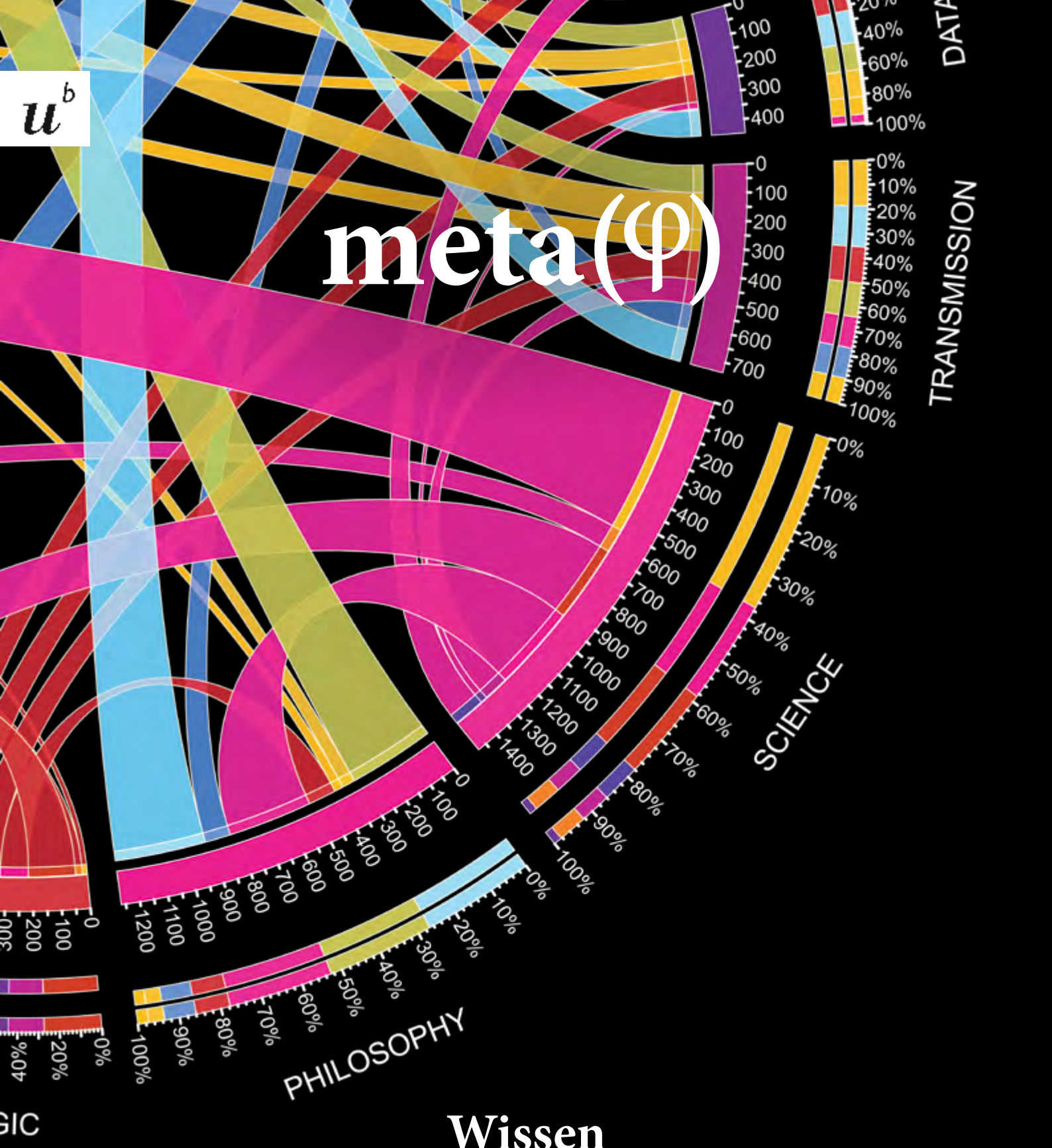


Wissen

Was können wir wissen?

Welchen Wahrheitsanspruch hat die Wissenschaft?

Was ist nicht-sprachliches Wissen?

Herausgeberin
Universität Bern, Institut für Philsophie
Länggassstrasse 49a, 3012 Bern, Schweiz

Redaktion
Meta(φ)
Fachschaft Philosophie, Universität Bern
metaphi.philo@lists.unibe.ch

Chefredaktor
Fabian Dali

Auswahl und Lektorat
Fabian Dali, Manuel Merki, Jasmine Kammermann

Produktion und Gestaltung
Manuel Merki

ISSN
Gedruckte Ausg.: Meta(φ) ISSN:2297-9948
Online-Ausg.: Meta(φ) [Elektronische Ressource]
ISSN:2297-9956

Meta(φ) erscheint halbjährlich.

Text-Einreichungen
Einreichungen zur Veröffentlichung sind erwünscht an metaphi.philo@lists.unibe.ch. Die Auswahl wird vom Redaktions-Team entsprechend dem Themenschwerpunkt, der Relevanz und des Schreibstils vorgenommen. Es besteht kein Anrecht auf Veröffentlichung, weder bei erfolgreicher Einreichung noch nach Anfrage durch ein Redaktions-Mitglied. In diesem Magazin veröffentlichte Texte gelten als wissenschaftliche Publikation und sind somit zitierfähig.

Erste Worte

Das philosophische Schaffen kann und darf ein einsames sein. Ein Gedanke mag in einer Person keimen und wachsen, später welken oder blühen, ohne sich jemals dem Licht der Öffentlichkeit gezeigt zu haben. Er kann gut sein und wahr und schön, ohne jemals in die Welt getragen worden zu sein. Er kann sogar wirken, im Innern, indem er ändert, wie das Aussen verstanden wird. Wohl kein Menschentier weiss das so gut wie das studierende, besteht dessen Bildung doch gerade darin, sich sinnsuchend der Produktion von Text zu verschreiben, von welchem es nicht im geringsten erwartet, dass er zu einem anderen Zweck als jenem der Benotung noch einmal gelesen wird. Es ist ein edles Bild und ein trauriges, vergossener Wein für die Götter sozusagen.

Das Wort aber wurde erfunden, um geteilt zu werden, und so befriedigt es nicht nur das Bedürfnis des denkenden Menschen, der Ungestalt seiner Vernunft feste Form zu verleihen, sondern auch jenes des Forschenden, das von ihm Entdeckte, bestenfalls Wissen genannt, anderen zugänglich zu machen. Es gehört nun vielleicht zu den zweifelhaften Vorzügen unserer Zeit, dass es an Möglichkeiten nicht fehlt, öffentlichen Platz zu finden, um auch die banalsten Ausscheidungen des eigenen Denkapparats redaktionell auszustellen, doch fühlen sich reflektierte Geister oftmals nicht dazu berufen, zwischen Katzenbildern und personalisierter Werbung noch ernst Gemeintes und wohl Durchdachtes anzubringen. Die Fachdiskussion, wenn auch sie zuweilen sich im Kreise dreht und schon per Definition nicht über einen festgelegten Tellerrand hinauszublicken vermag, bietet eine vorläufige Lösung für dieses Dilemma der Informationsgesellschaft. Hier, in diesem kollektiv-abstrakten Gefäss der Kommunikation, sammelt sich das Kondensat der aktuellen und globalen philosophischen Debatten zu jener intellektuellen Pfütze, in der sich der wissensliebende Mensch so gerne wühlt.

So habe ich im Namen der Fachschaft für Philosophie die ausserordentliche Ehre und das Privileg, kraft dieser Zeilen die Wiedergeburt eines hauseigenen Beitrags zu ebenjener Diskussion zu besingen. Nach zwei Semestern Stille treten wir mit Meta(φ) endlich das Erbe der altehrwürdigen ἀρετή an, die bereits seit 2013 einen exklusiven Einblick in das Werk unserer Studierenden bot. Dabei ist die drastische Reduktion der Anzahl griechischer Buchstaben im Namen natürlich erst der Anfang einer ganzen Reihe an Änderungen, mit der das neue Redaktionsteam der Publikation seinen Stempel aufdrückt. Als aufgrund ihrer fast metaphysischen Implikationen wohl gewichtigste Neuerung haben wir beschlossen, die ehemals ausschliesslich online verfügbare Schrift greifbar zu machen; in Kleinstauflage liegt die Meta(φ) im Philosophieinstitut erstmals auch in gedruckter Form zur Lektüre auf. Wie bisher setzt sich die Autorenschaft vornehmlich aus Major- und Minorstudierenden der Philosophie zusammen, neu wollen wir diese Chance aber auch nutzen, um der Fachschaft einen Einblick in das Schaffen der Dozierenden zu gewähren. Dr. Kevin Reuter macht hierfür den ersten Schritt und wir danken ihm wie auch allen anderen AutorInnen der ersten Ausgabe für ihren Beitrag und hoffen, uns in Zukunft abermals auf so grossartige Einreichungen verlassen zu dürfen.

Herzlichst, Fabian Dali
im Namen des Redaktionsteams.

Fabian Dali

Jasmine Kammermann

Manuel Merki

Friedrich Nietzsche
Über Wahrheit und Lüge im aussermoralischen Sinne

"Was ist also Wahrheit? Ein bewegliches Heer von Metaphern, Metonymien, Anthropomorphismen, kurz eine Summe von menschlichen Relationen, die, poetisch und rhetorisch gesteigert, übertragen, geschmückt wurden, und die nach langem Gebrauch einem Volke fest, kanonisch und verbindlich dünken: die Wahrheiten sind Illusionen, von denen man vergessen hat, daß sie welche sind, Metaphern, die abgenutzt und sinnlich kraftlos geworden sind, Münzen, die ihr Bild verloren haben und nun als Metall, nicht mehr als Münzen, in Betracht kommen."

Ludwig Wittgenstein
Tractatus logico-philosophicus

"2.19 Das logische Bild kann die Welt abbilden.

2.2 Das Bild hat mit dem Abgebildeten die logische Form der Abbildung gemein.

2.201 Das Bild bildet die Wirklichkeit ab, indem es eine Möglichkeit des Bestehens und Nichtbestehens von Sachverhalten darstellt.

2.202 Das Bild stellt eine mögliche Sachlage im logischen Raume dar.

2.203 Das Bild enthält die Möglichkeit der Sachlage, die es darstellt.

2.21 Das Bild stimmt mit der Wirklichkeit überein oder nicht; es ist richtig oder unrichtig, wahr oder falsch.

2.22 Das Bild stellt dar, was es darstellt, unabhängig von seiner Wahr- oder Falschheit, durch die Form der Abbildung.

2.221 Was das Bild darstellt, ist sein Sinn.

2.222 In der Übereinstimmung oder Nichtübereinstimmung seines Sinnes mit der Wirklichkeit, besteht seine Wahrheit oder Falschheit.

2.223 Um zu erkennen, ob das Bild wahr oder falsch ist, müssen wir es mit der Wirklichkeit vergleichen.

2.224 Aus dem Bild allein ist nicht zu erkennen, ob es wahr oder falsch ist.

2.225 Ein a priori wahres Bild gibt es nicht."

Inhaltsverzeichnis

Fundament

Vom Wahrheitsanspruch der Wissenschaft - Jonas Wittwer 6

Plausibilität der empirischen Unterbestimmtheit von Theorien als Argument gegen den Wissenschaftlichen Realismus - Emmanuel Téofilo Guizar Rosales 8

Fokus

Frank Jacksons Mary's Room Argument - Sebastian Drosselmeier 12

Qualia vs. Physikalismus / Funktionalismus - Sina Rast 17

Fachwort

Rationality and the Self-Attribution of Experiences - Dr. Kevin Reuter 22

Frontal

Willensfreiheit durch Kenntnis der Determinanten - Manuel Merki 28

Weisst du, wer du bist? - Franziska Wettstein 31

Fernsicht

Können wir die Zukunft der Gesellschaft kennen? - Martina Jakob 33

Philosophisterei

Wieso? Weshalb? Warum? - Jasmine Kammermann 42

Vom Wahrheitsanspruch der Wissenschaft

In der Frage nach dem Wahrheitsanspruch der Wissenschaft sind mindestens drei Subfragen enthalten.

1. Welchen Wahrheitsanspruch soll die Wissenschaft haben?
2. Welchen Wahrheitsanspruch kann die Wissenschaft überhaupt haben?
3. Wie kann dieser Wahrheitsanspruch umgesetzt werden?

Im Rahmen dieses Essays ist nur eine skizzenhafte Beantwortung dieser Fragen möglich. Hierbei soll der Schwerpunkt auf der ersten Frage liegen, ohne jedoch zu vergessen, dass zwischen diesen Fragen eine enge Verbindung besteht.

Wahrheitsanspruch der Wissenschaft

Wissenschaft generiert Wissen. Wissen wird in einer Standarddefinition aus der Erkenntnistheorie wie folgt definiert:

> x weiss, dass p, genau dann, (1) wenn es wahr ist, dass p, (2) wenn x glaubt, dass p, (3) und wenn x gerechtfertigt ist zu glauben, dass p.

Diese Definition enthält die Begriffe Wahrheit, Glauben und Rechtfertigung, die hier nicht weiter definiert werden können. Unter Idealbedingungen sollten die Begriffe Wahrheit und Glauben jeweils absolut erfüllt sein. Dieser ideale Zustand bleibt jedoch wenigen Fachgebieten vorbehalten.

Über viele Aspekte können nur gute Vermutungen gemacht werden, da verschiedene Grenzen es verhindern, einen Gegenstand genauer zu untersuchen. Mögliche Grenzen sind technischer oder historischer Natur. So kann es an Geräten mangeln, um Aufzeichnungen über einen Gegenstand zu machen, oder die Informationen sind schlicht nicht vorhanden, da es nie Quellen dazu gab. Diese Grenzen sind aber nicht absolut. So wurden beispielsweise immer wieder Geräte entwickelt, welche neue Erkenntnisse ermöglichten. Weitere Probleme ergeben sich aus der induktiven Methode, auf die viele Wissenschaften zurückgreifen. Denn gibt es keine theoretische Rechtfertigung dafür, davon auszugehen, dass sich Gegenstand A zum Zeitpunkt t2 ebenso verhält wie zum Messzeitpunkt t1.[1] Die Anzahl an möglichen Aussagen, denen ein Wahrheitswert von entweder 0 (falsch) oder 1 (wahr) zugeordnet werden kann, ist also sehr beschränkt.

Es gibt zwei Möglichkeiten, wie mit diesem Problem umgegangen werden kann. Erstens könnte sich die Wissenschaft auf die wenigen Aussagen beschränken, die tatsächlich beurteilt werden können. Dadurch verliert sie aber den grössten Teil ihrer Anwendungsgebiete, weshalb mir dies als keine befriedigende Lösung erscheint. Zweitens könnte die Wissenschaft mit fiktiven Wahrheitswerten von entweder 1 oder 0 operieren, die davon abhängen, wie hoch der Glaubensgrad ist, der einer These zugeordnet werden kann. Ich schreibe hier von fiktiven Wahrheitswerten, um zu verdeutlichen, dass eine Aussage als wahr beurteilt wird, ohne dass eine absolute Garantie besteht, dass die Aussage tatsächlich wahr ist. Die Höhe der Glaubensgrade sollte hierbei von der Rechtfertigung abhängen, auf welcher die Glaubensgrade beruhen.[2]

Eine Konsequenz, die sich aus diesem zweiten Ansatz ergibt, ist, dass man sich auf einen Wert für einen Glaubensgrad einigen muss, damit eine Aussage als wahr angesehen wird. Denn anders als bei der Frage, ob etwas wahr oder falsch ist, kann ein Glaubensgrad sämtliche Werte zwischen 0 und 1 annehmen. Dies ist abhängig davon, wie fest jemand glaubt, dass eine Aussage wahr ist oder nicht.

Der Anspruch an die Höhe dieser Glaubensgrade, welcher die Wissenschaft haben sollte, hängt indessen stark von ihrem jeweiligen Fach und dem erweiterten Kontext ab. Ein Glaubensgrad von 0.3, dass die theoretischen Grundlagen zu einer bestimmten Krebstherapie wahr sind, kann genügen, wenn der Patient kurz vor dem Exitus steht und diese Therapie die einzige Über-

1 Vgl. zur Induktion z.B. Hume (1740).

2 Vgl. zur Idee mit Wahrheitswerten die Schule des Bayesianismus z.B. Bartelborth (2013). In diesem Essay wird aber kein Bayesianismus vertreten, sondern nur eine darauf aufbauende Idee skizziert.

lebenschance ist. Wenn die Situation jedoch nicht bedrohlich ist und es Alternativen gibt, kann ein Wert von 0.3 nicht ausreichend sein.[3]

Ein weiterer relevanter Unterschied, welcher Glaubensgrad gefordert werden soll, hängt davon ab, ob ein Experiment möglich ist oder nicht. Wenn anhand eines Differenztests die kausale Abhängigkeit der Faktoren belegt werden kann, so sollte hier ein deutlich höherer Glaubensgrad gefordert werden, als wenn kein Experiment möglich ist. Denn ohne das Experiment ist es gar nicht möglich, einen bestimmten Glaubensgrad überhaupt zu überschreiten. So ist beispielsweise die Geschichte ein Fach, das auf Experimente verzichten muss. Dem Historiker stehen nur Quellen zur Verfügung, die im schlimmsten Fall sogar unentdeckt manipuliert wurden. Aus diesem Material muss er eine glaubhafte These ableiten, die aber niemals denselben Glaubensgrad erreichen kann wie die Schmelzpunktmessungen einer Chemikerin. So kann es beim Historiker sogar angemessen sein, auf das Zuordnen eines fiktiven Wahrheitswertes ganz zu verzichten und ausschliesslich den Begriff Glauben zu verwenden.[4]

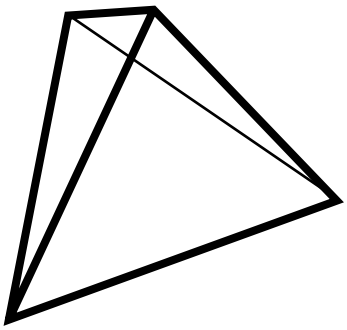
Doch nicht nur die Möglichkeit des Experimentes sollte Anlass dazu sein, hohe Glaubensgrade zu fordern. Weiter besteht auch ein Unterschied zwischen Fächern, in denen es möglich ist, einen Beweis mithilfe der Logik oder der Mathematik zu liefern, und Fächern, denen dieser Zugang nicht möglich ist. Diese Überlegung hat zur Folge, dass auch in Bereichen wie der theoretischen Physik sehr hohe Glaubensgrade zu fordern sind, ehe eine Aussage als wahr anzusehen ist.

Der letzte hier relevante Punkt findet sich in der Beweisführung mithilfe der Logik. So kann eine Konklusion nicht falsch sein, wenn die Prämissen wahr sind. Doch dies sagt nichts über die Wahrheit der Prämissen aus. Ich kann nur mit einem logischen Schema aufzeigen, durch welche Annahmen ich zu meiner Schlussfolgerung komme, und dass diese Schlussfolgerung zwingend sein muss, wenn den Prämissen der Wahrheitswert 1 zugeordnet wird. Durch ein solches Vorgehen bietet sich nun eine gute Methode an, um Theorien falsifizierbar zu machen. Diese Falsifizierbarkeit sollte eine weitere Auswirkung auf den Schwellenwert der Glaubensgrade haben, damit diese zu einem Wahrheitswert von 1 führen.[5]

3 Das Krebsbeispiel ist abgleitet aus einer Ideen von Peter Lipton. Vgl. dazu Lipton (2001) für den subjektiven Kontext vgl. auch Van Fraassen (1980).

4 Zu Überlegungen mit kausalen Abhängigkeiten und Differenztests vgl. (Mill, 1843).

5 Zur Idee der Falsifikation und den damit verbundenen mutigen Theorien vgl. (Popper, 1934).



Konklusion

Zusammenfassend ist festzuhalten, dass der Wahrheitsanspruch, den die Wissenschaft haben sollte, davon abhängt, welchen Wahrheitsanspruch Wissenschaft überhaupt haben kann. Dieser ist durch verschiedene Faktoren begrenzt. Diese Begrenzung kann aber nicht absolut angegeben werden. Schliesslich muss in der Wissenschaft auf Glaubensgrade zurückgegriffen werden. Die Frage, wie sehr wir von einer These überzeugt sind, sollte dabei einzig von der Rechtfertigung dieser These abhängen. Wie diese Rechtfertigung bewertet wird, hängt vom Fachgebiet und dem Kontext ab, in welchem eine Aussage als wahr oder als nicht wahr beurteilt werden muss. Eine detaillierte Antwort ist jeweils nur in Bezug auf die spezifische Fragestellung möglich und nicht global für die Wissenschaft als Ganzes.

Literatur

- Bartelborth, T. (2013). Sollten wir klassische Überzeugungssysteme durch bayesianische ersetzen? Freie Zeitschrift für wissenschaftliche Philosophie, 3.
- Hume, D. (1993[1739-1740]). Eine Untersuchung über den menschlichen Verstand. Felix Meiner. Lipton, P. (2001). Is Explanation A Guide to Inference? A Reply to Wesley C. Salmon. In Hon, G., editor, Explanation : theoretical approaches and applications, pages 93 – 120. Kluwer.
- Mill, J. S. (1973[1843]). A system of logic, ratiocinative and inductive. University of Toronto Press. Popper, K. (2005[1934]). Logik der Forschung. Mohr Siebeck.
- Van Fraassen, B. C. (1980). The scientific image. Clarendon Press.

Jonas Wittwer (28) ist Masterstudent in Wissenschaftsphilosophie und Geschichte im 1. Semester. Er interessiert sich für Wissenschaftsphilosophie und -Geschichte und die Philosophie der Biologie.

Plausibilität der empirischen Unterbestimmtheit von Theorien als Argument gegen den Wissenschaftlichen Realismus

Herauszufinden, was die Welt im Innersten zusammenhält, ist zweifellos der grösste Motivator für die Erkenntnisbemühungen eines jeden (Natur-) Wissenschaftlers, der sich mit Leib und Seele seiner Tätigkeit hingibt. Die Position des wissenschaftlichen Realismus erlaubt es einem jeden, der sie einnimmt, anzunehmen, dass dieser Motivator kein illusorisches Unternehmen darstellt. Gemäss dem wissenschaftlichen Realismus sind etablierte und erfolgreiche wissenschaftliche Theorien wahr und die in diesen Theorien angenommenen, meist unbeobachtbaren Entitäten haben real existierende Bezugsgegenstände. Das Argument der empirischen Unterbestimmtheit von Theorien (im Folgenden das AEUT genannt) stellt den wissenschaftlichen Realismus grundsätzlich in Frage. Ziel dieses Essays ist es, das AEUT als in seiner Plausibilität beschränkt zu entkräften. Dazu soll in einem ersten Teil das AEUT in einer seiner möglichen Formen als Argument dargestellt werden. In einem zweiten Teil werde ich zeigen, dass eine entscheidende Prämisse (und deren Folge) der AEUT unplausibel ist. In einem dritten Teil werde ich dafür argumentieren, dass eine Problemverschiebung in der Diskussion um das AEUT angestrebt werden sollte, und selbst einen Lösungsversuch zu skizzieren wagen. Dieser Versuch soll bloss einen Startpunkt für eine neu orientierte Diskussion bilden und erhebt nicht den Anspruch auf Vollständigkeit oder Unanfechtbarkeit.
Folgende Rekonstruktion des AEUT soll in diesem Essay als Ausgangspunkt der Debatte dienen (in Anlehnung an Beisbart, 2015, S. 3):

SU: Starke Unterbestimmtheitsthese: Zu jeder empirischen Theorie T lässt sich eine Alternativtheorie T' formulieren, die zu T empirisch äquivalent ist, jedoch T logisch widerspricht.

SUf: SU gilt unabhängig von der Grösse, Qualität und Präzision des empirischen Datensatzes.

EI: „Evidential Indistinguishability"-These: Zwischen zwei Theorien kann man nur aufgrund ihrer beobachtbaren Konsequenzen begründetermassen entscheiden.

K: Wir werden niemals begründetermassen entscheiden können, ob T oder T' wahr ist.

