



meta(Φ)

Urteilen, Werten, Handeln

Vollständige Syllogismen

• Marco R. Schori

Apriorität und Notwendigkeit

• Rafaela Schinner

Wie *man* spricht

• Luisa Pacozzi

What is Wrong about Crimes?

• Céline Hoog

How Should We Program Autonomous Driving Vehicles Under Moral Disagreement?

• Patrick Meinrad Müller

Wie entsteht wissenschaftlich fundierte Politik?

• Claus Beisbart

Ausgabe #9 2022

Herausgeberin

Fachschaft Philosophie der Universität Bern
Länggassstrasse 49a, 3012 Bern, Schweiz
mit Unterstützung des Instituts für Philosophie der Universität Bern

Redaktion

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

Chefredaktion

Vera Moser

Lektorat

Timo Junger, Isabel Käser, Vera Moser, Aleksandar Nikolić, Amina Roçi

Produktion und Gestaltung

Aleksandar Nikolić

Coverbild

Silja Bühlmann

Druck

Kromer Print AG Lenzburg

ISSN

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

Text-Einreichungen

Einreichungen zur Veröffentlichung werden gerne unter metaphi.philo@lists.unibe.ch empfangen. Die Auswahl der Texte erfolgt themenunabhängig, jedoch wird philosophische Relevanz des Themas vorausgesetzt. Für jede Ausgabe akzeptieren wir bis zu zwei Einreichungen pro Person. Es besteht kein Anrecht auf Veröffentlichung, weder bei erfolgreicher Einreichung noch nach Anfrage durch ein Redaktionsmitglied. In diesem Magazin veröffentlichte Texte gelten als wissenschaftliche Publikation und sind somit zitierfähig.

Spenden

Die Produktion des meta(φ) ist kostspielig. Sofern Ihnen das Journal gefällt, freut sich die Fachschaft Philosophie daher sehr über finanzielle Unterstützung. Spenden werden folgender Empfängerin zuteil:

Fachschaft Philosophie
Universität Bern, 3012 Bern
IBAN: CH53 0900 0000 3050 3703 2
Zahlungsgrund: «metaphi»

Erste Worte

Liebe Leser*innen

Neulich hatten sich mein Verlangen nach einem unbeschwertem Dasein und mein philosophischer Instinkt, alles zu hinterfragen, wieder einmal in die Haare gekriegt. In der Bäckerei, als mein Verlangen dem gedankenfreien Schlürfen meines Kaffees galt, erwachte mein Instinkt aus seiner schlummernden Existenz. Von der Ferne hörte ich die Worte einer Kundin: „Geben Sie mir bitte keine Tüte. So haben wir heute schon was Gutes getan.“ Unwillkürlich begann ich mich zu fragen: Ist es denn damit schon getan? Was wären die Bedingungen für gutes Handeln? Welche Rolle spielen dabei moralische Werte? Und gesellschaftliche Interessen? Wie können wir wissen, was gut ist? Und nach welchen Regeln sollen wir dies beurteilen?

Gerne hätte ich fundierte Antworten. Doch fundierte Antworten hängen nicht nur von den Fragen ab, die wir stellen, sondern auch vom Fundament, das wir legen. Um ein starkes Fundament zu bauen, müssen wir logischen Zusammenhängen folgen und wahre Prämissen aufstellen. Die Philosophie leistet dabei wichtige Arbeit. Um Ihnen einen Einblick in unsere Arbeit zu geben, haben wir für Sie auch in der neunten Ausgabe des *meta(φ)* eine Auswahl an Essays zusammengestellt. Unter dem Titel *Urteilen, Werten, Handeln* sind sechs Beiträge vereint, die von Logik und Erkenntnistheorie über Sprache bis hin zu Moral, Recht und Politik reichen. Im ersten Teil beschäftigen sich Marco R. Schori und Rafaela Schinner mit *Urteilen*. Schori greift bis tief in die Wurzeln der Logik und geht der Frage nach, was uns Aristoteles mit *τέλειος* sagen möchte. Er diskutiert verschiedene Interpretationen dieses Prädikats und kommt zum Schluss, dass Morisons Ansatz, der die Vollständigkeit als logische Eigenschaft von Syllogismen auffasst, am besten in das aristotelische Werk passt. Schinner befasst sich mit erkenntnistheoretischen Urteilen, indem sie dem Zusammenhang von Apriorität und Notwendigkeit nachgeht. Dabei setzt sie sich mit Kripkes Argumentation auseinander, die dafür spricht,

dass es mindestens eine Kategorie von a priori Propositionen gibt, die nicht notwendig sind.

Wenn wir Urteile fällen, kommt es auch auf Werte an. Zwei Beiträge zum Thema *Werten* zeigen auf, inwiefern unser Sprachgebrauch mit Wertungen verknüpft ist und weshalb gesellschaftliche Interessen bei der Bewertung von Handlungen eine Rolle spielen. Zum einen greift Luisa Pacozzi die Ansätze von Bourdieu und Young auf, die das Potenzial von legitimer Sprache kritisch betrachten. Dabei bringt sie zu Tage, wie unser Alltag mit solcher Sprache verwoben ist – was auf Anderssprechende unterdrückend wirken kann. Zum andern diskutiert Céline Hoog – ausgehend von Duffs Ansatz – den Unterschied von *mala prohibita* und *mala in se* und argumentiert dafür, dass die Abgrenzung weder sinnvoll zu ziehen ist noch einen praktischen Nutzen hat. Dabei wirft sie die Frage auf, wie in einer pluralistischen Wertegemeinschaft, mit verschiedenen Moralvorstellungen umgegangen werden soll. Diese Frage wird im dritten Teil von Patrick Müller aufgegriffen. Unter dem Leitwort *Handeln* fragt er danach, wie wir selbstfahrende Autos programmieren sollen, wenn wir uns nicht auf eine Moraltheorie einigen können. Sein Vorschlag ist, die Programmierung spieltheoretisch zwischen konkurrierenden Moraltheorien aushandeln zu lassen. Abgerundet wird der dritte und letzte Teil mit dem Fachwort von Claus Beisbart. Er erläutert, wie wissenschaftlich fundierte Politik entsteht, indem er die unterschiedlichen Funktionen von politisch und wissenschaftlich Handelnden unterstreicht.

Ich wünsche Ihnen eine spannende Lektüre und hoffe, die neunte Ausgabe des *meta(φ)* wird Sie zu fundierten Urteilen und gutem Handeln inspirieren.


Herzlichst, Vera Moser
im Namen des Redaktionsteams



Timo Junger



Isabel Käser



Vera Moser



Aleksandar Nikolić



Amina Roçi

“To look for a single general theory of how to decide the right thing to do is like looking for a single theory of how to decide what to believe.”

Thomas Nagel
Mortal Questions, 1979

“Understanding the value of something is not just a matter of knowing how valuable it is, but rather a matter of knowing how to value it—knowing what kinds of actions and attitudes are called for.”

T.M. Scanlon
What We Owe to Each Other, 1998

Inhaltsverzeichnis

Urteilen

Vollständige Syllogismen

- Marco R. Schori 4

Apriorität und Notwendigkeit

- Das kontingente Apriori gemäss Saul Kripke
- Rafaela Schinner 16

Werten

Wie *man* spricht

- Über das unterdrückende Potenzial legitimer Sprache
- Luisa Pacozzi 27

What is Wrong about Crimes?

- On the Distinction between *mala prohibita* and *mala in se* in Criminal Law
- Céline Hoog 36

Handeln

How Should We Program Autonomous Driving Vehicles under Moral Disagreement?

- Patrick Meinrad Müller 44

Wie entsteht wissenschaftlich fundierte Politik?

