intransparenz

Eigene Bewusstseinszustände sind dem Subjekt nicht als Folge physischer Ereignisse erklärbar. 7

3.2.2.2 Informationsintegration und Meinheit

Aus dieser Integrationsleistung ergibt sich als weitere relevante Eigenschaft von bewusstem Erleben etwas, das ich mit dem Begriff der "Meinheit" bezeichnen werde: Subjektives Erleben resultiert, wie oben hergeleitet, immer aus einem nicht separierbaren Amalgam von aussenweltlichen Faktoren und dem wahrnehmenden Subjekt zugehörigen kognitiven Integrationsleistungen. In dieser Bindung an das Subjekt unterscheiden sich subjektive Wahrnehmungszustände wesentlich von anderen, nicht bewussten oder objektivierbaren Arten von Information über bestimmte Zustände in der Welt. Sie sind als Produkt einer dem Subjekt eigenen Integrationsleistung an das Subjekt gebunden: Es gibt niemand als das Subjekt selbst, dem ein gegebenes Erleben als Folge eigener kognitiver Leistungen in einer bestimmten bewussten Form gegeben ist. Deshalb werden sie subjektiv als "eigene" Wahrnehmungen erlebt und attribuiert. Ich halte fest:

⁷Psychophysiologische Intransparenz kann man damit als Gegenpol zu dem sehen, was sich für Physikalist*innen als Qualiaproblem darstellt: So explanatorisch unklar es für Physikalist*innen ist, wie physische Zustände bestimmte mentale Qualitäten realisieren können, so explanatorisch unklar ist es einem Subjekt, dass und wie die eigenen mentalen Zustände physisch realisiert werden.

Fazit 3.2.2: Meinheit

Subjektive Wahrnehmung korrespondiert mit der Integration von dem Subjekt eigenen kognitiven Leistungen mit aussenweltlichen Gegebenheiten und wird deshalb als **eigen** erlebt.

3.2.2.3 Informations integration und Privatheit

Damit komme ich zum Begriff der "Privatheit": Die Hypothese, dass die Information, deren wir bewusst sind, das Produkt vieler Integrationsvorgänge ist, erklärt, warum das, was wir erleben, nicht bloss durch Ausseneinflüsse verursacht wird, sondern das Produkt der Interaktion dieser Ausseneinflüsse und den Leistungen eines Informationsintegrationssystems ist. Das Vorhandensein von bestimmten Erfahrungen bedeutet damit nicht nur das Vorhandensein bestimmter Ausseneinflüsse, sondern ebenso das Vorhandensein dieser Integrationsleistungen, die Teil dessen sind, was die Grundlage für die jeweils bewusst werdende Erfahrung bildet. So erklärt sich, warum subjektive Erfahrungen privat sind: Auch wenn es möglich wäre, für zwei Subjekte die gleichen Ausseneinflüsse herzustellen, so wäre das subjektive Erleben für jedes der beiden Subjekte jeweils das Produkt aus dem (gemeinsamen) Stimulus und der jeweils individuellen Integration der im Stimulus enthaltenen Information. Damit ergeben sich zwei Instanzen von subjektiven Erleben. Gleichzeitig lässt sich von einer Instanz des subjektiven Erlebens nicht schlüssig auf die andere Instanz eines solchen Erlebens schliessen, weil es, wie oben (Abschnitt 3.2.2.1) dargestellt, keinen eindeutigen Weg vom bewussten Erleben zurück zu den Sinnesdaten gibt. Genau diese Sinnesdaten sind aber wiederum Teil dessen, was es für die einer weiteren Instanz von subjektivem Erleben des Stimulus braucht. Jedes Subjekt hat damit immer nur Zugang zum eigenen Erleben. Damit ergibt sich eine weitere, für die spätere Argumentation hilfreiche Aussage:

Fazit 3.2.3: Privatheit

Subjektives Erleben ist aus informationstheoretischen Gründen notwendigerweise privat.

3.3 Subjektives Erleben, Repräsentation im Gehirn und Leiblichkeit

Nachdem wir uns im obigen Abschnitt mit der Frage des Zugangs zu mit Bewusstsein korrespondierenden Informationsintegrationsprozessen beschäftigt haben, wende mich nun der Frage zu, wie das Tragen und Integrieren von Information, das immer auch ein physischer Vorgang ist, mit dem Inhalt des subjektiven Erlebens des Subjekts zusammenhängt.

Psycho-physiologisch betrachtet scheint es, dass es beim Mensch für das Einsetzen von Erleben genügt, wenn es in bestimmten Hirngebieten zu genügend starken und in einer bestimmten Form strukturierten Erregungsmustern kommt. Dabei korrespondiert die Anatomie dieser Gebiete mit den verschiedenen Modalitäten dieses Erlebens (Tononi 2004, S. 6): Es sind konsistent unterschiedliche Gebiete, deren Erregung subjektiv etwa zu Erleben von Sehen, Hören, Berührung oder auch Emotionen wie Angst führt. Weiter sind Gebiete, die einer bestimmten Modalität entsprechen, räumlich weiter organisiert entsprechend der basalen repräsentationalen Struktur der Information, die die entsprechende Modalität enthält. So ist etwa der sogenannte somatosensorische Kortex, der für die Körperwahrnehmung notwendig ist, räumlich in einer Weise organisiert, die die Form des Körpers abbildet.

Im Falle der Wahrnehmung sind die entsprechenden kortikale Gebiete normalerweise über eine Reihe von Schaltstellen und Nervenbahnen mit an den entsprechenden Stellen (Augen, Ohren, Haut, etc.) sitzenden Sensoren verbunden und durch die genügende Stimulierung dieser Sensoren erregbar. Subjektiv wird die Erregung dieser kortikalen Gebiete in einer Qualität erfahren, die wir den mit dem jeweiligen Gebiet verbundenen Sensoren zuordnen: Die Erregung von primär mit dem Auge verbundenen Gebieten wird als "Sehen" erlebt, gleiches gilt für das "Hören" von neuronalen Erregungen in Gebieten, die mit dem Ohr verbunden sind, und für alle anderen Sinnesmodalitäten.⁸ Für die vorliegende Frage scheinen mir zwei Punkte relevant: Erstens sieht es empirisch so, dass solche zentralnervösen Erregungen in sensorischen Arealen notwendige und hinreichende Bedingungen für das Vorhandensein von entsprechendem subjektiven Wahrnehmungen sind. Dieses Wahrnehmen wird durch das Subjekt aber immer, und a-priori erklärt als ein dem Gehirn externes Ereignis: Es scheint also so, dass psychophysiologische Intransparenz durch repräsentationale Vorgänge überbrückt wird, die das, was das Subjekt wahrnimmt, immer als etwas anderes als Aktivität des Gehirn darstellen und somit externalisieren. Damit "übersetzen" solche psycho-physischen Brücken-

⁸Die Zuordnung von Erlebnisqualität und Sinnesmodalität muss nicht notwendigerweise 1:1 sein, wie zum Beispiel im Fall von Synästhesien. Das ist aber nicht im Widerspruch zu den Thesen, die ich hier verfolgen will.

prozesse, wie ungeklärt oder sogar unklärbar (Levine 1999) ihre Wirkweise auch sein mag, zwischen mentalen und physischen Zuständen und Ereignissen, wobei sich physischen Ereignissen korrespondierende mentale Zustände dem Subjekt in bestimmten Qualitäten (als Farbe, Klang, Schmerz..) manifestieren.

Als klassische Evidenz für solche Externalisierungen sollen hier die Beobachtungen von Wilder Penfield dienen. Penfield war ein Neurochirurg, der als erster während unter Lokalanästhesie durchgeführten Eingriffen am Gehirn von wachen Patient*innen mit schweren Epilepsien das Gehirn lokal stimuliert hat (Penfield und Boldrey 1937). In dieser Situation können die Patient*innen direkt berichten, was sie subjektiv erleben, während das Gehirn stimuliert wird. Tatsächlich erleben Patient*innen diese Stimulation nie als Stimulation am Gehirn, sondern immer als eine der sensorischen Zuordnung des stimulierten Gebiets entsprechenden Stimulation des Körpers. Eine der Patientinnen Penfields berichtete zum Beispiel von einem "Kribbeln im Finger", wenn eine bestimmte Stelle des somatosensorischen Kortex stimuliert wurde (Penfield und Boldrey 1937, S. 400).

Bestimmte zentralnervöse Ereignisse werden also subjektiv unwillkürlich als etwas erlebt, das nicht im Gehirn stattfindet, sondern durch den eigenen Körper erlebt wird. Ich werde für diese repräsentationalen Brückenprozesse den Begriff des Selbst-Modells verwenden, wie ihn Thomas Metzinger (Metzinger 2003) eingeführt hat, weil diese Brückenprozesse dem Subjekt nicht nur die Aussenwelt modellhaft als "selbst" Erlebtes repräsentieren, sondern auch, weil sie, wie wir im nächsten Abschnitt 3.4 sehen werden, selbst-repräsentational sind. Diese a-priori Projektionen der Ursachen von Erleben in die Peripherie, also das, was Metzinger als Selbst-Modell bezeichnet, überbrückt so die allgegenwärtigen, und kaum überwindbaren epistemischen Lücken, mit denen unser Bewusstsein aufgrund psychophysiologischer Intransparenzen sonst konfrontiert wäre. Damit sind wir beim zweiten Punkt, den ich hier festhalten will: Das Selbst-Modell, als eine unter üblichen Umständen funktionell sehr adäquate Form von Vorwissen, ersetzt die bewusst nicht zugänglichen Zusammenhänge innerhalb dessen, was die materiellen Grundlagen unseres subjektiven Erlebens zu bilden scheint. Ein Selbst-Modell, als transparenter und repräsentationaler Teil psycho-physischer Vorgänge, liefert das praktisch notwendige, aber anderweitig nicht verfügbare Bindeglied zwischen subjektivem Erleben und dem, was als "unser" Körper materiell ursächlich für dieses Erleben zu sein scheint.

Ich will in der Tradition von Maurice Merleau-Ponty (Merleau-Ponty 1945) und Gernot Böhme (Böhme 2008) für diese Form von einem über ein Selbst-Modell an unseren materiellen Körper gebundenen Erleben den Begriff der "Leiblichkeit" einführen. Der "Leib" ist damit der Inhalt dessen, was mental durch die Integration von sensorischer Information und unserem Selbst-Modell repräsentiert wird, während ich im Gegensatz dazu alles, was zum materiell definierten Organismus

gehört als "Körper" bezeichnen werde. Leib und Körper stehen natürlicherweise in einem engen Bezug zueinander, sind aber erfahrungsgemäss nicht identisch. Manifest wird diese Nicht-Identität in Form von empirisch beobachtbaren Dissoziationen zwischen Leib und Körper, die ich im Folgenden anhand mehrerer Beispiele illustrieren will.

So gibt es zum Beispiel in der neurologischen Praxis einerseits Fälle von Patient*innen, die einen Körperteil verloren haben, bei denen aber das Hirngebiet, das Wahrnehmungen aus diesem Körperteil abbildete, durch neuroplastische Vorgänge weiterhin erregt werden kann. Diese Patient*innen können etwa als Phantomschmerz immer noch ein subjektives, leibliches Erleben "in" diesem Körperteil haben, weil ihr Selbst-Modell den fehlenden Körperteil weiterhin enthält und die Aktivierung dieser sensorischen Gebiete hinreichend für das subjektive Erleben des Körperteils zu sein scheint. Umgekehrt führt ein Verlust von Hirngebieten, deren Erregung mit bestimmten subjektiven Erlebniskategorien einherging, auch zum Verlust des subjektiven Erlebens in dieser Kategorie, weil mit dem Verlust dieser Hirngebiete auch die entsprechenden repräsentationalen Fähigkeiten verloren gehen. Folglich kommt es zu einer leiblichen "Amputation", weil das so nicht mehr Repräsentierbare auch nicht mehr Teil des Leibes sein kann. Klinisch wird dies als Neglekt bezeichnet. Personen mit einem Neglekt können zum Beispiel das Bewusstsein für eine Körperseite verlieren, ohne sich diesem Verlust bewusst zu werden. Oliver Sacks beschreibt zum Beispiel sehr eindrücklich die dramatischen Konsequenzen, als er aufgrund eines neurologischen Ausfalls vorübergehend jeglichen repräsentationalen Zugang zu seinem Bein verlor (Sacks 2004). Für die gegebene Fragestellung besonders interessant ist, dass Personen, die einen Neglekt für einen bestimmten Körperteil haben, diesen Körperteil auch nicht mehr als "eigen" empfinden. Das daraus folgende Gefühl der "Befremdung" durch das Körperteil kann ein solches Leiden verursachen, dass die Person zum persistenten Wunsch kommt, den intakten "Fremdkörper" operativ entfernen zu lassen.⁹

Neben diesen dramatischen Bildern aus der Klinik gibt es auch alltägliche Ereignisse, in denen sich der üblicherweise transparente Charakter eines Selbst-Modells als repräsentationaler Zustand offenbart. Ein gutes Beispiel ist das sogenannte Eiscreme-Kopfweh: Ein Erleben von Schmerz in der Stirn, das sich einstellen kann, wenn man ein Eis zu schnell isst. Psycho-physiologisch erklärt sich dieses Erleben, weil der Nerv, der Schmerzsignale aus dem Stirnbereich weiterleitet, über dem Gaumen verläuft und durch Kälte reizbar ist. Die Kälte im Mund führt also dazu, dass dieser Nerv gereizt wird und einen Prozess anstösst, der mit einem bewussten Erleben korrespondiert. Aber auch wenn wir das wissen, erleben wir die Reizung des Nervs unmittelbar und unmissverständlich weiter als Schmerz in der Stirn, und in einer Qualität, die ganz anders ist als "Kaltsein". Diese äusserst konsistente

⁹Sogenanntes "Body integrity identity disorder", siehe z.B. Brugger und Lenggenhager (2014).

örtliche und qualitative Fehlattribution erklärt sich dadurch, dass unser Selbst-Modell, das den epistemisch versperrten materiellen Zugang zu den biologischen Zusammenhängen ersetzt, offensichtlich von Natur aus so angelegt ist, dass Signale von dem betroffenen Nerv unmittelbar als Schmerz und nicht als Kälte bewusst werden, der in der Stirn und nicht als im Gaumen lokalisiert wird.

Ein weiteres Beispiel für Zustände, in denen unser Selbst-Modell "durchschimmert", ist der Schlaf. Im Schlaf ist einerseits die Wirkung von Aussenreizen auf das Gehirn sehr eingeschränkt, gleichzeitig ist das Gehirn dabei nicht inaktiv. Die mit dieser Aktivität korrespondierenden Bewusstseinsinhalte bezeichnen wir als Träume. Wie es scheint, sind auch im Traum die Kategorien des Erlebens abhängig vom Erregungszustand der entsprechenden kortikalen Areale (Noreika u. a. 2020). Und weil wir auch beim Träumen über ein Selbst-Modell verfügen, erleben wir uns in unseren Träumen als jemanden mit einem Leib und machen im Traum Erfahrungen, die den Erlebensdimensionen dieses Leibs entsprechen. Träume sind damit leibliche Erfahrungen (Revonsuo 2005), die loslösbar sind vom biologischen Körper.

3.4 Metzingers These der transparenten Selbstrepräsentation

Die nächste Frage, die uns im Folgenden beschäftigen soll, ist, wie in diesen Prozessen etwas stattfindet, das als "Selbst" empfunden wird. Thomas Metzinger hat in diesem Kontext und in Übereinstimmung mit vielem des oben schon Erarbeiteten eine These entwickelt, das die besondere und unmittelbare Gegebenheit des eigenen Selbst als natürliche Eigenschaft eines Systems erklärt, das durch geeignete Informationsverarbeitungsprozesse zu repräsentationalen Leistungen fähig ist: Die These postuliert, dass es in einem System, in dem es zu repräsentationalem Bewusstsein kommt, auch zu einer bewussten Repräsentation des eigenen Selbsts kommt (Metzinger 2003; Metzinger 2007). Wie bei Tononi ist Metzingers These ebenfalls funktional, indem die Zustände der Systemelemente untereinander und mit der Aussenwelt in Abhängigkeit sind, und repräsentational, indem unterschiedliche Informationen mit gleicher Verhaltensrelevanz in semantisch äquivalente Zustände integriert wird. Ebenfalls analog zu Schlüssen, die oben schon aus Tononis These hergeleitet wurden, basiert Metzingers Modell auf der biologisch plausiblen Annahme, dass menschliche, und sicher auch nicht-menschliche Gehirne ständig möglichst kohärente Repräsentationen des "eigenen" Zustands bilden. Dabei ist der Satz an Zuständen, die mit "eigen" attribuiert werden, letztlich wiederum hauptsächlich durch die Struktur der Integrationsleistungen des peripheren und zentralen Nervensystems bestimmt (Metzinger 2007, S. 218). Weil diese SelbstRepräsentationen typischerweise global, hoch-integriert und kohärent sind, kommt es, so Metzinger, zu holistischen Gestaltphänomenen und zu einer um dieses holistische Selbsterleben zentrierten Erst-Personen-Perspektive (Metzinger 2007, S. 219, 222).¹⁰

Metzingers Vorschlag ist nun, dass der repräsentationale Charakter der Inhalte von bewusstem Erleben dem subjektiven Erleben selbst nicht als solcher zugänglich ist (Metzinger 2007, S. 237). Der repräsentationale Charakter von subjektivem Erleben ist, wie er es formuliert, "phänomenal transparent" und wird selbst nicht als solcher erlebt. Damit kommt es, so seine These, zu dem, was er als naiven Realismus bezeichnet: Die Inhalte von subjektivem Erleben erscheinen dem Subjekt so, wie sie erlebt werden, als unmittelbar real, und nicht als Produkt einer repräsentationalen Leistung. Nur die Inhalte von repräsentationalen Zuständen, nicht aber die Träger dieser Zustände sind also nach Metzinger introspektiv zugänglich (Metzinger 2007, S. 237). Konsistenterweise ist dieses Erleben von "Real-Sein" kompromittierbar, wenn die Passung von phänomenalem Erleben und Funktionieren gestört wird: Wenn wir zum Beispiel einen Drehschwindel erleben, kommt es zu Divergenzen zwischen somatosensorischen, vestibulären und visuellen Empfindungen, die subjektiv sehr schnell als ein Unreal-Sein der Aussenwelt empfunden werden.

Metzinger wendet nun seine These der phänomenalen Transparenz auch auf Selbst-Repräsentationen an. Daraus folgt, dass auch die kortikalen Träger von Selbst-Repräsentationen introspektiv nicht zugänglich sind, sondern sich dem Subjekt phänomenal transparent als direktes Erleben eines realen eigenen Selbst manifestieren (Metzinger 2007, S. 236-238). Einerseits ist deshalb die Weise, wie das Subjekt selbstrepräsentationale Zustände erlebt, wiederum durch eine durchgehende und "dicke Schicht" von psychophysiologischen Intransparenzen von der materiellen Basis dieser Zustände systematisch getrennt. Andererseits externalisiert das Selbstmodell auch solche selbstrepräsentationalen Zustände a-priori in etwas, das dem Subjekt "real" ist. Deshalb ist sich ein Subjekt seiner Selbst sicher.

3.5 Fehler im Selbsterleben

Im Abschnitt 3.3 haben wir bereits gesehen, dass uns der repräsentationale Charakter von durch ein Selbst-Modell vermitteltem Selbsterleben zu typischen Arten von Fehlern erst "befähigt". Mir scheint, dass diese Fehlerbereitschaft in repräsentationalen System gerade darauf hinweist, dass neben der sensorischen Evidenz das Vorwissen eine oft ebenso entscheidende Rolle für den Inhalt einer Repräsentation

¹⁰Dieses holistische Selbsterleben kann allerdings durch Pathologien oder geeignete experimentelle Eingriffe auch empfindlich gestört werden, siehe z.B. Brugger und Lenggenhager (2014) und Abschnitt 3.5.

hat. Nach Karl Fristons Predictive Coding Theory sind mentale Repräsentationen das Produkt eines Prozesses, bei dem kontinuierlich bestehende Erwartungen mit neu eintreffender sensorischer Information möglichst gut in Übereinstimmung gebracht werden (Friston 2005). Dieser Prozess ist notwendig, weil die eintreffende Information häufig unvollständig und mehrdeutig ist, sich aber mit der "richtigen" Vorinformation im Normalfall vervollständigen und eindeutig machen lässt. Sensorische Illusionen wie zum Beispiel das Sehen eines Kaniza-Dreiecks (Abb. 3.1) lassen sich so sehr gut erklären. Experimentell kann man die konkreten Eigenschaften dieses Vorwissens untersuchen, in dem man Versuchspersonen konfliktierende sensorische Informationen anbietet und dabei ihr Erleben misst. Was die Versuchspersonen berichten, reflektiert dann die vom Vorwissen abhängige unterschiedliche Gewichtung der verschiedenen Komponenten der sensorischen Information.

Weil sich auch unsere Selbstrepräsentation auf solches Vorwissen gehört, hat sich ein schillernder Forschungsbereich entwickelt, der anhand systematischer Fehler bei Selbstrepräsentationsleistungen die Funktionsweise dieser Leistungen untersucht. Klassische Beispiele von den entsprechenden experimentellen Paradigmen sind zum Beispiel die Rubber-Hand Illusion (Botvinick und Cohen 1998), propriorezeptive Drifts (Lenggenhager u. a. 2007), Körper-Transfer (Slater u. a. 2010) oder Out-of-body Erfahrungen (Blanke u. a. 2002). Ich denke nicht, dass es nötig ist, die entsprechenden Befunde hier auszubreiten. Wichtig ist mir hier vor allem festzuhalten, dass diese ganze Forschung darauf aufbaut, dass wir aufgrund einer Reihe von wahrscheinlich sehr früh angelegten repräsentationalen Eigenschaften unseres Selbst-Modells in einer sehr systematisierbaren Weise für Fehler anfällig sind, und dass diese Fehler nicht durch bewusste Einsicht in die Ursachen dieser Fehler korrigierbar sind. Fehlleistungen unseres Selbst-Modelle teilen damit eine wichtige Eigenschaft von repräsentationalen Zuständen im Allgemeinen.

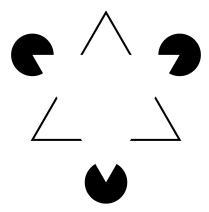


Abb. 3.1: Kanisza Dreieck: In der Figur erscheint ein illusionäres weisses Dreieck, weil das Gehirn unsere Wahrnehmung durch Vorannahmen ständig ergänzt. (©Fibonacci / Wikimedia Commons)

3.6 Zwischenfazit

Ziel dieses Kapitels war es, die Epistemik von Selbsterleben zu klären. Dabei hat sich herausgestellt, dass diese Epistemik wesentlich auf einem Selbst-Modell basiert, das zwischen physischen und mentalen Ereignissen vermittelt. Ich will, um dieses Kapitel abzuschliessen, nun noch kurz ein Zwischenfazit ziehen und die für die weitere Arbeit wichtigsten Punkte zur festhalten. Das sind die folgenden:

- 1. Selbsterleben ist ein dauernd aktiver repräsentationaler Prozess.
- 2. Selbsterleben steht, zumindest nach dem bis jetzt Gesagten, nicht im Widerspruch zu naturwissenschaftlich plausiblen raum-zeitlichen Folgebeziehungen. Materielle Systeme, deren Zustände mit Bewusstsein und Selbsterleben korrespondieren, unterscheiden sich im Bezug auf ihr Wirkvermögen also nicht notwendigerweise von Systemen, deren Zustände nicht mit Bewusstsein und Selbsterleben korrespondieren.
- 3. Die bis jetzt gefundenen Konklusionen basieren lediglich auf einem Satz von theoretischen Annahmen über wahrscheinlich notwendige Eigenschaften von Systemen, deren Zustände mit Bewusstsein und Selbsterleben korrespondieren. Die Konklusionen gelten deshalb auch, wenn wir keine hinreichenden Gründe für solche Eigenschaften angeben können.
- 4. Die getroffenen Annahmen sind biologisch-naturalistisch plausibel und werden durch empirische Beobachtungen gestützt.

- 5. Zwischen dem Erleben des Subjekts und den physischen Ereignissen besteht eine psychophysiologische Intransparenz: Das Subjekt hat aus informationstheoretischen Gründen systematisch keinen direkten Zugang zu den Ursachen von Selbsterleben.
- 6. Die durch psychophysiologische Intransparenz verursachten epistemischen Lücken werden durch dem Subjekt a-priori gegebene Annahmen überbrückt. Die Summe dieser Annahmen ist das, was Metzinger als Selbst-Modell bezeichnet.
- 7. Dieses Selbst-Modell ist dem Subjekt transparent: Das Subjekt hat keinen Zugang zum repräsentationalen und funktionalen Charakter seines Selbst-Modells.
- 8. Infolge der Transparenz des Selbst-Modells kommt es zu dem, was Thomas Metzinger als naiven Realismus bezeichnet. Erleben, und auch das Erleben des eigenen Selbst, ist für das Subjekt unmittelbar und realistisch.
- 9. Selbsterleben scheint unfehlbar: Weil dem Subjekt das eigene Erleben nicht als repräsentationalen Zustand, sondern a-priori als etwas Reales erscheint, ist ihm auch seine Qualität als erlebendes Subjekt a-priori real. Damit ist das Subjekt seiner Selbst sicher.
- 10. Selbsterleben manifestiert sich in spezifischen leiblichen Qualitäten, wie zum Beispiel Farbe, Klang, oder Schmerz.
- 11. Selbst-Modelle enthalten Abbildfunktionen zwischen Leib und Teilen des Körpers.
- 12. Die durch den Leib repräsentierten Teile des Körpers haben folglich relationale Eigenschaften, die dem Subjekt wiederum als leibliche Erfahrungen unmittelbar bewusst werden können.
- 13. Durch das Vorhandensein solcher leiblicher Erfahrungen ist das "Selbst" von "Nicht-Selbst" für das Subjekt unmittelbar und direkt erlebbar abgegrenzt.