Es gelte: Eine Theorie T' ist empirisch äquivalent zu einer Theorie T, wenn sie genau dieselben empirisch überprüfbaren Aussagen enthalten (Beisbart, 2006a, S. 3). Die Konklusion des AEUT ist unvereinbar mit der Position des wissenschaftlichen Realismus. Wenn T und T' zwei sich logisch widersprechende Theorien sind, dann können nicht beide zugleich wahr sein. Der wissenschaftliche Realismus vertritt die Position, dass etablierte und erfolgreiche wissenschaftliche Theorien wahr sind. Gemäss K können wir aber niemals entscheiden, ob T oder T' wahr ist, d.h. der Wahrheitswert von Theorien lässt sich nicht feststellen (siehe auch Wiltsche, 2013, S. 202). Daher ist der wissenschaftliche Realismus gemäss den Vertretern des AEUT falsch.
Das AEUT unterminiert den wissenschaftlichen Realismus und damit auch die Bemühungen von Wissenschaftlern, herauszufinden, was die Welt im Innersten zusammenhält. Abgesehen von der offensichtlichen Schwächung des Motivators wissenschaftlicher Tätigkeit würde das AEUT – wenn es konsequent zu Ende gedacht würde – wohl auch dazu führen, dass Tätigkeiten der empirischen Untersuchung eingestellt würden, weil eingesehen werden müsste, dass keine noch so grosse Menge an empirisch gewonnener Information den Wahrheitswert von Theorien endgültig entscheiden

könnte. Dies ist Grund genug, das AEUT anzugreifen. Mein Angriff auf das AEUT hat zum Ziel, zu zeigen, dass die SU – und im Speziellen die in ihr enthaltene Folge SUf – unplausibel sind. Es ist enorm unplausibel, dass mit der Erweiterung und Präzisierung des empirischen Datensatzes immer noch empirisch unterbestimmt ist, ob T oder T' wahr ist. Daher sind die SU und damit auch die SUf als zu stark zurückweisen. Ich werde zunächst eine historische Argumentation gegen die SU und SUf vorbringen, gefolgt von einer Zurückweisung der Strategie von Vertretern des AEUT, welche mittels Algorithmen oder Funktionen versuchen, Alternativtheorien zu generieren, um so SU und SUf zu beweisen. In der Geschichte der Naturwissenschaften habe es kaum Beispiele von Theoriepaaren T und T' gegeben, die empirisch exakt äquivalent waren, sich aber ansonsten so stark unterschieden, dass sie sich logisch widersprochen hätten (Beisbart, 2006a, S. 5). Die SU erhält also zumindest aus der Wissenschaftsgeschichte wenig Unterstützung. Für die seltenen Fälle von empirisch äquivalenten aber sich logisch widersprechenden Theorien, die wir aus der Wissenschaftsgeschichte kennen, ergab sich zudem bisher stets das Muster, dass sich die empirische Unterbestimmtheit der Theorien mit einer Erweiterung und Präzisierung des empirischen Datensatzes auflöste. So waren zum Beispiel zur Zeit von Kopernikus das geozentrische und das heliozentrische Weltbild empirisch äquivalent. Doch mit präziseren Messinstrumenten und genaueren und umfassenderen astronomischen Untersuchungen erwies sich bloss noch das heliozentrische Weltbild als mit dem erweiterten und präzisierten empirischen Datensatz kompatibel. So lässt sich in der Wissenschaftsgeschichte also auch keine direkte Evidenz für die SUf finden. Natürlich kann der Einwand geltend gemacht werden, dass die Wissenschaftsgeschichte nicht die prinzipielle Möglichkeit ausschliesst, dass sich eines Tages zwei empirisch äquivalente, sich aber logisch widersprechende Theorien entgegenstehen werden, die auch noch nach Jahrzehnten oder Jahrhunderten der Erweiterung und Präzisierung des empirischen Datensatzes den Status der empirischen Äquivalenz aufweisen und somit empirisch unterbestimmt wäre, welche der beiden Theorien wahr wäre. Die Überzeugung, dass auch zwischen zukünftigen empirisch äquivalenten Theorien mit dem Fortschreiten empirischer Untersuchungen begründet entschieden werden könne, findet ihren Ursprung in einem induktiven Schluss, der – da nicht deduktiv gültig – angreifbar ist. An dieser Stelle sei jedoch angemerkt, dass auch wenn die Wissenschaftsgeschichte die Möglichkeit von anhaltender empirischer Unterbestimmtheit bei erweiterten empirischen Datensätzen nicht logisch ausschliesst, sie doch gegen die Plausibilität einer solch andauernden Unterbestimmtheit spricht. Es lässt sich festhalten, dass aus der Perspektive der Wissenschaftsgeschichte sowohl die SU als auch die SUf unplausibel erscheinen.

Da die Wissenschaftsgeschichte keine starke Evidenz für die SU und SUf bietet, sind die meisten Argumente, welche für die Begründung dieser beiden Thesen vorgebracht werden, relativ abstrakter und gesucht konstruierter Natur.[1] So wurde vielfach versucht, einen Algorithmus zu konstruieren, der zu jeder Theorie T eine empirisch äquivalente, T aber logisch widersprechende Theorie T' formuliert. Wenn ein solcher Algorithmus gefunden würde, dann wäre die SU und ihre Folge SUf anzunehmen und das AEUT schlüssig. Kukla hat sich an der Formulierung eines solchen Algorithmus versucht (Kukla, 1998, S. 70f). Zu jeder Theorie T lasse sich eine Theorie T! konstruieren, die besage, dass wann immer eine Beobachtung stattfinde, T wahr sei, aber immer wenn keine Beobachtung stattfinde, T' wahr sei, die mit T logisch unvereinbar sei (siehe auch Wiltsche, 2013, S. 203). Eine ähnliche Strategie verfolgt Quine, wenn er mittels Proxy-Funktionen zu jeder Theorie eine empirisch äquivalente Theorie konstruieren möchte, die von ihr jedoch verschieden ist (Quine, 1990, S. 96). Das Vorgehen besteht in beiden Fällen im Wesentlichen darin, ausgehend von einer Theorie T eine Theorie T' zu formulieren, die eine Modifikation oder eine Funktion der Theorie T darstellt, sodass die empirische Äquivalenz der beiden Theorien garantiert ist, sich T und T' jedoch logisch widersprechen. Mittels solcher Algorithmen oder Funktionen lassen sich immer Alternativthe-

1 Hier ist anzumerken, dass auch einige weniger abstrakte und plausiblere Begründungsversuche vorgebracht wurden, die für die SU und SUf sprechen sollen. So konstruiert van Fraasen eine Theorie T, der gemäss die Newton'sche Physik und das Postulat gilt, dass unsere Sonne relativ zu einem absoluten euklidischen Raum ruht, und eine Theorie T', der gemäss die Newton'sche Physik und das Postulat gilt, dass unsere Sonne sich relativ zu demselben Raum mit einer konstanten Geschwindigkeit v bewegt (van Fraasen, 1980, S. 46f). Ob T oder T' wahr ist, ist empirisch unterbestimmt. Wiltsche schreibt: „[bei solchen T's] handle es sich offenkundig nicht um absichtlich erzeugte Artefakte, sondern um Theorierivalen, denen wir allem Anschein nach auch im naturwissenschaftlichen Alltag begegnen könnten." (Wiltsche, 2013, S. 205. Kursivsetzung von mir.) Ich denke, dass in diesem Satz dem Konjunktiv die grösste Aufmerksamkeit gelten sollte, was die gesamte Aussage stark relativiert. Diese empirische Unterbestimmtheit ergab sich in der Wissenschaftsgeschichte nicht – auch wenn sie theoretisch möglich gewesen wäre. Vielmehr war es tatsächlich aber so, dass Einstein in seiner speziellen Relativitätstheorie, welche die Newton'sche Mechanik ablösen sollte, die Unmöglichkeit, zu entscheiden, ob wir uns mit konstanter Geschwindigkeit relativ zu einem absoluten Raum bewegen oder uns in Ruhe relativ zu diesem befinden, zu einem Axiom erhoben hat. Dieses Wissen um die Information des Bewegungszustandes sei prinzipiell nicht erreichbar, da es in der Welt keine absoluten Bezugspunkte gebe, also alle Bewegungszustände immer relativ zu einem willkürlich gewählten Bezugspunkt (bzw. Inertialsystem) seien.

orien T' formulieren, sodass die Entscheidung zwischen T und T' empirisch unterbestimmt ist. Auch wenn mittels der Erweiterung und Präzisierung des empirischen Datensatzes evt. zwischen T und T' zum Beispiel zu Gunsten von T' entschieden werden könnte, so lässt sich durch den vorgeschlagenen Algorithmus zu T' eine Theorie T" formulieren, sodass die Entscheidung zwischen T' und T" abermals empirisch unterbestimmt ist. Dieses Spiel kann ins Unendliche weitergetrieben werden. Intuitiv möchten wir aber eine solche Argumentationsstrategie nicht akzeptieren. Zurecht, wie ich meine! Es stellt sich die Frage, ob T und T' tatsächlich in einem relevanten Sinne voneinander verschieden sind (Beisbart, 2006b, S. 2). Wir betrachten T' eher als mutwillig konstruierte, artifizielle Pseudotheorie. Zu diesem Eingeständnis wäre wohl auch Kukla bereit, wenn er zulässt, die nach seinem Algorithmus formulierte Theorie T' als eine „instance of logico-semantic trickery" zu bezeichnen (Kukla, 1993, S. 5). Auch Quine gibt zu, dass eine mittels einer Proxy-Funktion konstruierte Theorie T' nicht wirklich als eine eigenständige Theorie aufgefasst werden könne: „We hardly seem to be warranted in calling them [zwei durch eine Proxy-Funktion voneinander abhängige Theorien] two theories; they are two ways of expressing one and the same theory." (Qunie, 1990, S. 96) Laudan und Leplin umschreiben diesen Umstand so, dass T' in diesen Fällen kein „genuiner Rivale" für T sei, da T' parasitisch von T abhänge (Laudan & Leplin, 1991, S. 457). Wir wollen die Verteidigung der SU und SUf durch solche artifiziell konstruierte Pseudotheorien intuitiv nicht gelten lassen. Eine solche Zurückweisung der Argumentation für die SU und SUf setzt jedoch voraus, dass wir Kriterien formulieren, die zuverlässig zwischen artifiziellen Pseudotheorien und genuin entwickelten Theorien zu unterscheiden erlauben. Ich möchte dafür plädieren, dass in der Diskussion um das AEUT eine Problemverschiebung genau auf diesen Aspekt vollzogen werden soll. Im folgenden letzten Teil des Essays möchte ich nun noch versuchen, solche Kriterien skizzenhaft vorzuschlagen.

Damit T' ein genuiner Rivale für T darstellt, also keine artifizielle Pseudotheorie ist, soll T' nicht bloss eine Funktion oder Modifikation von T sein. T' soll nicht parasitisch von T abhängen, indem T' bloss das Resultat eines Algorithmus ist. T' soll eine eigenständige, von Grund auf von T unabhängig aufgebaute und konstruierte Theorie sein.[2] In T' sollen nicht bloss Entitäten aus T durch neue, aber funktional isomorphe Entitäten ersetzt werden. T' soll andere Gesetzmässigkeiten (die jedoch dasselbe Input-Output-Muster aufweisen können) postulieren. T' und T sollten sich nicht bloss in den angenommenen Entitäten voneinander unterscheiden, sondern auch in den Eigenschaften der Entitäten und der Struktur, in der die Entitäten zueinander in Beziehung stehen und so die beobachtbaren Phänomene hervorrufen. Falls wir uns eines Tages wirklich in der Situation wiederfinden würden, in der wir zwischen zwei Theorien T und T' entscheiden müssten, wobei T' die genannten Kriterien erfüllt, um keine Pseudotheorie zu sein, zu T empirisch exakt äquivalent ist, T aber logisch widerspricht, nachweislich der gesamte empirische Datensatz erschöpft ist[3] und zudem gezeigt werden könnte, dass sich immer eine solche Alternativtheorie T' formulieren liesse, dann hätten die SU und SUf und somit das AEUT ein ernstzunehmendes Gewicht. Aber auch dann muss sich der wissenschaftliche Realismus noch nicht geschlagen geben. So liesse sich auch die EI angreifen, wenn sich zeigen liesse, dass auch andere Kriterien als der empirische Datensatz dazu in der Lage wären, zwischen zwei sich widersprechenden Theorien begründetermassen[4] zu entscheiden. Man denke an den vielbesprochenen, für den einen plausiblen, für den anderen unplausiblen Kandidaten der Einfachheit von Theorien. Auch könnte man je nach Überzeugungssystem geneigt sein, im Falle von empirischer Unterbestimmtheit dem Bon Sens der Wissenschaftler zu vertrauen, die wahre von der falschen Theorie zu unterscheiden.

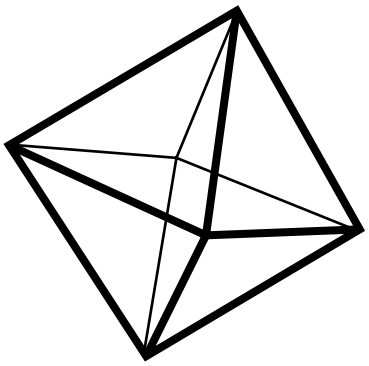
Ich glaube gezeigt zu haben, dass das AEUT ein gültiges und für den wissenschaftlichen Realismus auf den ersten Blick bedrohlich wirkendes Argument sein mag. Doch für seine Schlüssigkeit liegt – besonders in Hinblick auf SU und SUf – bloss sehr schwache Evidenz vor. So möchte ich dafür plädieren, dass die wahre Leistung des AEUT darin zu finden ist, dass Wissenschaftler da-

2 Diesem Kriterium würde beispielsweise die Theorie T', wie van Fraasen sie konstruiert hat, nicht gerecht (van Fraasen, 1980, S. 46f). Die Newton'sche Physik ist eine Theorie, die mit beiden Hilfshypothesen, dass die Sonne ruht oder sich mit konstanter Geschwindigkeit bewegt, kompatibel ist. Zugegebenermassen stellt sich dann aber wieder die Frage, wie entschieden werden soll, welche Hypothesen wesentlich zu einer gegebenen Theorie gehören, und welche Hypothesen bloss Hilfshypothesen darstellen.

3 Allein diese Bedingung lässt wenig Hoffnung auf Erfüllung zu: Wie sollen wir die theoretische und prinzipielle Grösse des möglichen empirischen Datensatzes bestimmen? Ist dieser überhaupt erschöpflich? Bestimmte Entitäten können wir mit unseren Sinnen nicht mehr wahrnehmen, aber ist dann der empirische Datensatz erschöpft? Kann die Messung solcher Entitäten durch indirekte aber hoch präzise Verfahren nicht als eine Erweiterung unserer Sinne verstanden werden? Gibt es einen Grund anzunehmen, dass eine Obergrenze der Messbarkeit von unbeobachtbaren Entitäten existiert? Das wären alles Fragen, die einer Beantwortung bedürften. Es ist fragwürdig, ob eine solche gelingen würde.

4 Begründungen sind im Wesentlicher relativer Natur: Was dem einen als eine Begründung reicht, reicht dem anderen nicht als Begründung hin. Begründungen sind immer relativ zu einem Überzeugungssystem. Die EI gilt im Überzeugungssystem des Empiristen, aber beispielsweise nicht in demjenigen des Rationalisten.

rauf aufmerksam gemacht werden, dass evt. auch andere Theorien mit dem aktuellen empirischen Datensatz kompatibel wären und dementsprechend mehrgleisig geforscht werden sollte. Die Sorge, dass sich – für den Fall, dass genuine Theorien T' tatsächlich formuliert würden – die empirische Unterbestimmtheit mit der Erweiterung und Präzisierung des empirischen Datensatzes nicht auflösen würde, kann getrost aufgegeben oder vertagt werden. In der Diskussion um das AEUT – sollte es durch tatsächlich genuin entwickelte T's wieder an Relevanz gewinnen – muss der Fokus eher auf die Formulierung von Kriterien für ernstzunehmende Alternativtheorien gelegt werden. Die Formulierung solcher Kriterien – welche wohl bloss durch eine Zusammenarbeit von Philosophen und Wissenschaftlern erreicht werden kann – würde einen Fortschritt nicht bloss in der Diskussion um das AEUT, sondern auch in der Wissenschaftsphilosophie und der Wissenschaft allgemein bedeuten.[5]



5 So könnte in der Wissenschaftsphilosophie die Frage, was eine gute Theorie ausmache, besser beantwortet werden. In der Wissenschaft würde der Fortschritt dadurch salient, dass evt. eher empirisch äquivalente, sich aber logisch widersprechende Theorien formuliert würden, welche zu einer fortschrittstreibenden Konkurrenz zwischen den Theorien führen würden.

Literatur

- Beisbart, C. 2006a. „Zur empirischen Unterbestimmtheit von Theorien (I)“ <http://www.claus-beisbart.de/teaching/su2006/realism/underdet.pdf> abgerufen am: 03.12.2015
- Beisbart, C. 2006b. „Zur empirischen Unterbestimmtheit von Theorien (II)“ <http://www.claus-beisbart.de/teaching/su2006/realism/underdet2.pdf> abgerufen am: 03.12.2015
- Beisbart, C. 2015. „Der wissenschaftliche Realismus und Objektivität“, EK Wissenschaftsphi-losophie 2015, Skript zur XIII. Vorlesung, Universität Bern.
- Kukla, A. 1993. „Laudan, Leplin, Empirical Equivalence and Underdetermination“. Analysis. Vol. 53(1), pp. 1-7.
- Laudan, L., Leplin, J. 1991. „Empirical Equivalence and Underdetermination“. Journal of Phi-losophy. Vol. 88(9), pp. 449-472.
- Qunie, W. V. O. 1990. Pursuit of Truth. Cambridge (MA)/USA: Harvard University Press.
- van Fraasen, B. C. 1980. The Scientific Image. Oxford/England: Clarendon Press.
- Wiltsche, H. A. 2013. Einführung in die Wissenschaftstheorie. Göttingen/Deutschland: Vandenhoeck & Ruprecht.

Emmanuel Téofilo Guizar Rosales (21) ist Bachelorstudent in Psychologie und Philosophie des Geistes im 5. Semester. Er interessiert sich für kognitive Psychologie und Neurowissenschaften, Physik, Wissenschaftsphilosophie und die Philosophie des Geistes und der Psychologie.

Frank Jacksons Mary's Room Argument

Das Mary's Room Argument (MRA) wurde von Frank Jackson in seinem 1982 erstmalig veröffentlichtem Paper „Epiphänomenale Qualia" in Form des Argumentes des unvollständigen Wissens (AUW) vorgebracht. Jackson bezeichnet sich selbst als „Qualia-Freund" und möchte dafür argumentieren, dass physikalisches Wissen nicht hinreicht, um bewusste Wahrnehmungszustände vollständig zu erklären.[1] Er lehnt den Physikalismus, also die Auffassung, dass alle Informationen physikalische Informationen seien, ab. Die entstehende Wissenslücke wird mit etwas gefüllt, das Jackson als „Qualia" bezeichnet - eine nicht-physikalische und rein mentale Eigenschaften besitzende Entität, die die persönliche Einsicht in die Wahrnehmung eines mentalen Zustandes beschreibt. Qualia ist nicht in physikalischen Begriffen erklärbar, beziehungsweise in einer starken Interpretation des MRA sogar distinkt von physikalischen Fakten, und beschreibt, wie es ist, etwas wahrzunehmen.[2] Da dieses Erleben, beispielsweise „die Schmerzhaftigkeit eines Schmerzes" oder „die Erfahrung, […] eine Rose zu riechen", nicht in physikalischen Begriffen erklärt werden könne, sei der Physikalismus abzulehnen.[3] Jackson verzichtet jedoch darauf, entscheidende Begriffe wie „physikalische Information" und weitere damit zusammenhängende Ausdrücke näher zu erläutern, da sich die zentralen Probleme des Physikalismus auch unabhängig von dessen genauer Erläuterung zeigen ließen.[4] Kern seiner Auffassung ist, dass bestimmte Aspekte der Wahrnehmung von keiner Menge physikalischer Information beschrieben werde können, sich also nicht auf diese reduzieren lassen. Für diese Auffassung argumentiert Jackson nun mit einem Argument, welches als „Intuition-Pump" dienen soll. Während jedoch einfache intuitive Argumente - wie, dass der Physikalismus falsch sei, da nichts, was aus physikalischer Sicht über eine Rose gesagt werden könne, ihren subjektiven Geruch erkläre - nach Jackson gültig und zum Teil auch überzeugend seien, müsse dennoch ein stärkeres Argument gefunden werden:

„Die Aufgabe der Qualia-Freunde besteht also darin, ein Argument zu finden, dessen Prämissen für alle, oder zumindest möglichst viele, intuitiv offensichtlich sind."[5]

Dieses Vorhaben verfolgt Jackson durch das AUW, während er die intuitive Gültigkeit der Prämissen durch das MRA zu plausibilisieren versucht.

Ziel dieser Arbeit ist es, Jacksons Argumentation zusammenzufassen und einige Einwände zu diskutieren. Dabei wird vor allem auf die Unterscheidung zwischen ontologischer und epistemischer Lesart des Argumentes eingegangen. Eigenständig soll ferner dafür argumentiert werden, dass das Mary's Room Argument als „Intuition-Pump" nur dann überzeugen kann, wenn man bereits die Falschheit des Physikalismus voraussetzt und es Jacksons Anspruch, ein möglichst überzeugendes Argument zu liefern, somit nicht gerecht werden kann. Abschließend soll ein einordnendes Fazit folgen und auf Jacksons aktuelle Position bezüglich seines Argumentes eingegangen werden.

Das Argument des unvollständigen Wissens

Jackson präsentiert das AUW in zwei Ausführungen. Zunächst beschreibt er das Farberlebnis von Fred, einer Person, deren Fähigkeit, Farben zu diskriminieren, weitaus genauer ist als üblich. Fred ist in der Lage, zwei Rottöne zu unterscheiden, die von allen anderen Menschen als identisch wahrgenommen werden. Nun sei es zwar möglich, alles über die Funktionsweise von Freds Gehirn festzustellen und zu erklären, warum er in der Lage ist, diese Farbtöne zu unterscheiden, seine interne Wahrnehmung ließe sich so jedoch nicht erklären. Dieses Erlebnis sei nur dann nachzuvollziehen, wenn das Gehirn einer Person mit normaler Farbdiskriminierungsfähigkeit an Freds Gehirn angepasst oder gar Freds Wahrnehmungsapparat transplantiert werden würde.[6] Doch es standen prinzipiell bereits vor einer solchen Operation alle physikalischen Informationen über die Funktionsweise seines Gehirnes zur Verfü-

1 Jackson, Frank: Epihänomenale Qualia, S. 83. in: Metzinger, Thomas (2006): Grundkurs Philosophie des Geistes. Band 1: Phänomenales Bewusstsein. Mentis: Paderborn.

2 Nida-Rümelin, Martine (2010): «Qualia: The Knowledge Argument», The Stanford Encyclopedia of Philosophy (Summer 2010 Edition).