- Claus Beisbart 51

Vollständige Syllogismen

1. Einleitung

Die aristotelische Syllogistik unterscheidet zwei Gruppen von Syllogismen. Eine Gruppe wird als τέλειος (gr. für „vollständig“), die andere als ἀτελής (gr. für „unvollständig“) bezeichnet. Wie dies zu verstehen ist, ist schon seit längerer Zeit Gegenstand von philologischen und philosophischen Diskussionen. In der letzten Zeit sind interessante Beiträge erschienen, die ein neues Licht auf das Verständnis dieser beiden Gruppen werfen. So hat Benjamin Morison in seinem Aufsatz „What is a Perfect Syllogism?“ einen neuen Ansatz vorgestellt. Er argumentiert dafür, diese Unterscheidung zu verstehen, wie die Unterscheidung zwischen einem Argument, welches direkt dem Schlusschema des *Modus Ponendo Ponens* folgt, und einem, welches noch weitere Schritte verlangt, um die Konklusion ableiten zu können. Dies würde bedeuten, dass Aristoteles in seiner Syllogistik ein solches Schlusschema einführt, was oft bezweifelt wird. Es wird meist angenommen, dass Aristoteles kein solches Schlussprinzip vorstellt und man sich daher bei der Beurteilung, ob ein Schluss gültig ist oder nicht, nur auf das verlassen kann, was er selbst in der *Analytica Priora* geschrieben hat. Laut Morison (2015) ist ein solches Schlussprinzip jedoch im *dictum de omni et nullo*¹ zu finden. Sollte dies zutreffen, können die Zweifel als unbegründet betrachtet werden. Daraus würde sich somit auch die Möglichkeit eröffnen, alle möglichen Kombinationen von Termini und deren Verhältnisse auf deren Gültigkeit zu prüfen, ohne sich dabei auf diejenigen Syllogismen beschränken zu müssen, die Aristoteles selbst als gültig erachtet hat. Zudem widerspricht die Position, für welche im Folgenden argumentiert wird, der gängigen Überzeugung, dass die aristotelische Syllogistik eine Art axiomatisches Logiksystem ist, wobei die ersten vier Modi die Axiome

des Systems bilden. Vielmehr handelt es sich bei der Syllogistik des Aristoteles um eine Regellogik. Aus diesen Gründen wird im vorliegenden Beitrag untersucht, wie für die verschiedenen Interpretationen dessen, was ein *vollständiger* Syllogismus ist, argumentiert wird und wie sich Morisons Vorschlag davon unterscheidet.

Um das Ziel dieses Beitrags zu erreichen, wird in einem ersten Schritt (Unterabschnitt 2.1) kurz die Textgrundlage für die Debatte vorgestellt und auf deren wichtigste Bestandteile hingewiesen. Darauf folgt eine Erläuterung der zentralen Begriffe und Darstellungsweisen von Syllogismen (Unterabschnitt 2.2). Die erste Gruppe von Interpretationen, welche im Abschnitt 3 vorgestellt wird, ist diejenige von Łukasiewicz (1957), Patzig (1963) und Nortmann und Ebert (2007). In Abschnitt 4 wird eine weitere Gruppe von Interpretationen vorgestellt, welche anhand von Malinks (2013) Ausarbeitung des *dictum de omni et nullo* illustriert wird. Daraufhin wird die Position von Morison (2015) zuerst skizziert und anschliessend argumentativ ausformuliert (Abschnitt 5), wobei mögliche Einwände mitberücksichtigt werden. Es wird aufgezeigt, dass auf diese Einwände Antworten formuliert werden können, welche die Vorzüge von Morisons Position hervorheben.

2. Begriffliche Erläuterungen

Dieser Abschnitt dient dem besseren Verständnis der folgenden Argumentation. Dazu werden die zentralen Textstellen von Aristoteles' *Analytica Priora* kurz vorgestellt und auf einige wichtige Aspekte des griechischen Textes und der deutschen Übersetzung hingewiesen. Danach erläutere ich die Verwendung der wichtigsten Begriffe und Formeln.

2.1 Textgrundlage

Im ersten Kapitel *der Analytica Priora (An. Pr.)* führt Aristoteles zentrale Begriffe für seine Syllogistik ein. Dazu gehört unter anderem auch, was es für einen

¹ Mit *dictum de omni et nullo* wird die Textstelle bezeichnet, um die es im vorliegenden Beitrag geht. Diese Bezeichnung kann mit „Aussage über alles und nichts“ übersetzt werden. Dadurch wird jedoch bereits die gängige Interpretation nahegelegt, gegen welche im Folgenden argumentiert wird.

συλλογισμός („syllogismos“, gr. für Deduktion oder Argument) bedeutet, vollständig, resp. unvollständig beschaffen zu sein (καὶ ποῖος τέλειος καὶ ποῖος ἀτελής, *An. Pr.* 1.1, 24^a 12-13). Die Definition dieser beiden Arten von Syllogismen findet sich in der folgenden Textstelle (*An. Pr.* 1.1, 24^b 22-26):

„Vollkommen‘ nenne ich einen Syllogismus, bei dem es über die Annahmen hinaus keines weiteren (Schrittes) bedarf, um das Notwendige einleuchtend zu machen. – ‚Unvollkommen‘ nenne ich einen Syllogismus, der | (dazu) eines oder mehrerer (Schritte) bedarf, die zwar aufgrund der (Verhältnisse zwischen den) gegebenen Termini notwendig, aber nicht mit den Prämissen angenommen sind.“ (Nortmann und Ebert 2007, 16)²

Der entscheidende Begriff dieser Definition, um *vollständige* (τέλειος) von *unvollständigen* (ἀτελής) Syllogismen zu unterscheiden, ist τὸ φανῆναι. In vorliegender Übersetzung wird dies mit „einleuchtend zu machen“ wiedergegeben. Dadurch wird eine Lesart nahegelegt, welche im folgenden Kapitel ausformuliert und kritisiert wird. An dieser Stelle sollte vorerst nur zur Kenntnis genommen werden, dass die Bezeichnung eines Syllogismus als *vollständig* darauf zurückzuführen ist, dass die Angabe der Prämissen ausreicht. Wie „das Notwendige einleuchtend zu machen (πρὸς τὸ φανῆναι τὸ ἀναγκαῖον)“ zu verstehen ist, wird in der Diskussion der folgenden Abschnitte erläutert.

Eine weitere zentrale Stelle für das Verständnis dessen, was einen vollständigen Syllogismus ausmacht, ist das sogenannte *dictum de omni et nullo* (*An. Pr.* 1.1, 24^b 28-30):

„Wir reden von ‚Von-jedem-Ausgesagtwerden‘, wenn man keines der unter den Subjektterminus fallenden Dinge herausgreifen kann, | von dem das andere nicht ausgesagt wird. Und bei ‚Von-keinem-(Ausgesagtwerden)‘ ebenso.“ (Nortmann und Ebert 2007, 16)

Diese Textstelle wird üblicherweise als Definition der affirmativen und privativen assertorischen Allaussage verstanden. Auch in der Übersetzung von Nortmann und Ebert wird dies anhand der Konstruktion „Von-jedem-Ausgesagtwerden“ für die affirmative Allaussage (Aa_xB) und „Von-keinem-(Ausgesagtwerden)“ für die privative Allaussage (Ae_xB) ersichtlich. Diese Lesart

wird der modalen Formulierung des *dictum de omni et nullo* im griechischen Text nicht gerecht. Daher argumentiere ich im Folgenden für die Ansicht von Morison, dass das *dictum* als Schlussregel verstanden werden sollte. Morison kritisiert, dass die Interpretation des *dictums* als *Erläuterung* von affirmativen und privativen Allaussagen der Modalität im griechischen Text in folgenden Punkten nicht gerecht wird: Das im Konjunktiv geschriebene ἢ kann auch „es ist möglich“ bedeuten. Die Modalpartikel ὅταν wird üblicherweise für Temporalsätze verwendet und bedeutet so viel wie „(immer) wenn, sooft“. Das Futur λεχθήσεται sollte als ‚schlussfolgerndes‘ Futur übersetzt werden (Morison 2015, 132–133). Solche Konditionalsätze zeichnen sich dadurch aus, dass, wenn die Bedingung angenommen wird, das Bedingte *notwendig* folgt. Wenn diese Feinheiten des griechischen Texts nicht in der Übersetzung ersichtlich werden, geht die tatsächliche Bedeutung des *dictum de omni et nullo*, für die ich in Abschnitt 5 argumentiere, verloren.