$_{\text{KAPITEL}}4$

Gestalt von Selbsterleben und eigener Urheberschaft

Im vorhergehenden Kapitel haben wir festgestellt, dass das, was mit "Selbst" bezeichnet wird, dem Subjekt durch ein phänomenal transparentes Selbst-Modell gegeben ist, das mentale und körperliche Zustände aufeinander abbildet. Das folgende Kapitel soll nun konkretisieren, was das Vorhandensein eines solchen Selbst-Modells für die subjektive Gestalt von Selbsterleben bedeutet. Ich werde mich für diese Konkretisierung der Gestalt von Selbsterleben auf einen Vorschlag von Shaun Gallagher (Gallagher 2000; Gallagher 2012) stützen, der besagt, dass sich ein sogenanntes minimales oder "atomisches" Selbst aus einen Erleben von Mein-Sein (Sense of Ownership) und einem Erleben von Urheberschaft (Sense of Agency) konstituiert.

Gallaghers Vorschlag lässt sich in systematischer Weise an die oben ausgeführten Punkte anknüpfen. Wir haben erstens festgestellt, dass die materiellen raum-zeitlichen Folgebeziehungen repräsentationaler Bewusstseinszustände subjektiv intransparent sind: Eine dicke Schicht von Informationsintegrationsschritten verunmöglicht systematisch, dass eine physische Kette von raum-zeitlichen Folgebeziehungen, in die solche Zustände eingebettet sind, aus den gegebenen Bewusstseinsinhalten rekonstruiert werden kann. Zweitens haben wir festgestellt, dass dieser fehlende Zugang zu den materiell realisierten Zusammenhängen durch ein Selbst-Modell ersetzt wird, das diese Zusammenhänge im Normalfall zuverlässig abbildet. Gallaghers Vorschlag kann man nun als eine Unterteilung der Phänomenologie von Selbsterleben anhand der beiden möglichen Wirkrichtungen in Selbst-Modellen auffassen: In Selbst-Modellen abbildbare Ereignisse, die von der Welt über ein Selbst-Modell auf das Erleben des Subjekts wirken, geben dem Erlebten die Eigenschaft des Mein-Seins. Umgekehrt führen vom Subjekt auf die Welt wirkende Ereignisse durch das Vorhandensein eines Selbst-Modells zu einem Erleben von Urheberschaft für das Ereignis. Von der Fragestellung her ist das Erleben von Urheberschaft am Anfang und im Zentrum des Interesses, weil es raum-zeitliche Folgebeziehungen vom Subjekt auf die Welt abbildet. Es wird sich aber zeigen, dass auch das Erleben von Mein-Sein für die gegebene Fragestellung relevante Aspekte beizutragen hat, weil Wirkungen vom Subjekt auf die Welt immer auch auf das Subjekt zurückwirken, und damit das Erleben von Urheberschaft auch Komponenten von Mein-Sein beinhaltet. Ich fange deshalb mit der Darstellung von Mein-Sein an.

4.1 Mein-Sein von Selbsterleben

Erleben von Mein-Sein ist die Erfahrung, dass bestimmte Ereignisse in der Welt bei mir zu einem leiblichen Erleben führen. Dieser Zugang verläuft biologisch betrachtet üblicherweise über die verschiedenen Sinnesorgane, mit denen ein Körper ausgestattet ist, manifestiert sich aber subjektiv in einer Qualität, die sich nicht einfach aus der Interaktion der physikalischen Eigenschaften des Ereignisses und dem Aufbau des entsprechenden Sinnesorgans herleiten lässt. Mein subjektives Erleben von Farbigkeit scheint im Normalfall objektiv dadurch auslösbar zu sein, dass elektromagnetische Wellen bestimmte Zustände in den Fotorezeptoren in meinem Auge verändern und dass es dadurch im Gehirn zu einer Kaskade weiterer neuronaler Ereignisse kommt. Gleiches gilt für Ereignisse, die Chemo-, Druck-, Schmerz-, Temperatur- oder andere Sensoren erregen, was subjektiv als Erleben von Geruch, Geschmack, Berührung, Geräusch, Schmerz, Wärme oder Ähnlichem einhergeht. Andererseits scheint es in einer fundamentalen Weise unmöglich, für solche biologischen Erklärungsversuche von Erleben brauchbare Brückenprinzipien zu finden, mit denen sich die jeweilige Qualität des subjektiven Erlebens auf die neurobiologischen Vorgänge reduzieren lassen (Levine 1999). 11 Im Erleben des Subjekts manifestiert sich dieses Fehlen objektiver Brückenprinzipien für raum-zeitliche Folgebeziehungen von physischen zu mentalen Ereignissen als psychophysiologische Intransparenz: Mögliche physikalisch-materielle Erklärungen der eigenen mentalen Zustände sind dem Subjekt in fundamentaler Weise unzugänglich. Gleichzeitig erlebt sich das Subjekt dauernd in regelhaften gegenseitigen Abhängigkeiten mit der physischen Welt. Diese Kombination von gegenseitigen regelhaften Abhängigkeiten mit der physischen Welt und der psychophysiologischen Intransparenz von subjektivem Erleben macht es erforderlich, dass dem Subjekt die vorhandene epistemische Lücke durch ein adäquates Selbst-Modell überbrückt wird. Als Folge dieser Überbrückungsleistung und der Tatsache, dass diese Leistung selbst nicht

¹¹Um das, was sich aus diesem subjektiven Erleben für Schlüsse zu dem Verhältnis zwischen mentalen und physischen Zuständen ergibt, läuft eine lange und häufig kontroverse philosophische Debatte (siehe z.B. Dennett (1988) und Chalmers (2007)). Die dabei strittigen Punkte müssen uns aber für das, was hier herleitet werden soll, nicht weiter beschäftigen. Wichtig für uns ist bloss, dass dieses subjektive Erleben im Gegensatz zu anderen Formen von Wissen und Überzeugungen an das gebunden ist, was aus der Sensorik des eigenen Körper zu kommen scheint.

bewusst wird, erlebt sich das Subjekt als direkt mit der Welt verbunden, und die Art und Weise, wie es die Welt erfährt, als real: Der mentale Zustand des Erlebens von zum Beispiel Rot ist, qua systematischer Unzugänglichkeit alternativer Erklärungen und qua Transparenz eines Selbst-Modells, subjektiv *identisch* mit denjenigen physischen Eigenschaften der materiellen Welt, die beim Subjekt zu Rot-Erleben führen.

Durch die Notwendigkeit eines an das Subjekt gebundenen Selbst-Modells für das subjektive Erleben eines Ereignisses ist dieses Erleben also an ein jeweils einzelnes Subjekt gebunden: Niemand anders als das Subjekt selbst kann in der gegebenen Form Zugang zum gegebenen Ereignis haben. Damit ist subjektives Erleben in diesen durch ein Selbst-Modell vermittelten sensorischen Qualitäten dem Subjekt eigen: Es ist mein Schmerz, mein Schmecken von Süsse, mein Erleben von Wärme, weil niemand ausser ich selbst zu einem jeweils bestimmten Token von Schmerz, Süss-Sein oder Wärme denselben Zugang als sensorische Erfahrung hat als ich.

Ich will diesen Zusammenhang von Mein-Sein und der Beteiligung eines Selbst-Modells anhand von zwei kontrastierenden Beispielen kurz illustrieren:

- Fall 1: Ich könnte glauben, die Erde sei eine Scheibe. Nun kann mir jemand eine Reihe von Gründen nennen, die für diese Person hinreichende Gründe für die Überzeugung sind, dass die Erde rund sei. Dadurch kann ich ebenfalls zur Überzeugung kommen, dass die Erde rund sei. Nachdem die Person mir ihre Gründe dargestellt hat, würde ich etwa sagen: "Jetzt, wo du mir das so dargestellt hast, glaube ich auch, dass die Erde rund ist."
- Fall 2: Ich habe keine Zahnschmerzen. Wenn nun jemand anders behauptet, dass ihr der Zahn wehtut, und auch wenn die Person das in der grösstmöglichen Ausführlichkeit tut und alle Gründe anführt, die sie hat, um zur Überzeugung zu kommen, dass sie Zahnschmerzen hat, wird das nicht dazu führen, dass ich im Ernst sagen kann: "Jetzt, wo du mir das so dargestellt hast, habe ich ebenfalls Zahnschmerzen."

Was unterscheidet die beiden Fälle? Im Fall 1 ist der Zugang zu dem, was die Aussage "Die Erde ist rund." in überzeugender Weise stützt, unabhängig von unseren jeweiligen Selbst-Modellen. Deshalb können wir aus gemeinsam zugänglichen Gründen zu denselben gemeinsamen Überzeugungen kommen. Im Fall 2 ist der Zugang zu dem, was die Aussage "Ich habe Zahnschmerzen." wahr macht, wesentlich abhängig von repräsentationalen Leistungen von dem uns jeweils nur alleine gegebenen Selbst-Modell und damit privat. Nur das Subjekt selbst kann durch sein eigenes Selbst-Modell ein bestimmtes Ereignis als Schmerz erleben und hat darum zum Wahrheitsgehalt der Aussage "Ich habe Zahnschmerzen" einen privilegierten

Zugang. Deshalb ist es nicht möglich, dieselben Zahnschmerzen gemeinsam zu haben. Es gibt, wie Hermann Schmitz es ausdrückt, immer nur genau eine Person, die einen Satz wie "Ich habe Zahnschmerzen." aussagen kann, während jemand anderes zwar darüber sprechen kann, aber zum Gemeinten, also zum gegebenen Auftreten von Erleben von Zahnschmerz, selbst keinen Zugang hat (Schmitz 2007, S.6). Wenn es nicht meine Zahnschmerzen sind, von denen ich spreche, spreche ich nicht über etwas, was sich mir als Schmerz manifestiert.

4.2 Selbst Ursache sein und basale Handlungen

Im vorhergehenden Abschnitt haben wir gesehen, wie Einflüsse von aussen auf das Subjekt zu einem Erleben führen, das dem Subjekt eigen ist. In umgekehrter Richtung führen, so Gallaghers These, Einflüsse vom Subjekt auf seine Aussenwelt zum Erleben von Urheberschaft: Erleben von Urheberschaft bezeichnet die Erfahrung, dass bestimmte Ereignisse in der Welt durch etwas verursacht zu sein scheinen, was ich als eigene Handlung leiblich erlebe. Äusserlich manifestiert sich das, was als Urheberschaft erlebt wird, offensichtlich als sogenannte Willkürmotorik, also in Form von koordinierten und typischerweise zielgerichteten Bewegungen des Körpers. Um den leiblich-mentalen Aspekt von Willkürmotorik besser zu charakterisieren können, ist es mir hilfreich, auf den Begriff der basalen Handlung zurückzugreifen, den Arthur Danto eingeführt hat: Als basale Handlungen (basic actions (Danto 1965)) definiert Danto Handlungen, die sich nicht mehr auf weitere Handlungen zurückführen lassen (Danto 1965, S. 142). Alle Handlungen sind also, so Danto, entweder selbst basale Handlungen, oder lassen sich als sogenannt komplexe Handlungen auf basale Handlungen zurückführen (Danto 1965, S. 142).

4.2.1 Was sind basale Handlungen?

Danto motiviert sein Konzept der basalen Handlung mit der Notwendigkeit, einen infiniten Regress zu unterbrechen, der droht, wenn jede Handlung immer nur durch eine vorhergehende Handlung erklärbar wäre. Ein solcher infiniter Regress würde es unmöglich machen, dass es zu einer Handlung kommen kann (Danto 1965, S.141), weil es dann keine Handlung gäbe, die am Anfang einer Handlungskette stehen könnte. Wenn es also überhaupt Handlungen gibt, muss es auch etwas geben, was zwar selbst eine Handlung ist, das aber nicht durch eine weitere und vorhergehende Handlung verursacht worden ist.

Basale Handlungen konstituieren damit für Danto einen Satz von Handlungsmöglichkeiten, bei denen die Person nichts weiter zu tun braucht, als "das zu tun", was der basalen Handlung entspricht. Basale Handlungen sind, so Danto, der Person a-priori gegeben (Danto 1965, S.145). Danto zieht in diesem Kontext eine

Analogie zwischen basalen Handlungen und basalen Sätzen (etwa: "Das sieht rot aus."): Auch deren Wahrheitsgehalt scheint der urteilenden Person nur entweder schon unmittelbar gegeben, oder dann unmöglich, nie aber Folge anderer Sätze. So wie Rot-Sehen nicht weiter begründet werden kann als dadurch, dass man rot sieht, gibt es, wie Danto argumentiert, keine weitere Handlung, die zu einer basalen Handlung führt (Danto 1965, S.146), als das Tun der basalen Handlung selbst.

Ein Beispiel: Die Handlung, ein Stück Kuchen zu essen, ist im Normalfall keine basale Handlung, weil sie sich auf eine Reihe von basalen Handlungen wie dem Greifen nach dem Kuchen zurückführen lässt. Die Greifbewegung selber ist dagegen wahrscheinlich eine basale Handlung, weil es keine weitere Handlung gibt, die sie erklärt, als das Greifen selbst.

Im Kontext der gegebenen Fragestellung ist es nicht nötig, diesen Satz von basalen Handlungsmöglichkeiten konkreter einzugrenzen, solange wir festhalten können, dass es für jede Person, die überhaupt handlungsfähig sein soll, einen nichtleeren Satz von basalen Handlungen geben muss. Wahrscheinlich ist es sogar unmöglich, einen allgemein gültigen Satz basaler Handlungsmöglichkeiten bestimmen zu wollen, weil es Handlungen gibt, die für eine Person eine basale Handlung darstellen, während sie für jemand anders die Folge weiterer Handlungen ist. So wäre vielleicht das Greifen eines Gitarrenakkordes für einen geübten Gitarristen eine basale Handlung, während jemand, der noch nie Gitarre gespielt hat, dieselbe Handlung durch eine Kombination von separat ausgeführten Fingerbewegungen herbeiführen würde.

Basale Handlungen sind aus zweierlei Gründen für die gegebene Fragestellung interessant. Einerseits sind komplexe Handlungen für die Klärung der Fragestellung wenig ergiebig: Indem sie das Eintreffen einer Handlung einfach durch eine vorhergehende Handlung erklären, gewinnen wir durch sie nichts an wesentlicher Erkenntnis zur Ursächlichkeit der handelnden Person. Zweitens tragen, wie wir sehen werden, basale Handlungen als der Person im Gegensatz zu komplexen Handlungen immer bereits unmittelbar gegebene Handlungsmöglichkeiten Eigenschaften, die qualitativ viel mit subjektivem Erleben von eigener Ursächlichkeit zu tun haben. Ich werde ich mich deshalb für die Klärung von Erleben von Urheberschaft im Allgemeinen vollständig auf die Klärung von Erleben von Urheberschaft von basalen Handlungen stützen.

4.2.2 Psychophysische Token-Token Identitäten bei basalen Handlungen

Wie manifestieren sich basale Handlungen subjektiv? In Abgrenzung zu anderen Körperbewegungen wie Reflexen haben basale Handlungen als leiblich-mentale Ereignisse nach Danto die Qualität des Wollens: Eine basale Handlung ist, wie Danto

es formuliert, not then the result of an act of will, it is an act of will" (Danto 1965, S.148). Wenn ich eine basale Handlung tun will, ist das für Danto also nicht Grund dafür, sondern subjektiv identisch damit, dass ich das auch tue. Damit begegnen wir in umgekehrter Wirkrichtung derselben subjektiven Identität von mentalen und materiellen Zuständen und Ereignissen, die wir schon beim Mein-Sein festgestellt haben: Ein gegebener mentaler Zustand des "Wollens" ist im Falle von basalen Handlungen, qua systematischer Unzugänglichkeit alternativer Erklärungen und qua Transparenz eines Selbst-Modells, subjektiv identisch mit denjenigen physischen Ereignissen in der materiellen Welt, die unmittelbar zu der entsprechenden willkürlichen Körperbewegung führen. Wie beim Erleben von Mein-Sein scheint es erneut so, dass objektivierbare Brückenprinzipien fehlen, über die sich die subjektive Qualität des Wollens in biologische Vorgänge wie Erregungen des motorischen Kortex führen, die dann zu den entsprechenden "gewollten" Muskelkontraktionen führen. Wiederum erklärt sich das Fehlen objektiver Brückenprinzipien für raum-zeitliche Folgebeziehungen zwischen mentalen und physischen Ereignissen durch psychophysiologische Intransparenz, nun aber von mentalem Erleben zu deren physischen Folgen: Während es subjektiv üblicherweise reicht, dass ich meinen Arm bewegen will, um meinen Arm zu bewegen, bleibt es mir in fundamentaler Weise verborgen, welche Hirngebiete, Nervenfasern, Muskeln etc. notwendig und hinreichend sind, um meinen Arm zu bewegen. Und wie gehabt machen es die fehlenden Brückenprinzipien von mentalen zu physischen Ereignissen notwendig, dass das Subjekt auch die vorhandene epistemische Lücke zwischen Wollen und Tun mit einem Selbst-Modell überbrückt.

Der individuelle Satz von möglichen basalen Handlungen, von dem Danto spricht, ist also der Satz von Körperbewegungen, die ein Selbst-Modell in Form möglicher Akte des Wollens dieser Körperbewegungen mental repräsentieren kann. Diese repräsentationale Eigenschaft von Selbst-Modellen macht es nun möglich, dass das Subjekt solche basalen Handlungen als frei erlebt: Eigenes Wollen ist für das Subjekt bereits hinreichend, um sich selbst als Ursache eines materiellen Ereignisses zu erleben. Gleichzeitig verhindert die psychophysiologische Intransparenz systematisch, dass das Subjekt sein Wollen als direkte Folge physischer Ereignisse erleben kann. In Abwesenheit möglicher alternativer Erklärungen erlebt das Subjekt folglich sein Wollen auch als notwendig für basale Handlungen. Auch all das, was dem Subjekt als mögliche Erklärungen für sein Wollen zugänglich ist (etwa Wünsche und Überzeugungen), stellt sich dem Subjekt in leiblich-mentaler Form dar: Es ist damit dem Subjekt eigen und nicht durch "fremde" Ursachen erklärbar. Subjektiv gibt es also nichts, was dem Subjekt "seine" basalen Handlungen erklären kann und das gleichzeitig unabhängig vom Subjekt ist. Ich halte fest:

Fazit 4.2.1: Das Subjekt erlebt eigenes basales Handeln als frei

Ein Selbst-Modell repräsentiert Akte des "Wollens" in einer Form, die sie subjektiv als notwendig und hinreichend für basale Handlungen erscheinen lässt. Deshalb erlebt sich das Subjekt in seinen basalen Handlungen als frei.

4.2.3 Mein-Sein, Epistemik und Materialisierung von basalen Handlungen

An diesem Punkt ist es interessant, epistemische Gegebenheiten und die materielle Realisierung basaler Handlungen einander gegenüberzustellen. Eine erste relevante Beobachtung für diese Gegenüberstellung können wir machen, indem wir feststellen, dass der Zugang zu den relevanten relationalen Eigenschaften einer basalen Handlung, wie alle über ein Selbst-Modell laufenden Zugänge, privat ist: Es ist immer nur das Subjekt selbst, und nie jemand anders, das über sein Selbst-Modell zu einem Token einer basalen Handlung den direkten Zugang als "Gewolltes" hat: Wie bei jedem leiblichen Erleben gibt es wiederum nur genau eine Person, die einen Satz wie "Ich bewege meine Hand, indem ich meine Hand bewegen will." aussagen kann, während jemand anderes zwar darüber sprechen kann, aber zum Gemeinten, also zum gegebenen Token einer Handbewegung als Akt des Willens, selbst keinen Zugang hat. Basale Handlungen sind damit wie alle anderen leiblichen Empfindungen dem Subjekt "eigen": Eine basale Handlung ist meine Handlung, weil es immer und nur ich selbst sein kann, der die basale Handlung gewollt und damit verursacht hat. Ich halte fest:

Fazit 4.2.2: Mein-Sein basaler Handlungen

Dem Subjekt sind die von ihm willentlich ausgehenden Handlungen eigen.

Daraus folgt als zweite relevante Beobachtung, dass die epistemischen Zugänge zu basalen Handlungen vom Subjekt und von ausserhalb des Subjekts her unabhängig voneinander sind: Während es bei der äusserlichen Beobachtung eines Ereignisses, das den Anschein einer basalen Handlung macht, grundsätzlich nicht feststellbar ist, ob das Ereignis vom Subjekt gewollt ist oder nicht – auch wenn bis auf die letzte synaptische Übertragung alles erklärbar wäre, enthielte diese Erklärung kein Erleben von Wollen –, ist es für das Subjekt einzig genau dieses Erleben von Wollen, das ein Ereignis zu einer basalen Handlung machen kann.

Gleichzeitig, und im Gegensatz zur epistemischen Situation, können wir drittens feststellen, dass sich die materielle Realisation einer basalen Handlung dadurch auszeichnet, dass sie in einem besonderen relationalen Verhältnis, nämlich dem des Gewollt-Werdens, zum Subjekt steht. Deshalb ist es möglich, bei der Ontologie von basalen Handlungen von einer Token-Token Identität von mentalen Zuständen des Wollens und physisch wirksamen materiellen Zuständen und Ursachen auszugehen, ohne dass sich die Beschreibung als Wollen auf die materielle Beschreibung reduzieren lässt. Diese Nichtreduzierbarkeit von relationalen Eigenschaften identischer Sachverhalte hat Frege in seiner Arbeit "Über Sinn und Bedeutung" festgestellt (Frege 1892, S. 25): So kann zum Beispiel ein und dieselbe Person für eine jemanden "mein Kind" für jemand anders aber "mein Geschwister" sein. Obwohl mit "Kind" und "Geschwister" dieselbe Person bezeichnet wird, wäre es offensichtlich falsch, wenn ich eine Person, die "mein Kind" ist, als "mein Geschwister" bezeichne. Frege nennt die Art des Gegebenseins eines Sachverhalts (also in diesem Beispiel die Gegebenheit aus Sicht z.B. des Elternteils im Gegensatz zu der des Geschwisters) den Sinn der Bezeichnung (Frege 1892, S. 26), und den (gemeinsamen) Sachverhalt dessen, was damit jeweils bezeichnet wird, dessen Bedeutung. Sätze, die sich in dieser Terminologie zwar im Sinn unterscheiden, aber dieselbe Bedeutung, und damit immer dieselben Wahrheitswerte haben, bezeichnen für Frege identische Sachverhalte (Frege 1892, S.36).

Offensichtlich ist es so, dass es relationale Eigenschaften gibt, die den Wahrheitswert des Sinns einer Aussage abhängig von der Relation zu einem epistemischen Agent ändern, während die Bedeutung einer Aussage referenzunabhängig ist. Das ist insbesondere dann der Fall, wenn die relationalen Eigenschaften einen sogenannt intensionalen Kontext bilden, wie das bei Verhältnissen von "wissen, dass", oder "erleben, dass" der Fall ist, die auch für unser Selbsterleben wichtig sind.¹² In gleicher Weise und durch den als "Erleben als eigen" aufgespannten intensionalen Kontext ist die Wahrheit der Beschreibung einer Körperbewegung als Akt des eigenen Wollens, nicht aber die Existenz der Bewegung selbst, abhängig von einer eindeutigen Beziehung zwischen dieser Bewegung und dem Erleben des einzelnen Subjekts: Für dieselbe Bewegung kann die Aussage, dass die Bewegung selbst gewollt ist, abhängig vom aussagenden Subjekt richtig oder falsch sein, ohne dass sich an der Existenz der Bewegung selbst irgendetwas ändert. Der Versuch, basale Handlung auf physische Ereignisse reduzieren zu wollen, muss also daran scheitern, dass das physische Ereignis auch ohne Bezug zum Subjekt bereits

¹²Ein klassisches Beispiel für einen solchen vom intensionalen Kontext abhängigen Wahrheitswert sind die beiden Aussagen "Lois Lane zweifelt an der Existenz von Superman." und "Lois Lane zweifelt nicht an der Existenz von Kent Clark.", die gemeinsam auch dann wahr sein können, wenn es der Fall ist, dass Kent Clark Superman ist. Das ist möglich, weil das Verb "zweifeln an" einen intensionalen Kontext bildet, der die Wahrheitswerte der Aussage bei identischen Extensionen verändern kann.

vollständig erklärt ist und dieser Bezug zum Subjekt in einem reduktionistischen Projekt damit notwendigerweise aus der Erklärung fällt. Gleichzeitig kann eine solchermassen reduzierte Erklärung von basalen Handlungen die vom subjektiven Erleben abhängigen Wahrheitswerte zur Bezeichnung der Bewegung "als eigenes Wollen" in keiner Weise abbilden, und bleibt damit grundsätzlich unvollständig. Mentale Zustände wie Zustände des Wollens können damit einerseits kraft einer Token-Token Identität mit physischen Zuständen materiell wirksam sein, und sind gleichzeitig kraft ihrer relationaler Eigenschaften, wie "einem Subjekt eigen sein" und "von einem Subjekt gewollt sein", nicht auf solche physischen Zustände reduzierbar (Koenig 2020). Ich halte fest:

Fazit 4.2.3: Identität von Zuständen des Wollens mit materiell wirksamen materiellen Zuständen bei basalen Handlungen

Jedes dem Subjekt als Wollen einer basalen Handlung gegebene Token eines mentalen Zustands ist mit einem physisch wirksamen Token eines materiellen Zustands identisch, ohne dass es darauf reduziert werden kann.