3 Jackson in Metzinger (2006), S. 84.

4 Ebd., S. 83.

5 Ebd., S. 84.

6 Ebd., S. 86.

gung, dennoch konnte nichts darüber gesagt werden, wie es ist, diese Rottöne unterscheiden zu können. Dieser Aspekt lasse sich nicht durch physikalisches Wissen erklären - der Physikalismus sei daher abzulehnen.[7]
Daraufhin eröffnet Jackson das Mary's Room Gedankenexperiment, auf welchem der Fokus dieser Arbeit liegen soll. Mary, eine brillante Wissenschaftlerin, ist gezwungen, ihr ganzes Leben in einem schwarz-weißen Raum zu verbringen, in welchem sie nie das Erlebnis hat, Farben zu sehen. Gegenstand ihrer Forschung ist jedoch die menschliche Farbwahrnehmung und so gewinnt sie mit der Zeit alle physikalischen Informationen über diese - von der Funktionsweise von Licht und Wellenlängen, bis zur Aufnahme und Verarbeitung dieser im Gehirn und der Projektion des Farberlebnisses. Nachdem sie ihre Forschung abgeschlossen hat und nun über vollständiges Wissen verfügt, ist es ihr gestattet, das schwarz-weiße Forschungsumfeld zu verlassen. Zum ersten Mal erlebt sie, wie es ist, die Farbe Rot zu sehen. Jackson fragt nun, ob Mary durch dieses Erlebnis etwas Neues lernt oder nicht. Denn wenn sie tatsächlich etwas Neues lernt, so muss der Physikalismus falsch sein, da sie bereits vor ihrer Freilassung über alle physikalischen Informationen verfügte. Dieses neu erworbene Wissen ist nach Jackson als „Qualia" zu bezeichnen.[8]

Einwände gegen das Argument des unvollständigen Wissens

Die epistemische und die ontologische Variante des Argumentes

Im folgenden Abschnitt sollen Martine Nida-Rümelins Eintrag im Stanford Encyclopedia of Philosophy folgend die zwei verschiedenen Varianten des MRA ausgearbeitet werden. Die schwächere Variante des MRA bezieht sich auf Marys epistemischen Status bezüglich der Wahrnehmung von Farben:

„(V1) The weaker version of the knowledge argument:

(1a) Mary has complete physical knowledge concerning facts about human color vision before her release.

(2a) But there is some kind of knowledge concerning facts about human color vision that she does not have before her release.

Therefore

(3a) There is some kind of knowledge concerning facts about human color vision that is non-physical knowledge."[9]

Die stärkere Variante bezieht sich dagegen auf den ontologischen Status von Fakten bezüglich der Wahrnehmung von Farben:

„(V2) The stronger version of the knowledge argument:

(1b) Mary knows all the physical facts concerning human color vision before her release.

(2b) But there are some facts about human color vision that Mary does not know before her release.

Therefore

(3b) There are non-physical facts concerning human color vision."[10]

Der zentrale Unterschied zwischen beiden Versionen findet sich in Prämisse 1: Während Mary in der epistemischen Variante über vollständiges physikalisches Wissen verfügt, sind ihr in der ontologischen Variante alle physikalischen Fakten bekannt. In (V1) könnte es der Fall sein, dass alle Fakten bezüglich der Farbwahrnehmung physikalische Fakten sind, das von Mary neu erworbene Wissen jedoch kein physikalisches Wissen ist. Es könnte sich dabei um eine andere (nicht-physikalische) Art des Wissens über die bereits bekannten physikalischen Fakten handeln, welches den „phänomenalen Charakter" der Farbwahrnehmung beschreibe. Mary könne zwar alle Fakten bezüglich des physikalischen Konzeptes der Farbwahrnehmung gekannt haben, nicht jedoch das phänomenale Konzept:
"All proponents of the view point out that, according to their proposal, physical concepts and phenomenal concepts are cognitively independent: it is impossible to see a priori that something that falls under a physical concept of a particular phenomenal character also falls under the corresponding phenomenal concept of that phenomenal character."[11]

7 Ebd., S.87.
8 Ebd., S. 88.
9 Nida-Rümelin (2010).
10 Ebd.
11 Nida-Rümelin (2010).

Jackson scheint jedoch vorauszusetzen, dass - sofern der Physikalismus wahr ist - die Kenntnis physikalischer Fakten Kenntnis über phänomenale Konzepte a priori beinhalte. Ebenso sei es jedoch möglich, dass ein a posteriori Verhältnis bestehe. Deutlich wird dies an dem Beispiel, dass zwar der Fakt, dass H20 die meisten Teile der Erde bedecke, auch den Fakt beinhalte, dass Wasser die meisten Teile der Erde bedeckt, die Identität von H20 und Wasser jedoch keinesfalls a priori ersichtlich ist. Ebenso wäre es möglich, zwar vollständiges Wissen über alle physikalischen Fakten bezüglich der Farbwahrnehmung zu besitzen, den phänomenale Charakter dieser jedoch nicht a priori ableiten zu können.[12] Diese Interpretation ermöglicht, dass Mary zwar etwas Neues lerne, nämlich den phänomenalen Charakter von Farbwahrnehmung, der Physikalismus der Form, dass alle Fakten physikalische Fakten seien, jedoch dennoch zutreffe.

Ein plausibler Intuition-Pump?

In diesem Kapitel möchte ich mich Moti Mizrahi und seiner Kritik an zirkulären (question-begging) und auf Intuitionen beruhenden Argumenten anschließen, um deutlich zu machen, dass das MRA kein konklusives Argument gegen den Physikalismus ist, sondern nur dann als überzeugend gewertet werden kann, wenn bereits anti-physikalistische Intuitionen geteilt werden.[13] Dazu soll zunächst die genaue Argumentationsstruktur des MRA und des daraus folgenden Argumentes gegen den Physikalismus herausgearbeitet werden:

Mary's Room:

P1)	Wenn Mary über alle physikalischen Informationen bezüglich der Farbwahrnehmung verfügt und etwas Neues bezüglich dieser lernt, dann sind nicht alle Informationen physikalisch.
P2)	Mary verfügt über alle physikalischen Informationen bezüglich der Farbwahrnehmung und lernt etwas Neues.
K1)	Nicht alle Informationen sind physikalisch.

beziehungsweise in aussagenlogischer Schreibweise:

P1)	A & N → ~J
P2)	A & N
K1)	~J

Prämisse 1 ist die Voraussetzung der Argumentation und Prämisse 2 ist die entscheidende Intuition des Gedankenexperimentes. Werden beide (zumindest for the sake of the argument) angenommen, so folgt die Konklusion 1. Aus dieser formt sich nun das Argument gegen den Physikalismus:

(II) Das Argument gegen den Physikalismus:

P3)	Wenn der Physikalismus stimmt, dann sind alle Informationen physikalisch.
K1)	Nicht alle Informationen sind physikalisch.
K2)	Der Physikalismus stimmt nicht.

beziehungsweise in aussagenlogischer Schreibweise:

P3)	P → J
K1)	~J
K2)	~P

Prämisse 3 ist hier die Voraussetzung und Konklusion 1 wurde aus dem Gedankenexperiment abgeleitet. Somit wurde gezeigt, dass wenn A&N vorliegt, der Physikalismus falsch ist. Doch überzeugt dieses Argument wirklich - und vor allem überzeugt es jemanden, der Physikalist ist?[14] Um dies zu überprüfen, soll das Argument aus der Perspektive eines Physikalisten betrachtet werden, der zunächst davon ausgeht, dass P wahr ist.

P1*)	P → J
P2*)	P
K1*)	J
P3*)	A & N → ~J

12 Alter, Torin: The Knowledge Argument Against Physicalism. The Internet Encyclopedia of Philosophy.

13 Siehe hierzu Moti Mizrahi in Corfield, David (2013). August 2013. The Reasoner, 7(8), 91–104 und Leng, Mary (2013). Juli 2013. The Reasoner, 7 (1), 1-14.

14 Wie aus Jacksons eingangs erwähnter Formulierung hervorgeht, möchte er „möglichst viele" Überzeugen. Dies kann sich sicherlich nicht auf nicht-Physikalisten beziehen. Warum sollten diese überzeugt werden? Er hat offensichtlich Personen im Blick, die dem Physikalismus zugeneigt sind.

K1) J

K2) ~(A & N)

Angenommen der Physikalismus ist richtig, dann ist auch wahr, dass alle Informationen physikalisch sind und damit muss A&N falsch sein. Ein Physikalist würde also bestreiten, dass es möglich ist, zugleich alle physikalischen Informationen zu besitzen und etwas Neues zu lernen. Zur Übersichtlichkeit sei das Argument nun abgekürzt wie folgt:[15]

P1**) P → ~(A & N)

P2**) P

oder P2**) (A & N)

K1**) ~(A & N)

oder K1**) ~P

Die entscheidende Frage lautet nun, ob hier ein modus tollens oder ein modus ponens angewendet werden sollte. Jackson favorisiert klarer Weise den modus tollens und möchte dies durch die intuitive Zustimmung zum Gedankenexperiment verdeutlichen: Wenn das Gedankenexperiment richtig ist, dann hat Mary alle Informationen und lernt trotzdem etwas Neues dazu. Intuitiv scheint ihm nun das Gedankenexperiment korrekt zu sein, also gilt:

P4) G → A&N

P5) G

K0) A&N

Somit wäre A&N nicht länger eine einfache Annahme, sondern zumindest dann gezeigt, wenn man die Intuition des Gedankenexperimentes teilt. Doch auch hier ergibt sich das gleiche Problem wie in der obigen Argumentation. Warum sollte die Gültigkeit des Gedankenexperimentes angenommen werden? Physikalisten würden ja gerade sagen, dass das Gedankenexperiment nicht gilt, da P → ~(A&N) und damit auch ~G. Als Physikalist ist man weder gezwungen G, noch A&N zu akzeptieren, vielmehr könne man sich ohne Probleme darauf berufen, dass entweder Mary nicht alle Informationen hatte oder sie nichts Neues lernt. Nur wenn man beiden Aussagen intuitiv zustimmt, so ist das gesamte Argument wirklich überzeugend.

Zusammenfassend lautet meinen Kritik, dass das MRA zwar nicht zirkulär ist[16], jedoch dem Anspruch, „möglichst viele" Personen anzusprechen, nicht gerecht werden kann. Gerade jene, die vom Physikalismus ausgehen - und welche es zu überzeugen gilt - sind an keiner Stelle gezwungen, den entscheidenen Annahmen zuzustimmen. Wie beispielsweise Dennett in „Consciousness Explained" (1991) teile ich Jacksons Intuitionen bezüglich Mary nicht.[17] Besonders unplausibel scheint mir die Annahme, dass Mary über alle physikalischen Informationen verfügt. Entscheidend ist hier jedoch die genaue Definition des Terminus „Physikalismus". Owen Flanagan unterscheidet beispielsweise zwischen metaphysischem und linguistischem Physikalismus:

„While metaphysical physicalism is the ontological claim that there are no non-physical individuals, properties or relations and no non-physical facts, linguistic physicalism says that "everything physical can be expressed or captured in the languages of the physical sciences."[18]

Vorstellbar wäre es, dass Mary über alle physikalischen Informationen im Sinne des linguistischen Physikalismus verfügt. Sie kennt alle relevanten Formeln, kennt die Funktionsweise von Wellenlängen und Farbrezeptoren, und erlangt nun nach ihrer Freilassung ein Erlebnis, dass sich nicht durch diese ausdrücken, beziehungsweise auf diese reduzieren lässt. Diesbezüglich teile ich die Intuition, dass ein so verstandener, reduktionistischer Physikalismus durch das MRA zumindest unplausibel wird. Bezüglich des ontologischen Physikalismus' scheint mir die Annahme, dass Mary über alle Informationen verfügt (oder überhaupt verfügen kann), jedoch nicht intuitiv zutreffend. Bei dem Wissen um Farbwahrnehmung könnte es sich um physikalisches Wissen handeln, welches sich nur durch direktes Erleben erlernen lässt und daher von Mary vor ihrer Freilassung gar nicht gewusst werden konnte. Auch wenn dieser Einwand ad hoc ist, schwächt er die Plausibilität einer der grundlegenden Prämissen stark ab und lässt sich daher als gerechtfertigte Kritik am Design des Gedankenexperimentes äußern.

15 Das Argument besitzt nun die gleiche Struktur wie Mizrahis Argument bezüglich der Gültigkeit von Zombie-Argumenten. Siehe hierzu Mizrahi in Corfield (2013), S. 94.

16 Anders als in Mizrahis Argumentation gegen Zombie-Argumente gehe ich nicht davon aus, dass ein Hilfsargument mit der Prämisse „das Antezendez von P1**) ist falsch" zur Begründung des modus tollens herangezogen wird. Siehe hierzu Mizrahi in Corfield (2013), S. 95.

17 Siehe hierzu Mizrahi in Leng (2013), S. 6.

18 Nida-Rümelin (2010).

Zusammenfassung

Insgesamt ist also wichtig zu betonen, dass eine genaue Explikation sowohl des Gedankenexperimentes selbst, als auch der zugrundeliegenden Debatte um die Kernaussage des Physikalismus unerlässlich ist. Ein intuitives Verständnis vorauszusetzen, scheint aufgrund der oben angesprochenen Schwierigkeiten wenig gewinnbringend. Zum einen da gezeigt werden konnte, dass durchaus mit dem Physikalismus kompatible Lesarten des MRA existieren und zum anderen, da sich ein ontologischer Physikalismus durch ein Gedankenexperiment prinzipiell kaum widerlegen lässt. Wenn dieser (grob vereinfacht) aussagt, dass alles für X relevante physikalisch ist, so ist ein Gedankenexperiment, welches plausibilisiert, dass es etwas für X relevantes und nicht-physikalisches gibt, stets dem Einwand ausgesetzt, dass es sich bei dem Gedankenexperiment schlicht um eine Fehlvorstellung handelt. Ebenso wie man sich nicht den Morgenstern als distinkt vom Abendstern vorstellen kann, da beide tatsächlich die Venus sind, so lassen sich keine nicht-physikalischen ontologischen Entitäten vorstellen. Mizrahi erläutert diesen Punkt bezüglich der metaphysischen (Un-)Möglichkeit von Zombies:

„Just as it is a misconception that the morning star is distinct from the evening star, since they are one and the same thing, i.e., the planet Venus, it is also a misconception that the mental is distinct from the physical, since they are also one and the same thing."

Daher scheint das MRA mehr zur eng verbundenen Debatte um Reduktionismus als zur eigentlichen Qualia-Debatte beizutragen. Während Jackson gegen den linguistischen (reduktionistischen) Physikalismus ein durchaus intuitives und konklusives Argument formulieren konnte, so scheint er auf ontologischer Ebene, auf welcher sich eine Entität wie „Qualia" verordnen ließe, wenig beizusteuern.[19] Dennoch muss anerkannt werden, dass die in dieser Arbeit vorgestellten Aspekte nur Teile der größeren Debatte um das MRA sind und daher ein endgültiges Fazit nicht möglich ist.

Abschließend sei erwähnt, dass Jackson die Gültigkeit des Gedankenexperimentes mittlerweile selbst bezweifelt. In Mind and Illusion (2003) argumentiert er für eine repräsentationalistische Theorie der Sinneswahrnehmung, bei welcher Qualia Repräsentationen mentaler (neuronaler) Zustände seien - eine Position, die mit einem ontologischen Physikalismus vereinbar ist.[20] Diese findet sich auch bei David Lewis, der davon ausgeht, dass „Qualia" oder „knowing what is is like to…" keine Form faktischen Wissens, sondern eine praktische Fähigkeit sei. Es handle sich bei Marys epistemischem Fortschritt also nicht um propositionales (knowing-that) Wissen, sondern nicht-propositionales (knowing-how):

„Knowing what it's like is the possession of abilities: abilities to recognize, abilities to imagine, abilities to predict one's behavior by means of imaginative experiements."[21]

Lewis begründet diese Ability-Hypothesis dadurch, dass sie mit dem Physikalismus vereinbar sei und über genauso viel Erklärungskraft wie ihre nicht dem Physikalismus vereinbare Gegenhypothese verfüge, die Hypothesis of Phenomenal Information. Nach dieser handle es sich bei Qualia um propositionales Wissen, welches durch die Elimination alternativer Möglichkeiten entstünde.[22] Aufgrund der Kompatibilität mit dem Physikalismus sei jedoch die Ability-Hypothesis vorzuziehen, beziehungsweise die Intuition des MRA entkräftet. Auf dieses wie auf andere Argumente kann jedoch im Rahmen dieses Essays nicht weiter eingegangen werden.

19 Dass Jackson sich auf ontologischer Ebene bewegt, wird beispielsweise auch von Nida-Rümelin (2010) bestätigt: „Although Jackson›s original formulation in terms of information is open to both interpretations it is clear that the second stronger version is what he had in mind".

20 Lycan, William, «Representational Theories of Consciousness», The Stanford Encyclopedia of Philosophy (Summer 2015 Edition).

21 Lewis, Daniel (1983): Mad Pain and Martian Pain. Oxford University Press, S. 131.

22 Nida-Rümelin (2010).

Literatur

Alter, Torin: The Knowledge Argument Against Physicalism. The Internet Encyclopedia of Philoso- phy, ISSN 2161-0002, http://www.iep.utm.edu/know-arg/, zuletzt aufgerufen am 29.02.2016

Corfield, David (2013). August 2013. The Reasoner, 7(8), S. 91–104.

Leng, Mary (2013). Juli 2013. The Reasoner, 7(1), S. 1-14.

Lewis, Daniel (1983): Mad Pain and Martian Pain. Oxford University Press, S. 122-132.

Lycan, William, "Representational Theories of Consciousness", The Stanford Encyclopedia of Phi- losophy (Summer 2015 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2015/entries/consciousness-representational/>.

Metzinger, Thomas (2006): Grundkurs Philosophie des Geistes. Band 1: Phänomenales Bewusst- sein. Mentis: Paderborn.

Nida-Rümelin, Martine (2010): "Qualia: The Knowledge Argument", The Stanford Encyclopedia of Philosophy (Summer 2010 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2010/entries/qualia-knowledge/>.

Sebastian Drosselmeier (24) ist Masterstudent in Political, Legal, and Economic Philosophy (PLEP) im 1. Semester. Er interessiert sich für politische Philosophie, (Meta-)Ethik und Wissenschaftsphilosophie.

Qualia vs. Physikalismus / Funktionalismus

In der Philosophie des Geistes gibt es ein sehr bekanntes Problem, welches sich mit der Interaktion zwischen dem menschlichen Geist mit seinem biologischen Gefäss, dem Körper, beschäftigt: das Leib-Seele-Problem. Die Frage dabei ist, ob die beiden Teile des menschlichen Daseins miteinander interagieren, und wenn ja, wie sie dies tun. Ich denke es leuchtet ein, dass mentale Zustände wie Gedanken oder Gefühle ohne ein biologisch intaktes Gehirn nicht zustande kommen könnten. Andersherum ist es jedoch nicht ganz so einleuchtend, wenn auch nicht undenkbar, dass ein biologisch intaktes Hirn geben könnte, dass keine mentalen Zustände besitzt. Das beste Beispiel dazu ist der klinische Hirntod, bei dem das Gehirn mechanisch noch mit Sauerstoff beliefert wird, man jedoch davon ausgehen kann, dass die Person kein aktives Bewusstsein mehr hat.

Vom Leib-Seele-Problem ausgehend entfaltet sich eine Debatte, die sich den Grundfragen des Problems widmet. Dabei sind zwei Positionen fundamental. Auf der einen Seite steht der Monismus, auf der anderen der Dualismus. Im Folgenden wird jedoch nur die monistische Position beleuchtet. Diese enthält wiederum zwei verschiedene Positionen. Jene der monistischen Theorien, die sich mit der Seite der Seele beschäftigen. Diese werden hier aber ebenfalls keine Rolle spielen. Dann gibt es monistische Theorien, die sich auf die Leibesseite fokussieren. Auf dieser Seite steht der Materialismus, bzw. der Physikalismus. Der Physikalismus sucht nach Lösungen, die auf rein physikalischen Eigenschaften basieren und durch sie erklärbar sind. In diese Sparte gehören daher auch Neurowissenschaftler, die z.B. das Phänomen des Schmerzes mit einer Antwort wie „das Feuern von C-Fasern" erklären könnten. Diese Herangehensweise ist der Tradition der Physik nach sehr objektiv geprägt, daher gelten diese Gesetzmässigkeiten bei jedem Menschen gleichermassen. Aus dem Physikalismus heraus entstand eine Art Unterkategorie, die auf demselben Prinzip beruht: der Funktionalismus. Die Grundidee ist dieselbe. Mentale Zustände sind hier allerdings nicht spezifisch neurologische Zustände, sondern an und für sich funktionale. Damit öffnet sich die Möglichkeit für weitere Annahmen, dass z.B. Computer, sofern sie funktional gleich organisiert sind wie ein menschliches Gehirn, ebenfalls mentale Zustände und somit Gefühle oder Gedanken realisieren können. Diese materialistischen Sichtweisen auf das Leib-Seele-Problem haben einige Kritik hervorgerufen. Anlass dafür geben unter anderem die sogenannten 'Qualia'. Am einfachsten umschrieben hat Thomas Nagel diese mit dem Ausdruck „what it is like" aus seinem bekannten Aufsatz „What is it like to be a bat?". Qualia meinen den subjektiven Erlebnischarakter mentaler Zustände. Das entscheidende Prädikat in diesem Satz ist „subjektiv". Das soll bedeuten, dass wenn eine Person A Schmerzen empfindet, dann erlebt sie diese auf eine Weise B. Und wenn eine Person C Schmerzen empfindet, dann erlebt sie diese auf die Weise D. Das Problem in der Qualiadebatte ist, dass weder die Person A, noch die Person C wissen können, wie es sich für die jeweils andere Person anfühlt jene Schmerzen zu empfinden.

Zusammengefasst lässt sich bereits der grundlegende Unterschied der beiden Konzepte deutlich erkennen: objektiv vs. subjektiv. Während der Physikalismus, sowie der Funktionalismus, bestrebt ist, mentale Zustände und Erlebnischarakter auf die Ebene der Objektivität und physikalischen Tatsachen zu reduzieren, versucht das Qualia-Konzept einen Zwischenraum für die Subjektivität zu schaffen. Charakteristisch für diesen subjektiven Raum ist allerdings, dass er naturwissenschaftlich nicht erklärbar ist. An dieser Stelle steht eine Erklärungslücke. Aber nicht nur diese Erklärungslücke lässt die Herangehensweise an das Leib-Seele-Problem lückenhaft erscheinen. Verschiedene Gedankenexperimente des Qualia-Konzeptes greifen insbesondere die weiterführenden Theorien des Funktionalismus an.

Im Folgenden werde ich zwei Argumente präsentieren, welche die Ansicht stützen, dass der subjektive Erlebnischarakter von mentalen Zuständen nicht auf physikalische Eigenschaften reduzierbar ist. Als

Erstes widme ich mich drei Kernpunkten aus Joseph Levines Argumentation zur Erklärungslücke. Als Zweites werde ich ein Gedankenexperiment von Ned Block vorstellen, welches sich der Spektrumsinversion widmet. Dabei versuche ich herauszuarbeiten, ob und wie erfolgreich diese beiden Argumente den Physikalismus bzw. Funktionalismus tatsächlich angreifen. Dazu wer-

de ich bei der Diskussion von Blocks Gedankenexperiment auch David J. Chalmers' Gegenargumente miteinbeziehen.

Die Erklärungslücke

Levine beschäftigt sich in seinen beiden Essays „Warum der Physikalismus Qualia auslässt"' und „Phänomenale Begriffe und der Materialismus" ausführlich mit der Erklärungslücke. Weil das Thema sehr ausführlich diskutiert ist, möchte ich im Folgenden drei Kernpunkte aus den beiden Texten herausgreifen und genauer besprechen.

Etwas, das als unbestreitbar angenommen werden muss, ist, dass unser Körper und alles was dazu gehört auf physikalischen Prinzipien beruht. Aufgrund moderner wissenschaftlicher Erkenntnis sollte dies allerdings nicht sonderlich schwierig fallen. Aufbauend auf diesem physikalischen Grundriss des menschlichen Körpers lässt Levine die Frage aufkommen, wie es denn sein kann, dass ein physikalisches System überhaupt in der Lage sei, etwas physikalisch nicht Erklärbares wie Qualia hervorzurufen[1]. Dieses Hauptproblem nennt Levine „Erklärungslücke". Der Grund für die physikalische Unerklärbarkeit von Qualia, so Levine, liegt im physikalischen Prinzip der Reduktion. Anders als beim bekannten Wasser/H2O-Beispiel[2] ist die Reduktion angewandt auf das Qualia Konzept nicht vollziehbar.