2.2 Verwendung von Begriffen und Formeln

Im vorliegenden Beitrag werden Begriffe verwendet, deren Übersetzung und Bedeutung in der Debatte teilweise unterschiedlich gehandhabt werden. Um Missverständnissen vorzubeugen, werden in diesem Unterabschnitt kurz die wichtigsten Begriffe und deren Verwendung in diesem Beitrag erklärt.

Unter Syllogistik wird gemeinhin ein logisches System verstanden, welches die Untersuchung von Syllogismen oder Schlussfiguren behandelt. Daher können auch stoische oder scholastische Logiksysteme dementsprechend bezeichnet werden. In diesem Beitrag beschränkt sich die Bedeutung jedoch auf das System von Aristoteles und die Interpretationen dieses Systems, welche hier besprochen werden.

Obwohl der Begriff des Syllogismus im Deutschen gut mit „Deduktion“ oder „Argument“ wiedergegeben werden kann, wird in diesem Beitrag meist nur „Syllogismus“ verwendet. Es soll darunter ein gültiges Argument verstanden werden, welches in der *Analytica Priora* beschrieben wird. Ein solches kann in einer von drei Figuren formuliert werden. Eine Figur beschreibt, welche Termini in welcher Prämisse Prädikat oder Subjekt sind. So ist bei der ersten Figur der Aussenterminus (*A*) der oberen Prämisse das Prädikat und der Mittelterminus (*B*) das Subjekt. Der Mittelterminus (*B*) zeichnet sich dadurch aus, dass er in beiden Prämissen vorkommt. In der unteren Prämisse ist der Mittelterminus bei der ersten Figur das Prädikat und der Aussenterminus (*C*) das Subjekt. Der Aussenter-

² Im Anhang befindet sich jeweils der griechische Text, auf den im Fließtext Bezug genommen wird.

minus der oberen Prämisse (A) ist in der Konklusion immer das Prädikat, derjenige der unteren Prämisse immer das Subjekt. Dies wird in Formeln stets anhand der folgenden Abfolge dargestellt, wobei der erste Grossbuchstabe jeweils das Prädikat, der zweite das Subjekt ist:

1. Figur

A ... B (Obere Prämisse)
 B ... C (Untere Prämisse)
 ∴ A ... C (Konklusion ∴)

2. Figur

M ... N (Obere Prämisse)
 M ... X (Untere Prämisse)
 ∴ N ... X (Konklusion ∴)

3. Figur

P ... S (Obere Prämisse)
 R ... S (Untere Prämisse)
 ∴ P ... R (Konklusion ∴)

Die Termini stehen zudem immer in einem qualitativen und quantitativen Verhältnis zueinander. So können Aussagen affirmativ oder privativ (qualitativ) und allgemein, partikulär oder unbestimmt (quantitativ) sein. Dies wird mit den Kleinbuchstaben *a*, *e*, *i* und *o* wiedergegeben. „ Φ “ steht in der folgenden Tabelle für das Prädikat und „ Ψ “ für das Subjekt einer Aussage:

	affirmativ	privativ
allgemein	<i>a</i> Φ wird von jedem Ψ ausgesagt	<i>e</i> Φ wird von keinem Ψ ausgesagt
partikulär	<i>i</i> Φ wird von einigen Ψ ausgesagt	<i>o</i> Φ wird von einigen Ψ nicht ausgesagt
unbestimmt	–	–

Da Aristoteles sowohl ein assertorisches (einfaches) als auch ein apodiktisches (modales) ‚Ausgesagtwerden‘ behandelt, werden hier die verschiedenen Modalitäten anhand von Indexikalien bei den Kleinbuchstaben angegeben. Das einfache oder assertorische ‚Ausgesagtwerden‘ wird mit einem X, das notwendige ‚Ausge-

sagtwerden‘ mit einem N und das beidseitig-mögliche ‚Ausgesagtwerden‘ mit einem Q dargestellt.

Wenn nun bei einer Figur die entsprechenden Bestandteile zusammengefügt werden, entstehen (gültige) Modi. Ein solcher Modus ist beispielsweise *Barbara QXQ*. Dieser kann schematisch folgendermassen aufgeschrieben werden:

Barbara QXQ

Aa_QB (Obere Prämisse)
 Ba_XC (Untere Prämisse)
 ∴ Aa_QC (Konklusion ∴)

Prämissen und Konklusionen werden auch als Aussagen bezeichnet. Prämissen heissen ausserdem Annahmen. Dies sind verschiedene Übersetzungsmöglichkeiten von *πρότασις*.

3. Vollkommene Syllogismen

Die erste Gruppe von Interpretationen, die nun betrachtet wird, übersetzt *τέλειος* mit ‚vollkommen‘ (vgl. Nortmann und Ebert 2007 oder Patzig 1963) und *φανῆναι* mit ‚einleuchtend‘ (Nortmann und Ebert 2007, 16) oder ‚evident‘ (Patzig 1963, 56–58). Es wird davon ausgegangen, dass die Bezeichnung eines Schlusses als *vollkommen* auf ein epistemologisches Phänomen, eine *evidente* Einsicht in die Gültigkeit der *vollkommenen* Syllogismen hinweise.

Diese Auffassung vertritt beispielsweise Łukasiewicz, wenn er schreibt, dass ‚vollkommene Syllogismen augenscheinliche [in der englischen Ausgabe ‚self-evident‘] Aussagen sind, welche weder einen Beweis brauchen noch haben‘ und ‚nicht-beweisbare, wahre Aussagen sind‘ (1957, 43). Er bezeichnet Syllogismen als Aussagen, da er sie als Konditionalaussagen betrachtet, wobei das Antecedens aus der Konjunktion der Prämissen und das Konsequens aus der Konklusion bestehe. Daraus ergebe sich für die vollkommenen Syllogismen, dass sie als die Axiome der aristotelischen Syllogistik betrachtet würden, die die Ableitungsbasis für die unvollkommenen Syllogismen bildeten (Łukasiewicz 1957, 43).

Dass die ersten vier Modi der assertorischen Syllogistik und die meisten Modi der ersten Figur der apodiktischen Syllogistik selbsterklärend sind, kann jedoch bereits mit dem Beweis des ersten Modus (*Barbara*: Aa_XB , Ba_XC ; Aa_XC) von Aristoteles widerlegt werden (*An. Pr.* 1.4, 25^b 37-40):

„Denn wenn das A von jedem B und das B von jedem C (ausgesagt wird), so wird notwendig auch das A von jedem C ausgesagt. **Oben ist**

nämlich erläutert worden, wie das ‚Von-jedem-(Ausgesagtwerden)‘ zu verstehen ist.“ (Nortmann und Ebert 2007, 19, Hervorhebung M.R.S.)

Mit dem Verweis auf eine bereits erbrachte Begründung für das notwendige Folgen der Konklusion des Modus *Barbara* XXX kann nicht die Rede davon sein, dass Aristoteles gedacht habe, die *vollkommenen* Syllogismen seien selbsterklärend und bedürften keines Beweises. Die Begründung der *Gültigkeit* dieser Modi ist das bereits angesprochene *dictum de omni et nullo*, worauf in dieser Textstelle deutlich verwiesen wird (Morison 2015, 112–113).