4.3 Alternative Handlungsmöglichkeiten

Im obigen Abschnitt haben wir festgestellt, dass ein Token eines mentalen Zustands des Wollens identisch sein muss mit einem physisch wirksamen Token eines materiellen Zustands, aber nicht auf einen solchen materiellen Zustand reduziert werden kann, weil er relationale Eigenschaften hat, die einen intensionalen Kontext schaffen. Die Gesamtheit dieser relationalen Eigenschaften ist, um ein bereits weiter oben (Abschnitte 3.3 und 3.4) eingeführtes Konzept wieder aufzunehmen, in dem Selbst-Modell des Subjekts enthalten: Alle materiellen Zustände, die in diesem Selbst-Modell enthalten sind, haben die relationale Eigenschaften, dem Subjekt eigen zu sein und dem Subjekt in einer besonderen und direkten Weise erfahrbar zu sein.

Bei der Untersuchung von Selbst-Modellen haben wir festgestellt, dass Selbst-Modelle physische und mentale Ereignisse systematisch aufeinander abbilden, dass aber diese Abbildungsleistung dem Subjekt nicht bewusst werden kann. Im gegebenen Fall von basalen Handlungen ersetzt ein Selbst-Modell so die gesamte, dem Subjekt völlig unzugängliche Kette von raum-zeitlichen Folgebeziehungen von einem subjektiv als Wollen erlebten repräsentationalen Zustand zu den koordinierten Kontraktionen und Relaxationen einer Vielzahl von Muskelfasern durch eine Token-Token Identität von Wollen und Bewegung. Diese erfahrungsgemäss sehr

zuverlässige und funktionale Abbildungsfunktion eines Selbst-Modells liefert damit das, was ein gutes Modell idealerweise leisten soll: Es abstrahiert einerseits so stark wie möglich von der Materialisierung dessen, was es abbilden soll, bleibt diesem aber funktional so ähnlich, dass es innerhalb eines interessanten praktischen Rahmens keine Rolle spielt, ob wir vom Modell oder dem Modellierten ausgehen. Innerhalb dieses praktischen Rahmens sind also die Wahrheitswerte aller möglichen Aussagen für das Modell und das Modellierte idealerweise die gleichen, sodass es, immer streng innerhalb dieses praktischen Rahmens, idealerweise zu einer Pseudo-Identität von Modell und Modelliertem kommt: Innerhalb des gegebenen Rahmens erlaubt ein gutes Modell, verschiedene mögliche Ereignisse innerhalb des Modellierten zu simulieren und zuverlässige Vorhersagen zu den entsprechend unterschiedlichen zu erwartenden Folgen zu machen. Ich werde deshalb diesen Rahmen, innerhalb dessen ein Selbst-Modell zuverlässige Aussagen macht, als seinen Gültigkeitsbereich bezeichnen.

In diesem Licht betrachtet besagt Metzingers These eines transparenten Selbst-Modells von subjektivem Erleben, dass jedes bewusste leibliche Erleben immer schon voraussetzt, dass es innerhalb des Gültigkeitsbereichs eines Selbst-Modells stattfindet, dass also ein Selbst-Modell das Modellierte in der Praxis zuverlässig genug abbildet: Nur unter dieser Voraussetzung kann uns das, was uns ein Selbst-Modell als Erfahrung abbildet, als "real", das heisst identisch mit den repräsentierten physischen Zuständen oder Ereignissen erscheinen. Gleichzeitig erklärt die Annahme einer a-priori Gültigkeit eines Selbst-Modells für leibliches Erleben auch dessen Transparenz: Weil alles Erfahren immer a-posteriori ist, die Existenz eines Selbst-Modells aber a-priori gegeben ist, kann es selbst gar nicht erfahrbar sein und kann damit auch nicht bewusst werden. Auch die Tatsache, dass wir Fehlrepräsentationen eines Selbst-Modells (wie bei der Hirnstimulation, beim Eiscreme-Kopfweh oder bei den Träumen, immer noch als "real" erleben, spricht dafür, dass eine Art von Identitätsbeziehung von mentalem Erleben und dadurch repräsentierten gehirn-externen Zuständen und Ereignissen dem Subjekt a-priori, und nicht a-posteriori gegeben zu sein scheint.

Nun erlaubt es ein Selbst-Modell als ein materiell vom Abgebildeten unabhängiges, aber funktional hinreichend äquivalentes System von raum-zeitlichen Folgebeziehungen, dass mehrere alternative und sich eventuell gegenseitig ausschliessende Ursachen und Folgen von Zuständen oder Ereignissen neben- oder hintereinander dargestellt werden können.¹³ Wie Metzinger (Metzinger 2007, S. 218) es formuliert: "An active self-model then can be seen as a subpersonal functional state: a set of causal relations of varying complexity that may or may not be realized at a given

¹³Theoretisch sind pathologische Fälle denkbar, denen die repräsentationale Möglichkeit fehlt, in Alternativen zu denken. Ich verzichte hier aber auf eine weitere Untersuchung solcher Spezialfälle, weil es mir scheint, dass solche Fälle auch nicht mit den Laienmeinungen zu Willensfreiheit kompatibel wären und deshalb wenig zur Fragestellung beitragen könnten.

point in time." Bei der Repräsentation von Ereignissen vom Typ basaler Handlungen sind dies immer jeweils mindestens zwei Alternativen, nämlich eine gegebene basale Handlung entweder auszuführen oder sie zu unterlassen. Ein Subjekt, das sich notwendigerweise durch ein solches Selbst-Modell erlebt, und das grundsätzlich zu basalen Handlungen fähig ist, erlebt sich damit notwendigerweise so, dass es in jedem gegebenen Kontext mindestens zwei Alternativen hat. Und weil diese Repräsentationen von möglichen basalen Handlungen im Selbst-Modell vom Subjekt direkt und nicht als Repräsentationen erlebt werden, erlebt das Subjekt diese Alternativen ebenfalls als real. Daraus folgt:

Fazit 4.3.1: Erleben von realen Wahlmöglichkeit basaler Handlungen

Ein intaktes Subjekt erlebt sich grundsätzlich so, dass es jederzeit die reale Wahl hat, eine basale Handlung auszuführen oder sie zu unterlassen.

4.4 Eigene Urheberschaft und Gültigkeitsbereichs des Selbst-Modells

Nachdem wir oben festgestellt haben, dass sich *innerhalb* des Gültigkeitsbereichs eines Selbst-Modells interessante Eigenschaften des subjektiven Erlebens von Handlungsmöglichkeiten als frei, als gewählt, als eigen und als ursächlich finden, will ich im Folgenden in der umgekehrten Richtung untersuchen, was das Erleben von Urheberschaft für den Gültigkeitsbereich eines Selbst-Modells bedeutet.

Zentral für die gegebene Fragestellung ist, dass es das Vorhandensein von Vorhersagen erlaubt, den Gültigkeitsbereich des eigenen Selbst-Modells empirisch zu validieren und gegebenenfalls anzupassen: Innerhalb des Gültigkeitsbereichs dieses Selbst-Modells ergeben eigene Willensakte zuverlässige Vorhersagen zu den Folgen dieser Willensakte, während ausserhalb des Gültigkeitsbereichs dieses Selbst-Modells Willensakte a-priori nicht möglich sind oder ohne vorhersehbare Folgen bleiben: Ereignisse, die dem Subjekt durch ein Selbst-Modell als Folgen von Willensakten erklärbar sind, werden deshalb natürlicherweise als selbst verursacht er-

¹⁴Eine weitere interessante Folge der Voraussagefähigkeit eines Selbst-Modells ist, dass sich das Subjekt dadurch als zeitlich extendiert erlebt: Einige der Ereignisse der Gegenwart erklären sich möglicherweise aus den dem Subjekt eigenen Handlungen in der Vergangenheit, und gegenwärtige eigene Handlungen führen zu in der Zukunft liegenden Erwartungen, die erfüllt werden können. Dadurch gewinnt das subjektive Selbsterleben eine zeitliche Extension, die möglicherweise den Kern einer personaler Identität über die Zeit bildet.

lebt (Beispiel in Tabelle 4.1).¹⁴ Ereignisse, die ausserhalb des Selbst-Modells stattfinden werden dagegen als *nicht-selbst verursacht* und damit als *fremd-verursacht* erlebt (Beispiele in Tabelle 4.2). In Einklang mit den neuropsychologischen Befunden aus Kapitel 3.3 kommt es auch dann nicht zu einem Erleben von eigener Urheberschaft, wenn der Körper des Subjekts materiell involviert ist, das Selbst-Modell aber keine entsprechenden Vorhersagen macht (siehe Beispiel in Tabelle 4.3).

| Aussage: | "Als Folge meines Wunsches nach Abkühlung bewegt sich mein Körper so, dass das Fenster aufgeht." |
|--|---|
| Vorhandensein eines gesetzmässigen Zusammenhangs von mentalen und physischen Ereignissen innerhalb eines Selbst-Modells: | Ja: Körperbewegungen, die zum Öffnen von Fenstern gehören, gehören üblicherweise auch zum Repertoire basaler Handlungen eines menschlichen Subjekts. Damit kann das Subjekt das Ereignis grundsätzlich durch einen eigenen Willensakt erklären. |
| Vorhersagekraft des eigenen Willens: | Das Erleben des Willensakts erlaubt es dem Subjekt, das Ereignis, "mein Körper bewegt sich so, dass das Fenster aufgeht" zuverlässig vorherzusagen. |
| Empirische Validierung eines Selbst-Modells: | Das subjektiv erlebte Ereignis, das Fenster öffnen zu wollen, koinzidiert, wie vom Selbst-Modell vorhergesagt, mit einer Körperbewegung, die zum Öffnen des Fensters führt. |
| Subjektives Fazit: | Der eigene Willensakt ist für das Subjekt eine hinreichende Erklärung für das Auftreten des entsprechenden Ereignisses, und das Subjekt identifiziert sich typischerweise als Urheber dieses Ereignisses. Es sagt von sich: "Ich habe das Fenster selbst aufgemacht". |

Tab. 4.1: Beispiel von üblichem Erleben von Selbstursächlichkeit

| Aussage: | "Als Folge komplexer physischer Vorgänge weht der Wind so, dass das Fenster aufgeht." |
|--|--|
| Vorhandensein eines gesetzmässigen Zusammenhangs von mentalen und physischen Ereignissen innerhalb eines Selbst-Modells: | Nein: "Mit-dem-Wind-Wehen" gehört nicht zum Repertoire von basalen Handlungen eines Menschen. Es ist dem Subjekt deshalb nicht möglich, etwas durch einen Willensakt des "Mit-dem-Wind-Wehen-Wollens" zu erklären. |
| Vorhersagekraft des eigenen Willens: | Es hat kein Willensakt stattgefunden, der etwas erklären könnte. |
| Empirische Validierung eines Selbst-Modells: | Fehlt. |
| Subjektives Fazit: | Das Öffnen des Fensters muss sowohl aus empirischen wie auch natürlichen Gründen durch etwas begründet werden, das ausserhalb dessen liegt, was dem Subjekt eigen ist. Das Subjekt sagt: "Es war "Nicht-Ich", das dazu geführt hat, dass das Fenster aufgegangen ist." |

Tab. 4.2: Beispiel von Erleben üblicher Fremdursächlichkeit

| Aussage: | "Als Folge komplexer physischer Vorgänge bewegt sich mein Körper so, dass das Fenster aufgeht." |
|--|---|
| Vorhandensein eines gesetzmässigen Zusammenhangs von mentalen und physischen Ereignissen innerhalb eines Selbst-Modells: | Nein: Die Aussage enthält kein mentales Ereignis. |
| Vorhersagekraft des eigenen Willens: | Es hat kein Willensakt stattgefunden, der etwas erklären könnte. |
| Empirische Validierung eines Selbst-Modells: | Fehlt. |
| Subjektives Fazit: | Das Öffnen des Fensters wird in dem gegebenen Beispiel durch etwas begründet, das ausserhalb dessen liegt, was dem Subjekt eigen ist. Es gibt für das Subjekt keinen Grund für die Aussage: "Ich habe das Fenster selbst geöffnet." |

Tab. 4.3: Beispiel von Fehlen von Erleben von Selbstursächlichkeit am eigenen Körper

Damit sind wir nun beim fehlenden Teil von dem, was neben dem Erleben von Mein-Sein in Gallaghers These zum minimalen Selbst gehört: Das *präreflektive Erleben von Urheberschaft* (Prereflective Sense of Agency) (Gallagher 2000; Gallagher 2012), das ich folgendermassen definieren will:¹⁵

Definition 4.4.1: Präreflektives Erleben von Urheberschaft

Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn ihm ein subjektiv erlebter eigener Willensakts das Eintreffen dieses Ereignisses als Wirkung zuverlässig vorhersagt.

Diese Vorhersage- und Evaluationsleistungen können wieder, wie im Rahmen der in Kapitel 3 erarbeiteten Epistemik und Biologie von Selbsterleben, auf verschiedenen Ebenen abgebildet werden. Eine dieser Ebenen ist die biologisch-materiellen Ebene, in der solche Leistungen als Resultat von neuronalen Informationsintegrationsprozessen verstanden werden. Parallel dazu existieren aber, so die Annahme, verschiedene mentale Ebenen, auf denen sich solche Vorgänge unmittelbar durch ganz bestimmte, dem Subjekt eigene und private Erlebensweisen manifestieren. Wie bereits in Abschnitt 4.2.2 begründet, werden wir davon ausgehen, dass diese verschiedenen Ebenen der Beschreibung materiell möglicherweise identisch sind, sich epistemisch aber grundsätzlich unabhängig voneinander verhalten. Abbildung 4.1 zeigt eine Darstellung aus einer Arbeit von Peter Haggard (Haggard 2008), die versucht, diese Schichtung von verschiedenen physiologischen und mentalen Beschreibungsebenen und deren funktionale Verbindungen bei solchen aktiven und zeitlich ausgedehnten Selbstwahrnehmungsprozessen zu schematisieren.

¹⁵Gallagher unterscheidet zwischen präreflektivem und reflektivem Erleben von Urheberschaft: Während präreflektives Erleben von Urheberschaft eine immediate und direkte Erfahrung bezeichnet (Gallagher 2000, S.16), meint reflektives Erleben von Urheberschaft ein rational deduktives Urteilen darüber, ob hinreichend Gründe für die Überzeugung vorhanden sind, sich als Ursache eines Ereignisses zu bezeichnen (Gallagher 2012, S. 18). Für die gegebene Fragestellung interessiert uns primär die präreflektive Form von Erleben von Urheberschaft. Im Folgenden meine ich mit Erleben von Urheberschaft, sofern nicht explizit anders bezeichnet, immer die präreflektive Form von Selbsterleben.

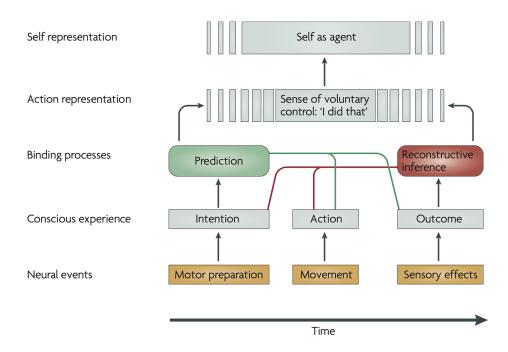


Abb. 4.1: Biologisch manifestiert sich eine Handlung als Aktivität von prämotorischen, motorischen und sensorischen Strukturen des Gehirns (Ebene "Neural events"), die subjektiv aber als eine Sequenz von Ereignissen mit einem jeweils unitären Bewusstsein erlebt werden (Ebene "Conscious experience"). Aus der Integration von erlebten Handlungsintentionen und damit kongruenten Folgen (Ebene "Binding process") erlebt die Person eigene Urheberschaft (Ebene "Action representation"), was die kausale Dimension des Selbst-Modells validiert (Ebene "Self representation"). Abbildung aus Haggard (2008).

Unter den psychologischen, neurowissenschaftlichen und philosophischen Forschenden gibt es einen breiten Konsens, dass solche Vergleichsvorgänge zwischen den vorhergesagten und den tatsächlich eintretenden Folgen von willentlichen Akten für das Erleben eigener Ursächlichkeit zentral sind und dass sich korrespondierende biologische Vorgänge auch systematisch feststellen lassen (Gallagher 2000; Blakemore und Frith 2003; Haggard 2008; Tsakiris 2017). Gleichzeitig gibt es einen recht grossen Spielraum bei der Frage, auf welcher Ebene diese Vorgänge stattfinden. Zur Debatte steht hier ein Bereich, der bei Vorgängen innerhalb des motorischen Systems anfängt, und der über den prädiktiven Wert von intentionalen Zuständen bis zu Vorhersagen geht, die sich aus sehr allgemeinen Überzeugungen und Wünschen der Person zu sich selbst ergeben (Gallagher 2007).

4.4.1 Pathologisches Erleben von Urheberschaft

Wir haben in Abschnitt 3.5 bereits die Beobachtung festgehalten, dass es typische und systematische Fehler im Selbsterleben geben kann. Das trifft auch für das Erleben von eigener Urheberschaft zu. Eine Personengruppe, bei der solche Fehler systematisch auftreten, sind Menschen mit einer Psychose. Die entsprechenden Symptome werden sinnigerweise als "Ich-Störungen" bezeichnet (Fuchs 2013) und z.B. akustisch-verbale Halluzinationen, über die solche Patienten häufig berichten, können als fehlende Selbstattribution von innerer Sprache interpretiert werden (Heinks-Maldonado u. a. 2007): Weil die betroffene Person einerseits Sprache wahrnimmt, sich dabei aber nicht als ursächlich erlebt, wird diese Sprachwahrnehmung als fremdverursacht erlebt. Die Person sagt dann, es sei, wie wenn da jemand spräche, ohne dass tatsächlich jemand spricht. Umgekehrt kann eine Person, die sich vielleicht gewünscht hat, dass einer anderen Person ein Unglück zustosse, und dieser Person unabhängig davon tatsächlich ein Unglück geschieht, sich in wahnhafter Weise als Ursache für dieses Ereignis erleben.

4.5 Problematik der kausalen Relevanz von erlebter eigener Urheberschaft

Als funktionales Modell repräsentiert ein Selbst-Modell nicht nur die verschiedenen dem Subjekt zur Verfügung stehenden Handlungsalternativen, es macht auch zu jeder dieser Alternativen eine mehr oder weniger präzise Vorhersage zu deren zu erwartenden jeweiligen physischen Folgen. Wenn diese Vorhersagen regelmässig und zuverlässig genug eintreffen, bildet ein Selbst-Modell offensichtlich Regelhaftigkeiten ab, die in Kombination mit einem eingetroffenen Willensakt das Vorhergesagte zu einem guten Grad zu erklären scheinen. Damit liefert das Selbst-Modell mögliche Erklärungen für die Folgen dessen, was das Subjekt als Wollen erlebt. Diese Erklärungen haben die Struktur eines deduktiv-nomologischen (DN) Modells: Es gibt offenbar einen Gültigkeitsbereich des Selbst-Modells, innerhalb dessen ein Satz von "Wenn X, dann Y"-Beziehungen allgemein gültig scheinen. Wenn es in Folge eines subjektiv zugänglichen Erlebens eines bestimmten Willensaktes als Ereignis X konsistent zu einem konkreten physischen Ereignis Y kommt, scheint es so, dass das Selbst-Modell eine logische Erklärung für diesen Zusammenhang liefert: Das Selbst-Modell, gemeinsam mit dem Eintreffen eines Willensaktes, sind gemeinsam hinreichende Bedingungen für das Eintreffen des Ereignisses.

Die Kombination von Selbst-Modell und entsprechendem Willensakt funktioniert damit strukturell analog zu naturwissenschaftlichen Erklärungen, die sich auch auf allgemein gültige Naturgesetze stützen (Tabelle 4.4). Damit scheint das regelhafte Auftreten von Y in Folge von X das Selbst-Modell $X \to Y$ zu bestätigen.

Das Subjekt hat durch diese Bestätigung der Gültigkeit seines Selbst-Modells möglicherweise einen Grund für die Überzeugung, Y selbst und willentlich verursacht zu haben. Verhalten, das entsprechende Bestätigungen sucht, ist jedenfalls beobachtbar, wenn zum Beispiel kleine Kinder mit dem Schalter einer Lampe spielen und dabei systematisch erkunden, ob das An- und Ausgehen des Lichts in einem regelhaften Bezug zur ihrer Manipulation des Schalters steht und sie so "selbst" das Licht an- und ausgehen lassen können.

Diese Begründung eigener Ursächlichkeit erweist sich jedoch als trügerisch. DN-Modelle sind nämlich aus einer Reihe von grundsätzlichen Einwänden schlecht geeignet, kausale Verhältnisse richtig abzubilden. Ich will hier auf einen dieser Einwände spezifisch eingehen, weil er auch in der gegebenen Fragestellung auf Probleme hinweist, mit denen ein Subjekt konfrontiert ist, dem qua faktisch gegebener systematischer Zusammenhänge von mental erlebten Willensakten und physischen Ereignissen die in seinem Selbst-Modell enthaltene Möglichkeit eigener Urheberschaft bestätigt zu sein scheint. Es geht um das Problem der kausalen Irrelevanzen:

Das kausale Irrelevanzproblem von Erklärungen durch DN-Modelle entsteht im Allgemeinen dadurch, dass in der logischen Struktur des DN-Modells irrelevante Elemente hinzugefügt werden können, und dass diese irrelevanten Elemente dann ebenfalls Erklärungskraft zu haben scheinen. Um ein klassisches Beispiel von Wesley Salmon (Salmon 1984) anzubringen, folgt aus den jeweils wahren Aussagen, dass (a) "alle Männer, die die Pille nehmen, nicht schwanger werden" und (b) "Jones ein Mann ist, der die Pille nimmt", dass (c) "Jones nicht schwanger wird". Die nomologische Aussage (a) und die Tatsache (b) enthalten aber mit dem Pille-Nehmen ein kausal offensichtlich irrelevantes Element, das gewissermassen im "Huckepack" der Aussagen (a'): "Männer werden nicht schwanger" und (b'): "Jones ist ein Mann" zu einer anscheinenden Erklärungskraft kommt, weil auch $(a' \wedge b')$ die Aussage (c) wahr macht.

Damit stellt sich die Frage, ob das Auftreten von Willensakten ein zwar möglicherweise unvermeidliches Begleitphänomen bestimmter physischer Ereignisse ist, dabei aber (wie das Pillenehmen von Jones) kausal irrelevant bleibt: Nehmen wir kurz an, dass es einen physischen Zustand X gibt, der hinreichend für ein weiteres physisches Ereignis Y ist, also $X \to Y$ nomologisch gültig ist. Wenn wir nun hinzufügen, dass das Subjekt X als eigenes Wollen erlebt, dann wäre dieses Wollen des Subjekts in gleicher Weise irrelevant für die Erklärung von Y, wie das Pillenehmen von Jones irrelevant ist für die Feststellung, dass er nicht schwanger wird. Das Wollen des Subjekts wäre damit epiphänomenal, also eine reine Begleiterscheinung, und keine Ursache für das Auftreten von Y. Diese Position wurde insbesondere auch von einigen modernen Naturwissenschaftlern prominent vertreten (Wegner 2003).

| Logische Struk- tur | Klassische Rolle | Logische Klassische Klassisches Beispiel Struk- Rolle tur | Beispiel Selbsterleben Rolle beim Selbsterleben | Rolle beim Selbsterleben |
|---------------------------|---------------------|---|---|--------------------------|
| $X \to Y$ | Gesetz | Immer, wenn Wasser kälter als 0°C wird, gefriert es. | Immer, wenn Wasser käl- Immer, wenn ich meine Selbst-Modell ter als 0°C wird, gefriert Hand bewegen will, bewegt sich meine Hand. | Selbst-Modell |
| X | Ereignis | Das Wasser ist kälter als 0°C | Das Wasser ist kälter als Ich will meine Hand be- Willensakt wegen. | Willensakt |
| Y | Wirkung | Das Wasser gefriert. | Meine Hand bewegt sich. Physisches Ereignis | Physisches Ereignis |

und einem Ereignis X, auf das sich dieses Gesetz bezieht (was zusammen das Explanans bildet) folgt logisch ein weiteres Ereignis Y (das Explanandum), das damit, so das DN-Modell, durch $(X \to Y) \land X$ erklärt wird. Das Auftreten von Y bestätigt **Tab. 4.4:** Selbsterleben im deduktiv-nomologischen Modell: Aus der Kombination eines allgemeingültigen Gesetzes $X \to Y$ nach dem Auftreten von X die Folgebeziehung $X \to Y$.