Levine unterstreicht diesen letzten Satz mehrmals in seinen Ausführungen. Die Frage, die sich mir jedoch stellt, bezieht sich auf den wissenschaftlichen Fortschritt. Ist es denn nicht denkbar, dass in ferner Zukunft Messgeräte entwickelt werden können, welche beim Vorgang des Feuerns von C-Fasern messen könnten, dass ein anderer Stoff, z.B. tba1, ausgeschüttet wird. Bei einem zweiten Messgang wird entdeckt, dass es noch viele weitere tba-Stoffe gibt. Daraus könnte man schliessen, dass beim Feuern von C-Fasern unterschiedliche Stoffe ausgeschüttet werden, welche das subjektive Erleben des Feuerns von C-Fasern darstellen. Und wenn man diese Stoffe entdecken würde, wäre die Reduktion durchführbar. Meines Erachtens ist diese Vorstellung zumindest denkbar. Schliesslich wurden mit der fortschreitenden Technik auch die Neuronen und deren Funktionen entdeckt. Weshalb sollte dies nicht auch für die Qualia möglich sein? Natürlich bleibt hier, so auch Levine, die Frage nach dem Warum noch bestehen. Warum sollten gerade diese tba-Stoffe für das subjektive Erleben mentaler Zustände verantwortlich sein? Aber könnte nicht auch hier der Fall eintreffen, dass das subjektive Erleben eine Funktion erfüllen soll, genauso wie z.B. eine Schmerzempfindung einen Schutzmechanismus darstellt? Dann könnte, in diesem Szenario, das subjektive Erleben von Schmerz als eine weitere Schutzstufe verstanden werden. Natürlich sind diese Überlegungen rein hypothetisch und vielleicht gar nicht realisierbar. Aber dennoch denkbar und daher, meines Erachtens, nicht ganz so unmöglich wie es Levine darstellt.

Einen zweiten Punkt in der Verteidigung von Qualia gegenüber dem Physikalismus macht Levine mit seinem Beispiel von philosophischen „Zombies"[3]. Levine meint, dass es doch vorstellbar sei, dass es menschliche Duplikate von uns geben könnte, welche physikalisch genau gleich aufgebaut seien wie wir. Der einzige Unterschied allerdings läge darin, dass diese „Zombies" nicht im Stande wären, mentale Zustände mit subjektivem Erlebnischarakter zu haben, obschon ihre C-Fasern genau gleich feuern wie die unseren. Durch dieses Gedankenexperiment und dadurch, dass es denkbar sei, sieht Levine nun die Erklärungslücke bestätigt. Seine Argumentation lehnt sich hier derjenigen der Fehlenden Qualia an, welche am deutlichsten durch das „Chinese Brain"-Gedankenexperiment von Ned Block[4] dargestellt wurde. Dem Experiment zufolge sei es nicht notwendig logisch, wenn eine Gruppe von Menschen die funktionale Organisation eines menschlichen Gehirns simuliert, sie tatsächlich dadurch ebenfalls ein gemeinsames Bewusstsein mit subjektiven Erlebnissen realisieren können.

Levines Schlussfolgerung führt dahin, dass der Physikalismus dennoch nicht aufzugeben sei. Das Problem der Erklärungslücke bestehe aber forthin, da der Physikalismus nicht im Stande sei, die Lücke zu schliessen. Bei dieser Schlussfolgerung leuchtet mir wenig ein, wieso Levine sich solch eine Mühe gibt, das Prinzip der Erklärungslücke auf jegliche Weise auszutesten, wenn er am Schluss doch nicht bereit ist, den Physikalismus als unzureichend abzustempeln. Die Tatsache, dass der Physikalismus momentan eine solche grosse, unüberwindbare Lücke aufweist, sollte doch beweisen, dass, solange nicht solche Messinstrumente existieren, wie oben beschrieben, der Physikalismus schlicht unzureichend ist.

Invertierte Erde

Im Essay „Invertierte Erde" konstruiert Ned Block ein Gedankenexperiment zur Spektrumsinversion, das seine Unterscheidung zwischen intentionalen und qualitativen Gehalten rechtfertigen und die funktionalistische Theorie qualitativen Gehalts widerlegen soll.

1 Levine (2001), S. 63
2 Levine (2001), S. 42ff
3 Levine (2002), S. 112
4 Erwähnt in Chalmers, Kapitel 2, Abschnitt 2

Die Argumentation um invertierte Spektren ist eine direkte Antwort auf den Funktionalismus und seine Behauptung, dass mentale Zustände funktionale Zustände seien. Invertierte Spektren widerlegen diese Behauptung, da sie beinhalten, dass zwei Personen zwei verschiedene mentale Zustände haben können, die exakt dieselbe funktionale Rolle spielen. Wenn also eine Person A einen roten Gegenstand der Aussenwelt wahrnimmt, hat sie den mentalen Zustand von Rot. Gleichzeitig könnte es aber eine Person B geben, die den gleichen roten Gegenstand wahrnimmt, aber den mentalen Zustand von Grün hat. Beide haben ein Rot-Erlebnis mit qualitativem Gehalt. Durch diese Inversionshypothese unterscheidet Block nun die zwei Arten des Gehalts von Erlebnissen[5]. Einerseits die Art und Weise wie das Erlebnis die Welt repräsentiert, intentionaler Gehalt genannt. Und andererseits wie es ist, dieses Erlebnis zu haben, qualitativer Gehalt genannt[6]. An einem Beispiel kurz verdeutlicht: Person A und Person B, beide funktional gleich organisiert, haben ein Erlebnis mit qualitativem Gehalt. Der intentionale Gehalt des Erlebnisses der Person A ist derselbe wie derjenige der Person B. Der Unterschied liegt einzig darin, dass die Person A den qualitativen Gehalt von Rot hat, während die Person B den qualitativen Gehalt von Grün hat[7]. Des Weiteren unterscheidet Block die beiden Arten des Gehalts von Erlebnissen dadurch, dass zwar jeder einen Begriff für Rot somit für sich auf der qualitativen Ebene besitzt, diese Unterscheidung im Begriff aber keinen Einfluss darauf hat, ob etwas tatsächlich Rot für eine der Personen erscheint. Folglich muss der intentionale Gehalt muss immer referenzabhängig sein. Das, was einen Begriff rot macht, ist seine Relation zu roten Dingen in der Aussenwelt. Somit ist es der qualitative Gehalt, worin sich die Erlebnisse der beiden Personen unterscheiden können[8].

Block glaubt nicht daran, dass Inversion möglich ist, obwohl er zugibt, dass er nicht beweisen kann, dass es sie nicht gibt[9]. Könnte es aber nicht möglich sein, dass bei so gut wie jedem Menschen eine Art Spektrumsinversion vorhanden ist und eine Farbwahrnehmung darum für jeden komplett anders ist? Dass keiner den gleichen qualitativen Gehalt von Rot besitzt wie ein anderer? Schliesslich haben wir alle andere Farbpräferenzen oder sogar Lieblingsfarben. Könnte es dann nicht sein, dass wir gerade deshalb eine Zu- oder Abneigung gegen eine bestimmte Farbe haben, im Gegensatz zu einer anderen Person, weil wir verschiedene qualitative Erlebnisse zu dieser Farbe haben? Wenn die Lieblingsfarbe der Person A nun Rot wäre und sie alles, was sie nur kann, in dieser Farbe gestaltet und die Person C deshalb nicht gerne zu Person A zum Tee vorbeischaut, weil die rote Wohnung der Person A für ihn als Giftgrün erscheint. Aber auch in diesem Szenario, und da hat Block recht, können wir ganz einfach nicht wissen, ob es so ist oder nicht.

Kommen wir nun zu Blocks Hauptargument, dem Gedankenexperiment der invertierten Erde. Die invertierte Erde unterscheidet sich von unserer Erde durch invertierte Farbausdrücke und invertierte Stimuli[10]. Das soll heissen, dass alle Farben auf der Invertierten Erde die Komplementärfarbe dessen besitzen, wie sie auf unserer Erde erscheinen. Somit besitzen Dinge in der Aussenwelt wirklich die Komplementärfarben. Der Himmel ist gelb, das Gras ist rot und die reife Erdbeere grün. Wie bereits erwähnt ist auch das Vokabular der Inventierten Erde invertiert. Sie benützen die gleichen Farbbegriffe wie wir für ihre Farben. Somit ist ihr gelber Himmel ist für sie „blau", wie unser blauer Himmel für uns „blau" ist. Daraus folgt, dass auch die intentionalen Gehalte der Einstellungen und Erlebnisse der Bewohner der Invertierten Erde invertiert sind. In Blocks Experiment wird eine Person unserer Erde entführt. Ihr wird von einer Gruppe von Wissenschaftlern eine invertierte Farblinse eingesetzt. Zum Schluss wird die Person mit ihrem Äquivalent auf der Invertierten Erde getauscht. Für die entführte Person hat sich farbentechnisch nichts geändert, da er durch die Linse die Komplementärfarben der invertierten Erde immer noch so sieht, wie er es auf der Erde getan hat. Blocks Argument liegt nun darin, dass die entführte Person an ihrem ersten Tag intentionale Gehalte invertiert wahrnimmt, ohne es zu wissen, und dennoch den gleichen qualitativen Gehalt besitzt, dank der invertierten Linse. Die Person liegt aber an sich falsch darin, weil sich sein qualitativer Gehalt „rot" auf den intentionalen Gehalt „grün" bezieht. Zumindest am Anfang. Blocks Argument jedoch sieht nun vor, dass sich nach 50 Jahren dieser Irrtum auflöst, weil sich über die Zeit durch die stetig invertierten intentionalen Gehalte auf der invertierten Erde auch die funktionalen Zustände invertieren, da der Zustand der zuvor durch rote Dinge in der Aussenwelt in ihnen hervorgerufen wurde, nun durch grüne Dinge der Aussenwelt hervorgerufen wird[11].

5 Block, S. 244

6 Der Einfachheit halber benütze ich in diesem Zusammenhang nur einen der von Block jeweils zwei vorgeschlagenen Begriffen.

7 Block, S. 245

8 Block, S. 247

9 Block, S. 249

10 Block, S. 258

11 Block, S. 260

Mit dieser Schlussfolgerung sieht Block nun die funktionalistische Theorie qualitativen Gehalts widerlegt und die Unterscheidung zwischen intentionalen und qualitativen Gehalten gerechtfertigt. Gegen Ende seines Essays geht Block auf einige Einwände gegen sein Invertierte-Erde-Gedankenexperiment ein und diskutiert diese. Einen Aspekt lässt er allerdings unbesprochen. Denn mir stellt sich hier die Frage, ob sich wirklich etwas ändert über diese 50 Jahre. Die entführte Person trägt ja immer noch diese invertierte Linse und hat somit immer noch nicht gemerkt, dass er oder sie auf der Invertierten Erde ist. Diese Person sieht die Aussenwelt also nicht plötzlich in den Komplementärfarben, sondern immer noch wie gewohnt. Sie hat sich also nicht daran gewöhnt, über die Jahre die Farbe Grün als Rot zu identifizieren, da sie die invertierte Farbe Grün ja nie als Grün gesehen hat. Von aussen her betrachtet, nehmen wir an, die Wissenschaftler, die diese Person entführt haben, beobachten diese Person immer noch, auch nach 50 Jahren, könnte man argumentieren, dass der funktionale Zustand invertiert ist. Für die entführte Person jedoch hat sich nie irgendetwas verändert und somit ruft für ihn immer noch die Farbe Rot den qualitativen Gehalt Rot hervor.

David J. Chalmers ist von Blocks Widerlegungsversuch des Funktionalismus nicht ganz überzeugt. Chalmers argumentiert in seinem Essay zur organisatorischen Invarianz, dass das Argument zur Invertierten Qualia, oder eben Spektrumsinversion, empirisch unmöglich sei. Dazu skizziert er eine leicht abgewandelte Variante zur Spektrumsinversion. Eine Person A und eine Person B sind funktional isomorph[12], mit dem einzigen Unterschied, dass Person B einen Teil ihres neuronalen Systems durch einen Silikonkreislauf ersetzt bekommen hat. Dies macht sich dadurch bemerkbar, dass für Person A eine Erdbeere wie gewohnt als rot, und somit mit einem roten qualitativen Gehalt, wahrgenommen wird, für die Person B aber als blau. Für das Beispiel der Tanzenden Qualia, wie er dieses Experiment nennt, wird bei der Person A ebenfalls ein solcher Silikonkreislauf eingesetzt. Sein natürliches neuronales Systems wird dadurch jedoch nicht ersetzt, sondern mittels eines Schalters so eingerichtet, dass durch Umlegen des Schalters entweder das natürliche System oder das silikonbasierte System eingeschaltet werden kann. Wird bei der Person A nun der Schalter umgelegt, wird auch diese die Erdbeere als blau wahrnehmen. Wenn nun der Schalter mehrmals hin und zurück umgelegt wird, „tanzt" diese Rot-Blau-Wahrnehmung wahrlich „vor den Augen" der Person A. Allerdings bemerkt dies die Person A gar nicht, weil der funktionale Zustand des silikonbasierten Systems genau gleich wie der des neuronalen Systems funktioniert. Darum ist die Tanzende Qualia nicht ausreichend, um den Funktionalismus auszuhebeln und dasselbe gelte daher für die Invertierte Qualia[13]. Aber auch Chalmers Ansicht ist nicht wirklich überzeugend. Sie erklärt nämlich nicht, wie etwas Physikalisches ein durch die Physik nicht erklärbares Phänomen hervorrufen kann.

Fazit

Blicken wir kurz auf die drei Autoren zurück, die sich, mehr oder weniger, alle mit der Frage beschäftigen, ob Qualia nun mit dem Funktionalismus kompatibel sind oder nicht. Während Levine und besonders Block eher dagegen argumentieren, ist Chalmers klar dafür. Alle drei haben meines Erachtens aber deutlich gezeigt, dass keine der drei Argumentationen, sei es für oder gegen Kompatibilität, eine alles umfassende Lösung bietet. Levine argumentiert deutlich gegen den strikten Physikalismus. Aber auch er hat Erklärungslücken in seiner Theorie. Er selbst argumentiert oft mit dem „es ist denkbar – daher nicht unmöglich", behauptet jedoch selber, dass gewisse Aspekte, wie die Reduktion des subjektiven Erlebnischarakters, unmöglich sind. Ich persönlich kann mir aber eine Alternative dazu vorstellen, daher sehe ich nicht ein, wieso es dann unmöglich sein sollte. Wenn ich nun selbst ein kleines Gedankenexperiment wagen dürfte: Wieso ist es so unmöglich, dass z.B. in der Zukunft ein Messinstrument entwickelt wird, das imstande ist bei der Schmerzempfindung im Gehirn die genauen Vorgänge aufzudecken? Sodass man schlussendlich sagen kann, dass wenn eine Person A Schmerzen empfindet, die C-Fasern in ihren Neuronen feuern und gleichzeitig die Kombination der Stoffe tba1 und tba2 ausschütten. Die Kombination von tba1 und tba2 fühlt sich auf die Weise B an. Die Versuche mit 1000 Probanden haben ergeben, dass jedes Mal, wenn diese Kombination ausgeschüttet wird, die Probanden es auf die Weise B empfinden. Für die Weise B werden ebenfalls durch die Untersuchungen mit 1000 Personen Begriffe entwickelt, die jeder Mensch, sobald er die Weise B erfährt (evtl. wird sie sogar durch Impulse künstlich erzeugt), versteht und auch ab sofort anwenden kann. Meines Erachtens ist diese Vorstellung denkbar, da die Wissenschaft es schliesslich auch schon bis in die molekulare Ebene geschafft hat. Wieso sollte sie es nicht noch weiter schaffen? Da dies aber alles bloss ein gedankli-

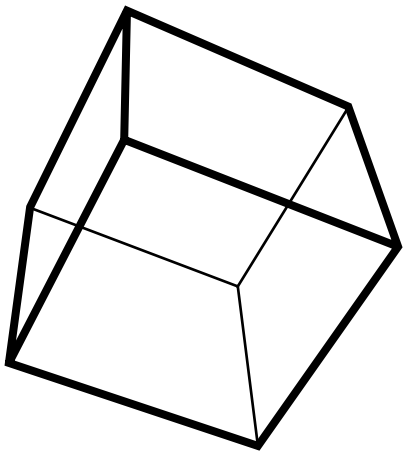
12 Heisst hier, dass sie funktional gleich organisiert sind.

13 Chalmers, Kapitel 4, Abschnitt 25

ches Konstrukt ist, müssen wir uns momentan wohl mit der Erklärungslücke von Levine abfinden. Auch bei Block gibt es Schwachstellen in seiner Theorie gegen den Funktionalismus. Zu Einigen nimmt er selbst Stellung in seinem Essay, und die kann er klar verwerfen; andere können durch beispielsweise Chalmers Tanzende Qualia ausgeschaltet werden. Diese Untersuchung diverser Argumente für oder gegen die Kompatibilität von Qualia und Funktionalismus hat schlussendlich bloss gezeigt, dass niemand eine befriedigende Antwort zu bieten hat.

Literatur

- Block, Ned (2001): Invertierte Erde, in: Heckmann/ Walter (Hrsg.): Qualia. Ausgewählte Beiträge. Paderborn: Mentis
- Chalmers, David (1995): Absent Qualia, Fading Qualia, Dancing Qualia, in: Metzinger, Thom-as (Hg.), Conscious Experience, Imprint Academic
- Levine, Joseph (2001): Warum der Physikalismus Qualia ‚auslässt'. in: Heckmann/ Walter (Hrsg.): Qualia. Ausgewählte Beiträge. Paderborn: Mentis
- Levine, Joseph (2002): Materialismus und Qualia. In: Pauen, Michael/ Stephan, Achim (Hrsg.): Phänomenales Bewusstsein-Rückkehr zur Identitätstheorie? Paderborn: Mentis



Sina Rast (23) ist Bachelorstudentin in Geschichte und Philosophie im 9. Semester. Sie interessiert sich für Ethik, philosophische Gedankenexperimente und nordamerikanische Studien.

Rationality and the Self-Attribution of Experiences

Advocates of the rationality account of self-knowledge hold that we often self-attribute mental states because it would be irrational not to do so. Two of the main proponents of this view, Gallois (1996) and Shoemaker (1996), extend the rationality account to include perceptual experiences. In this paper I show not only that this extension is problematic, but I also argue for the more extreme position that it is possible that a rational person never self-attributes perceptual experiences. In the first part, I introduce the rationality account of perceptual experiences and present the arguments of Gallois and Shoemaker in favour of this account. Their arguments rest on the premise that if people make appearance-reality distinctions, they self-attribute experiences. In using Austin's (1962) idea that appearance-statements are normally open to public confirmation, I argue in the second part, that many appearance-statements are not based on self-attributions of experiences. In the third part, I employ two examples - mirages and the moon illusion – to demonstrate that the intentions of a speaker determine whether an appearance-statement is intended to be a self-attribution of an experience, and that in principle, a rational person can always avoid self-attributing an experience.

The rationality account of self-knowledge

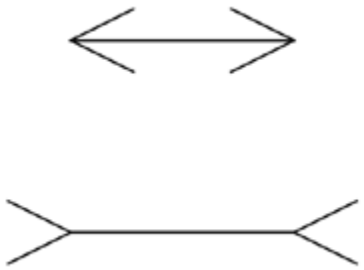
The rationality account of self-knowledge states that rationality requires of us to self-attribute mental states.[1] In order to illustrate the idea, take the following example from Gallois (1996, p.102): Imagine that yesterday you held the belief that life was roughly six thousand years old. But then you have a class on earth history and the evidence that the teacher presents, makes you believe that the world is in fact billions of years old. If you were not self-attributing beliefs, it would occur to you that today earth is billions of years old, whereas yesterday earth was six thousand years old. It seems therefore, that we have no choice but to self-attribute beliefs in order to avoid bizarre pictures of the world, e.g. a picture according to which the age of the world changed from one day to the next. According to Gallois, we are also faced with bizarre pictures of the world if we do not attribute experiences to ourselves. We can see what Gallois has in mind, by considering Sellars's (1956) famous example of John the shopkeeper, who looks at a blue necktie but which appears to be green inside the shop where it is illuminated by electric light. John offers two preliminary solutions to this puzzle: First, he states that the tie "was green in there, but now it is blue." (1956, p.37) Sellars is quick to point out that "neckties don't change their colour merely as a result of being taken from place to place." (ibid.) Second, John suggests that "electricity changes their colour and they change back again in daylight", but, so the response, "that would be a queer kind of change" (ibid.). It seems that the only way to avoid these two rather bizarre pictures of the world, is to distinguish the appearance of the colour of the tie from its reality, and hence, acknowledge that we have experiences that make things appear a certain way that do not necessarily correspond with reality. Thus, John should report this state of affairs by saying that although the tie is blue, it appears to be green. There can be no doubt that people often make appearance-reality distinctions, like the Müller-Lyer lines appear to be different in length, although they are the same in length; the tie looks green, but it is blue.

Furthermore, Sellars's story strongly supports the claim that if we do not make appearance-reality distinctions, we cannot avoid bizarre pictures of the world.[2] But is Gallois right in claiming that we self-attribute experiences when making these appearance-statements? Imagine that people did not refer to their inner states when talking about appearances, but rather to the actual properties of objects. Then, the above sentences could be translated into:

(a*) the Müller-Lyer lines are of different lengths, although they are the same in length;
(b*) the tie is green, but it is blue.

1 Many proponents of this account argue for its truth only when it comes to propositional attitudes. Some (e.g. Moran (2001) have even explicitly claimed that their rationality model does not apply to sensations and experiences. For a detailed overview of different versions of the rationality account, see Gertler (2011).

2 I accept Gallois's claim that it is rational to reject bizarre pictures of the world, if more plausible explanations are available.



Müller-Lyer lines

These statements are, of course, outright contradictions that manifest contradictory pictures of the world, e.g. in which lines are of different and of the same length at the same time. So it seems that when making an appearance-reality distinction we are committed to self-ascribing experiences.

Shoemaker (1996) presents a similar argument which differs only slightly from Gallois's reasoning. In the course of rejecting inner-sense theories of introspection, Shoemaker develops the notion of self-blindness which he defines as the inability to self-ascribe certain mental states although one has sufficient conceptual resources. He argues that it is impossible for a cognitively non-deficient creature to be self-blind to sensory experiences. While the literature has mostly focused on Shoemaker's arguments regarding the self-blindness to pain states, he also presents an argument for traditional perceptual experiences: If a person faces a red wall, he can identify the wall as red. However, if the wall is white and merely illuminated by red light, a cognitively non-deficient person should determine that the wall merely looks red, but is in fact white. Shoemaker then goes on to question whether a person who is self-blind could possibly theorise about a variety of perceptual conditions. In order to answer this question, he presents us with following thought experiment: Imagine that "we rig things so that [people] cannot see whether perceptual conditions are normal" (1996, p.230). In that case, both a normal and a self-blind person cannot determine the colour of the wall. However, the person who is not self-blind to her experiences, can at least theorise about the perceptual conditions because she can judge how the wall appears to her. In contrast, the self-blind person is unable to focus on how the wall appears to him and therefore cannot theorise about the perceptual conditions. Shoemaker concludes that we should deny "that the person, as described, has perceptual experiences in the sense we do" (1996, p.231), and hence, that self-attributing one's experiences is a necessary ability for a rational person with normal vision.

We can see that the arguments of Shoemaker and Gallois share the same structure. Both state that we can focus on the appearance of an object only if we self-attribute experiences. Let me state the argument in deductive form (hereafter Gallois-Shoemaker argument):

(GS1) If person P is rational, then P distinguishes appearance from reality. [3]

(GS2) If P distinguishes appearance from reality, then P self-attributes experiences.

From (GS1) and (GS2) it follows:

(GS3) If P is rational, then P self-attributes experiences.

It is important to note that (GS2) does not claim that a person distinguishes appearance from reality, if she self-attributes an experience. This claim should be accepted by everyone. In order for their argument to be sound, it needs to be true, that a person distinguishes appearance from reality, only if she self-attributes experiences. It is this conditional (GS2) that I will object to in the next two sections.[4] First (section 2), I show that many appearance-reality distinctions are made without self-attributing experiences. Second (section 3), I argue that even in situations in which we normally do self-attribute an experience, we need not do so.