Dass Aristoteles die *vollkommenen* Syllogismen als die Axiome seines Systems betrachtet hat, kann man ebenfalls in Zweifel ziehen. Denn in *An. Pr.* 1.7, 29^b 15–19 schreibt er, dass sich die Modi *Darii* (Aa_xB , Bi_xC ; Ai_xC) und *Ferio* (Ae_xB , Bi_xC ; Ao_xC) von *Barbara* XXX und *Celarent* XXX (Ae_xB , Ba_xC ; Ae_xC) ableiten lassen. Er nennt sie dennoch weiterhin *vollkommen*, was darauf schliessen lässt, dass es hinreichend ist, dass „es über die Annahmen hinaus keines weiteren (Schrittes) bedarf, um das Notwendige einleuchtend zu machen“, um einen Modus als *vollkommen* zu klassifizieren (Morison 2015, 158).

Sowohl Patzig (1963) als auch Nortmann und Ebert (2007) begründen die *Vollkommenheit* dieser Gruppe von Syllogismen anhand einer *evidenten Einsicht* in deren Gültigkeit. Gegenüber Łukasiewicz sind sie jedoch nicht der Auffassung, dass diese Evidenz keiner Begründung bedarf. Beide gehen davon aus, dass es sich dabei um eine Evidenz handelt, welche ersichtlich wird, wenn man die Syllogismen in einer gewissen *Form* wiedergibt. Das heisst, dass die Evidenz durch die *Abfolge* der Termini ersichtlich werde (Patzig 1963, 59; Nortmann und Ebert 2007, 295–296). Sie sind sich aber uneinig darüber, was evident wird, weshalb sie eine Art Prädikatenstafette vorschlagen. Durch die Abfolge der Termini in der geläufigen Formulierung der ersten Figur und insbesondere in der Formulierung bei 1. 4, 25^b 32–35 werde das Verhältnis deutlich, welches im *dictum de omni et nullo* für a_x - und e_x -Prämissen expliziert werde (Nortmann und Ebert 2007, 293–296).³ Patzig verweist hingegen auf

die Transitivität der Relation ‚Von-jedem-Ausgesagtwerden‘ im Fall von *Barbara* XXX. Für die restlichen *vollkommenen* Syllogismen führt er einen Vergleich mit einer modernen Relationslogik an (Patzig 1963, 61–66). Die Positionen von Patzig sowie Nortmann und Ebert gehen jedoch davon aus, dass die Gültigkeit dieser Syllogismen anhand einer Formulierung besonders *evident* wird (*An. Pr.* 1.4, 25^b 32–35):

„Wenn sich nun drei Termini so zueinander verhalten, daß der letzte im mittleren als einem Ganzen (enthalten) ist und der mittlere im ersten als einem Ganzen entweder (enthalten) ist oder nicht (enthalten) ist, kommt notwendig ein vollkommener Syllogismus hinsichtlich der Außentermini | zustande.“ (Nortmann und Ebert 2007, 19)

An dieser Stelle soll nicht der Auffassung widersprochen werden, dass die Formulierung der Prämissen und die Reihenfolge der Termini und Prämissen besonders deutlich hervorhebt, dass der Schluss gültig ist. Aristoteles verwendet diese Formulierung und Reihenfolge der Termini und Prämissen auch an einer anderen Stelle (*An. Pr.* 1.15, 33^b 33–36), um deutlich hervorzuheben, dass der modale Modus *Barbara* QXQ gültig ist (Nortmann und Ebert 2007, 295–296). Das Problem dieser Interpretation besteht vielmehr darin, dass eine Abstufung der *Vollkommenheit* suggeriert wird, welche im Text der *Analytica Priora* nicht wiedergefunden werden kann. Wenn ein Syllogismus als τέλειος bezeichnet wird, geschieht dies nicht abhängig von der Formulierung seiner Prämissen und es wird auch nicht näher gelegt, dass die *Vollkommenheit* in der oben aufgeführten Formulierung (C ist in B , B ist in A ; C ist in A) in irgendwelcher Weise stärker hervortritt als in einer anderen Formulierung (A kommt jedem B zu, B kommt jedem C zu; A kommt jedem C zu). Die Formulierung der Rechtfertigung für die Gültigkeit von *Barbara* QXQ scheint dieser Auffassung gar zu widersprechen (vgl. Morison 2015, 150–152): Die untere Prämisse wird zwar zuerst angegeben und ist in derselben Art wie in *An. Pr.* 1.4, 25^b 32–35 formuliert, die obere Prämisse wird jedoch in der üblichen Formulierung wiedergegeben. Daraus ergibt sich eine Abfolge der Termini ($C-B-A-B$), welche nicht für die ‚Prädikatenstafette‘ geeignet ist. Dennoch soll diese Formulierung greifbarer machen, weshalb *Barbara* QXQ gültig ist. Die Reihenfolge der Prämissen scheint auch aufgrund der Formulierung

den Syllogismen direkt und auf eine besondere Weise vom *dictum* garantiert werden (2007, 402).

3 Obwohl Nortmann und Ebert das *dictum de omni et nullo* als das Verhältnis bezeichnen, welches *evident* werde, können sie nicht derjenigen Gruppe von Interpretationen des τέλειος συλλογισμού zugerechnet werden, welche die Bezeichnung τέλειος als Hervorhebung einer logischen Eigenschaft verstehen. Der Unterschied besteht darin, dass Nortmann und Ebert die *Vollkommenheit* als evidente Einsicht in die Gültigkeit des Syllogismus aufgrund seiner Formulierung betrachten (2007, 293), wohingegen beispielsweise Barnes unter *Vollständigkeit* versteht, dass die entsprechen-

von *Felapton XXX* in *An. Pr.* 1.6, 28^a 26-28 keine besondere Wichtigkeit zu besitzen (Morison 2015, 111). Daraus ergibt sich, dass die Bezeichnung eines Syllogismus als *τέλειος* nicht darauf zurückzuführen ist, wie deutlich oder *evident* die Gültigkeit hervortritt. Es kann also davon ausgegangen werden, dass es sich bei der *Vollkommenheit* nicht um ein epistemologisches Phänomen oder eine *evidente* Einsicht handelt.

4. Vollständige Syllogismen

Die zweite Gruppe von Interpretationen, zu welcher auch die Position dieses Beitrages gezählt werden kann, betrachtet die Bezeichnung eines Syllogismus als *τέλειος* als die Hervorhebung einer *logischen* Eigenschaft. Um Missverständnissen vorzubeugen, wird, wenn von den *τέλειον συλλογισμὸν* aus der Sicht dieser Gruppe gesprochen wird, das Wort *τέλειον* mit „vollständig“ übersetzt. Auf die Bedeutung von *τὸ φανῆναι* muss an dieser Stelle noch nicht eingegangen werden. Dies wird im Abschnitt 5 erläutert. Wie im Abschnitt 3 bereits angesprochen, begründet Aristoteles die *Gültigkeit* der *vollständigen* Syllogismen mit einem Verweis auf das (bei den modalen *vollständigen* Syllogismen entsprechend *modale*) *dictum de omni et nullo*. Darin ist sich diese Gruppe von Interpretationen einig. Uneinigkeit herrscht hingegen darüber, wie das *dictum* verstanden werden soll. Entgegen der Ansicht von Morison verstehen es Malink (2013, 19) sowie Nortmann und Ebert (2007, 230) als *Erläuterung* von *a*- und *e*-Prämissen. Barnes (2007, 412) betrachtet es gar als Definition von affirmativen und privativen Allaussagen. Morison (2015, 131–136) kritisiert diese Interpretationen des *dictums*. Weiter werden in dieser Debatte zwei Semantiken unterschieden, welche auf dem *dictum* basieren. Sowohl Barnes als auch Nortmann und Ebert vertreten eine *orthodoxe* Lesart, Malink versteht sie hingegen *heterodox*. Diese Benennung geht auf Barnes (2007, 406–412) zurück.