Die Begründung eigener Willensfreiheit qua empirischer Gültigkeit eines Selbst-Modells, das diese Freiheit zulässt, ist deshalb unzureichend: Sie lässt offen, ob nicht andere, nicht-mentale Ereignisse die tragende kausale Rolle spielen, deren Eigenschaften im Widerspruch zum dem stehen, was sich aus der besonderen epistemischen Situation des Subjekts in Bezug zu den Qualitäten eigener Ursächlichkeit ergibt. Kandidaten für solche Ereignisse gibt es, weil mentale Ereignisse, die Teil einer basalen Handlung sind, zumindest token-token-identisch sind mit einzelnen physischen Ereignissen, die in der materiellen Welt in einer vorhersehbaren Weise kausal wirksam wären. Damit basierten diese für das Selbsterleben des Subjekts notwendigen raum-zeitlichen Folgebeziehungen in der materiellen Welt auf nomologischen Beziehungen, die jedoch im Gegensatz zum Selbst-Modell des Subjekts nicht repräsentational, möglicherweise aber deterministisch wären. Diesen nomologischen Beziehungen fehlte folglich, dass sie sich innerhalb eines modellhaften Abbilds der physikalischen Welt als Möglichkeiten darstellten und aus diesem Grund frei wären und Alternativen hätten.

Wir werden uns also die Frage nach der kausalen Relevanz von Erklärungen eigener Urheberschaft qua Gültigkeit eines Selbst-Modells und die Möglichkeit der darin enthaltenen Willensfreiheit genauer anschauen müssen, und dabei insbesondere die Rolle von Determinismus von materiellen raum-zeitlichen Folgebeziehungen innerhalb solcher Kausalerklärungen qua Selbst-Modell klären müssen. Darum wird es im folgenden Kapitel auch gehen.

KAPITEL 5

Erleben von Urheberschaft im Dilemma zwischen Freiheit und Determinismus

Wir haben in Kapitel 3 festgestellt, dass Selbsterleben, und darin eingeschlossen das Erleben von Ursächlichkeit des eigenen Selbst, durch ein Selbst-Modell vermittelte repräsentationale Zustände sind. Der repräsentationale Charakter dieses Erlebens wird dem Subjekt jedoch nicht phänomenal bewusst (Abschnitt 3.4), sodass das "Selbst" als etwas unmittelbar Reales erlebt wird (Metzinger 2003). Im gleichen Kapitel haben wir weiter festgestellt, dass ein solches Selbst-Modell statt rein materiell definierter raum-zeitlicher Folgebeziehungen gesetzmässige funktionale Beziehungen zwischen mentalen und physischen Ereignissen herstellt. Im Kapitel 4 wurde erarbeitet, welche Folgen diese epistemische Situation des Subjekts für das Erleben eigener Urheberschaft hat. Dabei haben wir festgestellt, dass, weil die für Erleben eigener Urheberschaft relevanten raum-zeitlichen Folgebeziehungen auf diesem Selbst-Modell beruhen (Abschnitt 4.2), das Subjekt sich als fähig erlebt, jederzeit frei handeln zu können (Abschnitt 4.3).

Gleichzeitig hat sich in Abschnitt 4.5 gezeigt, dass die tatsächliche kausale Rolle solcherart erlebter mentaler Zustände unklar ist. Noch dazu droht ein Widerspruch, weil das, was subjektiv als eigene freie Wahl erlebt wird, auf physischen Ereignissen beruht, die möglicherweise deterministisch sein müssen, wenn sie dem Subjekt eine vorhersehbare Urheberschaft erlauben sollen. Es wird daher in diesem Kapitel darum gehen, die Beziehung von Willensfreiheit und Determinismus unter dem Aspekt zu prüfen, dass sich diese Überzeugungen zu Willensfreiheit aus einem Erleben von Urheberschaft herleitet, das sich wiederum aus repräsentationalen Leistungen von sich (in unserem Fall primär in unserem Gehirn) materialisierten Vorgängen ergibt. Damit müssen wir uns mit der Frage beschäftigen, in welchem Verhältnis dieses Erleben eigener Urheberschaft zur Annahme eines in der Welt gültigen Determinismus steht. Es scheint mir, dass sich diese Prüfung, nach etwas Vorarbeit im Abschnitt 5.1, in drei Unterfragen aufteilen lässt:

1. Wie verhält sich die ontologische Gegebenheit von eigener Urheberschaft und Willensfreiheit zu Determinismus? Weil Erleben von

Urheberschaft beinhaltet, dass das Subjekt in vielerlei vorhersehbaren raumzeitlichen Folgebeziehungen zu seiner physischen Umwelt steht, müssen wir untersuchen, was es ist, das diese Folgebeziehungen ermöglicht, und in welchem Verhältnis das, was diese Folgebeziehungen vorhersehbar macht, zur Annahme eines Determinismus und zu Willensfreiheit steht. Diese Fragestellung wird im Abschnitt 5.2 bearbeitet.

- 2. Wie verhält sich die epistemische Gegebenheit eigener Urheberschaft Determinismus? Wir haben bereits festgestellt, dass das subjektive Erleben von eigener Urheberschaft fundamental durch einen repräsentationalen Zugang des Subjekts zur Welt geprägt ist, der über ein Selbst-Modell vermittelt wird. Es wird daher im Folgenden darum gehen müssen, das Verhältnis zwischen subjektiven epistemischen Gegebenheiten eigener Urheberschaft und Willensfreiheit und der Annahme eines Determinismus zu klären. Diese Fragestellung wird im Abschnitt 5.3 bearbeitet.
- 3. Welche kausalen Rollen spielen die ontologischen und epistemischen Gegebenheiten von Urheberschaft, und was folgt daraus für das Vorhandensein von Willensfreiheit? Gegeben der Fall, dass sich die ontologische und die epistemische Gegebenheit von eigener Urheberschaft in ihrem Verhältnis zu Determinismus signifikant unterscheiden, müssen wir klären, welche Rollen diese jeweiligen Gegebenheiten spielen, wenn es um die Beurteilung der Möglichkeit von Willensfreiheit geht. Diese Fragestellung wird im Abschnitt 5.4 bearbeitet.

Ich werde mich, um diese Fragen anzugehen, auf eine Reihe von Prämissen stützen, die ich dann aussagenlogisch zueinander in Bezug setzen werde. Einen Teil dieser Prämissen entnehme ich dem bereits Erarbeiteten, weitere Prämissen insbesondere zum Determinismus und zur Vorhersehbarkeit werde ich im folgenden Abschnitt einführen.

5.1 Determinismus, Vorhersehbarkeit und Erleben von Urheberschaft



Abb. 5.1: Determinimus, von Genis Carreras (Carreras 2013)

5.1.1 Physikalischer Determinismus

Um das Verhältnis von Erleben von Urheberschaft und Determinismus zu klären, brauchen wir eine Definition von Determinismus. Für das gegebene Projekt ist es von Vorteil, wenn wir dafür eine Definition wählen, die möglichst unkontrovers

und anspruchslos ist. Ein solches minimalistisches Verständnis von Determinismus macht es uns möglich, zusätzlich problematische Nebenschauplätze wie die Frage, ob eine bestimmte Form von Determinismus wahr ist und wie wir das wissen könnten (Keil 2018, S. 63) zu vermeiden. Ich werde daher im Folgenden eine Formulierung von Determinismus verwenden, die in Bezug zur Ontologie der für Willensfreiheit möglicherweise problematischen eindeutigen Bestimmtheit von Ereignissen neutral ist. Eine solche Formulierung entnehme ich aus Keil (2018, S. 53):

Prämisse 5.1.1: Deduktiv-nomologisches Verständnis von Determinismus

"Das Stattfinden eines Ereignisses wird erklärt, indem man Gesetze und Anfangsbedingungen auffindet, die, passend kombiniert, eine Beschreibung des Ereignisses zu deduzieren erlauben. Und der Weltlauf ist genau dann deterministisch, wenn aus deterministischen Gesetzen und der Beschreibung eines vollständigen Weltzustands Beschreibungen aller weiteren Weltzustände folgen."

Diese Definition von Determinismus ist neutral gegenüber dem, was die für Determinismus notwendigen Gesetzmässigkeiten ausmacht. Wir können uns diese Gesetzmässigkeiten als Gottes Wille¹⁶, als eine anderweitig nicht änderbare Weltordnung¹⁷ oder (modernerweise) als Naturgesetze¹⁸ denken, ohne dass wir in einen Widerspruch mit der obigen Definition von Determinismus kommen.

Des Weiteren soll unser Verständnis von Determinismus in der Lage sein, Folgebeziehungen in der physischen Welt vorwegzunehmen, weil wir einen Determinismus untersuchen wollen, der Bezug darauf nimmt, dass sich das Subjekt als in der physischen Welt ursächlich wirksam erlebt. Deshalb macht es Sinn, die Existenz eines physischen Determinismus zu postulieren, der Gesetzmässigkeiten in der physischen Welt beschreibt. Ich führe das mit der folgenden Prämisse ein:

 $^{^{16} {\}rm Theologischer~Determinismus:}$ "Gott führt alle Dinge gemäss seinem ewigen Willen und Sinn zusammen."

¹⁷Fatalismus: "Wenn Laios und Iokaste einen Sohn zeugen, so wird dieser seinen Vater erschlagen und seine Mutter heiraten."

¹⁸Physikalischer Determinismus: "Wenn bei einem Luftdruck von 1013.25 hPa Wasser kälter als 0 Grad wird, gefriert es."

Prämisse 5.1.2: Physikalischer Determinismus

Jedes physische Ereignis folgt logisch aus einem Satz von bestehenden Anfangsbedingungen und der gesetzmässigen Gültigkeit von allgemeinen und eindeutigen raum-zeitlichen Folgebeziehungen.

Zuletzt will ich noch eine weitere Prämisse einführen, die in der Debatte um mentale Verursachung eine grosse Rolle spielt. Die Prämisse beinhaltet die Geschlossenheit der physikalischen Welt (Kim 2005, S. 43) und besagt, dass alle physischen Folgen auch physische Ursachen haben. Sie formuliert damit, als Ablehnung der Möglichkeit von gänzlich-nichtphysischen Ursachen von physischen Ereignissen, eines der Kernprobleme von substanzdualistischen Theorien: Auf die Frage, wie denn, ausser als physisches Ereignis, irgendetwas ein physisches Ereignis verursachen kann, scheint es keine Antworten zu geben, die nicht ebenso mysteriös bleiben wie das, was zu beantworten wäre. Dieses Problem will ich vermeiden. Deshalb wird physikalische Geschlossenheit in der weiteren Argumentation eine Rolle spielen müssen. Ich führe sie in Form der folgenden Prämisse ein:

Prämisse 5.1.3: Physikalische Geschlossenheit der Welt

Jedes physische Ereignis in der Welt, hat eine hinreichende physische Ursache.

5.1.2 Definition von präreflektivem Erleben von Urheberschaft

Um präreflektives Erleben von Urheberschaft aussagenlogisch in Bezug zu den obigen Prämissen "Physikalischer Determinismus" und "Physikalische Geschlossenheit der Welt" setzen zu können, muss dieses ebenfalls definiert werden. Dazu kann ich die im Abschnitt 4.4.1 erarbeitete Definition als Prämisse einführen:

Prämisse 5.1.4: Präreflektives Erleben von Urheberschaft

Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn ihm ein subjektiv erlebter eigener Willensakts das Eintreffen dieses Ereignisses als Wirkung zuverlässig vorhersagt.

Diese Prämisse beinhaltet einerseits mit der Verursachung eines physischen Ereignisses den kausalen Aspekt der Fragestellung. Andererseits macht sie auch eine epistemische Aussage über den Zugang des Subjekts zum gegebenen Ereignis als vorhersehbare Folge eines dem Subjekt eigenen Willensaktes. Über diesen Zugang erlebt sich das Subjekt unmittelbar als Ursache des Ereignisses, wobei dieses Erleben, wie im Abschnitt 4.2.2 hergeleitet, Qualitäten hat, die Willensfreiheit zu stützen scheinen.

5.1.3 Vorhersehbarkeit der Folgen eigener Willensakte dank Gesetzmässigkeiten

Für die weitere Analyse des Erlebens von Urheberschaft müssen wir genauer untersuchen, was es bedeutet, wenn wir von vorhersehbaren Wenn-Dann-Beziehungen sprechen. Formaler ausgedrückt braucht es, wenn ein Ereignis U in der Gegenwart zuverlässig etwas über das zukünftige Eintreffen eines weiteren Ereignisses F aussagen soll, eine Aussage R, die, möglicherweise abhängig von weiteren Hintergrundbedingungen, eine logische Folgebeziehung zwischen den beiden Ereignissen U und F herstellt. Zwei wichtige Dinge lassen sich nun über Verbindung R sagen. Erstens muss sie regelhaft sein, weil sonst das Eintreffen oder zumindest die Wahrscheinlichkeit des Eintreffens von F aus dem Eintreffen von U nicht mehr zuverlässig vorhersehbar wäre, was subjektives Erleben von Urheberschaft verhindern würde. Weiter muss diese Aussage R allgemeingültig sein. Ohne die Allgemeingültigkeit von R würde drohen, dass der Anspruch von Vorhersehbarkeit in einen infiniten Regress fällt, der dazu führen würde, dass gegeben U keine eindeutige Aussage zum Eintreffen von F möglich wäre.

Dieser infinite Regress droht, weil unter der Annahme, dass R nicht allgemein gilt, für die Vorhersehbarkeit von F aus U und R eine weitere Aussage R' gelten müsste, die in Kombination mit einem weiteren Ereignis U' die Aussage R wahr machen würde. Wenn aber auch nicht vorhersehbar ist, wann U' eintrifft, bräuchte es eine weitere Kombination einer Aussage R'' und eines Ereignisses U'', dass das Eintreffen von U' nach sich ziehen würde, und so fort. Die Annahme, dass R (wieder gegeben, dass möglicherweise weitere Hintergrundbedingungen erfüllt sind) nicht

allgemein gültig ist, ist folglich inkompatibel mit der für Erleben von Urheberschaft notwendigen Vorhersehbarkeit von Ereignissen.

Damit bedingt die Vorhersehbarkeit bestimmter Folgen durch das Eintreffen bestimmter Ursachen also notwendigerweise die allgemeine, und damit gesetzmässige Gültigkeit bestimmter Wenn-Dann-Beziehungen. Ich will diesen Sachverhalt in der folgenden Prämisse festhalten:

Prämisse 5.1.5: Vorhersehbarkeit durch Gesetzmässigkeiten

Ein Ereignis ist genau dann zuverlässig vorhersagbar, wenn es aus dem Eintreffen eines bestimmten vorhergehenden Ereignisses und der gesetzmässigen Gültigkeit von bestimmten raum-zeitlichen Folgebeziehungen logisch folgt.

5.2 Physische Erklärbarkeit des Erlebens von Urheberschaft

Wir können uns nun, ausgehend von den oben dargestellten Prämissen, der ersten der drei Unterfragen zuwenden. Es geht um die Frage der Ontologie dessen, was dem Subjekt Wirksamkeit in der Welt ermöglicht, und in welchem Verhältnis das, was dieses Wirken vorhersehbar macht, zur Annahme eines Determinismus und zu Willensfreiheit steht.

5.2.1 Wirksamkeit dessen, was als eigene Ursächlichkeit erlebt wird

Wie bereits im Fazit 4.2.3 festgestellt, bezieht jeder Fall eines Willensaktes seine materielle Wirksamkeit daraus, dass er auch ein Fall eines physischen Ereignisses ist, das die entsprechende Wirkung nach sich zieht. Ich werde solche mit Willensakten wirkungsmässig identischen materiellen Ereignisse als *physische Realisationen* der Willensakte bezeichnen. Das Vorhandensein solcher Realisationen von Willensakten macht es möglich, dass das Wollen des Subjekts ohne Annahme von "magischen" Fähigkeiten eine materielle Wirkung entfalten kann und ohne dass wir gegen die Prämisse "Physikalische Geschlossenheit der Welt" verstossen. Formal können wir diese Realisierungsbeziehung aus der Konjunktion von "Präreflektives Erleben von Urheberschaft" und "Physikalische Geschlossenheit der Welt" im folgenden Argument formulieren:

Präreflektives Erleben von Urheberschaft (1)Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn ihm ein subjektiv erlebter eigener Willensakts das Eintreffen dieses Ereignisses als Wirkung zuverlässig vorhersagt. Physikalische Geschlossenheit der (2)Jedes physische Ereignis in der Welt, hat eine hinreichende physische Ursache. Physikalische Geschlossenheit der Das Subjekt erlebt sich als Ursache eines physischen Welt Präreflektives Erleben Ereignisses, wenn für dieses Ereignis eine physische Ursache vorhanden ist (die den subjektiv erlebten eigenen Willensakts realisiert) und wenn das Erleben dieses Willensakts das Eintreffen dieses Ereignisses zuverlässig vorhersagt.

Argument 5.1: Physische Realisierung des Willensakts

Damit das Subjekt ein präreflektives Erleben von Urheberschaft haben kann, müssen also jeweils zumindest Token-Token-Identitäten zwischen den Willen der Person repräsentierenden mentalen und in der Welt wirksamen physischen Ereignissen bestehen.

5.2.2 Vorhersehbarkeit dessen, was als selbst verursacht erlebt wird

Erleben von Urheberschaft bedingt, dass die Folgen eigener Willensakte vorhersehbar sind. Gegeben die Definition von präreflektivem Erleben von Urheberschaft aus der Prämisse "Präreflektives Erleben von Urheberschaft" und der Erkenntnis, dass Vorhersehbarkeit auf dem Vorhandensein von allgemeingültigen und eindeutigen Gesetzmässigkeiten beruht, lässt sich nun folgendes Argument machen:

(1) Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn ihm ein subjektiv erlebter eigener Willensakts das Eintreffen dieses Ereignisses als Wirkung zuverlässig vorhersagt. Präreflektives Erleben von Urheberschaft

(2) Ein Ereignis ist genau dann zuverlässig vorhersagbar, wenn es aus dem Eintreffen eines bestimmten vorhergehenden Ereignisses und der gesetzmässigen Gültigkeit von bestimmten raum-zeitlichen Folgebeziehungen logisch folgt.

Vorhersehbarkeit durch Gesetzmässigkeiten

(K₂) Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn ihm ein subjektiv erlebter eigener Willensakts und die gesetzmässige Gültigkeit von bestimmten raum-zeitlichen Folgebeziehungen dieses Ereignis als logische Folge vorhersagen.

Präreflektives Erleben von Urheberschaft Vorhersehbarkeit durch Gesetzmässigkeiten

Argument 5.2: Gesetzmässige Beziehungen sind notwendig für Erleben von Urheberschaft

5.2.3 Erleben von Urheberschaft und physische Gesetzmässigkeiten

Wir haben im Abschnitt 5.2.1 festgestellt, dass materiell wirksame Willensakte ihre Wirksamkeit dadurch erhalten, dass sie auch physische Zustände sind, die eine entsprechende Wirkung haben (K_1) . Weiter haben wir im Abschnitt 5.2.2 festgestellt, dass Erleben von Urheberschaft darauf angewiesen ist, dass es gesetzmässige raum-zeitliche Folgebeziehungen gibt (K_2) . Aus der Konjunktion von K_1 und K_2 ergibt sich folgender Schluss auf die Natur der Gesetzmässigkeiten, die für die Realisierungen dessen gelten müssen, was für präreflektives Erleben von Urheberschaft notwendig ist:

Aus der Konjunktion von K_1 und K_2 folgt:

(1) K₁: Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn für dieses Ereignis eine physische Ursache vorhanden ist (die den subjektiv erlebten eigenen Willensakts realisiert) und wenn das Erleben dieses Willensakts das Eintreffen dieses Ereignisses zuverlässig vorhersagt.

Physikalische Geschlossenheit der Welt Präreflektives Erleben von Urbeberschaft

(2) K₂: Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn ihm ein subjektiv erlebter eigener Willensakts und die gesetzmässige Gültigkeit von bestimmten raum-zeitlichen Folgebeziehungen dieses Ereignis als logische Folge vorhersagen.

Präreflektives Erleben von Urheberschaft Vorhersehbarkeit durch Gesetzmässigkeiten

(K₃) Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn für dieses Ereignis eine physische Ursache vorhanden ist, die den subjektiv erlebten eigenen Willensakts realisiert, und wenn ihm ein subjektiv erlebter eigener Willensakts und die gesetzmässige Gültigkeit von bestimmten raum-zeitlichen Folgebeziehungen dieses Ereignis als logische Folge vorhersagen.

Physikalische Geschlossenheit der Präreflektives Erleben von Urheberschaft Vorhersehbarkeit durch Gesetzmässigkeiten

Argument 5.3: Realisierungsbeziehung und Notwendigkeit von Gesetzmässigkeiten

Aus K₃ ergibt sich also, dass die für das Erleben von Urheberschaft notwendigen raum-zeitlichen Folgebeziehungen Beziehungen zwischen physischen Ereignissen sind. Gleichzeitig macht die Prämisse "Physikalischer Determinismus" eine Aussage zu diesen raum-zeitlichen Folgebeziehungen zwischen physischen Ereignissen, die sich in die Argumentation einbauen lässt: Wenn aus dem, was einen eigenen Willensakt realisiert, ein weiteres Ereignis logisch folgen soll, muss diese Realisation Teil der Anfangsbedingungen sein, aus denen, und aus gesetzmässig gültigen raum-zeitlichen Folgebeziehungen das Ereignis folgt. Aus der Konjunktion von K₃ und der Definition von physikalischem Determinismus ergibt sich also:

(1) K₃: Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn für dieses Ereignis eine physische Ursache vorhanden ist, die den subjektiv erlebten eigenen Willensakts realisiert, und wenn ihm ein subjektiv erlebter eigener Willensakts und die gesetzmässige Gültigkeit von bestimmten raum-zeitlichen Folgebeziehungen dieses Ereignis als logische Folge vorhersagen.

Physikalische Geschlossenheit der Welt Präreflektives Erleben von Urheberschaft Vorhersehbarkeit durch Gesetzmässigkeiten

(2) Jedes physische Ereignis folgt logisch aus einem Satz von bestehenden Anfangsbedingungen und der gesetzmässigen Gültigkeit von allgemeinen und eindeutigen raum-zeitlichen Folgebeziehungen. Physikalischer Determinismus

(K₄) Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn für dieses Ereignis physische Anfangsbedingung vorhanden ist, die die Realisation eines subjektiv erlebten eigenen Willensakts enthalten, und aus denen, und aus der gesetzmässigen Gültigkeit von allgemeinen und eindeutigen raum-zeitlichen Folgebeziehungen dieses Ereignis logisch folgt, und wenn das Erleben dieses Willensakts das Eintreffen dieses Ereignisses zuverlässig vorhersagt.

Physikalische Geschlossenheit der Welt Präreflektives Erleben von Urheberschaft Vorhersehbarkeit durch Gesetzmässigkeiten Physikalischer Determinismus

Argument 5.4: Erleben von Urheberschaft beruht auf Determinismus

Aus K₄ folgt, dass die Gesetzmässigkeiten, die für die Realisationen dessen gelten, was unser Erleben von Urheberschaft ausmacht, deterministisch sind, und dass das damit Erleben von Urheberschaft auf einem physischen Determinismus beruht. Wir haben mit "Vorhersehbarkeit durch Gesetzmässigkeiten", "Physikalische Geschlossenheit der Welt", "Präreflektives Erleben von Urheberschaft" und "Physikalischer Determinismus" einen Satz von Prämissen, der für die Realisierungen von subjektiv als selbst verursacht erlebten Ereignissen notwendig scheint.

Es lohnt sich hier, ein erstes Fazit im Hinblick auf die Hypothese zu ziehen, dass das Erleben von Urheberschaft Überzeugungen von Willensfreiheit zu begründen vermag. Mit Blick auf das Erleben von Urheberschaft fällt auf, dass dieses Erleben stark darunter leiden würde, wenn die physische Realisation einer basalen Handlungen nicht eindeutig bestimmte Folgen hätte, sondern Alternativen zuliesse: Als Folge meines Wunsches, zum Beispiel die rechte Hand zu bewegen, würde ich dann nicht notwendigerweise erleben, dass sich meine rechte Hand bewegt, sondern es wären auch alternative Ereignisse möglich, wie zum Beispiel, dass sich der rechte

Arm bewegt. Weil mein Erleben von Urheberschaft darauf beruht, dass als Folge meines Wunsches, die Hand zu bewegen, sich auch die Hand bewegt, und weil ich keinen Wunsch gehabt habe, der eine Bewegung des Arms vorhergesagt hätte, wäre mein Erleben von Urheberschaft empfindlich gestört. Auch wenn wir Determinismus als nur graduell und/oder probabilistisch gegeben annehmen, ändert sich nichts Wesentliches an diesem Befund: Je präziser, und damit deterministischer die physikalischen Gesetzmässigkeiten, desto weiterreichende und präzisere Vorhersagen sind möglich, und desto umfassender wird damit die Möglichkeit von subjektiven Erleben von Urheberschaft. Erleben von Urheberschaft kann also, unabhängig davon, wie "Determinismus" ausbuchstabiert wird, sinnvollerweise nicht im Gegensatz zum Bestehen von eindeutigen physikalischen Gesetzmässigkeiten stehen.