3 This premise is slightly simplified, but like Gallois and Shoemaker, I take it for granted that people are regularly confronted with situations in which there are differences between appearance and reality. Of course, if a rational person is never confronted with a situation in which it is rational to make an appearance-reality distinction, then that person has no rational obligation to make this distinction.

4 There is some well-known empirical data which supports the Gallois-Shoemaker argument: In the last few decades, our understanding of the cognitive development of young children has made major progress. Among the most celebrated studies are appearance-reality experiments (Flavell (1986), Gopnik & Astington (1988)), showing that whereas 3-year olds fail to distinguish appearance from reality, 5-6 year old children successfully pass the tests. The standard interpretation of the results of appearance-reality tasks is that "what helps children finally grasp the [appearance-reality] distinction is an increased cognizance of the fact that people are sentient subjects who have mental representations of objects and events" (Flavell, 1986, p.418). This interpretation has recently come under attack (Nudds, 2011), but is accepted by many philosophers and psychologists. If the standard interpretation is correct, then it provides empirical support for the Gallois-Shoemaker argument: (i) it connects the ability to distinguish appearance from reality with the capacity to self-attribute experiences, (ii) it links the mastering of the appearance-reality distinction with dissolving the bizarre picture of the world in which e.g. a toy car changes colour merely by being taken from one place to the other. Moreover, studies with autistic children (Baron-Cohen, 2000) show that children with cognitive impairment have greater difficulties in distinguishing appearance from reality.

Self-attributions and openness to public confirmation

When people distinguish appearance from reality, they often express this state of affairs by making an appearance-statement. However, it has been argued (Chisholm (1957), Jackson (1977)), that (i) there are many different senses of appearance-statements, e.g. the comparative sense, the epistemic sense, the phenomenal sense, and (ii) that not all appearance-statements are made on the basis of self-ascriptions of experiences. Unfortunately, these different senses of appearance-statements do not cut nature at its introspective joints. I therefore suggest that we first investigate the reasons for why people make appearance-reality distinctions. Austin (1962) makes a useful classification:

«But surely the plainest of men would want to distinguish (a) cases where the sense-organ is deranged or abnormal or in some way or other not functioning properly; (b) cases where the medium - or more generally, the conditions - of perception are in some way abnormal or off-colour;". (1962, p.13)

Let us call examples of the first category a-type cases – those include: hallucinations (e.g. Macbeth's dagger), illusions (Müller-Lyer lines), afterimages and dreams.

Cases that belong to the second group are characterised by abnormal or unfavourable perceptual conditions and which I call b-type cases: Sellars's tie example (bad lighting conditions), immersed stick in water (effects of media), etc.[5]

Austin is not only correct in claiming that people distinguish a-type from b-type cases, he is also right in focusing our attention on functional differences. It is simply wrong to argue that a person misrepresents his environment or even claim that he is deceived by his senses when looking at a stick immersed in water. Quite the opposite is the case: our senses would be malfunctioning if the stick immersed in water appeared straight. Despite these differences in both categories, respective appearance-statements have the same surface grammar. If we are aware of the distinction between the appearance and reality of the length of the Müller-Lyer lines, we say that 'the lines appear to be of different length, but they are of the same length' (illusion); John the shopkeeper claims that 'the tie appears to be green but it is blue' (unusual lighting conditions). Both of these statements and thus both a-type and b-type cases are considered by Gallois and Shoemaker to be statements that are based on self-attributions of experiences. However, by considering a further claim by Austin, we can see that this conclusion might be premature:

«It is perhaps even clearer that the way things look is, in general, just as much a fact about the world, just as open to public confirmation and challenge, as the way things are. I am not disclosing a fact about myself, but about petrol, when I say that petrol looks like water.» (1962, p.43)

Austin makes an important point here, and if we consider b-type cases, then Austin indeed seems to be correct. The case of a blue tie looking green because it is illuminated by electric light, and the example of a straight stick looking bent because it is immersed in water, seem to be facts about the world, and thus open to public confirmation. We can understand what Austin's idea of 'openness to public confirmation' amounts to in visual cases, by applying a photo test: The blue tie will yield a greenish tone on a photograph, and the stick immersed in water creates a bent line on a photograph. Colours and shapes that can manifest themselves on photographs are public properties, and thus can be publicly confirmed and challenged. However, experiences are private mental states, and assertions about private mental states are considered not to be open to public confirmation. Thus, a person who makes an appearance-statement which is open to public confirmation, does not base the appearance-statement on a self-attribution of an experience but reports a publicly challengeable state of affairs. Note, that I do not claim that people cannot think about their experiences when talking about how sticks look in water, or how the colours of ties appear in unfavourable lighting conditions. But if a person makes an appearance-statement which is open to public confirmation, then the self-attribution of an experience is only a contingent matter.[6] Thus, if advocates of the rationality account claim that we either need to self-attribute an experience or behave irrationally, it seems they do not present us with all the options. When John claims that the tie looks green but it is blue, he should not think that the tie is green and also blue (that would indeed be irrational). Nevertheless, John neither needs to think that he has an experience of a green tie, although the tie is blue. He probably thinks that although the tie is blue, mostly green light is reflected from the tie due to the bad lighting conditions. Similarly, if a person claims that the stick looks bent, but it is really straight, then her

5 I do not and need not claim that a-type and b-type cases exhaust all possibilities and are clear-cut categories.

6 What determines the status of an appearance-statement as open or closed to public confirmation, will be discussed in the next section.

choice is not between self-attributing an experience and being irrational. Instead, she will express knowledge of the stick's shape in normal conditions, and contrast it with how light is refracted in the current abnormal conditions.

Although Austin's claim seems to be supported by many b-type cases, is it really true that this feature of looks-statements is 'general', as he claims? The criterion of 'openness to public confirmation' does not seem to be applicable to many a-type cases: the way the Müller-Lyer lines look to a person is not subject to public confirmation and challenge. When you make a photograph of the lines, then it is not the case that the lines are of different length on the photograph (assuming of course a head-on photo). Hallucinations, dreams and afterimages are also not subject to public confirmation. Thus, it seems that Austin's claim that the way things look is in general open to public confirmation, does not hold if the senses are deceived, e.g. Müller-Lyer lines. Nonetheless, we have established a feature of appearance-statements, which normally holds for b-type cases but not for a-type cases, namely that the appearance of things is open to public confirmation and challenge. If this is correct, then Gallois and Shoemaker are wrong in thinking that a rational agent is required to self-attribute experiences every time she makes an appearance-reality distinction in a b-type scenario. But they might not be too worried about this. They can agree that we do not necessarily self-attribute experiences in circumstances in which we explain the occurrence of an appearance-reality distinction by reference to physical conditions (refraction, reflection of light, etc.). However, they can still claim that there are cases like illusions and hallucinations in which it is the experiences themselves that manifest an appearance-reality distinction, and in which appearance-statements are not open to public confirmation.

This counter-response from Gallois and Shoemaker only works, if in certain situations the status of an appearance-statement is always closed to public confirmation, and hence based on the self-attribution of an experience, e.g. an appearance-statement that is made when a person looks at the Müller-Lyer illusion is always an appearance-statement based on the self-ascription of an experience. However, as I will argue in the next section, this is not the case. The nature of appearance-statements is not correlated with the categories in which they fall in, i.e. a-type or b-type cases. What really matters, are the intentions of the person who makes an appearance-statement – whether she makes a statement she believes to be a fact about the world (open to public confirmation) or a statement which she believes to be grounded in her experiences. To illustrate this point, let us look at two examples: mirages and the moon illusion.

The importance of intentions for self-attributions

As adults, we are usually aware of the reasons for why appearance-reality differences occur. Nonetheless, this need not be the case. As an example, take the case of mirages which are most commonly associated with thirsty exhausted travellers, suddenly seeming to see an oasis in the midst of a desert. People have offered two distinct explanations to account for these appearances: they could be illusionary (a-type) or optical phenomena (b-type). Although a traveller might hallucinate an oasis, mirages are properly explained by the bending of light rays from distant objects and can be captured on camera - thus they are optical phenomena. The moon illusion, on the other hand, is a phenomenon that occurs when people look at the moon which is just above the horizon. The moon appears to be larger on the horizon than when it is high up in the sky. It was originally thought that due to light refraction in the atmosphere, the image of the moon does indeed occupy more space in the visual field than in normal circumstances. However, the moon illusion is not an optical phenomenon but is explained by the workings of our perceptual apparatus and is related to the Ponzo illusion; it cannot be captured on a photograph.

We can see from these two examples that people can be justifiably uncertain about whether the appearance of an object is different from its reality due to external physical conditions (b-type) or internal psychological conditions (a-type). Sometimes it is simply very difficult to ascertain the reasons for a difference in appearance and reality. Misattributions of this difference go both ways: mirages are often misattributed to deranged sensory perception, the moon illusion is often misattributed to the physical properties of the atmosphere. This, however, has the fascinating consequence that the status of an appearance-statement depends on the intentions of the speaker (Reuter, 2014), i.e. whether physical conditions or the perceptual apparatus are blamed for the conceived difference between appearance and reality. Although the appearance-statement 'there seems to be an oasis' is usually taken to be subject to public confirmation, a person can utter the same sentence, but uses the sentence in a different sense which precludes public confirmation. We can see that there are these two differ-

ent appearance-statements at play, by considering their communicative purposes in following dialogues.

Mirage talk 1:
S1: There appears to be an oasis over there.
S2: Yes, indeed, but how distant do you believe it really to be?
S1: Well, the air above the surface is very hot now, so the light rays could come from far away.

In 'Mirage talk 1' both speakers are aware of mirages as optical phenomena. They do not reflect on their experiences, but rather focus on the outside phenomena. In contrast, in 'Mirage talk 2', both speakers obviously blame S1's experience for the appearance-reality difference:

Mirage talk 2:
S1: There appears to be an oasis over there.
S2: You are starting to hallucinate. Take a sip of water, and rest for a while.
S3: That might be a good idea.

The third dialogue 'Mirage talk 3' is probably the most interesting, as it demonstrates, how appearance statements can have different communicative purposes, depending on the intentions of the speaker. S1 blames his experiences for what appears to him, and thus we need to classify his appearance-statement as self-attributive. However, the other person understands the same statement differently.

Mirage talk 3:
S1: There appears to be an oasis over there.
S2: Yes, indeed, but how distant do you believe it really to be?
S1: What? It appears to you too? I thought I had an hallucination.
S2: What are you talking about? The oasis is out there, but maybe not were we think it is.

The same dependence of appearance-statements on the intentions of the speaker can be shown to occur when it comes to the moon illusion.

Moon talk 1: (Appearance talk – not based on self-attribution of experiences)
S1: The moon appears to be larger on the horizon.
S2: Yes, it's amazing how the atmosphere bends the light rays.

Moon talk 2: (Appearance talk – based on self-attribution of experiences)
S1: The moon appears to be larger on the horizon.
S2: Yes, the illusion is astonishing.

Moon talk 3: (Two different senses of appearance talk)
S1: The moon appears to be larger on the horizon.
S2: Not to me. It looks as large as if it were high in the sky.
S1: What you cannot see it, but it is out there!
S2: No, it is an illusion that some people don't get.

These 'mirage' and 'moon talk' conversations highlight the differing intentions with which people make appearance-statements. If they consider appearance statements to be open to public confirmation, they do not focus on their mental states as representing the world in a certain way, but rather talk about how the world is, independent of their experiences. In contrast, they might use an appearance statement to talk about their experiences themselves - statements which then of course should be classified as self-attributive.[7] Austin was right to point out the importance of people categorically distinguishing a-type cases, e.g. illusions and hallucinations, from b-type cases, in which the external conditions of perception are abnormal. Furthermore, he correctly realized that it is 'normally' the case that when people talk about illusions &c. they state subjective facts, and if they talk about b-type cases they are 'normally' disclosing facts about the world. However, as we have just discussed, people might not know the reasons on which appearance-reality differences are based or deliberately refer to their experiences. Although we use the same appearance-words for both categories, appearance-statements that are made because people focus on their experienc-

7 Following objection to my claims about public confirmation might be raised: In the 'moon talk 2' - conversation, S1 and S2 agree with each other on the content of their experiences. Hence, it seems that S2 publicly confirms the content of S1's experience, in contrast to my claim. This objection rests on a misunderstanding of what it means to be open to public confirmation with what it means to talk about each others experiences. When two people are subject to the same illusion (e.g. the moon illusion) and their visual apparatus works in roughly the same way, then it should not surprise us, that they can agree with each other on the effect of being subject to an illusion. This, however, is very different from confirming or challenging the content of another person's experience, and can be most clearly seen, when one person's experience is different from another person's experience. Segall, Campbell and Herskovitz (1978) conducted cross-cultural studies on the effect of origin and culture on the strength of the Müller-Lyer illusion. It turned out that people who live in urban areas perceive the lines to have a greater difference in length than people from rural areas. People who are subject to a strong Müller-Lyer illusory experience would it no way be able to challenge or confirm other people's lack of perceiving the illusion. Whereas they can still talk about their experiences and 'confirm' that what they report is in line with what other people report, they cannot confirm or challenge another person's experience.

es, are intended not to be open to public confirmation or challenge, whereas appearance-statements that are made because the person believes a b-type case to be present, are intended to be open to such challenges. But if an appearance-statement is intended to be open to public confirmation, it is not based on a self-attribution of an experience.

In the second section of this essay I have argued that Gallois and Shoemaker should concede that there are many cases in which people make an appearance-reality distinction without self-attributing an experience, but that their account might still be valid for cases like illusions and hallucinations. This section has demonstrated that such a response does not work. The Gallois-Shoemaker argument does not state that there are cases in which we can or should self-attribute experiences to properly account for illusions and hallucinations. The argument states that if we are rational, we self-attribute experiences. And this is simply not the case, for it is always possible to blame the external world for disguising the true nature of an object or an object's property, e.g. 'the Müller-Lyer lines appear to be of different lengths because of abnormal lighting conditions'. This explanation is false, but it is not irrational to hold such a view. In fact, the moon illusion and mirages are typical cases in which many people wrongly interpret appearances.

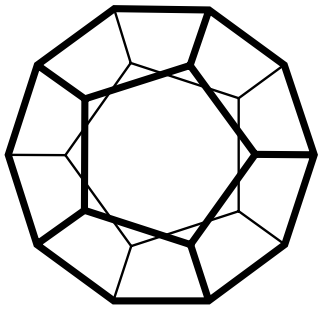
In order to avoid bizarre worlds, we should make appearance-reality distinctions. However, many appearance-statements are open to public confirmation and hence are not based on self-attributions of experiences. A rational person would only be obliged to self-attribute experiences if there were situations under which it is not possible to make an appearance-statement that is intended to be open to public confirmation. But, so I have argued, it is always possible for people to blame the world for appearance-reality differences and thus avoid any self-attributions of experiences. Although only true in principle, it is possible that a rational person always blames her environment and never self-attributes an experience.

Conclusion

No doubt, we often self-attribute experiences to account for appearances that conflict with reality. Gallois and Shoemaker argue that as rational agents we don't have a choice but to self-attribute experiences. In this essay, I have argued that this view is mistaken. We need to distinguish appearance-statements that are open to public confirmation and hence not based on self-ascriptions of experiences from those that are made because people reflect on their experiences. If people are subject to illusions and hallucinations, people should blame their experiences for appearance-reality differences but they need not. Even in those cases, it is rational to make a substantial appearance-reality distinction but not self-attribute experiences. All we need to do is to blame the external world for disguising the true nature of an object.

References

- Austin, J.L. (1962). Sense and sensibilia. G.J. Warnock (ed.). Oxford University Press.
- Baron-Cohen, S. (2000). Theory of mind and autism: a review. International Review of Research in Mental Retardation. 23, pp. 169-184.
- Chisholm, R. (1957). Perceiving: A philosophical study. Ithaca: Cornell University Press.
- Flavell, J. (1986). The development of children›s knowledge about the appearance-reality distinction. American Psychologist. 41(4).
- Gallois, A. (1996). The world without, the mind within: an essay on first-person authority. Cambridge University Press.
- Gertler, B. (2011). Self-Knowledge. In Routledge's New Problems in Philosophy series.
- Gopnik, A., Astington, J. (1988). Children›s understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. Child Development. 59(1).
- Jackson, F. (1977). Perception: a representative theory, Cambridge University Press.
- Moran, R. (2001). Authority and estrangement: an essay on self-knowledge. Princeton University Press.
- Nudds, M. (2011). Children›s understanding of perceptual experiences. In Perception, Causation, and Objectivity. Roessler, Lerman & Eilan (eds.).
- Reuter, K. (2014). The importance of intentions in introspection. In JL Austin on Language. B. Garvey (Ed.). Palgrave Macmillan, pp. 197-211.
- Segall, M., Campbell, D., Herskovitz, M. (1968). The influence of culture on vision. In Social Perception. Toch & Smith (eds.).
- Sellars, W. (1956 / 1997). Empiricism and the philosophy of mind. Brandom, R. (ed.). Harvard University Press. Cambridge, MA.
- Shoemaker, S. (1996). Self-Knowledge and ‹inner sense›: the broad perceptual model. In The first-person perspective and other essays. Cambridge University Press.



Dr. Kevin Reuter ist Dozent und Postdoktorand an der Universität Bern. Seine Interessengebiete liegen in der Philosophie des Geistes und der Psychologie, den Kognitions- und Sozialwissenschaften und der experimentellen Philosophie.

Willensfreiheit durch Kenntnis der Determinanten

Harte Deterministen sind harte Diskussionspartner. In Ihrem Weltbild ist alles vorherbestimmt und sie unterscheiden sich von Fatalisten vor allem darin, dass die einen die Naturgesetzte, die anderen etwas Gottähnliches als schicksalsbestimmend definieren. Dabei ist ihre Behauptung kein neutraler Standpunkt, da die Abrede der Willensfreiheit ein Gedankenkonstrukt voraussetzt, das keinesfalls bewiesen ist, auch wenn die einzelnen Prämissen empirisch messbar sind. Die Libertarier im Gegensatz dazu verteidigen den freien Willen mit solcher Inbrunst, dass sie oft rein intuitiv argumentieren und empirisch unbestimmte Prämissen voraussetzen. Zwar mögen sich Menschen durch die Abrede eines freien Willens in ihrem Selbst beleidigt sehen, die Regelmässigkeit in kausalen Ketten zu leugnen führt jedoch am Ziel vorbei.

Mir scheint, dass die Menschen ihre Determinanten[1] kennen möchten, um auf diesem Wissen eine freie Entscheidung treffen zu können. Im Folgenden will ich mein kompatibilistisches Modell der Willensfreiheit definieren. Natürlich sind auch diese Gedankengänge determiniert, durch Bücher, Vorlesungen, Diskussionen, meine Erfahrung auf diesem Planeten und Erkenntnisse, die Menschen lange vor unserer Zeit hatten. Die Kombination aus meinen Determinanten hat zu diesem Essay geführt und trotzdem habe ich entschieden, welche Determinanten das subjektiv beste Ziel ermöglichen. Dazu gleich mehr.

Determinanten

Determinismus verstehe ich im Folgenden als einen physikalischen Determinismus, der die kausalen Zusammenhänge[2] durch Naturgesetzte bestimmt. Die Betrachtung der Ereignisse reichen dabei bis zum Big Bang zurück, dessen Verteilung der Materie im Raum die ersten Determinanten vorgab. Je weiter die verschiedenen Materien zusammenschmolzen und damit die uns bekannten Elemente bildeten, umso komplexer und damit multikausaler wurde das sich ausdehnende Universum.[3] Das heisst: Immer mehr Ereignisse können zusammenwirken und damit exponentiell komplexere Kausalketten bilden.

Das Kausalitätsprinzip ist erster Teil dieser weichen Determinismus-Definition. Es ist jedoch nicht nur relevant, wie die einzelnen Determinanten zusammenhängen, wichtig ist auch, sie überhaupt zu kennen. Aus Sicht eines menschlichen Lebewesens lassen sich die Determinanten grob in internalistisch und externalistisch[4] unterteilen. Der Anfang der Determinierung liegt schon in den Genen und damit der körperlichen Entwicklung. Durch die Einflüsse von Aussen entstehen im kleinen Kind die ersten richtungsweisenden Determinanten. Es wird von unzähligen, ihm/ihr bis dahin unbekannten Determinanten geprägt, die es nicht frei wählen kann. Erfahrungen mit Menschen in der Familie, Freunden und Fremden sind genau so kausal bestimmend wie die demografischen oder genetischen Prägungen. Die Kenntnis über die eigene geistige Fähigkeit beginnt bei Level 0. Mit den Erfahrungen aus Sensorik, die der Erforschung der Umwelt dienen und bis zu wissenschaftlicher Arbeit mit standardisierten Methoden reicht, kann Wissen über Determinanten gewonnen werden. Es entwickeln sich Forschungsdrang und Verstandesgebrauch bzw. Wissbegierde. [5] Mit jedem graduell höherem Level kann komplexeres Wissen analysiert werden. Mit besseren Analyse- und Messverfahren wiederum können mehr Determinanten erkannt werden.

Level der Reflexion Determinanten-Kenntnis		**Erkennungs-aufwand**
5	unbekannte Determinanten	sehr hoch
4	verdeckte/verdrängte[6] D.	hoch
3	vermutete Determinanten	erhöht
2	teil-bekannte Determinanten	gering
1	bekannte Determinanten	kein

1 Determinanten sind Einflussfaktoren, die eine Entscheidung mitbestimmen. Sie grenzen sich von Dispositionen ab, indem sie rein durch die kausalen Zusammenhänge definiert sind.

2 Kausalität nicht als monokausale Ursache-Wirkungsbeziehung sondern als multikausale Abhängigkeiten verstanden.

3 Fachbegriff Nukleosynthese, vgl. Illiadis (2007): Nuclear physics of stars.

4 Begriffspaar, um Impulse von Innen und Aussen zu unterscheiden. Internatlistisches liegt dabei im Sachverhalt selber (hier im Menschen).

5 Vgl. Spiegeltest: Ab 14 bis 24 Monaten beginnt ein Kind sich selber im Spiegel zu erkennen.

Wäre vom Anfang des uns bekannten Universums bis jetzt jeder kausale Zusammenhang bekannt, würden wir zum Zustand der „Perfect Information“[7,8] gelangen. Mit dieser allumfassenden Information wäre erstmals ein wahrlich objektiver Handlungsentscheid möglich. Da diese Anforderungen an Informationen enorm sind und Entscheidungen oft mit bedeutend weniger Informationen gefällt werden können, verstehe ich „Perfect Information“ im folgenden als das objektiv maximal mögliche, zum Sachverhalt relevante Wissen. Dieses schliesst sowohl physikalische Information, als auch Wissen über die eigene Handlungsstruktur mit ein. Die Auswertung der Informationen geschieht durch Reflexion.

Reflexion

Unter Reflexion verstehe ich die Fähigkeit, über einen Sachverhalt aufgrund der bekannten Determinanten der Innen- und Aussenwelt nachzudenken und damit eine Entscheidung zu treffen. Diese Neukombination der vorhandenen Determinanten führt zu Erfahrung, die selber Bestandteil der subjektiv bekannten Determinanten wird. Dem reflektierenden Menschen zur Verfügung stehende Mittel sind unter anderem: Erkennen und Auswerten von Mustern, Logik, Kombinatorik, qualitative und quantitative Bewertung, u.v.a.m. Man spricht auch von den kognitiven[9] Fähigkeiten, um an mehr Wissen zu gelangen.

Durch die Reflexion werden Möglichkeiten bewusst, die ein Anderskönnen erlauben. Je mehr Optionen (ff. als Synonym zu Möglichkeiten verwendet) vorhanden sind, umso höher die Wahrscheinlichkeit, dass eine Entscheidung rational[10] ist. Um einen Entscheid zu einer Handlung objektiv rational zu fällen, ist also ein möglichst umfassendes Wissen nötig. Während durch Sammeln der Determinanten Wissen in Schrift, Bild, Ton etc. verfügbar gemacht wird, kann durch die Reflexion über deren Anwendungsmöglichkeiten nachgedacht werden. Dabei spielt das menschliche Gehirn alle Optionen im Rahmen der vorhandenen Informationen durch, um Voraussagen über die Konsequenzen einer Entscheidung machen zu können.