Der Grundgedanke dieser beiden Interpretationen wird nun anhand der ‚*abstrakten dictum Semantik*‘ von Malink (2013) illustriert. Demnach wird das *dictum* als notwendige und hinreichende Bedingung für affirmative (und privative) Allaussagen ausgelegt und „...keines ...herausgreifen kann“ (*μηδὲν ἧ λαβεῖν*, *An. Pr.* 1.1, 24^b 29) mit ‚da ist kein ...‘ gleichgesetzt (Malink 2013, 34–35). Daraus ergibt sich für das *dictum de omni* (Malink 2013, 35):

Aa_xB genau dann, wenn für jedes *Z*, wenn *Z* eines der *B* ist, dann wird *A* von *Z* ausgesagt

„[D]er unter den Subjekterminus fallenden Dinge“ (*τῶν τοῦ ὑποκειμένου*, *An. Pr.* 1.1, 24^b 29)⁴ weise auf eine Vielzahl von Dingen hin, welche unter das Prädikat fallen würden. Diese Relation gibt Malink mit der Formulierung ‚Teil der Vielzahl assoziiert mit‘ (fortan *tvam*) wieder. Die *orthodoxe* Interpretation unterscheidet sich von der *heterodoxen* nun darin, was unter dem ‚Teil der Vielzahl‘ verstanden wird. Die *orthodoxe* Interpretation geht davon aus, dass Aristoteles hier von Einzelgegenständen spricht und rekonstruiert die Argumente daher meist prädikatenlogisch (Nortmann und Ebert 2007, 229; Barnes 2007, 407). Die *heterodoxe* Interpretation versteht den ‚Teil der Vielzahl‘ hingegen nicht als Referenz auf Einzelgegenstände, die unter die Termini fallen, sondern schliesst auch die Möglichkeit von Gattungsbegriffen und Klassenbegriffen mit ein (Malink 2013, 47). Die ‚*abstrakte dictum Semantik*‘ bleibt bezüglich dieser Interpretationen neutral und kann daher von beiden Positionen akzeptiert werden. Malink formuliert die *Semantik* für das *dictum de omni et nullo* folgendermassen (2013, 37):

Aa_xB genau dann, wenn $\forall Z(Z\ tvam\ B \rightarrow Z\ tvam\ A)$
 Ae_xB genau dann, wenn $\forall Z(Z\ tvam\ B \rightarrow \neg Z\ tvam\ A)$

Mithilfe dieser Gleichsetzungen und den üblichen Regeln für die Junktoren und Quantoren der klassischen Prädikatenlogik lässt sich nun zeigen, wie beispielsweise *Barbara XXX* bewiesen werden kann (analog zum Beweis für *Celarent XXX* in Malink 2013, 38):

Barbara

1. Aa_xB (obere Prämisse)
2. Ba_xC (untere Prämisse)
3. $\forall Z(Z\ tvam\ B \rightarrow Z\ tvam\ A)$ (von 1; durch das abstrakte *dictum de omni*)
4. $\forall Z(Z\ tvam\ C \rightarrow Z\ tvam\ B)$ (von 2; durch das abstrakte *dictum de omni*)
5. $\forall Z(Z\ tvam\ C \rightarrow Z\ tvam\ A)$ (von 3,4)
6. $\therefore Aa_xC$ (von 5; durch das abstrakte *dictum de omni*)

Obwohl dies fruchtbare Ansätze für die Rekonstruktion in modernen formalen Systemen (*orthodox*: Mengenlehre-Semantik; *heterodox*: Quasiordnung-Semantik) liefert, hilft es aus den unten aufgeführten Gründen nicht weiter, wenn man verstehen möchte, weshalb Aristoteles einige Syllogismen als *vollständig* bezeichnet (Malink 2013, 39).

⁴ Malink (2013) und Morison (2015) präferieren die Textausgabe mit dem pluralen Artikel im Genitiv *τῶν*, weswegen der griechische Text an dieser Stelle auch so wiedergegeben wird. Die Auslassungen der verschiedenen Textausgaben werden anhand der Textstellen im Anhang ersichtlich.

Ein erstes Problem ergibt sich, da die ‚abstrakte Semantik‘ die Syllogismen *Darii XXX* und *Ferio XXX* nicht allein durch das *dictum de omni et nullo* beweisen kann. Malink führt daher zwei neue *dicta* ein: das *dictum de aliquo* und das *dictum de aliquo non*. Er leitet sie vom *dictum de omni et nullo* anhand dessen ab, was Aristoteles über die Widerspruchsverhältnisse der vier assertorischen Aussagemöglichkeiten sagt (Malink 2013, 36–37). Auch Barnes erwägt die Einführung des *dictum de aliquo et aliquo non*, verwirft diese Idee jedoch wieder, da Aristoteles selbst schreibt, dass das *dictum de omni et nullo* ausreiche, um auch *Darii XXX* und *Ferio XXX* zu beweisen (Barnes 2007, 403–404). Dies kann an der folgenden Textstelle deutlich gemacht werden (*An. Pr.* 1.4, 26^a 23–25):

„Denn einmal angenommen, daß das A jedem B, das B irgendeinem C zukommt. **Wenn nun das ‚Von-jedem-Ausgesagtwerden‘ das ist, als was wir es zu Beginn definiert haben**, so kommt | das A notwendig irgendeinem C zu.“ (Nortmann und Ebert 2007, 20, Hervorhebung M.R.S.)

Laut Morison ist die Einführung des *dictum de aliquo et aliquo non* zwar ein gangbarer Weg, doch ist sie nicht durch den Text von Aristoteles belegt (2015, 135). In eine ähnliche Richtung geht ein weiterer Kritikpunkt von Morison (2015, 133–136): Malink muss nicht nur zwei *dicta* ergänzen, um zu zeigen, dass einige *vollständige* Syllogismen gültig sind, sondern er wendet ein formales, modernes Logiksystem an. Wie auch Malink selbst eingesteht, erklärt dies nicht, weshalb Aristoteles diese Syllogismen *vollständig* nennt (Malink 2013, 39), da es beispielsweise auch möglich wäre, dieses Logiksystem zum Beweis der restlichen, nicht-*vollständigen* Syllogismen zu verwenden.

Ein dritter Kritikpunkt von Morison lautet, dass durch diese Interpretationen die bereits angesprochene Modalität von η und insbesondere in der Verbindung mit dem Futur $\lambda\epsilon\chi\theta\acute{\eta}\sigma\epsilon\tau\alpha$ nicht ausreichend Rechnung getragen wird (2015, 132–133). Dieser letzte Kritikpunkt trifft ebenfalls auf die erste Gruppe von Interpretationen zu. Im folgenden Abschnitt wird nun vorgestellt, wie die Position von Morison die Probleme der in Abschnitt 3 und 4 vorgestellten Gruppen von Interpretationen löst.