Die Alternativlosigkeit der Folgen von physischen Realisationen basaler Handlungen ist folglich unverzichtbar für ein funktionales Selbst-Modell und damit für ein intaktes Selbsterleben des Subjekts. Wer Erleben von Urheberschaft als Begründung von affirmativen Urteilen zu Willensfreiheit anführen will, findet in der Alternativlosigkeit der Folgen von physischen Realisationen basaler Handlungen eine mit den Naturwissenschaften kompatible Erklärung dafür, dass, wie in der Definition von präreflektivem Erleben von Urheberschaft (4.4.1) postuliert, die Folgen eines subjektiv erlebten Willensakts vorhersagbar sind.

Gleichzeitig, und wie mir scheint bemerkenswert, bezieht sich K_4 zweimal auf die Existenz von raum-zeitlichen Vorhersagen, macht aber keine Aussage dazu, wie diese zueinander stehen. Insbesondere macht sie keine Aussagen dazu, dass es die (deterministischen) raum-zeitlichen Folgebeziehungen zwischen physischen Ereignissen und Anfangsbedingungen sind, anhand derer dem Subjekt bestimmte physische Ereignisse vorhersehbar sind. K_4 ist damit gerade *keine* hinreichende Erklärung für Erleben von Urheberschaft, sondern eine Aussage über physische Ereignisse, auf denen dieses Erleben notwendigerweise beruht. ¹⁹ Ich werde auf diesen Punkt im in den Abschnitten 5.3 und 5.4 ausführlich zurückkommen.

¹⁹Das Argument liefert auch aus anderen Gründen keine hinreichenden Bedingungen für Erleben von Urheberschaft. Das Subjekt darf zum Beispiel bei der Vorhersage des Ereignisses und bei der Wahrnehmung der Folgen keine Fehler machen, weil sonst die wahrgenommenen Folgen der eigenen Handlung nicht den erwarteten Folgen dieser Handlung entsprechen, was definitionsgemäss zum Ausbleiben des Erlebens von Urheberschaft führt. Dieser Problematik sind wir bei den Ich-Störungen im Abschnitt 4.4.1 schon begegnet. Gleichzeitig scheint es unplausibel, dass solche Fehler auf einem fehlerhaften Verständnis der Naturgesetze beruhen. Sie wird uns aber für die gegebene Fragestellung nicht weiter beschäftigen, weil uns primär die Rolle von erfolgreichem Erleben von Urheberschaft für die Begründung von Willensfreiheit interessiert.

5.2.4 Inkompatibilismus von physikalischen Begründungen von Erleben eigener Urheberschaft und Willensfreiheit

Wir sind bei der Untersuchung der Frage, in welchem Verhältnis Erleben von Urheberschaft zu Determinismus steht, zum Schluss gekommen, dass dieses Erleben auf physischen Realisationen mentaler Zustände beruht, deren Einbindung in raum-zeitliche Folgebeziehungen deterministisch sein muss. Um die erste der oben gestellten Teilfragen zu beantworten, gilt es nun noch, dieses Ergebnis in Bezug zu den gegebenen Qualitäten von Willensfreiheit zu stellen. Ich stelle dazu folgendes Argument auf:

(1) K₁: Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn für dieses Ereignis eine physische Ursache vorhanden ist (die den subjektiv erlebten eigenen Willensakts realisiert) und wenn das Erleben dieses Willensakts das Eintreffen dieses Ereignisses zuverlässig vorhersagt.

Physikalische Geschlossenheit der Welt Präreflektives Erleben von Urheberschaft

(2) Jedes physische Ereignis folgt logisch aus einem Satz von bestehenden Anfangsbedingungen und der gesetzmässigen Gültigkeit von allgemeinen und eindeutigen raum-zeitlichen Folgebeziehungen. Physikalischer

(K₅) Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn für dieses Ereignis eine physische Ursache vorhanden ist, die den subjektiv erlebten eigenen Willensakts realisiert, und die aus einem Satz von bestehenden Anfangsbedingungen und der gesetzmässigen Gültigkeit von allgemeinen raum-zeitlichen Folgebeziehungen logisch folgt, und wenn das Erleben dieses Willensakts das Eintreffen dieses Ereignisses zuverlässig vorhersagt.

Physikalische Geschlossenheit der Welt Präreflektives Erleben von Urheberschaft Physikalischer Determinismus

Argument 5.5: Physikalisch eindeutige Bestimmtheit eigener Willensakte

Aus diesem Argument folgt, dass es für das Eintreffen des einen Willensakts realisierenden Ereignisses eine hinreichende und eindeutige physische Erklärung gibt. Dieser Befund ist nun im Widerspruch zu dem, was die relevanten Qualitäten von Willensfreiheit, so wie sie in der Laienmeinung bestehen, zu beinhalten scheinen. Um diesen Widerspruch zu formalisieren, führe ich noch eine an Peter Van Inwagen's Konsequenzargument (Van Inwagen 1983, S. 15) angelehnte weitere Prämisse ein, die diese aus dem Kapitel 2 entnommenen Qualitäten spezifizieren und in Gegensatz zu physischer Bestimmtheit setzen:

Prämisse 5.2.1: Eindeutige Bestimmtheit erlaubt keine zu Willensfreiheit zugehörigen Eigenschaften

Was eindeutig aus bestehenden Anfangsbedingungen und allgemeinen gesetzmässigen raum-zeitlichen Folgebeziehungen folgt, schliesst aus, das es dazu Alternativen gibt, dass das Ereignis frei ist, dass über das Ereignis bestimmt werden kann und dass das Ereignis kontrolliert werden kann.

Damit können wir folgendes Argument für die Inkompatibilität von physischen Realisationen von Willensakten mit den relevanten Qualitäten von Willensfreiheit machen:

(1) K₅: Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn für dieses Ereignis eine physische Ursache vorhanden ist, die den subjektiv erlebten eigenen Willensakts realisiert, und die aus einem Satz von bestehenden Anfangsbedingungen und der gesetzmässigen Gültigkeit von allgemeinen raumzeitlichen Folgebeziehungen logisch folgt, und wenn das Erleben dieses Willensakts das Eintreffen dieses Ereignisses zuverlässig vorhersagt.

Physikalische Geschlossenheit der Welt Präreflektives Erleben von Urheberschaft Physikalischer Determinismus

(2) Was eindeutig aus bestehenden Anfangsbedingungen und allgemeinen gesetzmässigen raum-zeitlichen Folgebeziehungen folgt, schliesst aus, das es dazu Alternativen gibt, dass das Ereignis frei ist, dass über das Ereignis bestimmt werden kann und dass das Ereignis kontrolliert werden kann.

Eindeutige Bestimmtheit erlaubt keine zu Willensfreiheit zugehörigen Eigenschaften

(K₆) Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn dieses Ereignis eine physische Ursache vorhanden ist, die einen subjektiv erlebten Willensakt realisiert, und zu dem es keine Alternativen gibt, das nicht frei ist, über das nicht bestimmt werden kann, und das nicht kontrolliert werden kann, und wenn das Erleben des Willensakts das Eintreffen des Ereignisses A zuverlässig vorhersagt. Physikalische
Geschlossenheit der
Welt
Präreflektives Erleben
von Urheberschaft
Physikalischer
Determinismus
Eindeutige
Bestimmtheit erlaubt
keine zu Willensfreiheit
zugehörigen
Eigenschaften

Argument 5.6: Inkompatibilismus für Erleben von Urheberschaft

Die Argumentationslinie, die Erleben von Urheberschaft auf "Vorhersehbarkeit durch Gesetzmässigkeiten", "Physikalische Geschlossenheit der Welt", "Präreflektives Erleben von Urheberschaft" und "Physikalischer Determinismus" abstützt, und die gesamthaft in Abbildung 5.2 dargestellt ist, führt also unter Einschluss der Prämisse "Eindeutige Bestimmtheit erlaubt keine zu Willensfreiheit zugehörigen Eigenschaften" zu einem Schluss, der auf der Realisierungsebene im Widerspruch zur These zu stehen scheint, dass Erleben von Urheberschaft Willensfreiheit erklären kann: Aus der Argumentation folgt, dass zu dem Grad, zu dem ein physischer Determinismus Erleben von Urheberschaft ermöglicht, dieser Determinismus gleichzeitig verhindert, dass das, was diese Urheberschaft realisiert, die relevanten Qualitäten von Willensfreiheit enthalten kann.

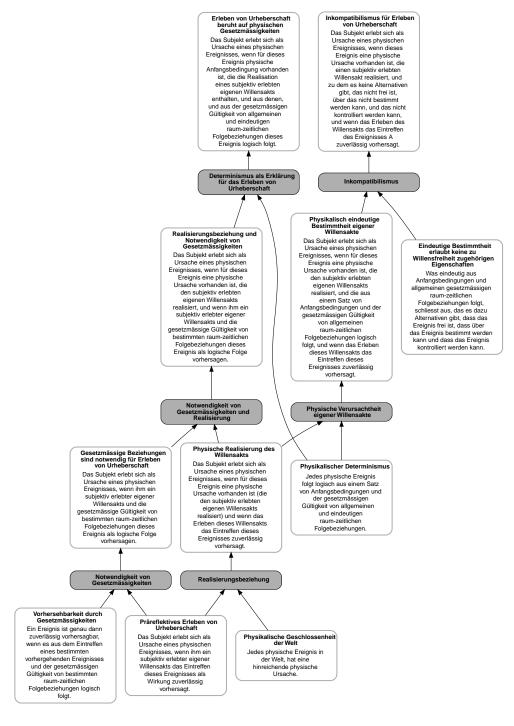


Abb. 5.2: Argumentationskarte für die Möglichkeit von Willensfreiheit, wenn die für Erleben von Urheberschaft notwendigen Gesetzmässigkeiten physikalischer Natur sind.

5.2.5 Erleben von Urheberschaft ohne Willensfreiheit

Die Konklusion aus dem Argument 5.6 impliziert, dass das, was Erleben von Urheberschaft realisiert mit Willensfreiheit inkompatibel sei und widerspricht damit möglicherweise den im Kapitel 2 dargestellten üblichen Intuitionen. Interessanterweise finden wir aber bei genauem Hinschauen in der Phänomenologie unseres Erlebens von Urheberschaft auch Ereignisse, die diesem Fazit entsprechen, denn nicht jedes Erleben von Urheberschaft hat die Qualitäten von Willensfreiheit: Empirisch belegt und subjektiv erlebbar scheint es bei Handlungen einen "Point of no Return" zu geben, bei dem es nicht mehr möglich ist, eine Handlung zu verhindern, obwohl das gewünscht ist. Mir selbst ist dieses Phänomen zum Beispiel bei Kartenspielen bekannt, bei denen es um Geschwindigkeit geht: Wenn mir in dem Moment, in dem ich eine Karte legte, bewusst wurde, dass ich dabei einen Fehler mache, musste ich erleben, dass ich das Ausspielen der falschen Karte nicht mehr verhindern konnte, obwohl ich diese Handlung im Moment des Ausspielens nicht mehr wollte. Ich musste erleben, wie die Ausspielbewegung ohne Kontrollmöglichkeit meinerseits bis zu ihrem ursprünglich intendierten, aber schon während der Tat bereuten Ende kam. Damit hatte ich zwar ein klares und der obigen Definition entsprechendes Erleben von Urheberschaft, dem aber die Qualitäten von Willensfreiheit wie Wahlfreiheit, Selbstbestimmung und Kontrolle völlig fehlten.

Wenn man dieses Phänomen des "Points of no return" wissenschaftlich untersucht, findet man, dass dieser Zeitpunkt ca. 200 ms vor dem Einsetzen einer Bewegung liegt: In einer Arbeit von Schultze-Kraft u.a. (2016) wurde ein System entwickelt, das die einer willentlichen Fussbewegung typischerweise vorhergehenden EEG-Potentiale detektieren konnte. Wenn dieser Detektor voraussagte, dass die Versuchsperson wahrscheinlich bald eine Fussbewegung ausführen würde, wurde der Versuchsperson ein Stopp-Signal präsentiert, auf das hin sie die Fussbewegungen unterdrücken sollte. Die Autoren fanden, dass praktisch alle Stopp-Signale, bei denen die Versuchspersonen die Bewegung nicht mehr verhindern konnte, weniger als 200 ms vor dem Einsetzen der Bewegung präsentiert worden waren. In anderen Worten: Bis zu 200 ms vor der Ausführung einer intendierten willentlichen Bewegung waren die Versuchspersonen noch in der Lage, eine intendierte Bewegung zu kontrollieren und abzubrechen (Schultze-Kraft u.a. 2016). Danach kam es zu einer, wie die Autoren es nannten, "ballistischen" Phase, in der sich die Ereignisse in einer Weise folgten, die von den Versuchspersonen nicht mehr beeinflusst werden konnte. Gleichzeitig ist diese ballistische Phase jedoch definitionsgemäss (siehe 4.4.1) kongruent mit dem Erleben eigener Urheberschaft.

Das Fazit, das wir aus K₆ gezogen haben, ist deshalb im Verhalten und Erleben der Person mit dieser ballistischen Phase während den letzten ca. 200 ms vor der Ausführung einer Bewegung kongruent und inkompatibel mit Willensfreiheit. Daraus folgt aber nur, dass nicht jedes Erleben von Urheberschaft eine

hinreichende Bedingung für Willensfreiheit ist, und dass für das Vorhandensein von Willensfreiheit offenbar noch weitere Bedingungen erfüllt sein müssen. Gleichzeitig scheint das, was mehr als ca. 200 ms vor einer Bewegung stattfindet, im Selbsterleben des Subjekts als Verursacher*in anders, und hat darum vermutlich etwas mit dieser weiteren notwendigen Bedingung zu tun. Nach der obigen Vorarbeit zur Epistemik, Gestalt und Biologie von Selbsterleben liegt es nahe, diese zusätzlich notwendige Bedingung für mit Willensfreiheit verbundenem Erleben von eigener Urheberschaft darin zu vermuten, dass diejenigen willentlichen Zustände, die das Subjekt als gewählt, selbstbestimmt und kontrollierbar erfährt, auch Zustände eines repräsentationalen Bewusstseins sind, die über ein Selbst-Modell an den materiell wirksame physische Zustände gekoppelt sind, während Prozessen, die sich innerhalb der ballistischen Phase abspielen, diese Eigenschaft fehlt. Der folgende Abschnitt wird darum die Rolle von repräsentationalen Zuständen und Selbst-Modellen beim Erleben in die Argumentation mit einbeziehen.

5.3 Erklärbarkeit des Erlebens von Urheberschaft durch ein Selbst-Modell

5.3.1 Kritik des physikalistischen Zugangs

Wir haben in Abschnitt 5.2 den Versuch unternommen, sowohl die Folgebeziehungen wie auch die epistemische Begründung von Erleben von Urheberschaft durch die Realisierungsbeziehung materiell wirksamer physischer Zustände und deren Abhängigkeit von physischen Gesetzmässigkeiten zu begründen. Das Fazit davon war, dass dieser Ansatz zwar den materiellen Teil des Problems gut zu erklären vermag, dass diese Begründung auf der Realierungsebene aber an dem vorbeigeht, was dem Subjekt die eigene Urheberschaft erklärt, und damit keine hinreichende Erklärung dieses Erlebens ist. Ich will diese Kritik anhand der folgenden Punkte noch etwas vertiefen:

• Fehlendes Mein-Sein: Die Realisierungsbeziehung ist unzureichend für das Erleben von Ursächlichkeit als "eigene" Ursächlichkeit: Sie begründet Ereignisse in einer Art und Weise, die zwar für deren Eintreffen hinreichend ist, die aber gleichzeitig dem Subjekt "uneigen" ist. Sie entspricht deshalb, wie in Tabelle 4.3 dargestellt, nicht der Art und Weise, wie das Subjekt Zugang zu dem hat, was die eigene Ursächlichkeit subjektiv begründet, und zu der das Subjekt keine Alternative hat. Damit kann die Realisierungsbeziehung für die Erlebenskomponente der subjektiven Urheberschaft keine verständliche Erklärung liefern.

- Problem der Erklärungslücke: Das Erleben von Urheberschaft bedingt gemäss Definition einen regelhaften Zusammenhang zwischen mentalen Ereignissen und physischen Ereignissen. Die Realisierungsbeziehung impliziert zwar, dass es im Einzelfall über Token-Token-Identitäten einen solchen Zusammenhang gibt. Diese Form von Zusammenhang lässt aber offen, aufgrund welcher Gesetzmässigkeiten diese Zusammenhänge regelhaft sein sollen. Auch die Annahme von physischen Gesetzmässigkeiten macht keine Aussagen zu Beziehungen zwischen mentalen und physischen Zuständen. Ausgehend einzig von Token-Token Identitäten von mentalen und physischen Zuständen, oder von gesetzmässigen Zusammenhängen nur zwischen Paaren von physischen Ereignissen ist es also gerade nicht möglich, ein Erleben von Urheberschaft zu begründen. Entsprechend enthält die Konklusion K₄ zwei erst einmal nicht voneinander abhängige Vorhersagebeziehungen.
- Multiple Realisierbarkeit repräsentationaler Zusammenhänge: Es ist unklar, welche Rolle physische Gesetzmässigkeiten haben, wenn es um Zusammenhänge zwischen multipel realisierbaren Zuständen geht. So gehorcht zum Beispiel auch die Reihe von Buchstaben, aus denen sich dieser Text zusammensetzt, Regeln der Orthografie, die etwas darüber aussagen, wie eine Reihe von Buchstaben ein Wort repräsentiert. Diese Regeln begründen sich aber erst einmal nicht durch die physischen Gesetze, denen die Realisierungen diese Buchstabenfolge unterliegen, denn diese Buchstabenfolge kann in beinahe beliebiger Form realisiert werden. Gleichzeitig folgen Buchstabenfolgen, unabhängig von ihrer Realisation, in der Weise, wie sie ein Wort repräsentieren, denselben orthografischen Regeln. Es gibt also Gesetzmässigkeiten, deren Gültigkeit auf der Repräsentationsebene unabhängig von den Gesetzmässigkeiten auf der Realisierungsebene dieser Repräsentationen besteht. Solch ein Fall liegt wohl auch vor, wenn das Subjekt anhand des Erlebens eigener Willensakte Vorhersagen über deren Wirkung macht: Erstens können wir davon ausgehen, dass solche Willensakte multipel realisierbar sind, zweitens gibt es, analog zu den oben erwähnten Regeln der Orthografie, in Form eines Selbst-Modells gesetzmässige Zusammenhänge zwischen multipel realisierbaren Willensakten und physischen Ereignissen. Weil für das Erleben von Urheberschaft also keine eindeutige physische Ursache und damit auch keine eindeutigen Anfangsbedingungen nötig sind, geraten die Konklusion "Physikalisch eindeutige Bestimmtheit eigener Willensakte" und "Inkompatibilismus für Erleben von Urheberschaft" unter Druck, denn es ist unklar, ob sie in Gegenwart eines Selbst-Modells und multipler Realisierbarkeit von Willensakten gültig bleiben.

5.3.2 Vorhersehbarkeit durch ein Selbst-Modell

Abschnitt 5.2.5 schloss mit dem Fazit, dass als eigen erlebte Urheberschaft zwar nicht notwendigerweise als willensfrei erlebt wird, wenn die materiell wirksamen Zustände nicht Teil eines repräsentationalen Bewusstseins sind, sonst aber typischerweise schon. Weiter haben wir im Abschnitt 5.3.1 gesehen, dass eine auf Naturgesetzen basierende Erklärung von Erleben eigener Urheberschaft systematische Lücken aufweist und in Gegenwart von durch ein Selbst-Modell gegebenen Zusammenhängen zwischen multipel realisierbaren Zuständen angreifbar wird.

Es gilt also im Folgenden, das Verhältnis von Willensfreiheit und präreflektivem Erleben eigener Ursächlichkeit unter Einbezug eines Selbst-Modells und multipler Realisierbarkeit von dem Subjekt bewussten repräsentationalen Zuständen aussagenlogisch noch weiter zu klären. Diese Klärung geht, analog zu Abschnitt 5.2.1, von den beiden Grundpfeilern des präreflektiven Erlebens eigener Ursächlichkeit, also eigener Kausalität und Vorhersehbarkeit der Folgen eigener Willensakten aus. Sie stützt sich deshalb ebenfalls auf das bestehende Argument 5.1 ("Physische Realisierung des Willensakts") und 5.2 ("Gesetzmässige Beziehungen sind notwendig für Erleben von Urheberschaft"). Weiter kann ich das Selbst-Modell als regelhafte Brücke zwischen repräsentationalen mentalen Zuständen und physischen Zuständen einführen:

Prämisse 5.3.1: Repräsentation von allgemein gültigen Gesetzmässigkeiten im Selbst-Modell

Es gibt allgemein gültige und eindeutige raum-zeitliche Folgebeziehungen zwischen Willensakten und physischen Ereignissen, die dem Subjekt durch ein Selbst-Modell gegeben sind.

Aus dieser Ausgangslage lässt sich die für das Erleben eigener Urheberschaft notwendige Vorhersagefähigkeit statt durch physische Gesetzmässigkeiten durch die Gültigkeit eines Selbst-Modells erklären:

(1) K₂: Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn ihm ein subjektiv erlebter eigener Willensakts und die gesetzmässige Gültigkeit von bestimmten raum-zeitlichen Folgebeziehungen dieses Ereignis als logische Folge vorhersagen.

Präreflektives Erleben von Urheberschaft Vorhersehbarkeit durch Gesetzmässigkeiten

(2) Es gibt allgemein gültige und eindeutige raumzeitliche Folgebeziehungen zwischen Willensakten und physischen Ereignissen, die dem Subjekt durch ein Selbst-Modell gegeben sind.

Repräsentation von allgemein gültigen Gesetzmässigkeiten im Selbst-Modell

(K₇) Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn ihm ein subjektiv erlebter eigener Willensakts und die gesetzmässige Gültigkeit eines Selbst-Modells dieses Ereignis als logische Folge vorhersagen. Präreflektives Erleben von Urheberschaft Vorhersehbarkeit durch Gesetzmässigkeiten Repräsentation von allgemein gültigen Gesetzmässigkeiten im Selbst-Modell

Argument 5.7: Vorhersehbarkeit durch ein Selbst-Modell

Diese in Argument 5.7 gewonnene Konklusion K₇ ist interessant, weil sie es einerseits ermöglicht, viele der in Abschnitt 5.3.1 formulierten Probleme des physikalistischen Zugangs zu vermeiden.²⁰ Ich will darauf Punkt für Punkt eingehen, wobei ich auf die oben formulierten Probleme explizit Bezug nehme:

- Das generelle Fehlen der Qualität von Willensfreiheit bei eigener Urheberschaft ist kontraintuitiv: In Argument 5.7 wird Erlebens von Urheberschaft durch repräsentationale Zustände und nicht durch physische Zustände begründet. Die für Erleben von Urheberschaft relevanten repräsentationalen Zustände werden aber, wie wir in Kapitel 3 erarbeitet haben, begründbarerweise mit den für Willensfreiheit notwendigen Eigenschaften erlebt. Das vermeidet vorerst den kontraintuitiven Schluss, dass sich Erleben von eigener Urheberschaft und Willensfreiheit gegenseitig ausschliessen.
- Fehlendes Mein-Sein: Das Mein-Sein des Erlebens von Urheberschaft ist gegeben, indem das Subjekt, wie in Abschnitt 4.1 hergeleitet, die Dinge durch ein Selbst-Modell hindurch in einer ihm eigenen Qualität erlebt.
- Problem der Erklärungslücke: Die regelhafte Verbindung von mentalen und physischen Zuständen, wie sie für Erleben von Urheberschaft nötig ist,

²⁰Aus dem Argument 5.7 folgt auch, dass Fehler im Selbst-Modell zu dysfunktionalem Erleben von Urheberschaft führen. Entsprechenden empirischen Befunden dazu sind wir in Abschnitten 3.5 und 4.4.1 bereits begegnet. Die Konklusion K₇ erfährt damit eine empirische Bestätigung.

bleibt zwar objektiv immer noch unerklärbar. Subjektiv ist sie in der Form eines in seiner Gültigkeit empirisch überprüfbaren Selbst-Modells dem Subjekt aber direkt zugänglich und unerlässliche Basis seines Erlebens. Damit muss, und kann sich, Erleben von Urheberschaft innerhalb der Gültigkeit dieses Selbst-Modells auf solche regelhaften Zusammenhänge zwischen mentalen und physischen Ereignissen stützen.