Das Anderskönnen ist klassischerweise eine indeterministische Position, die von Libertariern vertreten wird. Nicht wenige Libertarier akzeptieren jedoch das Vorhandensein einer physischen Begrenzung, in deren Rahmen die Entscheidung frei ist. So spricht Keil in seinem Fähigkeitsbasierten Libertarismus von „auf starken charakterlichen Dispositionen beruhende[n] Handlungen“, die jedoch „überlegungszugänglich“ bleiben.[11] Damit akzeptieren so argumentierende Libertarier kausale Zusammenhänge, gestehen deren Einflussnahme auf die Entscheidung jedoch zu wenig Spielraum zu, wenn sie jedes Motiv für aktualisierbar und damit bekannt halten. Empirisch gesehen spricht vieles für determinierte Regelmässigkeiten, solange dem Subjekt diese Regelmässigkeit nicht bekannt ist (später unter „self-fulfilling prophecies“ beschrieben).

Determinanten können bewusst oder unbewusst im Hirn abgespeichert sein.[12] Mit steigender Reflexionsfähigkeit, die mit verschiedenen Methoden trainiert werden kann, ist es möglich, immer komplexeres Wissen auszuwerten und damit bewusst zu machen, bzw. in einer Entscheidung zu berücksichtigen. Dadurch entstehen für das Subjekt neue Optionen. Unbewusste Determinanten sind physischer Art, bei denen die Kausalketten der Determinanten eine Entscheidung auslösen, da die benötigten Informationen dem Bewusstsein nicht zugänglich sind, oder zur Vereinfachung automatisiert wurden. Je stärker eine Determinante im menschlichen Hirn abgespeichert ist, umso stärker beeinflusst sie die Entscheidung. Das ändert sich jedoch mit Bekanntwerden dieser Determination, denn sobald dem Subjekt bewusst ist, wie es in einer Situation handelt, wird es diesen Faktor mit berücksichtigen.

In den Sozialwissenschaften werden diese Einflüsse zu den „self-fulfilling prophecies“ gezählt. Vorher unbekannte Determinanten werden einer bestimmten Menge an Personen bekannt gemacht, wodurch diese ihre Entscheidungen anpassen. Klassische Beispiele sind Gerüchte in der Wirtschaft, ein Börsen-Guru der ein Börsenereignis vorhersagt, da er bestimmte Determinanten kennt, was innert kürzester Zeit das Marktverhalten massiv verändert, oder Versagens- und Konfrontationsängste, die bei Leistungstests dazu führen, dass Prüflinge mit entsprechender Vorhersage tatsächlich schlechter

6 Verdrängung gilt als Schutzmechanismus vor unakzeptablen Gedanken, Gefühlen oder Wünschen in der Psychoanalyse, die aus dem Bewusstsein verbannt werden. Colman (2014): A Dictionary of Psychology. Es könnte gesagt werden, dass diese Determinanten unbewusst weiter wirken, wenn Reflexion < Level 4.

7 Begriff aus der Spieltheorie, der alle relevanten Informationen für eine Entscheidung umfasst. Vgl. Osborne und Rubinstein (2001): A course in game theory.

8 In religiösen Betrachtungen auch als Allwissenheit bezeichnet, mit Göttlichem gleichgesetzt.

9 Kognition (lateinisch ‚cognoscere‘): erkennen, erfahren, kennenlernen. Kognitive Prozesse sind verantwortlich für Wissen und Bewusstsein. Blackburn (2014): The Oxford Dictionary of Philosophy.

10 Rationalität definiert als die beste Handlung im Rahmen der bekannten Information zum erreichen eines subjektiven Ziels.

11 Keil (2013): Willensfreiheit. S. 176.

12 Neuronale Korrelate des Bewusstseins beschreiben neuronale Prozesse (Gehirnaktivitäten), die minimal hinreichend für den Bewusstseinszustand sind. Wikipedia (2015).

oder besser abschneiden. Auch der umgekehrte Fall der „self-destroying prophecies“ tritt häufig auf. So kann die Vorhersage einer Katastrophe dazu führen, dass genügend Massnahmen und Verbesserungen getroffen werden, die ein Eintreffen verunmöglichen oder ein hohes Zwischenergebnis dem Erstplatzierten später den Sieg kosten, da er sich zu sicher schien.

Wenn die Determinanten also einer unbestimmten Menge an Wesen bekannt sind, rechnen sie diese entweder weg oder sie verändern die „Vorhersehung“. Wenn wir davon ausgehen, dass alle Determinanten bekannt sind, also „Perfect Information“ besteht, wüsste das Subjekt auch über dieses Verhalten Bescheid. Damit würde das erwartete Ergebnis wiederum komplexer und bedarf noch höherer Fähigkeiten der Reflexion.

Harte Deterministen wollen hier einwenden, dass auch die menschliche Fähigkeit zur Reflexion determiniert sei. Dem muss nicht grundsätzlich widersprochen werden, wenn der Wille zur Reflexion nicht ausschliesslich durch Determinanten bestimmt ist. Jedes menschliche Wesen hat die Determinante der Neugierde in sich, die es von Anfang an die Welt entdecken lässt, wodurch immer neue Determinanten erlernt werden. Mit fortschreitender Entwicklung der Reflexionsfähigkeit, bei der das Kind bereits in frühen Jahren anfängt über sich und die Aussenwelt nachzudenken, steigt die Wahrscheinlichkeit, das Reflexions-Level zu erhöhen. Durch das wachsende Wissen über den Prozess des Reflektierens selber erwächst ein freier Wille, der durch die Betrachtung der externalistischen und internalistischen Determinanten auf der Metaebene nicht mehr von Abhängigkeiten bestimmt ist. Ein Mensch, der alles Wissen über den Kosmos und über sich selbst besitzt ist nicht mehr determiniert.

Willensfreiheit

Ich habe dafür argumentiert, dass sowohl kausale Zusammenhänge, als auch die Fähigkeit der Reflexion die Willensfreiheit beeinflussen. Zusammen betrachtet ergibt sich ein Modell der Willensfreiheit, die graduell mit dem Wissen um Determinanten steigt. Je mehr Informationen bis zur „Perfect Information“ über einen Sachverhalt bekannt sind, um so mehr Optionen für die bewusste Entscheidung ergeben sich. Anders gesagt: Je mehr bekannte kausale Ursachen in den Prozess der Entscheidung einfliessen, umso eher lässt sich die Wirkung vorhersagen.

Wenn einer Person 20% aller Informationen über sich und seine Welt bekannt sind, hat sie die Möglichkeit zu 20% willensfrei zu handeln. 80% der Determinanten bleiben im Verborgenen. Eine subjektiv rationale Entscheidung ist also die, welche innerhalb der bewussten Möglichkeiten am besten zum Ziel führt. Eine objektiv rationale Entscheidung wäre die Kenntnis aller relevanten Determinanten über die Aussenwelt, als auch über die subjektive Determinierungen.

Zusammengefasst: Wenn alle Determinanten unbekannt sind, kann kein freier Wille entstehen. Wir wüssten dann nicht was uns beeinflusst und wären von der Lenkung durch die Determinanten abhängig. Durch die Verbesserung der Erkennungfähigkeit können Determinanten jedoch immer besser erkannt werden. Dies geschieht durch die Schulung der Sinne über Beobachtungen aller Art. Mit der Reflexion über äussere und innere Determinanten und der Reflexion über die Reflexion selber können neue Determinanten entstehen oder aufgelöst werden. Das Nachdenken über eigene Determinierungen lässt Menschen Vorhersagen machen über ihre eigenen Entscheidungen. Wären folglich alle Determinanten bekannt, könnte von vollständiger Willensfreiheit gesprochen werden. Willensfreiheit ist damit graduell abhängig von der Erkenntnisfähigkeit über die Umwelt und über die Einflüsse auf das Subjekt selber.

Literatur

- Blackburn, Simon (2014): The Oxford Dictionary of Philosophy, Cognition. Oxford University Press
- Colman, Andrew M. (2014): A Dictionary of Psychology, Repression. Oxford University Press
- Illiadis, Christian (2007): Nuclear physics of stars. Wiley-VCH Verlag, Weinheim.
- Osborne, Martin J. and Rubinstein, Ariel (2001): A course in game theory. The MIT Press, Cambridge
- Wikipedia (2015): Neuronales Korrelat des Bewusstseins

Manuel Merki (29) ist Bachelorstudent in Philosophie im 5. Semester. Er interessiert sich für die Philosophie des Geistes, die experimentelle Überprüfung philosophischer Thesen und wissenschaftsphilosophische Themen.

Weisst du, wer du bist?

Von einigen Philosophinnen und Philosophen wird die Haltung vertreten, dass es ein wichtiger Teil davon ist eine Person zu sein, dass man sich ständig selbst interpretiert.[1] Das Produkt dieser Selbstinterpretation ist das sogenannte „Narrativ". Das Narrativ ist eine Art Geschichte, die eine Person über das eigene Leben erzählen kann, welche beschreibt, wie es innerhalb des Lebens von einem Ereignis zu einem anderen kam.[2] Auf Grund dieses Narratives ist eine Person in der Lage „narrative Erklärungen" über gewisse Handlungen zu geben; d.h. sie im eigenen Leben in einen Kontext zu stellen. Z.B.: Vera fährt mit dem Velo zur Uni, weil ihr die Umwelt am Herzen liegt. Die Narrativbildung ist dabei jedoch kein bewusster Prozess wie das Schreiben eines Romans, sondern eine Art Selbstinterpretation, die wahrscheinlich bewusste wie auch unbewusste Teile enthält.[3] Die Frage, die ich in diesem kurzen Text betrachten möchte ist die, ob diese Selbstinterpretation – die Narrativbildung – Wissen über die eigene Person generiert. Um diese Frage zu beantworten möchte ich zunächst betrachten, wie man überhaupt feststellen kann, ob es sich dabei um Wissen handelt oder nicht.

Verifikation

Was ist die Relevanz der Verifikation? Damit das Narrativ Wissen darstellen kann, muss es auch wahr sein, und wenn es wahr[4] ist, dann muss es dafür auch eine Methode der Verifikation geben. Unter dem Begriff „Verifikation" verstehe ich auch nichts weiter als eine Methode um herauszufinden, ob etwas wahr ist, das heisst, ob etwas „der Fall" ist. Unter „der Fall sein" verstehe ich wiederum nichts anderes, als dass die Realität spezifische physikalische Eigenschaften besitzt. Damit es beispielsweise der Fall ist, dass es Eis in meinem Gefrierfach hat, muss Wasser in meinem Gefrierfach einen gewissen Aggregatszustand haben. Dies ist natürlich keine besonders detaillierte Erklärung, denn vieles, was „der Fall ist", ist physikalisch kompliziert und kann nicht zur Gänze in ein paar Sätzen beschrieben werden. Dieses Beispiel soll jedoch darauf hindeuten, wie man sich das „der Fall sein" vorstellen kann.

Was ist das Narrativ? Wie kann es verifiziert werden?

Die Beschreibung des Narratives als teils unbewusster Prozess macht es nur schwer greifbar, was wiederum die Frage nach der Verifikation verkompliziert. Um es „greifbarer" zu machen, möchte ich vorschlagen, dass sich das gesamte Narrativ potentiell in Sätzen ausdrücken lässt. Das Narrativ ist zwar nicht identisch mit einer Menge an Sätzen, doch ist die „Übersetzbarkeit" in sprachliche Formulierungen ein wichtiger Teil seiner Natur (man denke an die „narrativen Erklärungen").

Doch wie helfen Sätze bei der Bestimmung der „Wahrheit" des Narratives? Dazu braucht es eine weitere Annahme, nämlich dass Sätze mit Sachverhalten in der Welt korrespondieren (können). Unter Sachverhalten verstehe ich „Ausschnitte aus der Welt", wie beispielsweise das gefrorene Wasser in meinem Gefrierfach oder auch der Weg eines Wassertropfens vom Eiszapfen bis zum Boden, als weiteres Beispiel. Die Grenze der möglichen Sachverhalte ist die Grenze der menschlichen Abstraktionsfähigkeit.

Doch korrespondieren Sätze wirklich mit Sachverhalten? Bezieht sich „der Weg des Wassertropfens vom Eiszapfen bis zum Boden" auf reale physikalische Eigenschaften? Meiner Meinung nach bestimmt nicht direkt, denn die Formulierung ist viel zu vage, um eine klare Identifikation zuzulassen. Hängt dieser Eiszapfen am Dach meines Hauses, oder am Dach des Nachbarhauses? Die Formulierung hilft mir bei dieser Frage nicht. Beim Satz „Vera fährt mit dem Velo zur Uni, weil ihr die Umwelt am Herzen liegt" ist es nicht anders: Welche Vera ist gemeint?

Eine potentielle Lösung für dieses Problem könnte sein, dass Sätze und Formulierungen wie die obigen lediglich Abkürzungen für viel kompliziertere Sätze sind und sich die Auswahl der möglichen Sachverhalte durch den Kontext einschränkt. (Wenn ich sage, dass Vera mit dem Velo zur Uni fährt, weil ihr die Umwelt am Herzen

1 Jeanine Weekes Schroer und Robert Schroer, „Getting the story right: a Reductionist narrative account of personal identity," Philosophical Studies (2014), doi: 10.1007/s11098-014-0278-z, S. 450.

2 Schroer und Schroer, „Getting the story right: a Reductionist narrative account of personal identity," S. 445.

3 Franziska Wettstein, „Die Bedeutung des Selfies für deine eigene Identität," Philosophischer Blog – Alles Rund um den Mensch, http://blogs.philosophie.ch/mensch/2016/06/27/die-bedeutung-des-selfies-fur-deine-eigene-identitat/ (Zugriff: 2.9.2016).

4 Mir ist nicht bekannt, dass es Philosophinnen oder Philosophen gibt, welche als Bedingung für „Wissen, dass x" annehmen, dass x nicht wahr ist, jedoch wäre aus einer pragmatischen Perspektive denkbar, dass Wahrheit als Bedingung für Wissen zu schwierig zu erreichen ist und darum abgelehnt wird.

liegt, während Vera an uns vorbeifährt, so ist aus dem Kontext abzulesen, um welche Vera es sich handelt.) Die Verifikation von Sätzen ist also fehleranfällig, aber dennoch möglich. Man denke an einen Satz wie „der Himmel ist blau“[5]; da bei diesem Satz so klar ist, auf was er sich bezieht, ist seine Wahrheit leicht festzustellen.

Probleme bei der Verifikation

Bei der Verifikation des Narrativs treten zwei Probleme auf. Das erste folgt aus obigen Überlegungen: Wie identifizieren wir den richtigen Sachverhalt um herauszufinden, ob das der Fall ist, was ein narrativer Satz[6] aussagt? Nehmen wir den Satz „Ich ging heute ins Fitnesscenter, weil ich sportlich bin“ als Beispiel. Was genau muss der Fall sein, dass ich sportlich bin, d.h. was ist die Definition von „sportlich“? Reicht dafür der einmalige Gang ins Fitnesscenter oder ist die Schwelle, die überschritten werden muss, um von Unsportlichkeit zu Sportlichkeit zu gelangen, bei anderthalb wöchentlichen Fitnesscenterbesuchen und 22.5 Minuten Jogging sowie 37 täglichen Liegestützen?

Oder was wäre, wenn Patrick auf Facebook posted „Ich ging heute ins Fitnesscenter, weil ich sportlich bin“, nachdem er zum ersten Mal in seinem Leben im Fitnesscenter war, sich jedoch vorgenommen hat, von nun an regelmässig zu gehen, und sich rein aufgrund dieser Intention als sportlich ansieht? Man kann sich mögliche Antworten auf diese Frage denken, aber es ist zumindest keine einfach zu beantwortende Frage.

Das zweite Problem bei der Verifikation tritt auf, wenn bei der sprachlichen Formulierung des Narratives Fehler geschehen. Man stelle sich vor, Patrick werde nach seinem Besuch im Fitnesscenter im Bus in Smalltalk verwickelt. „Bist du sportlich?“, fragt seine Sitznachbarin, auf seine Sporttasche deutend. Vielleicht antwortet Patrick mit einem „Ja“, da er nicht erklären möchte, dass es sein erster Besuch im Fitnesscenter war und er seine Freizeit bis anhin vor dem Fernseher verbracht hatte. Patricks „Ja“ kann aufgrund seiner Form[7] als narrative Aussage betrachtet werden, aber im Kontext des Beispiels ist denkbar, dass die narrative Aussage kein Ausdruck von Patricks tatsächlichem Narrativ ist (wenn Patrick nicht von sich denkt, dass er sportlich ist). Die einfachste Lösung für dieses Problem wäre natürlich, in diesem Fall nicht von einem narrativen Satz zu sprechen, aber dann müssen wir dafür erklären, wie wir erkennen, wann ein Satz ein tatsächlich narrativer Satz ist, und wann nicht.

Der Irrtum über das Narrativ

Ein weiteres Hindernis für den Wissensgewinn aus der Narrativbildung stellt die Möglichkeit des Irrtums dar. Beispielsweise könnte es sein, dass Vera über sich selbst denkt, dass sie Velo fährt, weil ihr die Umwelt wichtig ist, doch eigentlich würde Vera Auto fahren, wenn sie nur das Geld für ein Auto hätte, ohne dass ihr dieser Umstand bewusst wäre.

Dass solche Irrtümer geschehen können, ist nicht verwunderlich. Wir mögen uns an das Beispiel der Sportlichkeit zurückerinnern, wenn ich die genaue Definition eines Begriffes nicht kenne, so kann ich leicht den Fehler machen, den Begriff fälschlicherweise auf mich anzuwenden. Und vor allem wenn ich den starken Wunsch habe, dass ich eine bestimmte Eigenschaft habe, dann mag ich geneigt sein, den Begriff schneller auf mich anzuwenden, als ich sollte. Wer kennt das nicht?

Wissen wir also, wer wir sind?

Ich habe in diesem kurzen Text darauf hingewiesen, dass dem Wissensgewinn aus der Narrativbildung einige Hindernisse im Weg liegen. Bei der Übersetzung des Narratives in sprachliche Äusserungen (sprich: Sätze) können Fehler geschehen und die Identifikation von für die Verifikation relevanten Sachverhalten ist ebenfalls nicht einfach. Dennoch können wir eine pragmatische Haltung einnehmen und sagen, dass wir doch eine recht gute Idee davon haben, wer wir sind. Denn auch wenn nicht in jedem Fall klar festzumachen ist, worauf genau sich sprachliche Äusserungen beziehen, so ist dies auch nicht gänzlich unklar. Wäre dies der Fall, so wäre menschliche Kommunikation unmöglich. Typischerweise würden wir uns selbst nicht als sportlich bezeichnen, wenn wir nur ins Fitnesscenter gehen, um den anderen Besucherinnen und Besuchern beim Trainieren zuzusehen. Und auch wenn wir uns irren; die Selbstinterpretation, die das Narrativ ausmacht, ist ein ständiger Prozess, der auch eine Revision von Urteilen beinhaltet. Im Grunde verhält es sich beim Selbstwissen wie wohl in jedem Bereich des Wissensgewinns: Je mehr man investiert, um den Fehlerquellen zu entgehen, desto näher gelangt man der Wahrheit, und somit dem Wissen.

5 Oder vielleicht besser: „Unter normalen Umständen (wolkenlos, normale Lichtverhältnisse) ist der Himmel blau“.

6 Unter einem narrativen Satz verstehe ich den Ausdruck eines Teils des Narratives in sprachlicher Form (beispielsweise der Satz über Vera).

7 Er sagt etwas über die eigene Person aus.

Franziska Wettstein (24) ist Masterstudentin in Philosophie und Strafrecht im 11. Semester. Sie interessiert sich für Metaphysik, Theorien der Bedeutung und Rechtsphilosophie.

Können wir die Zukunft der Gesellschaft kennen?

Im Zuge der grossen Fortschritte der Wissenschaft ist man verleitet zu glauben, dass dem menschlichen Wissen kaum Grenzen gesetzt sind. Die modernen Wissenschaften haben Licht auf Dinge geworfen, die man zuvor jenseits der Grenzen des Wissbaren geglaubt hatte. Es wurden überraschende Erkenntnisse gewonnen, die uns nicht nur zu einem besseren Verständnis der Welt verhelfen, sondern es uns auch ermöglichen, beeindruckende Prognosen zu machen. Und die Zukunft scheint weitere solche Erkenntnisse zu versprechen, ja, unser Wissen scheint sich geradezu grenzenlos auszubreiten. Nicht „Grenzen", sondern „Wachstum" oder „Fortschritt" sind die Schlagworte unserer Zeit.

Doch es gibt ein Gebiet innerhalb der Wissenschaften, das kaum vom gegenwärtigen Fortschrittsglauben betroffen zu sein scheint: die Sozialwissenschaften. In Bezug auf Wissen über uns selber, unsere Gesellschaft und unsere Zukunft sind wir vielmehr sehr bescheiden geworden. Während frühere Soziologen noch glaubten, die notwendigen sozialen oder historischen Gesetze erkannt zu haben und den Lauf der Geschichte prognostizieren zu können, erhebt heute kaum ein Sozialwissenschaftler mehr einen derartigen Anspruch. Im Nachhinein scheinen zwar stets alle zu wissen, warum es zum Börsencrash oder zur Revolution kommen musste, aber voraussagen, so ist man sich mittlerweile schon fast einig, kann man solche Ereignisse nicht. Warum ist es in den Sozialwissenschaften so schwierig, zuverlässige Prognosen aufzustellen? Warum können wir die Zukunft der Gesellschaft nicht vorhersagen? Werden wir jemals in der Lage sein, dies zu tun? Oder liegen die Antworten auf Fragen über unsere Zukunft jenseits der Grenzen des Wissbaren?

Die vorliegende Arbeit beschäftigt sich mit diesen Fragen. Sie untersucht also eine Wissensgrenze, die für die Sozialwissenschaften häufig postuliert wird: Wir können die Zukunft der Gesellschaft nicht kennen. Wir wollen versuchen zu verstehen, inwiefern und warum diese Grenze besteht. Im Zentrum des Erkenntnisinteresses steht dabei nicht die Grenze unseres momentanen sozialwissenschaftlichen Wissens, sondern jene unseres möglichen Wissens: die Grenze des Wissbaren. Dazu werden drei mögliche Erklärungsansätze für die postulierte Unmöglichkeit sozialer Prognosen untersucht und beurteilt:

These 1:

> Die Geschichte unserer Gesellschaft lässt sich nicht voraussagen, weil sie nicht determiniert ist.

These 2:

> Die Geschichte unserer Gesellschaft lässt sich nicht voraussagen, weil sie nicht durch soziale Gesetze determiniert ist.

These 3:

> Die Geschichte unserer Gesellschaft lässt sich nicht voraussagen, weil sie zu unregelmässig ist.

Ich werde diese Thesen der Reihe nach besprechen, um schliesslich eine Antwort auf die Fragen zu finden, inwiefern und warum wir die zukünftige Geschichte der Menschheit nicht kennen können, und wo in diesem Fall die Grenzen des Wissbaren verlaufen könnten.[1]

1 Die drei Thesen, die ich untersuchen werde, stellen natürlich nicht die gesamte Menge an möglichen Beschränkungen für sozialwissenschaftliches Wissen über die Zukunft dar. So könnte man zusätzlich etwa behaupten, dass Prognosen in den Sozialwissenschaften nicht möglich sind, weil wir selbst Teil unseres Forschungsgegenstandes sind. Ein Argument, das in diesem Zusammenhang vorgebracht wurde, stammt von Karl Popper. Es besagt, dass die Geschichte der Menschheit stark beeinflusst wird durch unser aktuelles Wissen. Dieses scheint aber ständig anzuwachsen, und da wir nicht wissen können, was zukünftige Generationen (zusätzlich) alles wissen werden (sonst hätten dieses Wissen ja bereits heute), können wir auch die gesellschaftliche Entwicklung nicht kennen (Popper 1997). Diese und ähnliche Paradoxien, die sich daraus ergeben, dass Sozialwissenschaftlerinnen selbst Teil des Systems sind, dessen Entwicklung sie betrachten, halte ich für sehr interessant und wichtig. In diesem Zusammenhang werde ich sie aber ausklammern und den Schwerpunkt auf die drei erwähnten Thesen legen. Mit meiner Analyse erhebe ich daher auch keinen Anspruch auf Vollständigkeit. Ich werde nicht versuchen, die Grenzen des Wissbaren in den Sozialwissenschaften genau zu vermessen, sondern lediglich in bestimmten Regionen genauer untersuchen, wo sie verlaufen könnten.