5. Das *dictum de omni et nullo* verstanden als Schlussregel

Wie bereits angesprochen, kann die Interpretation von Morison ebenfalls derjenigen Gruppe zugerechnet werden, die $\tau\acute{\epsilon}\lambda\epsilon\iota\omicron\varsigma$ als Hervorhebung einer *logischen* Eigen-

schaft von Syllogismen betrachtet. Aus diesem Grund wird auch hier $\tau\acute{\epsilon}\lambda\epsilon\iota\omicron\varsigma$ mit „vollständig“ übersetzt. $\tau\omicron$ $\phi\alpha\nu\eta\nu\alpha\iota$ wird hingegen als „sichtbar zu werden“ übersetzt (Morison verwendet im Englischen „to become apparent“ (2015, 109)). Dies soll darauf hinweisen, dass die *Gültigkeit* der (*vollständigen*) Syllogismen sichtbar gemacht wird. Anders ausgedrückt, kann man sagen, dass sich laut Morison die *vollständigen* von den *unvollständigen* Syllogismen dadurch unterscheiden, dass das notwendige Folgen der Konklusion aus den Prämissen auf unterschiedliche Weise sichtbar gemacht werden muss (Morison 2015, 114). Die *vollständigen* Syllogismen sind *vollständig*, weil die Prämissen und die Konklusion derart zueinander in Beziehung stehen, wie es vom *dictum de omni et nullo* beschrieben wird und keine weiteren Annahmen (durch Konversion oder *reductio ad impossibile*) abgeleitet werden müssen (Morison 2015, 112). Im Sinne dieser Lesart wird das *dictum de omni et nullo* nicht als Explikation von den Konditionen verstanden, die herrschen müssen, um *a-* oder *e-*Prämissen als ‚wahr‘ oder ‚falsch‘ zu bewerten. Vielmehr gibt es an, was jemand tun kann, wenn eine *a-* oder *e-*Prämisse in einem Argument vorkommt. Das *dictum de omni et nullo* kann somit als *Schlussregel* verstanden werden, die eine vergleichbare Rolle wie die Schlussregel *Modus Ponendo Ponens* in Kalkülen der klassischen Aussagenlogik einnehmen kann (Morison 2015, 140–141).

Im Folgenden wird nun untersucht, wie das *dictum de omni et nullo* interpretiert werden muss, um es als *Schlussregel* verstehen zu können. Zudem werden Probleme dieser Lesart deutlich gemacht und zusätzliche Gründe angegeben, diese Interpretation anzunehmen. Wenn die Interpretation von Morison als annehmbare Alternative in Betracht gezogen werden soll, muss sie in der Lage sein, das Vorgehen von Aristoteles zu erklären, ohne die Probleme der bereits besprochenen Interpretationen aufzuweisen. Aus Platzgründen beschränkt sich die Verteidigung der Position von Morison (2015) weitgehend auf die Behandlung derjenigen Textstellen der *Analytica Priora*, welche problematisch für seine Position sind.

Entgegen der Ansicht von Łukasiewicz wird hier davon ausgegangen, dass Aristoteles die *Gültigkeit* der *vollständigen* Syllogismen erklärt, beziehungsweise anhand des *dictum de omni et nullo* sichtbar werden lässt. Die Angabe des *dictums* als *Aufweis* der Gültigkeit wird in der im Abschnitt 4 zitierten Textstelle (*An. Pr.* 1.4, 26^a 23–25) sehr deutlich. Auch bei der Besprechung von einigen modalen *vollständigen* Syllogismen (*An. Pr.* 1.8, 30^a 2–3⁵ oder

5 An dieser Stelle bespricht Aristoteles die *e-*Konversion mit einem Hinweis auf das *dictum de omni et nullo*.

1.14, 32^b 40-33^a 1 und 33^a 24-25) kann man einen mehr oder weniger deutlichen Hinweis auf das (entsprechend modale) *dictum de omni et nullo* feststellen. Ein weiterer wichtiger Hinweis ist, dass Aristoteles bei der Behandlung von *Ferio QQQ* davon spricht, dass der *Beweis* derselbe sei (ἀπόδειξις δ' ἡ αὐτή, *An. Pr.* 1.14, 33^a 27) wie bei *Darii QQQ* (Morison 2015, 125–126). Sowohl Nortmann und Ebert als auch Patzig beharren hingegen nicht darauf, dass *vollständige* Syllogismen im selben Sinn wie bei Łukasiewicz als selbsterklärende Schlüsse zu verstehen sind. Die evidente Einsicht komme vielmehr aufgrund der Abfolge der Termini zustande (Nortmann und Ebert 2007, 295–296). Deutlich werde dies sowohl bei der Behandlung von *Barbara QXQ* als auch der ersten Formulierung von *Barbara XXX* und *Celarent XXX* (*An. Pr.* 1. 4, 25^b 32-35). Morison führt noch zwei weitere Textstellen an, die für diese Interpretation zu sprechen scheinen. Dies ist einerseits die Rechtfertigung für *Barbara NXN* (*An. Pr.* 1. 9, 30^a 15-23) und andererseits der namenlose und *unvollständige* Syllogismus der ersten Figur *aoi-QQQ* (*An. Pr.* 1. 14, 33^a 27-34). Vor allem die Wortwahl bei der Figur *aoi-QQQ* stellt laut Morison einen besonderen Fall dar (2015, 156). Aristoteles schreibt, dass „aufgrund der angenommenen Prämissen kein **evidenter Syllogismus** zustande [kommt]“ (διὰ μὲν τῶν εἰλημμένων προτάσεων οὐ γίνεται φανερός συλλογισμός, in Nortmann und Ebert 2007, 39, Hervorhebung M.R.S.). Die Verwendung von φανερός συλλογισμός lässt vermuten, dass Aristoteles tatsächlich evidente Syllogismen meint, wenn er von *vollständigen* Syllogismen spricht. Morison merkt jedoch an, dass auf die Definition von *vollständigen* Syllogismen verwiesen wird, wenn Aristoteles schreibt, dass es „*aufgrund der angenommenen Prämissen* nicht zu einem evidenten Syllogismus“ kommt (Morison 2015, 156–157). Die Emphase dieser Textstelle liegt demnach auf der Tatsache, dass die angegebenen Prämissen nicht ausreichen, um das notwendige Folgen der Konklusion sichtbar werden zu lassen (Morison 2015, 157). Dies legt die Interpretation von Morison nahe und nicht diejenige von Nortmann und Ebert oder Patzig. Die Übersetzung der Definition von *vollständigen* Syllogismen kann somit folgendermassen angepasst werden (*An. Pr.* 1. 1, 24^b 22-24):

„Vollständig‘ nenne ich einen Syllogismus, bei dem es über die Annahmen hinaus keines weiteren (Schrittes) bedarf, *damit* das notwendige (Folgen der Konklusion) sichtbar wird.“

Nachdem nun klar ist, wie diese Definition verstanden werden sollte, muss noch aufgezeigt werden, wie das *dictum de omni et nullo* verstanden werden sollte, um

daraus eine, beziehungsweise zwei, Schlussregeln gewinnen zu können.

Anhand der Rechtfertigung der Gültigkeit von *Barbara NXN* und *Celarent NXN* wird laut Morison (2015, 136–137) deutlich, wie das *dictum de omni et nullo* verstanden werden sollte (*An. Pr.* 1. 9, 30^a 21-23):

„Denn da jedem *B* das *A* mit Notwendigkeit zukommt oder nicht zukommt, **das *C* aber einen Teil der *B*-Dinge** darstellt, so ist klar, daß auch für das *C* das eine oder das andere davon mit Notwendigkeit gelten wird.“ (Nortmann und Ebert 2007, 31, Hervorhebung M.R.S.)

Die Neuformulierung der unteren Prämisse als „das *C* aber einen Teil der *B*-Dinge darstellt“ (τὸ δὲ Γ τι τῶν Β ἐστὶ) erinnert laut Morison an die von ihm und Malink (2013, 34, Fussnote 2) bevorzugte Version des *dictum de omni et nullo* (Morison 2015, 137). Nortmann und Ebert übersetzen das *dictum* folgendermassen (2007, 16): „[W]enn man keines der unter den Subjektterminus fallenden Dinge herausgreifen kann“ (ὅταν μηδὲν ἧ λαβεῖν [τῶν] [τοῦ ὑποκειμένου], *An. Pr.* 1. 1, 24^b 28-30). Beide Textstellen verwenden „etwas + Genitiv“, um über das zu sprechen, was unter *B* (den Subjektterminus der oberen Prämisse *AaB*) fällt. Die Formulierung der Prämisse soll demnach eine Erinnerung an das *dictum* hervorrufen und somit auf die Anwendung der *Schlussregel* verweisen. Anders gesagt: Aristoteles *wendet* die *Schlussregel* an dieser Stelle *an*.