• Multiple Realisierbarkeit repräsentationaler Zusammenhänge:

Selbst-Modelle scheinen grundsätzlich in der Lage zu sein, Zusammenhänge zwischen multipel realisierbaren Zuständen abzubilden. Im Gegensatz zu physischen Gesetzmässigkeiten sind sie damit geeignet, regelhafte Zusammenhänge zwischen Paaren von jeweils multipel realisierbaren Zuständen abzubilden, wie sie bestehen zwischen repräsentationalen Zustände des Wollens und Zuständen, die dem Eintreffen des Gewollten entsprechen. Diese Abbildungsbeziehung basiert, wie im Abschnitt 4.4 ausführlich dargestellt, nicht auf Gesetzmässigkeiten, sondern auf Erfahrung und ist auch durch Erfahrung modifizierbar. Ich werde diesen Aspekt der sogenannten realisierungsunabhängigen Abhängigkeitsverhältnisse im Abschnitt 5.4.1 vertiefen.

Gleichzeitig ergeben sich eine Reihe von Problemen, die mir für die weitere Arbeit bedeutsam scheinen: Erstens stellt sich die Frage nach der Kausalität repräsentationaler Zustände. Es ist nämlich vorerst unklar, ob wir die Kausalität des eigenen Willensakts auf der Realisierungs- oder auf der Repräsentationsebene festmachen sollen: Einerseits ist jedes Token eines Willensaktes qua seiner Realisierung physisch wirksam, was dafür spricht, auch dessen Kausalität in dieser Realisierungsebene zu verorten. Umgekehrt ist es so, dass ein Willensakt, egal in welcher Realisierung er auftritt, durch sein Auftreten einen Unterschied macht und somit in einem kontrafaktischen Verständnis von Kausalität ein guter Kandidat für eine Ursache ist. Ich werde auf diese Frage der Relevanz von Kausalerklärungen auf diesen beiden Ebenen im Abschnitt 5.4 noch zurückkommen.

Zweitens, und verwandt mit dem obigen Problem, drängt sich die Frage auf, inwiefern, wenn wir Erleben von eigener Urheberschaft statt in einer physischen in einer repräsentationalen Weise begründen, diese eigene Urheberschaft bloss eine Repräsentation eines anderweitig begründeten Kausalgeschehens ist. Die Annahme, dass eigene Urheberschaft real sei, wäre damit, um Thomas Metzingers Begrifflichkeit wieder aufzunehmen, naiv. Diese Feststellung scheint mir auch darum naheliegend, weil wir tatsächlich auch in sehr unrealen Kontexten wie in Träumen eigene Urheberschaft erfahren können. Erleben eigener Urheberschaft wäre zwar noch nicht notwendigerweise falsch, aber wir können uns, weil wir uns nur auf unsere eigenen repräsentationalen Zustände stützen können, ihrer Realität nicht sicher sein. Erleben eigener Urheberschaft qua Selbst-Modell wäre damit in diesem Sinn prekär.

Zusammenfassend macht die Einführung des Selbst-Modells als Grundlage für unser Erleben von Urheberschaft dessen Erklärung also intuitiver, vermeidet eine Reihe von Problemen und wird teilweise auch durch empirische Befunde gestützt. Gleichzeitig öffnet sich die Frage nach der kausalen Relevanz von repräsentationalen Zuständen, und wir verlieren mit dem Wechsel der Erklärungsbasis von physischen zu repräsentationalen Zuständen die Sicherheit, dass das, was wir als eigene Urheberschaft erleben, notwendigerweise real ist.

5.3.3 Multiple Realisierbarkeit repräsentationaler Zustände und Determinismus

Der vorhergehende Abschnitt hat dargestellt, dass sich eine Reihe von Problemen und kontraintuitiven Schlüssen vermeiden lässt, wenn die für das Erleben eigener Urheberschaft notwendige Fähigkeit zur Vorhersage durch ein gültiges Selbst-Modell und nicht durch die Gültigkeit physischer Gesetzmässigkeiten erklärt wird. Dieses Selbst-Modell lässt es explizit zu, dass die für eine Vorhersage ausschlaggebenden mentalen Zustände multipel realisierbar sind. Die Annahme, dass Erleben eigener Urheberschaft auf multipel realisierbaren repräsentationalen Zuständen beruht, hat aber auch Folgen für das Verhältnis von Erleben eigener Urheberschaft zu Determinismus. Darum wird es in den nächsten Argumenten gehen. Zunächst führe ich dazu multiple Realisierbarkeit als weitere Prämisse ein:

Prämisse 5.3.2: Multiple Realisierbarkeit mentaler Zustände

Subjektives Erleben, das einen gegebenen Inhalt repräsentiert, ist multipel realisierbar.

Damit sind die für Erleben von Urheberschaft notwendigen Willensakte multipel realisierbar, was das folgende Argument abbildet:

Präreflektives Erleben von Urheberschaft (1)Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn ihm ein subjektiv erlebter eigener Willensakts das Eintreffen dieses Ereignisses als Wirkung zuverlässig vorhersagt. Multiple (2)Subjektives Erleben, das einen gegebenen Inhalt re-Realisierbarkeit mentaler Zustände präsentiert, ist multipel realisierbar. Physikalische Geschlossenheit der (K_8) Das Subjekt erlebt sich selbst als Ursache eines phy-Welt Präreflektives Erleben sischen Ereignisses, wenn ihm ein subjektiv erlebon Urheberschaft Multiple ter und multipel realisierbarer eigener Willensakt das Realisierbarkeit mentaler Zustände Eintreffen dieses Ereignisses als Wirkung zuverlässig vorhersagt und wenn die Realisation des Willensaktes auch eine physische Ursache dieses Ereignisses ist.

Argument 5.8: Willensakte als multipel realisierbare Ursachen

Weiter sollen die aus Argument 5.8 erhaltenen alternativen Formen eines Willensakts physisch wirksam sein. Dazu dient wiederum die Realisierungsbeziehung aus dem Argument 5.1. Damit lässt sich folgendes Argument zur multiplen Realisierbarkeit von Willensakten machen:

(1) K₁: Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn für dieses Ereignis eine physische Ursache vorhanden ist (die den subjektiv erlebten eigenen Willensakts realisiert) und wenn das Erleben dieses Willensakts das Eintreffen dieses Ereignisses zuverlässig vorhersagt.

Physikalische Geschlossenheit der Welt Präreflektives Erleben von Urheberschaft

(2) K₈: Das Subjekt erlebt sich selbst als Ursache eines physischen Ereignisses, wenn ihm ein subjektiv erlebter und multipel realisierbarer eigener Willensakt das Eintreffen dieses Ereignisses als Wirkung zuverlässig vorhersagt und wenn die Realisation des Willensaktes auch eine physische Ursache dieses Ereignisses ist.

Physikalische Geschlossenheit der Welt Präreflektives Erleben von Urheberschaft Multiple Realisierbarkeit mentaler Zustände

(K₉) Das Subjekt erlebt sich selbst als Ursache eines physischen Ereignisses, wenn ihm ein subjektiv erlebter und multipel realisierbarer eigener Willensakt das Eintreffen dieses Ereignisses als Wirkung zuverlässig vorhersagt und wenn die Realisation des Willensaktes auch eine physische Ursache dieses Ereignisses ist.

Physikalische Geschlossenheit der Welt Präreflektives Erleben von Urheberschaft Multiple Realisierbarkeit mentaler Zustände

Argument 5.9: Alternative Realisierbarkeit eigener Willensakte

Zuletzt können wir aus dieser multiplen Realisierbarkeit von Willensakten und der Epistemik von Erleben eigener Urheberschaft qua Selbst-Modell eine Aussage zur multiplen Realisierbarkeit eigener Urheberschaft bilden:

(1) K₇: Das Subjekt erlebt sich als Ursache eines physischen Ereignisses, wenn ihm ein subjektiv erlebter eigener Willensakts und die gesetzmässige Gültigkeit eines Selbst-Modells dieses Ereignis als logische Folge vorhersagen.

Präreflektives Erleben von Urheberschaft Vorhersehbarkeit durch Gesetzmässigkeiten Repräsentation von allgemein gültigen Gesetzmässigkeiten im Selbst-Modell

(2) K₉: Das Subjekt erlebt sich selbst als Ursache eines physischen Ereignisses, wenn ihm ein subjektiv erlebter und multipel realisierbarer eigener Willensakt das Eintreffen dieses Ereignisses als Wirkung zuverlässig vorhersagt und wenn die Realisation des Willensaktes auch eine physische Ursache dieses Ereignisses ist.

Physikalische Geschlossenheit der Welt Präreflektives Erleben von Urheberschaft Multiple Realisierbarkeit mentaler Zustände

(K₁₀) Das Subjekt erlebt sich selbst als Ursache eines physischen Ereignisses, wenn ihm ein subjektiv erlebter und multipel realisierbarer eigener Willensakt und die gesetzmässige Gültigkeit eines Selbst-Modells das Eintreffen dieses Ereignisses als logische Folge vorhersagen und wenn die Realisation des Willensaktes auch eine physische Ursache dieses Ereignisses ist.

Physikalische
Geschlossenheit der
Präreflektives Erleben
von Urheberschaft
Vorhersehbarkeit durch
Gesetzmässigkeiten
Repräsentation von
allgemein gültigen
Gesetzmässigkeiten im
Selbst-Modell
Multiple
Realisierbarkeit
mentaler Zustände

Argument 5.10: Eigene Urheberschaft ist multipel realisierbar

Das in Argument 5.10 gefundene Fazit, dessen Herleitung in Abbildung 5.3 grafisch dargestellt ist, impliziert also, ausgehend von der Gültigkeit eines Selbst-Modells und der multiplen Realisierbarkeit mentaler Zustände, dass das, was dem Subjekt als eigene Ursächlichkeit qua Willensakt erfahrbar ist, physisch multipel realisierbar ist und damit nicht eindeutig definiert sein muss. Das Fazit stützt sich auf die Prämissen "Physikalische Geschlossenheit der Welt", "Präreflektives Erleben von Urheberschaft", "Gesetzmässige Beziehungen sind notwendig für Erleben von Urheberschaft", "Repräsentation von allgemein gültigen Gesetzmässigkeiten im Selbst-Modell" und "Multiple Realisierbarkeit mentaler Zustände", nicht aber auf die Prämisse "Physikalischer Determinismus". Es widerspricht damit einerseits Erklärungen von Ursächlichkeit mentaler Zustände, wie sie durch zum Beispiel substanzdualistische Thesen gegeben werden, weil diese die Prämisse "Physikalische Geschlossenheit der Welt" verletzen. Gleichzeitig ist das Fazit gegenüber der Prämisse "Physikalischer Determinismus" indifferent: Das Fazit 5.10 ist unabhängig davon gültig, ob die Prämisse "Physikalischer Determinismus" wahr ist oder falsch. Die physische Unbestimmtheit dessen, was dem Subjekt als eigene Ursächlichkeit erfahrbar ist und was die relevanten Qualitäten von Willensfreiheit enthält, ist damit mit einem physikalischen Determinismus kompatibel.

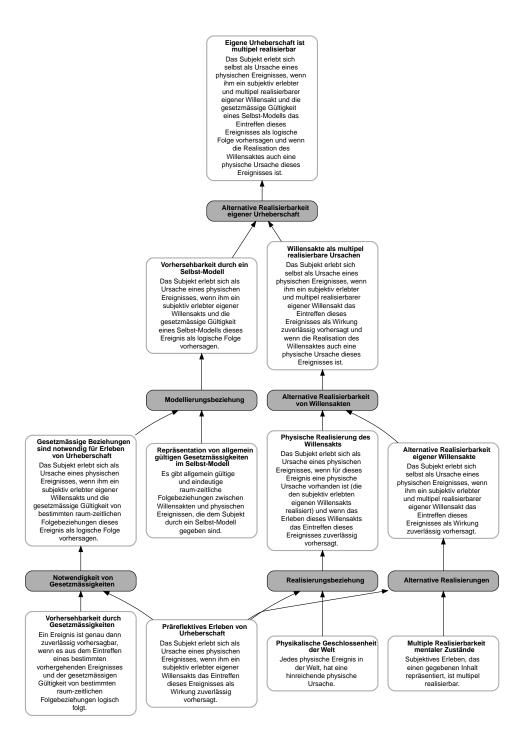


Abb. 5.3: Argumentationskarte für die Möglichkeit von Willensfreiheit, wenn die für Erleben von Urheberschaft notwendigen Gesetzmässigkeiten durch ein Selbst-Modell gegeben sind.

5.3.4 Folgen der multiplen Realisierbarkeit von eigener Urheberschaft

Was beinhaltet Willensfreiheit, wenn sie in der hier vorliegenden Weise begründet werden soll? Gegeben ein Verständnis von Willensfreiheit, das sich nur durch ein Selbst-Modell und damit immer nur über repräsentationale Leistungen begründen kann, ist das Subjekt in dem Sinn willensfrei, als dass es nicht nötig ist, dass ein spezifisches physisches Ereignis eintreten muss, um zu erleben, dass es das tut, was es will. Vielmehr ist es das Eintreffen eines mentalen Zustands des Wollens, der für das entsprechende Erleben von Urheberschaft ausschlaggebend ist, wobei es nicht relevant ist, in welcher Weise dieser mentale Zustand realisiert wird, solange es in einer physisch wirksamen Form ist. Willensfreiheit, solchermassen durch Erleben von Urheberschaft begründet, widerspricht aber der Auffassung, dass Willensfreiheit bedeutet, in genau der gleichen Situation etwas anderes tun zu können. Diese Auffassung wäre nämlich für das Erleben eigener Urheberschaft deswegen problematisch, weil es für Subjekte, die sich in identischen Situationen unterschiedlich verhalten würden, unmöglich wäre, über komplexere und zeitlich extendierte Sequenzen von zielgerichteten Handlungen ihr eigenen Tuns vorherzusagen und sie damit in solchen Kontexten ihr Erleben eigener Urheberschaft unterminieren würden.

Das gefundene Fazit assoziiert mit dem Begriff der Willensfreiheit also primär das Erleben, trotz unterschiedlicher Situationen das Gleiche wollen und entsprechend handeln zu können, auch wenn dieses Wollen und Handeln immer wieder anders realisiert werden kann. Ich werde das, was diesen über einen Zeitraum verteilten verschiedenen Realisationen von Willensakten gemeinsam ist, als "Gründe" des Subjekts bezeichnen. Es sind diese subjektiven Gründe, die es uns erlauben, zu einem gewissen Grad unabhängig von äusseren Umständen auch ein längerfristiges Ziel zu erreichen und die uns, um eine Sichtweise von Harry Frankfurt aufzunehmen, erst zur Person machen (Frankfurt 1988).

Ich werde auf diese Punkte im abschliessenden Fazit noch einmal zurückkommen. Bevor ich aber dazu kommen kann, bedarf es einer weiteren Klärung der kausalen Relevanz der durch physische Gesetzmässigkeiten oder durch ein Selbst-Modell vermittelten eigenen Ursächlichkeit. Der Erklärungsweg über physische Gesetzmässigkeiten ist problematisch, weil er keinen Bezug zum Erleben des Subjekts herstellt und den üblichen Intuitionen zu Willensfreiheit widerspricht. Umgekehrt spricht der Erklärungsweg über ein Selbst-Modell nur von Erleben von Urheberschaft, und nicht von Ursächlichkeit selbst. Damit fehlt in diesem Erklärungsweg noch ein Schritt vom Erleben, für das, wie wir bereits im Unterkapitel 4.5 festgestellt haben, deduktiv-nomologische Beziehungen hinreichen mögen, zu ursächlicher Relevanz. Dieser Schritt wird aber im nächsten Abschnitt folgen.

5.4 Kausale Relevanz von repräsentationalen und physischen Zuständen

In Abschnitt 5.2 habe ich untersucht, wie sich Determinismus zur physischen Gegebenheit von eigener Urheberschaft verhält. Dabei haben wir festgestellt, dass auf der Ebene der Realisierung von mentalen Zuständen ein Widerspruch besteht zwischen dem Erleben von Urheberschaft und dem Vorhandensein von Wahlmöglichkeiten, Willensfreiheit, Selbstbestimmung und Kontrolle. Es scheint so, dass es nur dort und dann, wo ein angenommener Determinismus gültig ist, dem Subjekt möglich ist, die für Erleben von Urheberschaft notwendigen erfolgreichen Vorhersagen von eigenen Handlungen zu machen. Gleichzeitig scheint es problematischerweise auch so zu sein, dass das Subjekt überall dort und dann, wo ein solcher Determinismus gilt, keine selbst kontrollierbaren und frei bestimmbaren Handlungsalternativen hat.

In Abschnitt 5.3 habe ich festgestellt, dass ein funktionales Selbst-Modell ebenfalls und alternativ zu physischen Gesetzmässigkeiten in der Lage ist, dem Subjekt die für Erleben von Urheberschaft notwendigen Fähigkeit zu erfolgreichen Voraussagen von Handlungsereignissen zu geben. Weil aber die mentalen Zustände des Wollens, die diese vorhersagbaren Handlungen verursachen, multipel realisierbar sind, ist es in dieser Erzählweise der kausalen Zusammenhänge nicht nötig, dass ein ganz bestimmtes externes Ereignis eintritt, dem die vorgesagte Handlung folgt. Es reicht, wenn irgendeines der möglichen Ereignisse eintritt, das den über das Selbst-Modell mit der Handlung assoziierten mentalen Zustand realisiert. Es besteht also eine gewisse, aber keine vollständige Unabhängigkeit zwischen dem, was das Subjekt als eigene Urheberschaft erlebt, und physischen Ereignissen.

Die beiden oben hergeleiteten Arten, eine kausale "Geschichte" einer Handlung zu erzählen, unterscheiden sich offensichtlich in der Ebene, auf der sie stattfindet. In der ersten der Erzählweisen ist diese Ebene materiell und beinhaltet primär die physischen Realisationen der mit einer Handlung assoziierten mentalen Zustände. Die zweite Erzählweise spielt vorwiegend in der mentalen Ebene und bezieht sich primär auf multipel realisierbare, durch die physischen Zustände realisierten mentale Eigenschaften. Beide Arten der Erzählweise hängen miteinander zusammen, indem die repräsentationalen mentalen Zustände auf physisch realisierte Zustände supervenieren. Gleichzeitig, und für die gegebene Fragestellung zentral, unterscheiden sich die beiden Erzählweisen in dem, was sie für die Möglichkeit von Willensfreiheit bedeuten. Während in der ersten, materiellen Erzählweise Erleben von Urheberschaft mit Willensfreiheit inkompatibel ist, liefert die Erzählweise auf der supervenierenden Ebene Erklärungen für subjektive Überzeugungen, dass unser willentliches Handeln frei wählbar und selbstbestimmt ist.

Indem wir diese beiden Erzählweisen kritisch gegenüberstellen und auf ihre Relevanz prüfen, kommen wir also vielleicht zu interessanten Aussagen zu Willensfreiheit. Darum werde ich mich im Folgenden bemühen und mich dabei auf bestehende Argumente zur Kausalität in Supervenienzbeziehungen stützen. Daraus wird sich ergeben, dass uns Überlegungen, was eine "gute" Kausalerklärung ausmacht, Gründe geben, die mentale, supervenierende Erzählweise zu bevorzugen, und darauf basierend übliche Überzeugungen zu Willensfreiheit zu stützen.

5.4.1 Kausalerklärungen und realisierungsunabhängige Abhängigkeitsverhältnisse

Die Gründe, die ich für die Bevorzugung der mentalen Erzählweise anführen möchte, ergeben sich aus der Frage, was wir damit meinen, wenn wir von "Verursachung" sprechen. Verursachung ist nach David Lewis (Lewis 1986, S. 632) etwas, was durch seine Anwesenheit einen Unterschied macht. Verursachung ist damit, wie auch schon im Abschnitt 4.5 dargestellt, in seiner logischen Struktur etwas anderes als eine deduktiv gültige Folgebeziehung (Lewis 1986, S. 636). Im Kontext von Supervenienzbeziehungen hat das relevante Folgen für die Frage, ob wir Verursachung auf der Ebene der Realisation mentaler Zustände oder auf der Ebene der supervenierenden mentalen Eigenschaften verorten sollen. James Woodward illustriert das am Beispiel einer Taube, die darauf trainiert wurde, auf einen roten Knopf zu picken (Woodward 2008). In einem gegebenen Fall ist der Knopf scharlachrot. Damit superveniert die Eigenschaft des Knopfes, rot zu sein, auf das Scharlachrotsein des Knopfes. Woodwards stellt nun fest, dass sich bei der Frage, ob (a) das Scharlachrotsein des Knopfes als Realisation des Rotseins oder (b) das Rotsein des Knopfes als supervenierende Eigenschaft des Scharlachrotseins das Picken der Taube verursachen, die Antworten signifikant voneinander unterscheiden. Woodward argumentiert nämlich, dass es im gegebenen Fall die supervenierende Eigenschaft des Rotseins ist, die verursacht, dass die Taube auf den Knopf pickt, weil es das Rotsein ist, das einen Unterschied macht gegenüber dem Fall, in dem der Knopf nicht rot ist. Das Gleiche trifft nun auf der Ebene der Realisation des Rotseins nicht notwendigerweise zu: Das Verhalten der Taube würde sich nicht ändern, wenn das Scharlachrotsein des Knopfes mit einer anderen Realisation der Eigenschaft Rotsein kontrastiert wird, indem zum Beispiel der Knopf nun himbeerrot wäre. Es ist damit die supervenierende Eigenschaft des Rotseins, im Kontrast zu einem Nicht-Rotsein, und nicht dessen Realisation als Scharlachrotsein, im Kontrast zu einem Nicht-Scharlachrotsein, die das Picken der Taube verursacht, während das Scharlachrotsein des Knopfes in Woodwards Argumentation eine unbrauchbare Kausalerklärung ("defective from the point of view of explanation", Woodward (2008, S. 234)) darstellt.

Das Vorhandensein einer Supervenienzbeziehung impliziert nicht notwendigerweise, dass damit assoziierte kausale Beziehungen auf der supervenierenden Ebene verstanden werden müssen. Damit es die supervenierende Ebene ist, die die relevanten Kausalbeziehungen darstellt, braucht es gemäss Woodward sogenannte realisierungsunabhängige Abhängigkeitsverhältnisse (realisation independent dependency relations, RIDR, Woodward (2008, S. 241)): Nur wenn solche realisierungsunabhängigen Abhängigkeitsverhältnisse (wie das Irgendein-Rotsein des Knopfes für das Picken des Knopfes) bestehen, liefert die Ebene der supervenierenden Eigenschaften, und nicht die Ebene der Realisierungen diejenigen kontrafaktischen Kontraste, die für die Beurteilung des Bestehens von ursächlichen Zusammenhänge die richtigen Antworten liefern. Dem ist so, weil es das Vorhandensein der supervenierenden Eigenschaften, nicht aber das Vorhandensein bestimmter Realisationen dieser Eigenschaften sind, die in der kontrafaktischen Analyse notwendigerweise einen Unterschied machen.

Gleichzeitig trifft nun das klassische Irrelevanzproblem von nomologisch hinreichenden Bedingungen als Kausalerklärungen auch auf die von Woodward als unbrauchbar bezeichnete kausale Beschreibung der Ereignisse auf der Ebene der Realisierung zu. Ursprünglich sind wir diesem Problem in Abschnitt 4.5 begegnet, als sich die Frage stellte, ob Erklärungen eigener Ursächlichkeit qua Selbst-Modell und mentalem Akt ein Epiphänomen sein könnten. Das Irrelevanzproblem besagt, kurz rekapituliert, dass nomologisch hinreichende Bedingungen irrelevante Elemente enthalten können, die dann fälschlicherweise als Teil der Kausalerklärung erscheinen. Wie bei Wesley Salmons Beispiel von Jones, der die Pille nimmt und nicht schwanger wird, können solche nomologisch hinreichende Beziehungen also Elemente enthalten, die keinen Unterschied machen und die deshalb kausal irrelevant sind: Um wieder zu Woodwards Beispiel zurückzukommen: Während das Scharlachrotsein des Knopfes natürlich eine nomologisch hinreichende Bedingung für das Picken der Taube ist, ist die Art des Rotseins für die Kausalerklärung irrelevant. Woodwards Argument impliziert damit, dass kausale Beziehungen, die sinnvollerweise als realisierungsunabhängige Abhängigkeitsverhältnisse dargestellt werden, auf Ereignisse supervenieren, deren spezifische Eigenschaften in Teilen irrelevant sind.

Im Kontext der Frage nach den relevanten Kausalerklärungen für subjektiv erlebte Urheberschaft scheint es mir, dass bei diesem Erleben solche realisierungsunabhängigen Abhängigkeitsverhältnisse in Form eines funktionalen Selbst-Modells zwingend gegeben sind. Wenn wir weiter auf die in Abschnitt 4.4 gemachte Feststellung zurückkommen, dass Erleben von Urheberschaft notwendigerweise immer innerhalb des Gültigkeitsbereichs eines solchen funktionalen Selbst-Modells stattfindet, und dass konkretes Erleben von Urheberschaft damit die Gültigkeit einzelner Bereiche des Selbst-Modells immer wieder empirisch validiert, folgt daraus,

dass für alles aus einem gültigen Selbst-Modell folgende Erleben eigener Ursächlichkeit solche realisierungsunabhängigen Abhängigkeitsverhältnisse zwischen dem subjektiven Erleben und dem verursachten Ereignis bestehen. Insofern sich also das Subjekt, wie in Abschnitt 5.3 dargestellt, immer nur durch dieses Selbst-Modell als Ursache der Handlung erfährt, und insofern damit die Kausalität qua realisierungsunabhängigen Abhängigkeitsverhältnisse sinnvollerweise auf der supervenierenden Ebene beschrieben werden muss, sind es dann die mentalen und nicht die physischen Ereignisse, die die für den kontrafaktischen Test informativen Unterschiede machen und damit kausal relevant sind.