Wissen über die Zukunft unserer Gesellschaft

Es gibt wohl kaum einen wissenschaftlichen Bereich, dessen Forschungsgegenstand uns so nah und so unzugänglich zugleich ist wie jener der Sozialwissenschaften. Hier geht es um uns, und unter anderem um unsere Zukunft. Jeder fühlt sich hier als Experte, jeder glaubt so einiges zu wissen, aber selbst die besten Wissenschaftlerinnen scheinen nicht allzu viel zu wissen. Wir alle wissen, dass auch morgen die Leute noch zur Arbeit gehen werden, dass es auch übermorgen noch Staaten geben wird und dass in der nächsten Woche in der Schweiz keine Revolution stattfinden wird. Doch wer wird in drei Jahren den Nobelpreis für Chemie gewinnen? Wie wird sich der Dow-Jones-Index in den nächsten fünf entwickeln? Und in den nächsten 50 Jahren, 100 Jahren? Wird es einen 3. Weltkrieg geben? Wird es die Menschheit in 5000 Jahren noch geben?

Eine verlockende Antwort auf diese Fragen ist die folgende: Das können wir nicht (so genau) wissen, weil es von uns und unseren freien Entscheidungen, die wir noch nicht gefällt haben, abhängt. Können wir vielleicht die Zukunft nicht kennen, weil diese aufgrund unserer eigenen Freiheit noch gar nicht festgelegt ist?

These 1: Die Geschichte unserer Gesellschaft lässt sich nicht voraussagen, weil sie nicht determiniert ist.

Viele Leute teilen die Intuition, dass der Mensch über einen freien Willen verfügt und dass sich menschliches Handeln daher fundamental vom Verhalten anderer Dinge unterscheidet. Menschliches Handeln sei frei und damit undeterminiert und nicht prognostizierbar. Wenn man den freien Willen konsequent (d.h. libertarisch und nicht kompatibilistisch) denkt, würde dies bedeuten, dass der soziale Weltlauf nicht determiniert ist, da ihn Menschen als freie Wesen jederzeit beliebig gestalten können. Da wir allerdings nicht wissen, welche freien Entscheidungen Menschen in Zukunft fällen werden – so müsste man dann bezüglich sozialwissenschaftlicher Prognosen folgern – können wir auch den zukünftigen Verlauf der Geschichte nicht kennen (vgl. dazu bspw. Geert 2007) .

Eine solche These steht natürlich in einem offensichtlichen Widerspruch zu einer anderen metaphysischen These: jener des Determinismus. Libertarische Verfechter der Willensfreiheit berufen sich daher oft darauf, dass die moderne Quantenphysik nahelege, dass die Welt nicht determiniert sei. Ob unsere Welt auf Quantenebene deterministisch oder indeterministisch ist, kann und muss in diesem Zusammenhang nicht entschieden werden. Bei genauerer Betrachtung scheint die Antwort auf diese Frage für die Willensfreiheit nämlich ziemlich unerheblich zu sein. Denn sowohl in einer deterministischen als auch in einer indeterministischen Welt ist die Annahme einer kausalen Sonderstellung des Menschen als wenig plausibel einzustufen. Warum sollte die ganze Welt nach bestimmten (probabilistischen oder deterministischen) Regeln funktionieren, während ausgerechnet wir uns über all diese hinwegsetzen und gewissermassen als unbewegte Beweger unsere eigenen Kausalketten anstossen können? Als biologische Organismen unterliegen auch Menschen den Naturgesetzen – seien sie nun deterministisch oder probabilistisch. Falls unsere Welt deterministisch ist, müssten wir diese unter der Annahme eines freien Willens partiell ausser Kraft setzen – eine wenig überzeugende Annahme. Und auch wenn in der Welt so etwas wie Zufall existieren sollte, ist unserer Freiheit damit kaum geholfen. Zufall alleine macht uns nicht freier. Man müsste vielmehr behaupten, dass dieser (quantenphysikalische) Zufall unter gewissen Umständen der Macht unseres freien Willens unterliegt. Letzterer müsste dann die Indeterminiertheit auf Quantenebene irgendwie zu seinen Gunsten nutzen und die Quantenteilchen innerhalb des Spielraums der jeweiligen Wahrscheinlichkeiten frei nach seinen Bedürfnissen lenken. Auch diese Annahme scheint mir sehr gewagt. Es spricht mehr für die entgegengesetzte Annahme, dass Menschen genauso determiniert oder indeterminiert sind wie alles andere auch und lediglich unser Gefühl von bzw. Wunsch nach Freiheit dafür verantwortlich ist, dass manche glauben, menschliches Handeln unterscheide sich diesbezüglich fundamental von den Verhaltensweisen anderer Dinge.

Wenn wir die Annahme eines freien Willens und einer kausalen Sonderstellung des Menschen zurückweisen, bedeutet dies für unsere erste These folgendes: Es ist durchaus möglich, dass die Geschichte unserer Gesellschaft nicht determiniert ist – nämlich dann, wenn es (1) auf Quantenebene echten Zufall gibt und sich dieser (2) auf höhere Ebenen (und insbesondere gesellschaftliche Phänomene) auswirkt. Sollten diese beiden Bedingungen zutreffen, haben wir allerdings wenig Grund zur Annahme, dass lediglich die Sozialwissenschaften betroffen wären. Vielmehr wäre unser Wissen über die Zukunft insgesamt nur noch probabilistischer Natur. Somit mag unsere erste These zwar stimmen. Sie ist aber für unsere Fragestellung nicht von besonderem Interesse, da sie sich ganz allgemein auf Prognosen und nicht nur auf gesellschaftliche Prognosen bezieht. Unsere soziale Zukunft ist genauso determiniert oder indeterminiert wie alles andere auch.

Aber irgendwie scheint ja trotzdem ein Unterschied

zwischen Prognosen in der Physik und solchen in der Soziologie zu bestehen. Die Physik – so stellt man sich das als Aussenstehende zumindest vor – kann Gesetze formulieren, die ausnahmslos immer und überall gelten, und diese benutzen, um etwa die zukünftige Bewegung von Teilchen oder die Bahn von Himmelkörpern vorherzusagen. Aber gibt es auch soziale Gesetze, die wir als Grundlage für soziale Prognosen verwenden können? Das führt uns zur zweiten These:

These 2: Die Geschichte von Gesellschaften lässt sich nicht voraussagen, weil sie nicht durch soziale Gesetze determiniert ist.

Die Sozialwissenschaften operieren mit Variablen wie Geschlecht, Alter, Herkunft und Einkommen und verwenden Konzepte wie Staaten, Geld, Revolutionen, Arbeitslosigkeit etc. Derartige Variablen oder Konzepte werde ich soziale Variablen oder soziale Konzepte nennen.

Beginnen wir nun die Diskussion der zweiten These anhand eines kleinen Gedankenexperiments: Angenommen, wir hätten einen hochintelligenten Computer, der für alle Individuen, Institutionen, Staaten etc. die Ausprägung aller sozialen Variablen kennt und die gesamte Geschichte der Menschheit zur Verfügung hat, um zu berechnen, wie diese Variablen zusammenhängen. Unser Computer hat dabei nicht nur das vollständige empirische Wissen, sondern ist auch der perfekte Theoretiker – er verfügt also über ein perfektes sozialwissenschaftliches Wissen. Ausserdem weiss er alles, was man über externe Einflüsse auf das soziale System wissen kann, verfügt also auch über perfektes Wissen über die Entwicklung von Klima, Wetter, über zukünftige Erdbeben etc. Weiss dieser hypothetische Computer alles, was man über die Zukunft der Gesellschaft wissen kann? Könnte er in einer deterministischen Welt perfekte Vorhersagen (und in einer indeterministischen genaue Wahrscheinlichkeitsprognosen) machen?

Ich bin der Überzeugung, dass er dies nicht könnte, denn zwischen den Entitäten, die unsere sozialen Konzepte herausgreifen, bestehen keine (direkten) Kausalzusammenhänge. Unsere sozialen Variablen sind zu ungenau, um die soziale Welt vollständig zu erfassen und zu prognostizieren. Zwischen Dingen wie Alter, Herkunft, Geschlecht und Einkommen bestehen keine perfekten Zusammenhänge. Es kann folglich keine sozialen Gesetze geben. Dieses Problem ist meines Erachtens nicht ein praktisches, das sich nur durch unglückliche Fehlspezifikation der Konzepte ergibt, sondern ist prinzipiell und unvermeidbar. Ich werde es im Folgenden genauer erläutern. Dazu muss aber zunächst eine metaphysische Annahme klargestellt werden, auf der meine Argumentation beruht:

Stellen wir uns dazu einen etwas anderen Computer vor. Dieser Computer verfügt über keinerlei Informationen zu sozialen Konzepten und Regelmässigkeiten. Er kennt lediglich alle Gesetze der Physik und den momentanen Zustand aller Teilchen. Weiss dieser Computer alles, was man über den zukünftigen Zustand von Teilchen wissen kann? Oder könnte er diesen besser vorhersagen, wenn er auch über chemisches, biologisches und soziologisches Wissen verfügen würde? Ich werde hier die Position vertreten, dass unser Computer seine Prognosen zum zukünftigen Zustand von Teilchen nicht verbessern könnte, indem er chemisches, biologisches oder soziologisches Wissen mit einbezieht. Diese Position ist natürlich kontrovers, denn sie impliziert insbesondere die Verneinung der von einigen Wissenschaftlerinnen postulierten Abwärtskausalität – der Ansicht, dass emergente Systeme[2] (bspw. chemische, biologische, psychische, soziale etc.) eigene kausale Kräfte entwickeln, mit welchen sie auf darunterliegende Ebenen und somit letztendlich auch auf das Verhalten von Teilchen einwirken (Nagel 1970, O'Connor et. al 2015). Meine Position wird hingegen Folgende sein: Genuine Kausalkräfte gibt es nur auf der physikalischen Ebene. Sie legen fest, welche Muster und Strukturen – ich werde diese als phänomenologische Regelmässigkeiten bezeichnen - auf höherliegenden (z.B. biologischen oder sozialen) Betrachtungsebenen entstehen. Ich behaupte also, um es noch im philosophischen Fachjargon auszudrücken, dass alles über das Physikalische superveniert (vgl. dazu: Stoljar 2015; O'Connor et al 2015; Van Riel et al. 2015; Cat 2014).

Doch was bedeutet diese These für die Sozialwissenschaften und für soziale Prognosen? Lassen sich alle Wissenschaften, und damit auch die Sozialwissenschaften, auf die Physik reduzieren? Dies folgt keineswegs aus der dargelegten Position und scheint auch nicht der Fall zu sein. Unsere sozialen Konzepte weisen nämlich keine gemeinsame physikalische Kausalstruktur auf. Die Entitäten (wie Staaten, Wertehaltungen, Geld etc.), die unser soziales Leben ausmachen, besitzen keine einheitliche physikalische Konstitution. Ein Staat ist nicht einfach eine bestimmte Anzahl von Elektronen und Protonen in

2 Ab hier werde ich ab und zu die Begriffe „System", "Umwelt" und „Erklärungsebene" verwenden. Unter einem System verstehe ich zunächst nichts anderes als eine frei definierbare Menge von physikalischen Teilchen, unter Umwelt all jene Teilchen, die nicht zum System dazugehören. Ein System lässt sich nun auf unterschiedlichen Ebenen betrachten. Dabei bestimmt man stets unterschiedliche Erklärungseinheiten (oder Konzepte), für deren Zusammenwirken man sich dann interessiert. So kann man etwa das soziale System mithilfe sozialer Konzepte wie Individuen, Staaten, Wertehaltungen etc. beschreiben.

bestimmter Anordnung, die dann stets eine bestimmte physikalische Wirkung hat. Soziale Konzepte sind also auf physikalischer Ebene unterschiedlich realisierbar (vgl. Papineau 2009, Little 1993, Stoljar 2015). Wenn wir nun aber akzeptieren, dass nur auf physikalischer Ebene fundamentale Kausalbeziehungen bestehen, so bedeutet dies, dass die Entitäten, welche unter ein soziales Konzept fallen, stets unterschiedliche kausale Wirkungen haben können. Und dies gilt nicht nur für die sozialen Konzepte, über die wir zurzeit verfügen, sondern für alle möglichen sozialen Konzepte. Egal wie geschickt wir bei der Kategorienbildung vorgehen: Es wird uns nicht gelingen, mit sozialen Konzepten eine Menge von Entitäten herauszugreifen, deren Elemente eine homogene Kausalstruktur aufweisen. Die Entitäten werden vielmehr immer gewisse kausal relevante physikalische Gemeinsamkeiten, aber auch viele kausal relevante physikalische Unterschiede aufweisen. Dies legt folgenden Schluss nahe: Wir können keine sozialen Gesetze formulieren, denn wir können nicht erwarten, dass es in der sozialen Welt etwas gibt, das immer und überall in gleicher Weise gilt. Revolutionen, Staaten, Arbeitslosenraten etc. haben keine einheitlichen Kausalwirkungen, und die Beziehung, in der sie zueinander stehen, weist zwar gewisse Regelmässigkeiten auf, die aber – aufgrund der oben erläuterten Problematik – nicht perfekt sind. Es gibt also, so lässt sich das Zwischenergebnis unserer Überlegungen zusammenfassen, keine sozialen Kausalgesetze (vgl. dazu: Fay 1994; für Gegenpositionen: Kincaid 1990, Hempel 1942).

Doch legt unsere Argumentation nicht auch nahe, dass es keine biologischen oder chemischen Gesetze geben kann? In gewisser Weise gilt die erläuterte Problematik natürlich auch für andere Wissenschaften. So gibt es ja unseren Überlegungen zufolge auch etwa in der Biologie oder der Chemie keine genuinen Kausalkräfte, sondern lediglich phänomenologische Regelmässigkeiten. Es lässt sich allerdings davon ausgehen, dass biologische und insbesondere chemische Konzepte auf physikalischer Ebene viel homogener (und in gewissen Fällen vielleicht sogar vollkommen homogen) realisiert sind und damit auf diesen Erklärungsebenen auch vollkommenere Regelmässigkeiten zu beobachten sind. Je weiter wir aber in den Erklärungsebenen emporklettern, desto physikalisch heterogener werden auch die jeweils interessanten Phänomene oder Einheiten, und desto uneinheitlicher die Zusammenhänge zwischen ihnen. Die sozialwissenschaftliche Ebene ist mit Konzepten wie Revolutionen oder Börsencrashs äusserst weit von der fundamentalen physikalischen Ebene entfernt. Soziale Regelmässigkeiten, so könnte man daraus schliessen, sind besonders unregelmässig.

Doch was bedeutet dies nun für das Wissen über die Zukunft der Gesellschaft? Kann man Revolutionen nicht vorhersagen, weil es keine sozialen Kausalgesetze gibt? Dieser Schluss scheint nicht so ganz zu stimmen. Stellen wir uns dazu noch einen dritten Computer vor. Er weiss alles, was auch bereits unser zweite Computer wusste (physikalische Gesetze, momentaner Zustand aller Teilchen). Ausserdem haben wir ihm beigebracht, was wir mit Revolutionen, Börsencrashs etc. meinen. Solche Ereignisse sind zwar physikalisch unterschiedlich realisierbar und weisen keine einheitliche physikalische Kausalstruktur auf. Daraus folgt aber nicht, dass es sich in physikalischen Mustern oder Abläufen nicht erkennen liesse, ob es sich um eine Revolution oder einen Börsencrash handelt. Wenn alles über das Physikalische superveniert, müssten soziale Ereignisse prinzipiell auch physikalisch „definierbar“ sein. Über solche Definitionen verfügt unser Computer. Nun berechnet/simuliert er aufgrund der Gesetze der Physik auf Teilchenebene den zukünftigen Zustand aller Teilchen. Wenn dabei ein Muster entsteht, das von einer unserer sozialen Definitionen abgedeckt wird, sagt er uns, dass ein Phänomen der entsprechenden sozialen Kategorie (z.B. eine Revolution) eintreffen wird. Ein solcher Computer müsste nun, so würde ich meinen, die Geschichte der Menschheit (soweit sie metaphysisch determiniert ist) voraussagen können.

Somit führt die Tatsache, dass es keine sozialen Gesetze gibt, nicht direkt dazu, dass die soziale Zukunft ungewiss ist. Sie bedeutet lediglich, dass vollkommen exakte und umfassende Prognosen über die Zukunft der sozialen Welt mit genuin sozialwissenschaftlichen Methoden und Konzepten nicht möglich sind. Es wäre aber prinzipiell immer noch denkbar, dass wir unsere soziale Zukunft – analog zum Vorgehen unseres dritten Computers – mithilfe physikalischer Gesetze vorhersagen können. Auch hier scheint bei genauerer Betrachtung keine Wissensgrenze zu liegen.

Wir können also die Zukunft unserer Gesellschaft nicht aufgrund sozialer Gesetze vorhersagen, aber können wir dies aufgrund der Gesetze der Physik? Können wir vielleicht herausfinden, was uns erwartet, indem wir wie unser Computer die Bewegungen von Elektronen, Protonen und Quarks betrachten, um dann zu schauen, ob sie zu einer Revolution führen? Ein derartiges Unterfangen hört sich vollkommen absurd an, denn unsere Welt scheint viel zu komplex und zu chaotisch zu sein, um es erfolgreich durchführen zu können. Liegt der Schlüssel zur Unmöglichkeit sozialer Prognosen vielleicht in dieser Komplexität, in diesem Chaos?

These 3: Die Geschichte unserer Gesellschaft lässt sich nicht voraussagen, weil sie zu unregelmässig ist.

Damit ein hypothetischer allwissender Computer eine perfekte (bzw. bestmögliche) Prognose über den Weltverlauf machen kann, muss er über zwei Arten von Wissen verfügen: Er muss (1) die Gesamtheit aller (fundamentalen) Kräfte oder Gesetze kennen und (2) die Anfangsbedingungen (bzw. die momentanen Bedingungen) exakt feststellen können. Im vorigen Abschnitt haben wir uns mit dem ersten Wissensbereich beschäftigt und behauptet, dass diese Gesetze nicht sozial, sondern physikalisch seien. Nun wollen wir uns jenen Problemen für sozialwissenschaftliche Prognosen widmen, die vor allem mit dem Feststellen der Anfangsbedingungen zusammenhängen. Ich werde argumentieren, dass sich diese auf einer Teilchenebene niemals perfekt bestimmen lassen, da wir dazu stets den Zustand aller Teilchen des Universums mit unendlicher Genauigkeit feststellen müssten. Daher sind wir gezwungen, auf andere Erklärungsebenen auszuweichen, um das Verhalten der infrage stehenden Systeme auf diesen zu approximieren. Wie viel prognostischen Erfolg wir mit einem solchen Vorgehen haben können, hängt davon ab, wie regelmässig das System auf der von uns gewählten Erklärungsebene ist. Da soziale Systeme weitgehend unregelmässig sind, sind diesbezügliche Prognosen äusserst schwierig und ungenau. Diese Überlegungen werde ich in den folgenden Abschnitten etwas genauer ausführen und begründen.

Welches sind die Anfangsbedingungen, die wir kennen müssten, um die soziale Entwicklung voraussagen zu können? Nehmen wir an, wir würden alle fundamentalen Kräfte kennen und hätten die Zeit kurz „eingefroren", um alle Messungen vorzunehmen, die wir für die Voraussage der künftigen Menschheitsgeschichte brauchen. Welche Teilchen würden in unsere Messung eingehen? Sicherlich einmal alle, aus denen wir Menschen bestehen, denn wir müssen ja wissen, welche Vorgänge bei allen Personen im Gehirn ablaufen werden, was sie sagen werden, wohin sie sich bewegen werden etc. Allerdings sind Menschen keine geschlossenen Systeme. Sie reagieren auf Umwelteinflüsse wie Wetterereignisse, Klimaentwicklungen oder Erdbeben, aber auch ganz allgemein auf die Landschaft, die sie umgibt, also beispielsweise auf die Sterne, die sie am Himmel sehen. Das heisst, dass wir auch alle Teilchen erfassen müssen, die unsere Umwelt bestimmen. Das wären dann sicher einmal die Teilchen auf der Erde, aber auch jene der Sonne und all der Sterne und Planeten, die wir sehen oder deren Wirkung (z.B. Wärme) wir fühlen. Da aber das Verhalten all dieser Teilchen (wenngleich minim) von anderen Teilchen anderer Sterne beeinflusst wird (z.B. durch Gravitation), müssten wir auch diese erfassen. Man sieht schnell, wohin das führt: Für die perfekte und absolut sichere Prognose müssten wir alle Teilchen im ganzen Universum erfassen. Zudem können sich minime Veränderungen in den Anfangsbedingungen zu grösseren Veränderungen in der Entwicklung des Systems aufschaukeln (Chaos, vgl. dazu: Greschik 1999). Wenn wir uns bei einigen Teilchen in unseren Messungen leicht vertun, könnten wir zu einem ganz anderen Resultat gelangen. Wollen wir dieses Problem umgehen, müssten wir den Zustand jedes Teilchens mit absoluter Genauigkeit feststellen. Um die Zukunft der Gesellschaft sicher voraussagen zu können, müssten wir also den Zustand aller Teilchen mit absoluter Genauigkeit messen. Niemand würde ernsthaft in Betracht ziehen, ein solches Projekt in Angriff zu nehmen. Seine Realisierung ist für uns Menschen vollkommen undenkbar.

In dieser Hinsicht stellt aber das soziale System noch keinen Sonderfall dar. In jedem System scheint jedes Teilchen von jedem anderen Teilchen im Universum beeinflusst zu werden. Und für jedes System ist damit eine vollständige und präzise Beschreibung und Prognose auf Teilchenebene vollkommen undenkbar. Sind also Prognosen ganz allgemein nicht möglich? Oder: Wie können wir etwas über die über zukünftige Revolutionen oder über die Bahn von Himmelskörpern wissen, ohne den Zustand aller involvierten Teilchen zu kennen? Ich würde behaupten, dass unsere aktuelle und wohl auch die sinnvollste Herangehensweise jener unseres ersten Computers gleicht. Damit wir überhaupt etwas sehen können, betrachten wir die Welt nicht auf Teilchenebene, sondern vereinfachend auf höherliegenden Erklärungsebenen. Wir sprechen nicht von physikalischen Teilchen und fundamentalen physikalischen Gesetzen, sondern verwenden Konzepte wie Himmelskörper, Revolutionen, Staaten etc. Statt uns um fundamentale Gesetze und Teilchen zu kümmern, untersuchen wir nun die phänomenologischen Regelmässigkeiten, die in Systemen auf höherliegenden Betrachtungsebenen entstehen. Damit büssen wir natürlich einiges an Genauigkeit ein und können – wenn die verwendeten Konzepte physikalisch nicht reduzierbar sind – das Verhalten des Systems nicht mehr ganz exakt erfassen. In manchen Systemen sind aber die Regelmässigkeiten zwischen Einheiten höherer Ebenen sehr ausgeprägt oder näherungsweise perfekt. In solchen Fällen lässt sich die Entwicklung eines Systems sehr gut approximieren, ohne auf die Teilchenebene zurückgreifen zu müssen. Wollen wir etwa die Laufbahn der Erde vorhersagen, so müssen wir nicht alle Teilchen im Universum

erfassen. Vielmehr können wir uns Sonne und Erde und die anderen Planeten in unserem Sonnensystem als Punkte mit einer bestimmten Masse vorstellen und den Einfluss aller Teilchen ausserhalb unseres Sonnensystems vernachlässigen. Wenn wir nun die Erdlaufbahn beispielsweise mithilfe einiger (nicht-fundamentaler) physikalischer „Gesetze" berechnen, werden wir zu einem recht guten Resultat kommen. Diese Methode basiert somit gewissermassen auf Modellbildung. Wenn wir Systeme auf höherliegenden Erklärungsebenen betrachten, bilden wir in gewisser Hinsicht ein Modell, das um vieles einfacher, aber dafür auch etwas ungenauer und unvollständiger ist als die physikalische Realität. Wenn wir nun auf dessen Grundlage Prognosen machen, werden sie zwar nicht vollkommen exakt sein, können jedoch eine sehr gute Approximation darstellen. Die allermeisten „Gesetze", die in den Wissenschaften formuliert werden, sind solche Approximationen. Wie gut sich die Entwicklung eines Systems auf diese Weise approximieren lässt, hängt davon ab, wie regelmässig dieses ist. Wie exakt eine Einzelwissenschaft sein kann, ist damit dadurch bestimmt, wie regelmässig sich das System, das sie beschreiben will, auf der Erklärungsebene, die sie verwendet, verhält. Doch was ist mit einem regelmässigen System genau gemeint?