Die Übersetzung von Nortmann und Ebert (2007, 16) muss entsprechend angepasst werden, um dem Verständnis von Morison gerecht zu werden:

Wir reden von ‚Von-jedem-Ausgesagtwerden‘, immer wenn (ὅταν) es möglich ist (ἧ) keines⁶ (μηδὲν) [der unter den Subjektterminus fallenden Dinge] ([τῶν] [τοῦ ὑποκειμένου]) herauszugreifen (λαβεῖν), von dem das andere nicht ausgesagt werden wird (οὐ λεχθήσεται). Und bei ‚Von-keinem-(Ausgesagtwerden)‘ ebenso.

Mit dieser Übersetzung wird sowohl der Modalität von ἧ als auch dem Futur λεχθήσεται Rechnung getragen. Auch Barnes bemerkt diese Modalitäten, gibt jedoch vier Gründe an, sie in der Übersetzung nicht miteinander zu müssen (Barnes 2007, 389):⁷ Erstens könne

6 Diese Formulierung orientiert sich an der englischen Übersetzung von Morison (2015, 137).

7 Auf die Argumentation von Barnes bezieht sich auch Malink,

„wird nicht (aus)gesagt werden“ verstanden werden als „wird nicht prädiert werden“.⁸ Zweitens sei klar, dass es ein ‚schlussfolgerndes‘ Futur ist. In diesem Punkt sind sich Morison und Barnes einig (Morison 2015, 133). Die dritte Pedanterie betrifft nun die Modalität von ὅταν μηδὲν ἧ λαβεῖν in „...keines ...herausgreifen kann“.⁹ Damit Aristoteles die Unterscheidung zwischen ‚Von-jedem(keinem)-Ausgesagtwerden‘ (Aa_xB) und ‚Notwendig-von-jedem(keinem)-Ausgesagtwerden‘ (Aa_NB) aufrechterhalten könne, dürfe man die Modalität von ὅταν μηδὲν ἧ λαβεῖν nicht ernst nehmen (Barnes 2007, 389). Denn wenn man die Modalität ernst nehmen würde, wäre auch schon das assertorische Verhältnis zwischen zwei Begriffen ein notwendiges, was ‚Von-jedem(keinem)-Ausgesagtwerden‘ (Aa_xB) und ‚Notwendig-von-jedem(keinem)-Ausgesagtwerden‘ (Aa_NB) zu ein und demselben Verhältnis zwischen zwei Begriffen werden lasse (Barnes 2007, 389). Dies trifft jedoch nur zu, wenn man das *dictum de omni et nullo* als Definition oder Explikation von ‚Von-jedem(keinem)-Ausgesagtwerden‘ versteht. Wenn man das *dictum* hingegen als Schlussregel versteht, betrifft die Modalität hingegen das Schlussverfahren und nicht die Definition von ‚Von-jedem(keinem)-Ausgesagtwerden‘. Es ist nicht mehr das Verhältnis zwischen zwei Begriffen als ein notwendiges definiert, sondern das Verhältnis zwischen zwei Prämissen und der Konklusion. Dies passt auch zu der berühmten Definition von Syllogismen, die Aristoteles kurz vor dem *dictum* formuliert (*An. Pr.* 1. 1, 24^b 18-20):

„Ein Syllogismus ist eine Rede, in der, wenn bestimmte (Sachverhalte) gesetzt sind, ein von den gesetzten (Sachverhalten) verschiedener (Sachverhalt) **sich mit Notwendigkeit** dadurch **ergibt**, daß | die gesetzten (Sachverhalte) vorliegen.“ (Nortmann und Ebert 2007, 16, Hervorhebung M.R.S.)

Hier ist deutlich die Rede davon, dass das Verhältnis zwischen den Prämissen und der Konklusion ein *notwendiges Verhältnis* ist. Der letzte Punkt betrifft im spezifischen das Wort λαβεῖν. Barnes schreibt, dass dies nicht wörtlich genommen werden sollte, da es sonst möglich wäre, zu sagen, dass *A* von allen *B* ausgesagt werde,

wenn er schreibt, dass „...keines ...herausgreifen kann“ könne man mit „da ist kein ...“ ersetzen (2013, 35), weswegen hier die Begründung von Barnes besprochen wird.

8 Im Englischen steht: „will not be said‘ may be replaced by ‚will not be predicated“ (Barnes 2007, 389).

9 Barnes übersetzt dies ins Englische mit „you can’t take“ und setzt es mit „you won’t take“ gleich (vgl. Barnes 2007, 389).

wenn wir aus irgendwelchen Gründen nicht in der Lage wären, diejenigen *B* herauszugreifen, von denen nicht *A* ausgesagt wird (2007, 389). Dies würde die Syllogistik kontingent machen, da sie abhängig von der Fähigkeit zu erkennen wäre, welche *Bs* unter die *As* fallen. Versteht man das *dictum* als Definition von *a*- und *e*-Aussagen, so kann dieses Problem auftauchen. Es ergibt sich jedoch kein Problem für die Interpretation von Morison, da bei den vollständigen Syllogismen durch die zweite Prämisse genau *festgesetzt* wird, welche *Bs* unter die *As* fallen. Das Herausgreifen der *Bs* muss dementsprechend wörtlich genommen werden. Es handelt sich dabei nämlich um das Angeben der zweiten Prämisse. Wenn sich herausstellt, dass es einige von den *Bs* gibt, von denen *A* nicht ausgesagt wird, hat man schlicht den falschen Syllogismus verwendet und muss die zweite Prämisse entsprechend anpassen. Die Probleme, die Barnes und Malink mit der modalen Formulierung im *dictum* haben, stellen sich dementsprechend der Interpretation von Morison nicht.

Wie steht es um die anderen Probleme, die bereits angesprochen wurden? Muss ebenfalls ein formales, modernes Logiksystem verwendet werden und müssen zwei zusätzliche *dicta* postuliert werden, um einen Beweis für die vollständigen Syllogismen zu erhalten? Um diese Fragen zu beantworten, muss noch erläutert werden, wie Morison sich vorstellt, wie das *dictum de omni et nullo* genau zu verstehen ist.

Er formuliert das *dictum* zweimal, um sowohl die Schlussregel des *dictum de omni* als auch die des *dictum de nullo* zu erhalten (vgl. Morison 2015, 138 und 139):

dictum de omni

Gegeben ist die Prämisse Aa_xB . Immer wenn man eine andere Prämisse hat, die eine Subklasse von *B* postuliert, folgt notwendigerweise, dass *A* auch von der Subklasse ausgesagt wird.

dictum de nullo

Gegeben ist die Prämisse Ae_xB . Immer wenn man eine andere Prämisse hat, die eine Subklasse von *B* postuliert, folgt notwendigerweise, dass *A* nicht von der Subklasse ausgesagt wird.