Es sind also, wenn wir Woodwards Argumentation hier folgen und wenn wir zusätzlich annehmen, dass die für sein Argument notwendigen realisierungsunabhängigen Abhängigkeitsverhältnisse in Form eines gültigen Selbst-Modells hinreichend gegeben sind, die mentalen Zustände, die die relevanten Unterschiede machen und die damit als die relevanten Kausalerklärungen verstanden werden sollen. Gleichzeitig, so mein Vorschlag, ist das, was wir als "Unterschied machen" von bewussten mentalen Zuständen verstehen, im Erleben des Subjekts etwas, was als "Wahl" erlebt wird. Damit ist, so schliesst die Argumentation hier, ein willentliches Tun des Subjekts immer eine bewusste Wahl dieses Tuns, das als Gewähltes einen Unterschied macht und damit eine Alternative darstellt zum Nicht-Gewählten.

5.4.2 Verursachung als Eigenschaft zweiter Ordnung

Man kann nun einwenden, dass die These, Determinismus sei nicht mit Willensfreiheit kompatibel, gar keinen Bezug nimmt zu einem Vorhandensein von Kausalzusammenhängen, sondern nur auf nomologisch hinreichenden Bedingungen aufbaut. Weil diese inkompatibilistische These also gar keine Kausalzusammenhänge beinhaltet, spielt es dann, so kann man weiter einwenden, auch keine Rolle mehr, welche der verschiedenen Erklärungsebenen "relevante" kontrafaktische Aussagen liefert und ob diese "relevante" Ebene mit Willensfreiheit kompatibel ist. Umgekehrt scheint es möglich, dass Kausalzusammenhänge selbst auf nomologische Beziehungen supervenieren (Lewis 1994, S. 479): Gegeben physikalische Geschlossenheit, wäre das, was eine kontrafaktische Analyse darstellt, dann einfach ein Herausgreifen eines bestimmten "lokalen" Ausschnitts aus einem globalen Netz von nomologischen und damit deterministischen Abhängigkeiten und Anfangsbedingungen. Das impliziert umgekehrt nicht, dass die aus einer solchen kontrafaktischen Analyse folgenden Aussagen unabhängig sind von einem solchen globalen und deterministischen Netzwerk nomologischer Beziehungen.

Zusätzlich sind die Voraussetzungen für Erleben von Urheberschaft stärker als das reine Vorhandensein von kausalen Zusammenhängen zwischen einzelnen Instanzen von Ereignissen: Neben einer korrekten Darstellung der kontrafaktischen Verhältnisse ist Erleben von Urheberschaft auf systematische Vorhersehbarkeit von

Ereignissen angewiesen.²¹ Das bedingt, wie in Abschnitt 5.1.3 hergeleitet, zusätzlich die Gültigkeit bestimmter gesetzmässiger Zusammenhänge zwischen verschiedenen Typen von Ereignissen.

Diese inkompatibilistische Position kann sich also darauf berufen, dass der Begriff der Kausalität für die Erklärbarkeit des Weltgeschehens möglicherweise überflüssig ist, und findet dabei auch Unterstützung von prominenter Seite (z.B. Russell (1912)). Wenn sich nun gleichzeitig affirmative Aussagen zu Willensfreiheit auf das Vorhandensein von Kausalität stützen, unterminiert diese Verzichtbarkeit auf den Begriff der Kausalität auch solche affirmativen Aussagen zu Willensfreiheit, unabhängig davon, auf welcher Ebene die "relevanten Kausalbeziehungen" verortet werden. Gleichzeitig kann sich die These, dass sich Überzeugungen zu Willensfreiheit durch Erleben von Urheberschaft begründen lässt, grundsätzlich nicht vom Vorhandensein von nomologischen Beziehungen verabschieden, weil gewisse nomologischen Beziehungen den Kausalzusammenhängen selbst und damit auch der für Erleben von Urheberschaft notwendigen Vorhersehbarkeit zugrunde liegen.

Wenn wir diesen Einwand gelten lassen, impliziert das, dass die Annahme von Kausalität nicht notwendigerweise eine These über Möglichkeiten des Weltgeschehens ist, sondern erst einmal eine Sichtweise auf diese Welt von einem Subjekt, das nun einfach nur kein Laplacescher Dämon²² ist. In diesem Sinne müssten wir folglich eine mit dem Kausalbegriff operierende kontrafaktische Herangehensweise an das Weltgeschehen erst einmal als einen an bestimmten Typ von Subjekten gebunden Prozess verstehen, die als Akteure bestimmte epistemische Dispositionen bereits mitbringen. Das ist eine These, die in der Diskussion um Kausalität auch vertreten wird. Ein klassischer Text, der eine solche, sogenannt Akteur-basierte Definition von Kausalität ("agency accounts of causation") vertritt, ist die Arbeit "Causation as a Secondary Quality" von Menzies und Price (1993). In der Arbeit definieren die Autoren Kausalität folgendermassen:

"... an event A is a cause of a distinct event B just in case bringing about the occurrence of A would be an effective means by which a free agent could bring about the occurrence of B (Menzies und Price 1993, S.187)."²³

²¹Es gibt Fälle, in denen ein Subjekt kontrafaktisch betrachtet Ursache eines Ereignisses ist, ohne dabei eigene Urheberschaft zu erleben, weil es das verursachte Ereignisses aus dem eigenen Wollen nicht oder unzutreffend vorhersagen konnte (wenn es zum Beispiel die Kontrolle über das Velo verliert), oder das Ereignis nicht oder fehlerhaft wahrnimmt.

²²Der Begriff des Laplaceschen Dämons kommt aus einem Gedankenexperiment von Pierre-Simon Laplace, das besagt, dass ein genügend intelligenter Dämon, der zu einem gegebenen Augenblick alle Naturgesetze der Welt und die Lage aller Dinge in der Welt kennen würde, daraus auf die gesamte Vergangenheit und die gesamte Zukunft der Welt schliessen könnte (Laplace 1816, S. 2).

Menzies und Price begründen diese als interventionistisch bezeichnete Theorie von Kausalität interessanterweise, indem sie sich auf den Standpunkt stellen, dass Kausalität eine sogenannt sekundäre Qualität eines Sachverhalts sei. Sekundäre Qualitäten sind dadurch definiert, dass sie im Gegensatz zu sogenannt primären Qualitäten von bestimmten epistemischen Dispositionen der Beobachtenden abhängen. Sekundäre Qualitäten stehen damit, im Gegensatz zu primären Qualitäten, notwendigerweise in einer Relation zu einem bestimmten Typ von Subjekt, das zu gewissen Erfahrungen fähig sind. Es braucht, um beim Beispiel der Autoren zu bleiben, ein Subjekt mit der Fähigkeit des Farbsehens, das basierend auf dem Vorhandensein bestimmter elektromagnetischer Wellen die Aussage "Die Blume ist rot" machen kann. "Rotsein" ist damit eine sekundäre Qualität, während die Länge der elektromagnetischen Wellen, die von der Blume ausgehen, als primäre Qualität unabhängig ist von einem Subjekt mit der Fähigkeit, "Farbe" zu sehen. Kausalität wäre damit, in Menzies und Prices Verstehen, eine Qualität von Zusammenhängen zwischen Ereignissen, zu deren Wahrnehmung alle und nur Subjekte des Typs "freie*r Akteur*in" ("free agent") fähig sind. Umgekehrt zeichnen sich "freie Akteur*innen" dadurch aus, bestimmte Zusammenhänge zwischen Ereignissen als "kausal" zu erkennen.

Es gibt nun zwei Dinge, die ich für die weiteren Überlegungen hier festhalten will:

• Primäre und sekundäre Qualitäten unterscheiden sich für das Subjekt in fundamentaler Weise durch das, was sie für das Subjekt wahr macht: Der Wahrheitsgehalt von sekundären Qualitäten braucht für das Subjekt keine weitere Begründung, sondern ergibt sich direkt durch Anschauung (Menzies und Price 1993, S.194). Er ist der Person nur entweder direkt evident, oder es ist ihr, falls sie nicht vom richtigen Typ ist, grundsätzlich unmöglich, aufgrund von sekundären Eigenschaften eine Wahrheitsaussage zu machen. Das Wahrnehmen von Farbe liefert dafür wieder ein gutes Beispiel: Ein Mensch, der seit seiner Geburt blind ist, hat keinen Zugang zur Begründung eines Satzes wie "Die Blume ist rot." durch die Aussage, "Die Blume ist rot, weil ich sie als rot sehe." Gleichzeitig kann ein solcher Mensch basierend auf Wissen über die primären Qualitäten dessen, was ein zu Farbsehen fähiges Subjekt als "Farbe" erlebt, richtige Aussagen über das Verhalten von "Farben" herleiten (Jackson 1986). Sekundäre Qualitäten sind gegenüber primären Qualitäten also insofern ausgezeichnet, in dem sie einem entsprechenden Subjekt zuerst, direkt, und bereits mit einem starken Wahrheitsanspruch zugänglich sind. Subjekte, die vom Typ "freie*r Akteur*in" sind, sind damit in der Lage, be-

²³Diese Formulierung lässt zu, dass eine entsprechende kausale Struktur möglich ist, ohne das Akteure vorhanden sind.

stimmte Verhältnisse zwischen A und B direkt in einer Form zu erfassen, die besagt, dass sie B hervorrufen können, indem sie A tun.

• Es scheint wiederum so, dass sekundäre Qualitäten auf entsprechende primäre Qualitäten supervenieren. Primäre und sekundäre Qualitäten verweisen deshalb auf dieselben Sachverhalte, sind aber, um in Freges Terminologie zu bleiben, in ihrem Sinn für das wahrnehmende Subjekt unterschiedlich. Der Sinn der Aussage: "Die Blume ist rot." ist damit ein anderer als jener der Aussage: "Die Blume emittiert elektromagnetische Wellen mit einer Länge von ca. 750nm", obwohl beide Sätze auf denselben Sachverhalt verweisen und damit in Freges Terminologie dieselbe Bedeutung haben. Auch wenn analog Sachverhalte, die sich Subjekten vom Typ "freie*r Akteur*in" als Kausalzusammenhänge präsentieren, auf Sachverhalte supervenieren, die sich ohne Regress auf einen Kausalitätsbegriff darstellen lassen, ist der Sinn der supervenierenden Kausalaussage doch ein anderer als das, was sich aus Aussagen über den Sachverhalt auf der Ebene der Supervenienzbasis ergeben kann.

Sekundäre Qualitäten als eine besondere "Erlebensweise" von Sachverhalten durch ein mit den entsprechenden Fähigkeiten ausgestattetes Subjekt entsprechen damit genau dem, was ich in Kapitel 4 eingeführt habe; die subjektiv erlebte Gestalt von Sachverhalten, wie sie durch die repräsentationalen Eigenschaften des Selbst-Modells abgebildet werden. Gleichzeitig stehen sekundäre Qualitäten, als auf primäre Qualitäten supervenierende Eigenschaften, in einem notwendigen und gesetzmässigen Bezug zu diesen primären Qualitäten. Analog zu Gallaghers Vorschlag, dass die Gesamtheit von subjektivem Selbsterleben in eine sensorische ("sense of ownership") und eine motorische ("sense of agency") Domäne zerfällt, postulieren nun auch Menzies und Prices, dass wir das Vorhandensein von sekundären Qualitäten auch bei der Erklärung von Handlungen zulassen müssen. Kausalität wäre damit etwas, was sich definiert als ein besonderer Zugang eines zur Wahrnehmung von Kausalität fähigem Subjekts. Menzies und Price formulieren das folgendermassen:

"We want to show that in admitting action on a par with perception as a means of access to the world, it also has the resources to meet the various objections that have seemed to show that agency cannot, in principle, play a constitutive role in an account of causation. (Menzies und Price 1993, S.191-192)."

Erleben von Urheberschaft als freie*r Akteur*in wäre damit auf derselben Stufe ("on par with") wie Perzeption und unterscheidet sich von ihr nur in ihrer Wirkrichtung. Ich will im Weiteren nicht auf die Einwände eingehen, mit denen sich die Autoren in der Arbeit weiter beschäftigen, weil mich hier primär interessiert,

zu welchen Schlüssen wir kommen, wenn wir andere, nicht an freie Akteur*innen gebundene Verstehensweisen von Kausalität ablehnen. Denn auch falls für unser Verständnis von Kausalität nichts anderes bliebe als das, was uns Menzies und Prices Definition anbietet, dann leiten sich daraus eine Reihe von Dingen ab, die mir im gegebenen Kontext relevant scheinen:

- Wenn wir Kausalität als sekundäre Qualität verstehen, dann ist die Basis des Selbsterlebens einer/es freie*r Akteur*in ("as free agent" (Menzies und Price 1993, S.187)) als aus dem Subjekt in die Welt gerichtete Ursächlichkeit erst einmal nicht anders gegeben als die der aus der Welt in das Subjekt gerichtete Ursächlichkeit, die die Basis und die besondere Phänomenologie unserer bewussten Wahrnehmung ausmacht. Infolgedessen wären Erleben von Urheberschaft und das damit verbundene Erleben von Willensfreiheit in derselben wie Metzinger sagen würde naiven Weise "wahr", unwahr, wirklich, illusorisch oder ..., wie eine Blume rot ist, wie Zucker süss schmeckt oder wie ein Eiswürfel im Mund in der Stirn schmerzen kann. Insofern also das bewusste Erleben des Subjekts als Akteur*in die Qualität des Freiseins beinhaltet, wäre Willensfreiheit Teil eines "Gesamtpakets" von bewusstem Erleben, und liesse sich nicht einfach einzeln aus diesem Paket herauslösen. Wer Willensfreiheit im Kontext von akteur*innen-basierter Verursachung als illusorisch ablehnt, müsste damit auch alle anderen Formen von subjektivem Erleben illusorisch betrachten und wäre mit der Frage konfrontiert, was dieses subjektive Erleben denn überhaupt noch abbilden kann. Das scheint mir ein hoher Preis zu sein.
- Direktes Erleben von sekundären Qualitäten kommt oft mit einem immediaten und starken subjektiven Wahrheitsanspruch. Das erklärt, warum wir typischerweise intuitiv starke Überzeugungen zur eigenen Willensfreiheit haben. Gleichzeitig sind, wie zum Beispiel beim Eiscreme-Kopfweh, sekundäre Qualitäten als repräsentationale Qualitäten anfällig für Täuschungen. Infolgedessen ist es, kongruent mit der Alltagserfahrung, auch möglich, dass sich freie Akteur*innen in dem, was sie als Kausalität erfahren, täuschen.
- Menzies und Prices Analyse von Kausalität impliziert, dass Kausalität relationale Eigenschaften zu einem sich als Akteur*in erlebenden Subjekt hat. Die Notwendigkeit dieser relationalen Eigenschaften von Kausalbeziehungen zu als bewusstes Erleben zugänglichen mentalen Zuständen impliziert nun, dass interventionistisch verstandene Kausalzusammenhänge nicht einfach auf ihre Supervenienzbasis reduziert werden können, weil diese Supervenienzbasis solche relationalen Eigenschaften nicht abbilden kann. Dem ist so, wie schon in Abschnitt 4.2.3 und in einer eigenen Seminararbeit hergeleitet (Koenig 2020), weil sich materiell identische Sachverhalte im Kontext von

relationalen Eigenschaften wie "wissen, dass" oder "kennen als" unterscheiden können. Das Vorhandensein von mentalen Zuständen, mit denen interventionistisch verstandene Kausalität notwendigerweise in Beziehung steht, beinhaltet als leibliches Erleben eine Fülle solcher relationaler Eigenschaften, die sich in Formen wie "kennen als" oder "erfahren als" manifestieren (siehe auch Abschnitt 4.1). Eine interventionistisch verstandene Kausalität kann damit auf Sachverhalte supervenieren, ohne dass sie sich darauf reduzieren lässt, und sogar ohne dass das, worauf sie supervenieren, notwendigerweise selbst Kausalzusammenhänge sein müssen. Auch eine interventionistisch verstandene Kausalität sperrt sich also analog zu Woodwards Konklusionen zu mentaler Verursachung aus grundsätzlichen und allgegenwärtigen Gründen systematisch gegen eine Reduzierbarkeit auf ihre materielle Basis.

5.4.3 Willensfreie Personen und willenlose Zombies

Um dieses Kapitel abzuschliessen, finde ich es interessant, noch kurz in ein Thema zu gehen, dass die Philosophie des Geistes im Kontext von sekundären Qualitäten und bewusstem Erleben schon seit langem beschäftigt. Es geht um den Begriff des "philosophischen Zombies" (Chalmers 1996), einem hypothetischen sensomotorischen Automaten, der physisch von Wesen mit bewusstem Erleben ununterscheidbar ist und sich auch gleich verhält, dem aber jedes bewusste Erleben fehlt. Ob solche philosophischen Zombies möglich sind und was daraus folgt, ist Thema einer lebhaften philosophischen Debatte (z.B.Dennett (1995) und Lanier (1995)), soll uns aber hier nicht weiter beschäftigen. Was mir für die gegebene Fragestellung aber interessant scheint, sind die Intuitionen, die sich aus der Vorstellung solcher Zombies ergeben: Wenn die Laienmeinung zum Vorhandensein solcher deterministischen humanoiden Automaten untersucht wird, findet man nämlich, dass solchen Zombies Willensfreiheit üblicherweise abgesprochen wird (Björnsson und Shepherd 2020). Es gibt, um diese Intuition noch zu erweitern, aus der experimentellen Philosophie auch Evidenz, die besagt, dass in Laienüberzeugungen das Vorhandensein von Willensfreiheit an das Bestehen von Bewusstsein gebunden ist und dass Personen, die ihr Bewusstsein verlieren, auch keine Willensfreiheit mehr besitzen (Shepherd 2012; Shepherd 2015). Parallel dazu würden sich wohl auch die Begriffe ändern, mit denen das Wirken eines solchen Zombies erklärt würde: Während sein Wirken mechanistisch, also nicht anders als das eines überzeugend gebauten humanoiden Roboters zu erklären wäre, erklären wir unser eigenes Wirken durch Begriffe wie "Wünschen" und "Handeln", deren Bedeutung über die Mechanik der diesen Ereignissen zugrunde liegenden materiellen Vorgänge hinausgeht. Diese Nähe der Frage nach Willensfreiheit zum Zombie-Gedankenexperiment scheint mir, besonders auch in Kombination mit Menzies und Prices akteur-basiertem Verständnis von Kausalität, aufschlussreich über die zentrale Rolle der Phänomenologie von

Bewusstsein, wenn wir unsere Überzeugungen zu Willensfreiheit zu verstehen versuchen.

5.5 Supervenierende Kausalerklärungen und Laienüberzeugungen zu Willensfreiheit

Nach den eher theoretischen Überlegungen und Argumenten der vorhergehenden Abschnitte will ich zum Abschluss dieses Kapitels noch einmal kurz auf die konkretere Bedeutung von Willensfreiheit kommen. Einerseits haben wir ja bereits in Kapitel 2 gesehen, dass Laienüberzeugungen zu Willensfreiheit mit bestimmten Aussagen verbunden sind. Ich möchte im folgenden Abschnitt noch einmal kurz auf sie zurückkommen und versuchen, ausgehend von den vorhergehenden Argumenten und Überlegungen zu Erleben von Urheberschaft mögliche philosophische Begründungen für diese Überzeugungen anzubieten. Andererseits ergibt sich aus den obigen Überlegungen und Argumenten auch ein genaueres Bild dessen, was wir unter dem Begriff der Willensfreiheit sinnvollerweise verstehen können. Auch das will ich im Folgenden kurz darstellen.

Um die bekannten Laienaussagen zu Willensfreiheit durch die speziellen Eigenschaften des Erlebens eigener Urheberschaft zu begründen, werde ich die fünf Kernaussagen zur Bedeutung von Willensfreiheit aus Abschnitt 2.4 hier noch einmal kurz aufnehmen und summarisch Möglichkeiten darstellen, wie sich jede dieser Aussagen als Folge der oben hergeleiteten philosophischen Überlegungen zu Erleben von Urheberschaft verstehen lassen könnte. Ich will es aber hier offen lassen, ob diese Erklärungen letztendlich in der Lage sind, als formale Beweise für diese Laienaussagen zu dienen, weil das nicht Ziel dieser Arbeit ist und deren Rahmen sprengen würde.

5.5.1 Wahlmöglichkeiten

Um die erste Aussage zu den vorherrschenden Überzeugungen zu Willensfreiheit gemäss dem Free Will Inventory (Nadelhoffer u. a. 2014) aus dem gegebenen Kontext zu verstehen, müssen wir die Aussage aus Prämisse 2.4.1 überprüfen:

"Die Form des subjektiven Zugangs zu eigener Urheberschaft kann erklären, dass das Subjekt überzeugt ist, die eigenen Handlungen wählen zu können."

Die Begründung für diese Aussage durch die Form des subjektiven Zugangs zu eigener Urheberschaft lässt sich folgendermassen formulieren: Was das Subjekt präreflektiv als Urheberschaft erlebt, sind multipel realisierbare mentale Zustände, die auf physische Zustände supervenieren und damit kausal wirksam sind. Die kontrafaktisch relevante kausale Erklärung findet sich aber auf der supervenierenden mentalen Ebene in Form eines über ein Selbst-Modell vermittelten repräsentationalen Gehalts. Allen physischen Zuständen, die diesen mentalen Gehalt realisieren, ist gemeinsam, dass sie dem Subjekt seine eigenen kausalen Rollen immer erst als eigene Möglichkeiten, und das Realisieren einer dieser Möglichkeiten als eigene und freie Wahl gegeben sind. Das, was das Subjekt als eigene Urheberschaft erlebt, erklärt darum eine Handlung durch ein mentales Ereignis, das immer die Qualität einer eigenen und freien Wahl hat.

5.5.2 Willensfreiheit

Um die zweite Aussage zu den vorherrschenden Überzeugungen zu Willensfreiheit aus dem gegebenen Kontext zu verstehen, müssen wir die Prämisse 2.4.2 überprüfen:

"Die Form des subjektiven Zugangs zu eigener Urheberschaft kann erklären, dass das Subjekt überzeugt ist, dass der eigene Wille frei sei."

Begründung: Was das Subjekt präreflektiv als Urheberschaft erlebt, sind multipel realisierbare mentale Zustände, die sich gemeinsam dadurch auszeichnen, dass ihr Gehalt notwendigerweise eine Komponente des Wollens beinhaltet. Das Vorhandensein eines solchen Zustands erklärt sich darum nicht aus dem Vorhandensein einer spezifischen Realisation dieses Zustands, sondern aus dem Bestehen von dem Subjekt eigenen Wünschen. Diese eigenen Wünsche sind multipel realisierbar und in diesem Sinn unabhängig vom Vorhandensein bestimmter Realisationen dieses mentalen Zustands. Der Wille des Subjekts ist damit insofern frei, als sein Auftreten nicht auf das Auftreten eines bestimmten physischen Zustands angewiesen ist.

5.5.3 Selbstbestimmung

Um die dritte Aussage aus dem gegebenen Kontext zu verstehen, müssen wir die Prämisse 2.4.3 überprüfen:

"Die Form des subjektiven Zugangs zu eigener Urheberschaft kann erklären, dass das Subjekt überzeugt ist, über sich selbst bestimmen zu können."

Begründung: Was das Subjekt präreflektiv als Urheberschaft erlebt, sind über das Selbst-Modell vermittelte mentale Zustände, die dem Subjekt eigen sind und

die es in einer ihm eigenen und leiblichen Weise erfährt. Gleichzeitig sind es, wie in Abschnitt 5.4.1 begründet, diese mentalen Zustände und nicht deren spezifischen Realisierungen, die in einer kontrafaktischen Analyse die Kausalbeziehungen richtig abbilden. Damit sind es dem Subjekt eigene mentale Zustände, und nicht dem Subjekt fremde physische Zustände, die das Handeln des Subjekts bestimmen.

5.5.4 Kontrolle

Um die vierte Aussage aus dem gegeben Kontext zu verstehen, müssen wir die Prämisse 2.4.4 überprüfen:

"Die Form des subjektiven Zugangs zu eigener Urheberschaft kann erklären, dass das Subjekt überzeugt ist, die eigenen Entscheidungen und Handlungen kontrollieren zu können."

Begründung: Was das Subjekt präreflektiv als Urheberschaft erlebt, beinhaltet immer ein eigenes und willentliches Tun. Weil im Kontext der Supervenienzbeziehung zwischen dem Willensakt und der Handlung ein realisierungsunabhängiges Abhängigkeitsverhältnis besteht, ist dieses willentliche Tun kontrafaktisch nicht abhängig davon, dass ein bestimmter physischer Zustand eintrifft, sondern davon, dass es das Subjekt will. Kontrafaktisch ist es, mit Ausnahme eines Subsekundenbereichs vor einer aktuellen Bewegung (siehe Abschnitt 5.2.1), damit hinreichend, dass das Subjekt nicht mehr will, um eine Entscheidung oder Handlung in dem Sinn zu kontrollieren, dass sie nicht mehr gilt oder stattfindet.