Damit sich die Entwicklung eines Systems gut approximieren lässt, müssen zwei Bedingungen erfüllt sein: Das System muss (1) auf der entsprechenden Erklärungsebene eine hohe innere Regelmässigkeit aufweisen und sollte (2) in einem regelmässigen Verhältnis zu seiner Umwelt stehen. Die erste Bedingung verbietet, dass Systeme auf unserer Betrachtungsebene (zu) komplex, emergent und chaotisch sind. Ihr Verhalten muss sich auf der entsprechenden Erklärungsebene durch möglichst wenige und einfache Regeln (geringe ebenenspezifische Komplexität) möglichst vollständig (geringe ebenenspezifische Emergenz[3]) und robust hinsichtlich der genauen Anfangsbedingungen (geringes ebenenspezifisches Chaos) ausdrücken lassen (vgl. dazu: Greschnik 1999).[4] Die zweite Bedingung fordert etwas Ähnliches für die Beziehung zwischen dem infrage stehenden System und seiner Umwelt: Optimalerweise hat die Umwelt einen möglichst geringfügigen und über die Zeit gleichbleibenden Einfluss auf unser System.[5] Systeme, bei denen diese Bedingungen erfüllt sind, nenne ich idealtypisch regelmässige Systeme[6].

Unser Sonnensystem scheint ein sehr regelmässiges System zu sein. Für die Bewegungen der Planeten in unserem Sonnensystem sind beide Bedingungen recht gut erfüllt. Die Bahn der Erde ist zwar gewissermassen davon beeinflusst, ob sich ein bestimmtes Proton am Nord- oder am Südpol der Erde befindet, allerdings ist dieser Einfluss vernachlässigbar klein. Unabhängig vom genauen Zustand aller beteiligten Teilchen bewegt sich die Erde immer in einer ähnlichen Bahn um die Sonne. Diese Bahn folgt auf der Erklärungsebene von Himmelskörpern einfachen Regeln und ist durch diese ziemlich genau determinierbar. Unser Sonnensystem ist also weder besonders komplex, chaotisch oder emergent und weist damit eine sehr grosse innere Regelmässigkeit auf. Ausserdem ist die Bahn der Planten nur gering und immer in etwa gleicher Weise von Masse und Teilchenanordnung anderer Sterne und Planeten ausserhalb unseres Sonnensystems beeinflusst. Die Umwelt (alle Teilchen, die nicht zum System gehören) hat also nur einen kleinen Einfluss auf das Verhalten unseres Sonnensystems.

Unser soziales System scheint hingegen ein Paradebeispiel für ein unregelmässiges System darzustellen. Wenn wir die Zusammenhänge zwischen sozialen Einheiten möglichst gut beschreiben wollen, scheinen wenige einfache Regeln nicht auszureichen. Die sozialen Zusammenhänge sind komplex. Und auch ein Modell, das tausende von sozialen Variablen in komplexen Operationen miteinander verbindet, würde nicht genügen, um die Entwicklung des sozialen Systems gut zu approximieren. Auch hier würden noch entscheidende Entwicklungen stattfinden oder emergieren, die wir mit den Konzepten/Einheiten unserer Erklärungsebene nicht vorhersehen können. Ausserdem scheint das soziale System äusserst chaotisch zu sein. Es ist durchaus vorstellbar, dass ein kleiner Unterschied in den Anfangsbedingungen zu einer vollkommen anderen historischen Entwicklung führt. Wenn Hitlers Mutter in der Zeitung

3 Der Begriff „ebenenspezifische Emergenz" mag verwirren, zumal ja vorher argumentiert wurde, dass im ontologischen Sinne keine Emergenz existiert. Emergenz ist hier aber kein ontologischer, sondern ein epistemischer Begriff, der stets mit einer bestimmten Erklärungsebene verbunden ist. Er bezieht sich darauf, dass wir in Abhängigkeit der Konzepte, die wir auf bestimmten Erklärungsebenen verwenden, das Verhalten des Systems nicht vollständig erfassen können. Da unsere Konzepte „ungenau" sind, ist auf deren Grundlage das Gesamtverhalten eines Systems nicht vollständig prognostizierbar.

4 Wie stark die erste Bedingung erfüllt ist, hängt natürlich wiederum davon ab, wie homogen die ebenenspezifischen Entitäten in ihrer kausalen Struktur sind. Ebenenspezifische Komplexität, Emergenz und Chaos resultieren aus geringer ebenenspezifischer Homogenität.

5 Andernfalls müssten wir diesen Einfluss in unser Modell mit aufnehmen, in einen systematischen Zusammenhang mit unserem System bringen und modellieren, wie sich die Umwelt entwickeln wird. De facto hätten wir dann aber nicht mehr unser ursprüngliches System modelliert, sondern dieses um die relevanten Umweltfaktoren erweitert.

6 Natürlich handelt es sich bei Regelmässigkeit bzw. Unregelmässigkeit nicht um ein binäres, sondern um ein kontinuierliches Merkmal von Systemen. Um Komplexität zu vermeiden, werde ich aber dennoch idealtypisch von (un-)regelmässigen Systemen sprechen.

nicht die Werbung für diesen interessanten Film gesehen hätte, wäre sie nicht ins Kino gegangen, hätte Hitlers Vater nicht kennengelernt, Hitler wäre nicht gezeugt worden, der zweite Weltkrieg hätte nicht stattgefunden etc. Dazu kommt, dass die zukünftige Entwicklung der Gesellschaft stark von ihrer Umwelt beeinflusst ist, d.h. etwa davon, an welchem Tag wir welches Wetter haben werden, ob es in bestimmten Regionen Erbeben geben wird etc. Aber derartige geographische Phänomene sind selbst wiederum komplex, chaotisch und emergent. Wir müssten nun auch diese komplexe Umwelt modellieren und berechnen, welchen Einfluss sie auf die gesellschaftliche Zukunft hat. Insgesamt betrachtet scheint unser soziales System das perfekte Beispiel für ein unregelmässiges System darzustellen. Unregelmässige Systeme lassen sich, wie wir argumentiert haben, nicht gut modellieren. Man kann sie nicht vereinfachend beschreiben, ohne dass man dabei Dinge weglässt, die ihre weitere Entwicklung massgeblich beeinflussen. Diese Entwicklung lässt sich daher nur äusserst ungenau vorhersagen. In sozialen Systemen, so kann man es zusammenfassen, sind also keine (genauen) Prognosen möglich, weil es sich um unregelmässige Systeme handelt.

Haben wir hier also eine klare Wissensgrenze und den Grund für ihr Bestehen gefunden? Ist die soziale Welt tatsächlich durch und durch unregelmässig und unvorhersehbar? Bei genauerer Betrachtung stellt man fest, dass eine derartige Konklusion zu radikal ist. An gewissen Stellen scheint selbst in unserem sozialen System eine gewisse Ordnung zu herrschen. Auch wenn wir nicht prognostizieren können, ob eine bestimmte Frau in der Schweiz in fünf Jahren ein Kind bekommen wird, so können wir doch ungefähr voraussagen, wie hoch in fünf Jahren die Schweizer Geburtenrate ist. Auch wenn wir nicht voraussagen können, ob eine bestimmte Person einen bestimmten Job erhalten wird, können wir ungefähr abschätzen, wie viele Menschen in Deutschland im nächsten Jahr arbeitslos sein werden. Und sowohl die Geburtenrate als auch die Anzahl Arbeitsloser scheint nicht massgeblich davon beeinflusst zu sein, ob eine bestimmte Person an einem bestimmten Abend entscheidet ins Kino zu gehen oder nicht. Es sieht so aus, als würde sich manchmal auf der gesellschaftlichen Makroebene doch eine gewisse Ordnung und Regelmässigkeit einstellen. Auch wenn es uns auf der sozialen Erklärungsebene Komplexität, Chaos und Emergenz verunmöglichen, einzelne Handlungen von Menschen sicher vorherzusehen, und damit auch den unfehlbaren Blick auf die Zukunft versperren, können wir gewisse Aspekte ebendieser durch eine etwas unscharfe Brille erkennen. Wenn sich nämlich Einzelhandlungen additiv aggregieren, wenn also aus einzelnen Mutterschaftsentscheidungen Geburtenraten oder aus einzelnen erfolglosen Jobsuchen Arbeitslosenraten werden, wird im Durcheinander wieder eine gewisse Ordnung erkennbar. In der sozialen Welt der Summen und Durchschnitte können wir einiges sehen. Ein etwas ungenauer und unsicherer Blick in unsere gesellschaftliche Zukunft ist also trotz Unregelmässigkeit möglich.

Die Tatsache, dass das soziale System auf der uns zugänglichen Beschreibungsebene sehr unregelmässig ist, scheint also eine (wenn nicht die) bedeutende Grenze für das Wissen über die Zukunft der Gesellschaft darzustellen. Unsere diesbezüglichen Prognosen werden stets etwas ungenau, unvollständig und unsicher sein. Wo allerdings die genauen Grenzen des Wissbaren verlaufen, oder wie viel Genauigkeit, Vollständigkeit und Sicherheit wir doch noch erlangen können, lässt sich nur schwer beurteilen. Dies ist vor allem deshalb der Fall, weil es insbesondere auch von Fähigkeiten der Maschinen, die wir in Zukunft entwickeln werden, abhängt, wie weit wir hier kommen werden. Da wir aber kaum je eine Maschine entwickeln werden, welche den Zustand aller Teilchen im Universum mit perfekter Genauigkeit feststellen kann, werden wir – insbesondere in Systemen, die so komplex, chaotisch, emergent und umweltabhängig sind wie menschliche Gesellschaften - wohl immer mit der Ungenauigkeit nicht-fundamentaler Erklärungsebenen zu kämpfen haben. Es besteht für uns hier irgendwo eine Grenze des Wissbaren

Diskussion und Fazit

Wo verlaufen also die Grenzen des Wissbaren in Bezug auf Prognosen in den Sozialwissenschaften? Zunächst einmal scheint keine Grenze durch eine vermeintliche fundamentale Besonderheit des menschlichen Handelns zu entstehen. Wie alles andere in unserer Welt unterliegen Menschen den Naturgesetzen. Sie verfügen über keinen (libertarisch gedachten) freien Willen, der ihr Verhalten und damit die Zukunft der Gesellschaft unvorhersehbar machen würde. Unsere erste These – Die Geschichte unserer Gesellschaft lässt sich nicht voraussagen, weil sie nicht determiniert ist – scheint also kein Licht auf die Unmöglichkeit sozialer Prognosen zu werfen. Wenn unsere Welt auf der Quantenebene indeterministisch ist, sind (potentiell) sämtliche Wissenschaften davon betroffen. Ist sie deterministisch, so gilt dies auch für die soziale Welt.

Wie sieht es mit der zweiten These – Die Geschichte unserer Gesellschaft lässt sich nicht voraussagen, weil sie nicht durch soziale Gesetze determiniert ist – aus? Ich habe sehr ausführlich dafür argumentiert, dass es

tatsächlich keine sozialen Gesetze gibt. Fundamentale Kausalbeziehungen, so die Annahme, bestehen nur auf physikalischer Ebene. Was wir auf anderen Erklärungsebenen beobachten, sind lediglich phänomenologische Regelmässigkeiten, die in der Regel nicht perfekt sind. Unsere sozialen Konzepte beziehen sich nun auf Entitäten, die zwar (auch physikalisch gesehen) gewisse Ähnlichkeiten aufweisen, jedoch über keine vollkommen homogene Kausalstruktur verfügen. Wenn man die Zusammenhänge zwischen diesen sozialen Entitäten betrachtet, wird man gewisse Regelmässigkeiten, jedoch niemals universelle soziale Gesetze feststellen können. Liegt also eine Grenze prognostischen sozialen Wissens in der Inexistenz sozialer Gesetze? Dies wäre nur dann der Fall, wenn sich soziale Prognosen nur auf der Basis von sozialen Gesetzen machen liessen. Dies muss allerdings nicht der Fall sein, da prinzipiell auch noch Prognosen aufgrund physikalischer Gesetze in Frage kommen. Auch unsere zweite These liefert uns noch keine klaren Anhaltspunkte zum Verlauf der Grenzen des Wissens über die Zukunft der Gesellschaft.

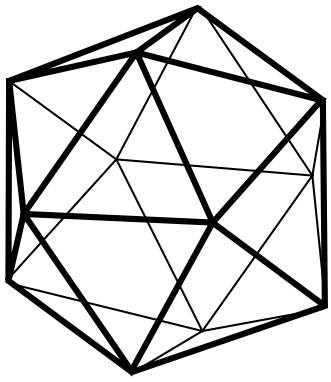
Aufschlussreicher scheint hingegen die These, dass die Unregelmässigkeit des sozialen Systems für die Unmöglichkeit oder Ungenauigkeit sozialer Prognosen verantwortlich ist. Wir können unser soziales System nicht auf Teilchenebene erfassen, um seine Entwicklung zu simulieren, da wir dazu den Zustand jedes Teilchens im Universum mit vollkommener Genauigkeit erfassen müssten. Daher sind wir gezwungen, auf uns zugängliche Erklärungsebenen mit den dazugehörigen physikalisch heterogenen Konzepten auszuweichen. Für die Beschreibung unseres sozialen Systems kommen beispielweise soziale, psychologische, biologische, für dessen Umwelt geographische Konzepte in Frage. Manche Systeme sind erstaunlich regelmässig und ihre zukünftige Entwicklung lässt sich auf höheren Erklärungsebenen gut approximieren. Das soziale System aber, ist äusserst komplex, emergent, chaotisch und umweltabhängig, oder anders gesagt: unregelmässig. Die entsprechenden Konzepte lassen sich nicht durch einfache Regeln verbinden, und können das Verhalten des Systems nur sehr unvollständig erfassen. Kleine Veränderungen in den Anfangsbedingungen können zu bedeutenden Unterschieden in der Entwicklung führen. Zudem ist die Gesellschaft von einer ebenfalls komplexen Umwelt umgeben und stark von dieser beeinflusst. All dies führt dazu, dass genaue und umfassende Prognosen in den Sozialwissenschaften nicht möglich sind. Wir haben es hier mit einer Grenze des Wissbaren zu tun. Ihr genauer Verlauf lässt sich allerdings nicht ausmachen. Es ist anzunehmen, dass wir immer kompliziertere soziale Modelle bilden und ausserdisziplinäres Wissen immer besser einbeziehen können. So wird es uns hier wohl gelingen, den Bereich des Wissens in Zukunft weiter auszubauen. Allerdings scheinen wir dies nicht unbegrenzt tun zu können. Da wir niemals sämtliche Teilchen des Universums mit perfekter Genauigkeit erfassen können, werden wir auf einer Erklärungsebene bleiben, auf welcher keine perfekten Zusammenhänge bestehen. Wir können nur emergierende Regelmässigkeiten beobachten – und diese sind in der sozialen Welt äusserst unregelmässig.

Vom alten Traum der exakten und sicheren Prognose von Revolutionen und Börsencrashs müssen wir uns also verabschieden. Ob und wann die Zukunft solche Ereignisse für uns bereithält, liegt jenseits der Grenze des für uns Wissbaren. Und nicht prinzipielle Probleme, wie etwa ein (vermeintlicher) freier Wille des Menschen oder die Unmöglichkeit sozialer Gesetze, scheinen dafür verantwortlich zu sein. Das ist vielmehr praktischer Natur: Unsere soziale Welt ist – wie vieles andere auch – schlicht zu komplex, zu gross, zu unübersichtlich und zu chaotisch, als dass wir vollständig vorhersehen könnten, was in ihr passieren wird. Das heisst aber nicht, dass wir nichts über künftige Revolutionen und Börsencrashs wissen können. Ungefähre und stets mit Unsicherheit behaftete Abschätzungen zukünftiger Trends und Ereignisse sind uns durchaus möglich. Wir werden nie mit ungetrübtem Blick in die Zukunft schauen, sondern stets eine unscharfe Brille tragen, welche uns nur einige Konturen und Farben erraten lässt. Aber auch wenn wir diese Brille niemals ablegen werden, können wir sie immer weiter schärfen. Und darin besteht meines Erachtens eine der zentralen Aufgaben der Sozialwissenschaften.

Martina Jakob (24) ist Bachelorstudentin in Sozialwissenschaften und Philosophie im 9. Semester. Sie interessiert sich für quantitative Soziologie, Development Economics und soziale Ungleichheit.

Literatur

- Cat, Jordi, «The Unity of Science» (2014): The Stanford Encyclopedia of Philosophy (Winter 2014 Edition), Edward N. Zalta (ed.), URL =http://plato.stanford.edu/archives/win2014/entries/scientific-unity/.
- Fay, B. (1994): General Laws and Explaining Human Behaviour, in: Changing Social Science (Sabia, D. R. & Wallulis, J., eds.), SUNY Press, Albany, 1983, pp. 103-128, reprinted in Martin & McIntyre (1994), pp. 91-110.
- Greschik, Stefan (1999): Das Chaos und seine Ordnung: Einführung in komplexe Systeme. München:Deutscher Taschenbuch Verlag.
- Hempel, Carl G. (1942). The function of general laws in history. Journal of Philosophy 39 (2):35-48.
- Keil, Geert (2007): Willensfreiheit, Berlin: de Gruyer.
- Kincaid, H. (1990): Defending laws in the social sciences. Philosophy of the Social Sciences 20 (1): 56-63.
- Little, Daniel (1993). On the scope and limits of generalizations in the social sciences. Synthese 97 (2):183 - 207.
- Nagel, Ernest (1970): Über die Aussage „Das Ganze ist mehr ist mehr als die Summe seiner Teile“ in Logik der Sozialwissenschaften / hrsg. von Ernst Topitsch. Königstein/Ts.: Verlagsgruppe Athenäum.
- O›Connor, T. und Wong, H. Y. (2015): «Emergent Properties», The Stanford Encyclopedia of Philosophy (Summer 2015 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/sum2015/entries/properties-emergent/.
- Papineau, David (2009): “Physicalism and the Human Sciences” in Mantzavinos, C. (ed.) (2009). Philosophy of the Social Sciences: Philosophical Theory and Scientific Practice. Cambridge University Press.
- Popper, Karl R. (1997): Prediction and Prophecy in the Social Sciences, Vortrag auf dem zehnten International Congress for Philosophy, Amsterdam 1984, hier nach der dt. Übersetzung: ders., Prognose und Prophetie in den Sozialwissenschaften, in: ders., Vermutungen und Widerlegungen: Das Wachstum der wissenschaftlichen Erkenntnis. Teilband II: Widerlegungen. Übersetzt von G. Albert, Tübingen: J. C. B. Mohr/Paul Siebeck, S. 487-503.
- Stoljar, Daniel (2015): «Physicalism», The Stanford Encyclopedia of Philosophy (Spring 2015 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/spr2015/entries/physicalism/.
- Van Riel, R. und Van Gulick, R. (2015): «Scientific Reduction», The Stanford Encyclopedia of Philosophy (Fall 2015 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/fall2015/entries/scientific-reduction/.



Wieso? Weshalb? Warum?

Über die Motivation von Philosophiestudenten[1]

Es ist so weit. Das Universitätsgebäude füllt sich wieder mit summenden Köpfen. Darunter auch die neuen Philosophiestudenten, manche frisch aus dem Ei geschlüpft, manche eher vom Typ Midlife-Crisis.

Die Studienwahl ist ein grosser Schritt. Zugegeben, heute nicht mehr ganz so gross wie er einst war, da Quereinsteiger und Karrierewandel zum Standard gehören, aber dennoch plant man zumindest die nächsten fünf Jahre mit einem Thema zu verbringen und viel Zeit investieren. Fünf Jahre, das sind, wenn man von einer Lebenserwartung von achtzig Jahren ausgeht, immerhin 6.25%. Das ist nicht wenig. Wenn man bedenkt, dass die meisten bereits mindestens zwanzig sind, bevor sie ihr Studentenleben beginnen, sind es sogar achtkommaperiodedrei Prozent des verbleibenden Lebens.

Wenn man Medizinstudenten nach der Motivation für ihre Studienwahl fragt, kriegt man eine ziemlich uniforme Antwort; sie wollen Menschen helfen, Leben retten, Schmerz reduzieren. Selbst die Wirtschaftler können sich noch eine gute Begründung aus dem Finger saugen; das Geschäft des Vaters, der gute Lohn. Doch bei den Geisteswissenschaftlern sieht es schwieriger aus, wenn man mal die zahlreichen anstrebenden Lehrpersonen ausblendet. Und ich wage es sogar zu behaupten, dass Philosophie, was die unklare Studienmotivation angeht, ein Spitzenreiter ist.

Klar ist aber, dass Philosophie ein Studienfach ist, das niemand ohne hinreichend Eigenmotivation durchhält. Neben den hoffentlich informativen Vorlesungen und Seminaren ist der bedeutendste Teil der Wissensaneignung selbstständiges Lesen und Verfassen von Texten; oftmals ohne klare Abgabedaten. Dass doch immer wieder Leute das Studium abschliessen, ohne, dass viel externe Motivation im Studienplan enthalten ist, lässt darauf schliessen, dass sie also gute Gründe dafür haben, sich mit Philosophie zu beschäftigen.

Philosophie gibt die Möglichkeit, von Grund auf zu hinterfragen. Es gibt niemals ein «Deshalb!» aufs «Weshalb?». «Wieso?», «Weshalb?» und «Warum?» zeichnen den guten Philosophiestudenten aus! Wer nicht fragt, der ist dumm! Soziale Normen sind zum Überprüfen da, Begriffe zum akribisch definieren und Theorien sind mit Gegenbeispielen zu bombardieren und Inkonsistenzen aufzudecken!

Dementsprechend ist die Dichte an Freigeistern und Querdenkern besonders hoch unter den Philosophiestudenten. Man erkennt sie an fehlendem Schuhwerk, viel schwarzer Kleidung und dem allgemeinen Konsens, dass THC legalisiert gehöre. Polyamoristen, Veganer, Buddhisten, Yogis, Musiker und zukünftige Geschäftsführer; kaum ein Völkchen ist so unterschiedlich. Doch auch kaum ein Völkchen übt sich so sehr in Toleranz und Akzeptanz.

Leute entscheiden sich dafür, Philosophie zu studieren, nicht, weil sie lernen wollen, sondern weil sie verstehen und ändern wollen. Sie wollen verbessern, sie wollen hinterfragen, sie akzeptieren nicht einfach, was ihnen eingetrichtert wurde. Und sie alle haben ihre Eltern nicht nur in der üblichen kleinkindlichen Fragephase zum Wahnsinn getrieben, sondern tun dies noch heute.

1 Der Einfachheit und Eleganz der Formulierungen halber wird in diesem Text konsequent die «geschlechtsneutrale» Form verwendet. Ausserdem vermissen wir noch immer ein gutes Suffix für wahre Geschlechterneutralität, um alle Leute in- oder ausserhalb des Geschlechterspektrums einzuschliessen. Das Redaktionsteam nimmt gerne Vorschläge entgegen.

Jasmine Kammermann (22) ist Bachelorstudentin in Philosophie und Musikwissenschaft im 5. Semester. Sie interessiert sich für Ästhetik, Philosophie der Musik und Logik.