5.5.5 Ausserliche Unantastbarkeit

Um die letzte Aussage zu verstehen, müssen wir die Prämisse 2.4.5 überprüfen:

"Die Form des subjektiven Zugangs zu eigener Urheberschaft kann erklären, dass das Subjekt überzeugt ist, selbst dann einen freien Willen zu haben, wenn seine Wahlmöglichkeiten durch äussere Umstände vollständig eingeschränkt sind."

Begründung: Was das Subjekt präreflektiv als Urheberschaft erlebt, beinhaltet die mentale Repräsentation von Handlungsmöglichkeiten, während ihm deren materielle Realisierung als solche wegen der psychophysiologischen Intransparenz grundsätzlich unzugänglich bleibt. Das Subjekt ist damit frei, solche mentalen Repräsentationen von Handlungsmöglichkeiten zu wollen, selbst wenn die materielle Realisierung der Handlung unmöglich ist. So hat zum Beispiel eine Person, die etwa einen Arm verloren hat, typischerweise immer noch eine mentale Repräsentation dieses Arms, die auch die Möglichkeit repräsentiert, diesen zu bewegen.

Damit kann sie auch immer noch den Arm bewegen wollen, auch wenn jede reale Armbewegung grundsätzlich unmöglich ist.

5.6 Eigenschaften einer durch Erleben von Urheberschaft erklärbaren Willensfreiheit

Was bedeutet es nun für unser Verständnis zum Verhältnis von Willensfreiheit und Determinismus, wenn wir das Vorhandensein von Willensfreiheit durch die Epistemik eigener Urheberschaft erklären wollen?

Um bei der negativen Seite anzufangen, widerspricht ein solchermassen begründetes Verständnis von Willensfreiheit der libertarischen Position, dass es möglich ist, unter den exakt gleichen Umständen auch anders zu können (Van Inwagen 1983, S. 16). Dem steht im Weg, dass Erleben von Urheberschaft, und damit ein wesentlicher Teil unseres Selbsterlebens, zwingend auf regelhafte Beziehungen zwischen mentalen und physischen Ereignissen angewiesen ist: Nur wenn bei genügend ähnlichen Umständen und Ereignissen genügend ähnliche Dinge geschehen, ist es überhaupt möglich, die Folgen eigenen Handelns vorherzusehen. Solche Regelhaftigkeiten sind damit, wie auch in Abschnitt 5.2 gesehen, eine Voraussetzung dafür, dass das Subjekt erst die Möglichkeit hat, sich selbst als Ursache zu erleben. Damit bedingt Erleben von Urheberschaft mindestens eine lokale und probabilistische Vorhersehbarkeit, die mit einem jederzeit anders Können in Konflikt steht.

Im Umkehrschluss scheint es also so, dass weniger ein Determinismus als ein Indeterminismus für unser Selbstverständnis als freie Akteur*innen problematisch wäre: Wenn das Subjekt für sein Erleben von Urheberschaft auf die Vorhersehbarkeit zukünftiger Ereignisse angewiesen ist, und wenn gleichzeitig keine zuverlässige Ursachenkette vom Moment der Vorhersage bis zum vorhergesagten Ereignis führt, müsste ein solches Erleben von Urheberschaft ein eher schwaches, seltenes und zufälliges Ereignis sein. Das ist aber offensichtlich nicht der Fall, denn wir empfinden uns in unserem Alltag meist zu einem überwiegenden Teil und oft auch über lange Zeiträume als Ursache von Ereignissen, die wir gewollt haben, ohne dass es dazu notwendig scheint, dass wir Kenntnis von genügend eindeutigen materiellen raumzeitlichen Folgebeziehungen zwischen dem Moment der Vorhersage und dem Eintreten des gewünschten Ereignisses haben. Die Erklärung für diese Fähigkeit liegt offensichtlich in der multiplen Realisierbarkeit der Ursachen für das erwünschte Ereignis, die die mentale Indeterminiertheit konkreter materieller raum-zeitlicher Folgebeziehungen kompensiert. Damit ist das Subjekt in der Lage, in unterschiedlichen Gegebenheiten dem gleichen Wunsch entsprechend zu handeln, und erlebt dies als Freiheit von äusseren Umständen. Willensfreiheit, solchermassen als ein "bei seinen Gründen bleiben können" verstanden, ist damit etwas, was dem Subjekt langfristig Urheberschaft überhaupt erst erlaubt. In diesem Sinn bedingen sich ein Erleben von Urheberschaft und das Vorhandensein von einer in einem so verstandenen Willensfreiheit gegenseitig. Nur wer einigermassen unabhängig von den momentanen Umständen gemäss einigermassen persistenten Gründen handeln kann, wird in einem weiteren Rahmen Urheberschaft erleben können, und nur wer sich als Urheber seines Tun empfindet, wird die eigenen Taten als frei empfinden. Das mag, um hier abzuschliessen, auch dem entsprechen, was René Descartes in seiner vierten Meditation festhielt:

"Um frei zu sein, ist es nämlich nicht nötig, dass ich mich nach beiden Seiten hin gleich verhalten kann, sondern im Gegenteil: Je mehr ich mich der einen Seite zuneige, entweder weil ich auf dieser Seite den Grund des Wahren und Guten evident einsehe, oder weil Gott mein innerstes Denken so angelegt hat, desto freier wähle ich sie. (Descartes 2011, S. 64)"

KAPITEL 6

Fazit

Die vorliegende Arbeit hatte zum Ziel, Laienintuitionen zu Willensfreiheit daraufhin zu überprüfen, ob sie sich durch präreflektives Erleben von Urheberschaft begründen lassen, und was eine derart begründete Willensfreiheit für philosophisch interessante Aussagen mit sich bringt. Dieser Versuch schliesst natürlich nicht aus, dass Willensfreiheit auch anders verstanden, begründet oder verneint werden kann. Sofern er gelingt, kann er nicht nur für sich beanspruchen, die überwiegenden Intuitionen abzubilden, sondern deren Begründung auch in einer Weise herzuleiten, die sich auf Beobachtungen stützt, die uns in unserem alltäglichen Erleben als praktisch Handelnde immer schon vertraut sind.

Das Fazit der aussagenlogischen Analyse der Fragestellung kommt zum Schluss, dass das, was wir als eigene bewusste Urheberschaft erleben, zu begründen vermag, dass wir in unserem Handeln Alternativen haben, dass unser Wollen frei ist, dass wir über uns bestimmen können, dass wir unsere Handlungen kontrollieren können und dass unser Wille auch dann noch frei ist, wenn wir äusserlich keine Wahlmöglichkeiten haben. In den Prämissen, die dieses Fazit stützen, kommt Determinismus nicht vor. Willensfreiheit, in dieser Art hergeleitet, ist darum der Gültigkeit eines physikalischen Determinismus gegenüber neutral: Statt auf Naturgesetze stützt sich die Begründung für dieses Fazit auf die Abhängigkeit dieses Erlebens von auf materielle Zustände supervenierenden, multipel realisierbaren Bewusstseinszuständen. Über ein dem Subjekt transparenten Selbst-Modell repräsentieren diese mentalen Zustände bestehende Wirkungen vom Subjekt auf dessen Aussenwelt zwangsläufig so, dass ihr phänomenaler Gehalt die relevanten Qualitäten von Willensfreiheit immer schon beinhaltet.

Diese Unabhängigkeit des eigenen Willens impliziert umgekehrt, dass Willensfreiheit, solcherart verstanden, kein materieller Sachverhalt ist. Als eine funktionale, repräsentationale und phänomenale Gegebenheit ist sie dem Bewusstsein zwar offensichtlich, ein Realismus dieses Erlebens entpuppt sich aber, um bei Thomas Metzinger Begriff zu bleiben, als naiv: Wir können uns der materiellen Gegebenheit dessen, was wir willentlich verursachen, nie sicher sein, weil die repräsentationalen Vorgänge, über die uns materielles Geschehen zugänglich ist, Fehler enthalten

können. Das ist etwas, was wir auch erleben, zum Beispiel, wenn wir im Traum meinen, Dinge zu "tun". Das Verhältnis zwischen dem, was wir als unseren eigenen freien Willen erleben und den materiellen Realisationen dieses Willens ist damit immer prekär. Trotzdem sind solche repräsentationalen Gegebenheiten auch nicht einfach falsch, weil sie im Normalfall mit der materiellen Welt in einem systematischen Bezug stehen und diese materielle Welt zuverlässig abbilden. Was wir selbst verursacht haben und was wir möglicherweise in Zukunft verursachen könnten, ist darum etwas, was wir laufend überprüfen und gegebenenfalls auch revidieren müssen.

Mit Woodwards Begriff des realisierungsunabhängigen Abhängigkeitsverhältnisses habe ich weiter argumentiert, dass es diese supervenierende mentale Ebene und nicht die Ebene der materiellen Realisation ist, die die für willentliches Handeln notwendigen Kausalbeziehungen sinnvoll abbildet. Solche realisierungsunabhängigen Abhängigkeitsverhältnisse sind aber, in Form eines Selbst-Modells, die einzige Möglichkeit, über die das subjektive Bewusstsein der Person mit seiner Aussenwelt in Beziehung tritt. Gleichzeitig ist diese mentale Ebene pervasiv geprägt durch eine Reihe von a-priori vorhandenen relationalen Eigenschaften, die sich dem Subjekt als leibliches Erleben manifestieren. Das verunmöglicht es systematisch, dass sich das, was subjektiv als eigene Urheberschaft wahrgenommen wird, vollständig auf eine materielle Ebene reduzieren lässt, die keine Willensfreiheit mehr zuliesse.

Auch wenn wir davon ausgehen, dass es für die Erklärung von Zusammenhängen auf der materiellen Ebene gar keine Kausalerklärungen braucht, können wir subjektiv erlebte Urheberschaft als eine Disposition von Subjekten des Typs "freie*r Akteur*in" verstehen, bestimmte Zusammenhänge als "kausal" zu erkennen. Auch in diesem Fall habe ich argumentiert, dass die Reduktion solcher Dispositionen auf die materielle Ebene grundsätzlich problematisch ist: Wenn es richtig ist, dass diese Disposition auf gleichem Fuss wie unsere Fähigkeiten zu perzeptuellem Bewusstsein steht, wäre ein nicht-reduzierbares Bedeuten von Selbsterleben als "frei*e Akteurin" ebenso schwer bestreitbar wie das nicht-reduzierbare Bedeuten der Erfahrung von zum Beispiel Farbsehen oder Schmerz.

Was bedeutet Willensfreiheit, wenn sie durch Erleben von eigener Urheberschaft begründet werden soll? Wenn sich bewusstes Erleben von Ursächlichkeit dadurch begründet, dass das Wollen und Tun einer Person zu einem gewissen Grad unabhängig von den äusseren Umständen ist, dann geht es bei einer solcherart verstandenen Willensfreiheit weniger um die Frage der (In-)Determiniertheit von Handlungen, sondern, wie es auch Geert Keil betont, primär um die Frage der geeigneten Determination von Handlungen (Keil 2018, S. 85). Die Determination von Handlungen über multipel realisierbare, repräsentationale mentale Zustände lässt insofern Freiheit zu, als sie das Subjekt in die Lage versetzt, mehrere Handlungsal-

ternativen bewusst zu repräsentieren und eine dieser Repräsentationen als eigene Wahl, und zu einem gewissen Grad unabhängig vom Vorhandensein bestimmter materieller Voraussetzungen, zu realisieren.²⁴ Damit wird das Subjekt "Herr*in im eigenen Haus": Was dieses Effektiv-Machen einer Handlungsrepräsentation verursacht, ist wieder unabhängig von dessen konkreter Realisierung, vielmehr ist es bestimmt durch die dem Subjekt eigenen Gründe (Kane 2007, S. 25).

Willensfreiheit, so verstanden, ist also die direkt erlebbare Fähigkeit, qua eigenem Wollen unter unterschiedlichen und gegebenenfalls sogar widrigen Umständen sinngemäss Gleiches zu bewirken. Geert Keil verwendet in diesem Kontext den Begriff der hindernisüberwindenden Fähigkeit zur Willensbildung (Keil 2018, S.32), die sich fundamental von einer in vielerlei Hinsicht problematischen Disposition unterscheidet, auf bestimmte Umstände in einer bestimmten Weise zu reagieren (Keil 2018, S.31): Das Subjekt handelt damit so, wie es handelt, nicht als passive Folge von materiellen Ursachen, sondern aktiv und aus zeitlich exendierten eigenen Gründen, die in immer wieder unterschiedlichen Supervenienzbasen und damit letztendlich beinahe beliebigen Weisen realisiert werden können. Im direkten Erleben des Subjekts ist das, was empirisch im Bereich dieser Fähigkeiten liegt, eigen, während sich alles, was ausserhalb dieser Fähigkeiten liegt, fremd anfühlt.

Vom theoretischen Standpunkt her ist der Begriff der Willensfreiheit, wie er hier hergeleitet wurde, kaum mit "metaphysischem Gepäck" belastet. Willensfreiheit, wie wir sie erleben, stützt sich in dieser Herangehensweise auf ein aus unserer alltäglichen Praxis nicht wegzudenkendes, theoretisch aber schwer vollständig durchdringbares Verhältnis zwischen dem leiblich erlebenden Subjekt und seiner Umwelt: Wenn die Überlegungen in Kapitel 3 richtig sind, ergibt sich Erleben von Urheberschaft notwendigerweise bereits, wenn wir von wenigen und wie mir scheint philosophisch relativ unkontroversen Annahmen ausgehen. Weiter würde man die in diesen Annahmen enthaltenen Funktionen und die daraus resultierenden Fähigkeiten aus einer evolutionsbiologischen Sichtweise wohl als generell vorteilhaft für die biologische Fitness eines Organismus betrachten, sodass wir für die Erklärung des Ursprungs dieser Fähigkeiten zumindest nicht schon a-priori im Widerspruch zu einer recht erfolgreichen biologischen Theorie stehen.

Zuletzt will ich hier festhalten, dass Willensfreiheit, auch wenn sie sich in der obigen Weise erklären liesse, durch ihre Abhängigkeit an phänomenales Bewusstsein so schwer erklärbar bleibt, wie dieses Bewusstsein selbst. Ohne dieses mysteriöse, aber immer schon dagewesene, nicht abstreitbare Bewusstsein wären wir willenlose Zombies. Umgekehrt erklärt sich mit dieser Bindung von Willensfreiheit an ein realisierungsunabhängiges Bewusstsein und unter der Annahme, dass "ein vernünftiges, mithin zur intelligibelen Welt gehöriges Wesen" ein Wesen mit Be-

²⁴Harry Frankfurt hat argumentiert, dass es das Vorhandensein dieser Art von Willensfreiheit ist, die uns zur Person macht (Frankfurt 1988).

wusstsein ist, auch das Anfangszitat von Immanuel Kant (Kant 1785, S. 525) aus seinen Grundlagen der Metaphysik der Sitten:

"Als ein vernünftiges, mithin zur intelligibelen Welt gehöriges Wesen kann der Mensch die Causalität seines eigenen Willens niemals anders als unter der Idee der Freiheit denken; denn Unabhängigkeit von den bestimmenden Ursachen der Sinnenwelt (dergleichen die Vernunft jederzeit sich selbst beilegen muss) ist Freiheit."

.

Literatur

- Beckermann, Ansgar (2008). Analytische Einführung in die Philosophie des Geistes. Walter de Gruyter.
- Björnsson, Gunnar und Joshua Shepherd (2020). "Determinism and attributions of consciousness". In: *Philosophical Psychology* 33.4, S. 549–568.
- Blakemore, Sarah-Jayne und Chris Frith (2003). "Self-awareness and action". In: Current Opinion in Neurobiology 13.2, S. 219–224.
- Blanke, Olaf u. a. (2002). "Stimulating illusory own-body perceptions". In: *Nature* 419.6904, S. 269–270.
- Böhme, Gernot (2008). Ethik leiblicher Existenz: Über unseren Umgang mit der eigenen Natur. Suhrkamp.
- Botvinick, Matthew und Jonathan Cohen (1998). "Rubber hands 'feel' touch that eyes see". In: *Nature* 391.6669, S. 756–756.
- Brugger, Peter und Bigna Lenggenhager (2014). "The bodily self and its disorders: neurological, psychological and social aspects". In: *Current Opinion in Neurology* 27.6, S. 644–652.
- Carreras, Genis (2013). *Philographics*. BIS Publishers.
- Chalmers, David (1996). The Conscious Mind: In Search of a Fundamental Theory. Oxford Paperbacks.
- (2007). "The hard problem of consciousness". In: *The Blackwell Companion to Consciousness*, S. 225–235.
- Danto, Arthur C (1965). "Basic actions". In: American Philosophical Quarterly 2.2, S. 141–148.
- Dennett, Daniel C (1984). Elbow room: The varieties of free will worth wanting. MIT University Press Cambridge, MA Press.
- (1988). "Quining qualia". In: Consciousness in Contemporary Science. Hrsg. von A Marcel und E Bisiach. Oxford University Press, S. 42–77.
- (1995). "The unimagined preposterousness of zombies". In: *Journal of Consciousness Studies* 2.4, S. 322–26.
- Descartes, René (2011). Meditationen. Mit sämtlichen Einwänden und Erwiderungen. Hrsg. von Christian Wohlers. Bd. 598. Philosophische Bibliothek. Felix Meiner Verlag.
- Feldman, Gilad, Roy F Baumeister und Kin Fai Ellick Wong (2014). "Free will is about choosing: The link between choice and the belief in free will". In: *Journal of Experimental Social Psychology* 55, S. 239–245.

- Fischer, John Martin u.a. (2009). Four Views on Free Will. John Wiley & Sons. Frankfurt, Harry G (1988). "Freedom of the Will and the Concept of a Person". In: What is a Person? Springer, S. 127–144.
- Frege, Gottlob (1892). "Über Sinn und Bedeutung". In: Zeitschrift für Philosophie und philosophische Kritik, S. 25–50.
- Friston, Karl (2005). "A theory of cortical responses". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 360.1456, S. 815–836.
- Fuchs, Thomas (2013). "The self in schizophrenia: Jaspers, Schneider, and beyond". In: One century of Karl Jaspers' General Psychopathology. Hrsg. von G Stanghellini und T Fuchs. Oxford University Press, S. 245–257.
- Gallagher, Shaun (Jan. 2000). "Philosophical conceptions of the self: Implications for cognitive science". In: *Trends in Cognitive Science* 4.1, S. 14–21.
- (2007). "The natural philosophy of agency". In: *Philosophy Compass* 2.2, S. 347–357.
- (2012). "Multiple aspects in the sense of agency". In: New Ideas in Psychology 30.1, S. 15–31.
- Haggard, Patrick (2008). "Human volition: Towards a neuroscience of will". In: *Nature Reviews Neuroscience* 9, S. 934–946.
- Heinks-Maldonado, Theda H u.a. (2007). "Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations". In: Archives of General Psychiatry 64.3, S. 286–296.
- Jackson, Frank (1986). "What Mary didn't know". In: *The Journal of Philosophy* 83.5, S. 291–295.
- Kane, Robert (2007). "Libertarianism". In: Four Views on Free Will. Hrsg. von JM Fischer u.a. Malden, MA: Oxford: Blackwell Publishers, S. 5–43.
- Kant, Immanuel (1785). Grundlegung zur Metaphysik der Sitten. Bd. IV. AA.
- Keil, Geert (2018). Willensfreiheit und Determinismus. Reclam Verlag.
- Kim, Jaegwon (2005). "The supervenience argument motivated, clarified and defended". In: *Physicalism, or Something Near Enough*. Princeton University Press. Kap. 2, S. 32–92.
- Koenig, Thomas (2020). "Yes, I did it! Eine Replik auf Kims These der physikalischen Reduktionierbarkeit von mentalen Kausalerklärungen". Arbeit zum Seminar Das Problem der mentalen Verursachung (455296-HS2019) von Prof. Dr. Vera Hoffmann-Kolss.
- Lanier, Jaron (1995). "You can't argue with a zombie". In: Journal of Consciousness Studies 2.4, S. 333–45.
- Laplace, Pierre Simon de (1816). Essai philosophique sur les probabilités. Courcier. Lenggenhager, Bigna u.a. (2007). "Video ergo sum: Manipulating bodily self-consciousness". In: Science 317.5841, S. 1096–1099.

- Levine, Joseph (1999). "Conceivability, identity, and the explanatory gap". In: Toward a Science of Consciousness III: The Third Tucson Discussions and Debates (Complex Adaptive Systems), S. 3–12.
- Lewis, David (1986). "Causation". In: *Philosophical Papers*. Hrsg. von Adler J.E. und L.J. Rips, S. 632–638.
- (1994). "Humean supervenience debugged". In: *Mind* 103.412, S. 473–490.
- Menzies, Peter und Huw Price (1993). "Causation as a secondary quality". In: *The British Journal for the Philosophy of Science* 44.2, S. 187–203.
- Merleau-Ponty, Maurice (1945). Phénoménologie de la perception. Gallimard.
- Metzinger, Thomas (2003). Being No One. Cambridge, Ma: MIT Press.
- (2007). "Empirical perspectives from the self-model theory of subjectivity: a brief summary with examples". In: *Progress in Brain Research* 168, S. 215–278.
- Monroe, Andrew E, Kyle D Dillon und Bertram F Malle (2014). "Bringing free will down to Earth: People's psychological concept of free will and its role in moral judgment". In: *Consciousness and Cognition* 27, S. 100–108.
- Nadelhoffer, Thomas u. a. (2014). "The free will inventory: Measuring beliefs about agency and responsibility". In: Consciousness and Cognition 25, S. 27–41.
- Nahmias, Eddy, Jason Shepard und Shane Reuter (2014). "It's OK if 'my brain made me do it': People's intuitions about free will and neuroscientific prediction". In: Cognition 133.2, S. 502–516.
- Nichols, Shaun (2011). "Experimental philosophy and the problem of free will". In: *Science* 331.6023, S. 1401–1403.
- Noreika, Valdas u. a. (2020). "Modulating dream experience: Noninvasive brain stimulation over the sensorimotor cortex reduces dream movement". In: *Scientific Reports* 10.1, S. 1–19.
- Penfield, Wilder und Edwin Boldrey (1937). "Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation". In: *Brain* 60.4, S. 389–443.
- Platon (1987). Parmenides (Ubersetzt von Ekkehard Martens). Reclam.
- Revonsuo, Antti (2005). "The self in dreams". In: *The Lost Self*. Oxford University Press, S. 206–219.
- Russell, Bertrand (1912). "On the notion of cause". In: *Proceedings of the Aristotelian Society*. Bd. 13. JSTOR, S. 1–26.
- Sacks, Oliver (2004). Der Tag, an dem mein Bein fortging. Reinbek: Rowohlt.
- Salmon, Wesley C (1984). Scientific explanation and the causal structure of the world. Princeton University Press.
- Schmitz, Hermann (2007). Der unerschöpfliche Gegenstand. Grundzüge der Philosophie. Bonn: Bouvier Verlag.

- Schultze-Kraft, Matthias u.a. (2016). "The point of no return in vetoing self-initiated movements". In: *Proceedings of the National Academy of Sciences* 113.4, S. 1080–1085.
- Searle, John (2000). "Consciousness". In: Annual Reviews of Neuroscience 23, S. 557–578.
- Shepherd, Joshua (2012). "Free will and consciousness: Experimental studies". In: Consciousness and Cognition 21.2, S. 915–927.
- (2015). "Consciousness, free will, and moral responsibility: Taking the folk seriously". In: *Philosophical Psychology* 28.7, S. 929–946.
- Slater, Mel u. a. (2010). "First person experience of body transfer in virtual reality". In: *PloS One* 5.5, e10564.
- Sytsma, Justin und Jonathan Livengood (2015). The Theory and Practice of Experimental Philosophy. Broadview Press.
- Tononi, Giulio (2004). "An information integration theory of consciousness". In: *BMC Neuroscience* 5.1, S. 1–22.
- Tsakiris, Manos (2017). "The multisensory basis of the self: from body to identity to others". In: *The Quarterly Journal of Experimental Psychology* 70.4, S. 597–609.
- Van Inwagen, Peter (1983). An Essay on Free Will. Oxford University Press.
- Wegner, Daniel M (2003). "The mind's best trick: How we experience conscious will". In: *Trends in Cognitive Sciences* 7.2, S. 65–69.
- Wisniewski, David, Robert Deutschländer und John-Dylan Haynes (2019). "Free will beliefs are better predicted by dualism than determinism beliefs across different cultures". In: *PloS one* 14.9, e0221617.
- Wittgenstein, Ludwig (1958). The blue and brown books. New York: Harper & Brothers.
- Woodward, James (2008). "Mental causation and neural mechanisms". In: Being Reduced: New Essays on Reduction, Explanation, and Causation. Hrsg. von Jakob Hohwy und Jesper Kallestrup. Oxford University Press Oxford, S. 218–262.