Mit diesen beiden Regeln kann man sehr schnell einsehen, weshalb die vollständigen assertorischen Syllogismen *Barbara* und *Celarent* gültig sind. Sie verlangen nur die Anwendung einer Regel auf die vorgebrachten Prämissen:

Barbara

1. Aa_xB (obere Prämisse)
2. Ba_xC (untere Prämisse)
3. $\therefore Aa_xC$ (von 1 und 2; durch das *dictum de omni*)

Celarent

1. Ae_xB (obere Prämisse)
2. Ba_xC (untere Prämisse)
3. $\therefore Ae_xC$ (von 1 und 2; durch das *dictum de nullo*)

Die Prämissen werden somit nicht zuerst in einer modernen formalen Sprache rekonstruiert und es wird auch nicht anhand der Regeln der klassischen Prädikatenlogik ein Schluss abgeleitet, der wiederum in der Sprache der Syllogistik formuliert werden muss. Es ergeben sich laut Morison jedoch zwei Probleme: Erstens erkläre Aristoteles nicht, was ‚Von-jedem(keinem)-Ausgesagtwerden‘ bedeute und zweitens seien Aa_xB und Ae_xB keine kontradiktorischen Aussagen, was man aufgrund der Formulierung der *dicta* jedoch denken könne (Morison 2015, 139). Diese Probleme löst er, indem er ‚A wird von jedem B ausgesagt‘ als ‚B ist A‘ und ‚A wird von keinem B ausgesagt‘ als ‚B ist nicht A‘ versteht. Dies fällt auch mit dem üblichen Verständnis von Aa_xB und Ae_xB zusammen. Diese Aussagen stehen sich ebenso als *konträre* Gegensätze gegenüber, wie dies auch von Aristoteles selbst gesagt wird (Morison 2015, 140). Dasselbe Ergebnis erhält man jedoch, wenn man den Satz direkt vor dem *dictum de omni et nullo* als *Erläuterung* von ‚Von-jedem-Ausgesagtwerden‘ (Aa_xB) versteht (An. Pr. 1. 1, 24^b 26-28):

„Daß etwas in etwas als einem Ganzen ist und daß letzteres von jedem ersteren ausgesagt wird, ist dasselbe.“ (Nortmann und Ebert 2007, 16)

In dieser Art und Weise kann man nun auch die anderen möglichen Quantifizierungen und Qualifizierungen der Prämissen formulieren:

Ae_xB

Dass etwas in etwas als einem Ganzen *nicht* ist und dass Letzteres von *keinem* Ersteren ausgesagt wird, ist dasselbe.

Ai_xB (und Ao_xB)

Dass etwas in etwas *nicht als einem Ganzen* ist und dass Letzteres von *einigen* Ersteren (nicht) ausgesagt wird, ist dasselbe.

Auch hier erhalten wir einen *konträren* Gegensatz für Aa_xB und Ae_xB , da beide Aussagen falsch sind, wenn der dritte Satz wahr ist. Dies würde dazu auch noch das Problem der ‚fehlenden‘ *Explikation* für das ‚Ausgesagtwerden‘ lösen. Was es jedoch bedeutet, „[d]ass etwas in etwas als einem Ganzen ist“, ist hiermit noch nicht gelöst. Die Semantik kann nicht mehr *orthodox* verstanden werden, da durch das *dictum de omni et nullo* Subklassen von B beschrieben werden, was eine *heterodoxe* Lesart nahelegt (Morison 2015, 139). Durch die *Erläuterung* von a -, e -, i - und o -Aussagen als Verhältnisse zwischen dem Ganzen und seinen Teilen kommt jedoch auch ein *mereologisches* Verständnis der Semantik in Frage. Auf eine Antwort darauf kann im Rahmen dieses Beitrags leider nicht weiter eingegangen werden.¹⁰

Um einen Beweis für *Darii XXX* und *Ferio XXX* zu erhalten, macht sich Morison (2015, 149) zunutze, dass Aristoteles im *dictum de omni et nullo* das Verhältnis zwischen der Subklasse und B wie eine *unbestimmte* Aussage formuliert (An. Pr. 1. 1, 24^a 18-22). Anders gesagt, enthält sich Aristoteles der Bestimmung des quantitativen Verhältnisses zwischen den Termini der zweiten Prämisse. Dementsprechend kann die zweite Prämisse entweder eine a - oder eine i -Aussage sein. Die Beweise für *Darii XXX* und *Ferio XXX* könnten somit analog zu den Beweisen für *Barbara XXX* und *Celarent XXX* verlaufen:

Darii

1. Aa_xB (obere Prämisse)
2. Bi_xC (untere Prämisse)
3. $\therefore Ai_xC$ (von 1 und 2; durch das *dictum de omni*)

Ferio

1. Ae_xB (obere Prämisse)
2. Bi_xC (untere Prämisse)
3. $\therefore Ao_xC$ (von 1 und 2; durch das *dictum de nullo*)

Nun bleibt noch zu sehen, weshalb eine direkte Anwendung der *dicta* nicht funktioniert, wenn man einen *unvollständigen* Syllogismus beweisen möchte. Wie bereits angesprochen, liegt der Unterschied darin, dass ausser dem Angenommenen auch noch anderes hinzugezogen werden muss. Dies kann entweder mit Hilfe der drei Konversionen (e -, i - oder a -Konversion) oder anhand der *reductio ad impossibile* gemacht werden, welche Aristoteles in An. Pr. 1.2 einführt. Dies soll noch am Beweis für *Camestres XXX* illustriert werden (An. Pr. 1. 5, 27^a 9-14):

¹⁰ Für eine eingehende Besprechung einer *mereologischen* Alternative siehe Vlasits (2019).

Camestres

1. Ma_xN (obere Prämisse)
2. Me_xX (untere Prämisse)
3. Xe_xM (von 2; durch e -Konversion)
4. Xe_xN (von 3 und 1; durch das *dictum de nullo*; Celarent)
5. $\therefore Ne_xX$ (von 4; durch e -Konversion)

Da die untere Prämisse (Me_xX) nichts über eine Subklasse des N der oberen Prämisse aussagt, kann keines der *dicta* angewendet werden. Wenn jedoch durch die e -Konversion der unteren Prämisse Xe_xM abgeleitet wird, kann analog zu *Celarent* Xe_xN geschlossen werden, da die obere Prämisse eine Subklasse von M postuliert. Mit einer erneuten Anwendung der e -Konversion erhält man dann die gewünschte Konklusion Ne_xX . Dies zeigt auf, dass dieser Syllogismus *nicht vollständig* ist, weil er noch um weitere Annahmen *ergänzt* werden muss.

6. Fazit

Im vorliegenden Beitrag wurde untersucht, wie verschiedene Autoren für deren Verständnis von *vollständigen* respektive *vollkommenen* Syllogismen argumentiert haben, inwiefern sich Morisons Vorschlag von denen der anderen Autoren unterscheidet und welche Vorzüge sein Vorschlag bietet. Hierzu wurde in einem ersten Schritt die Textgrundlage vorgestellt (Abschnitt 2.1). Es wurden dazu die relevanten Merkmale der Definition von *vollständigen* respektive *unvollständigen* Syllogismen hervorgehoben, die für die weitere Diskussion von Bedeutung waren. Daraufhin wurden noch einige Bemerkungen zur Verwendung von Begriffen und der Notation von Syllogismen gemacht (Abschnitt 2.2). Die Diskussion wurde mit der Besprechung der Positionen begonnen, die *vollkommene* Syllogismen als diejenigen verstehen, die *evident* oder *selbsterklärend* sind (Abschnitt 3). Diese Positionen konnten jedoch abgelehnt werden, da Aristoteles selbst *Erklärungen* für die *Gültigkeit* dieser Syllogismen unabhängig von ihrer genauen Formulierung vorbrachte. Dies führte zur Diskussion von Interpretationen, welche die *Vollständigkeit* von Syllogismen als eine *logische* Eigenschaft betrachten (Abschnitt 4). Doch durch die Anwendung moderner Semantiken werden diese Interpretationen der Definition von *vollständigen* Syllogismen nicht mehr gerecht. Dem gegenüber konnte aufgezeigt werden (Abschnitt 5), dass Morisons Verständnis des *dictum de omni et nullo* eine Alternative bietet, welche die Probleme der vorangehenden Positionen vermeiden kann. Demnach sollte das *dictum de omni et nullo* als eine Schlussregel verstanden werden.

Eine Schwierigkeit, welche sich für Morisons Ansatz ergeben hat, war, dass das *dictum de omni et nullo* nicht mehr als *Erläuterung* oder *Definition* von ‚Von-jedem(keinem)-Ausgesagtwerden‘ dienen konnte. Es wurde jedoch vorgeschlagen, die vage Formulierung im dem *dictum* vorangehenden Satz als eine solche *Erläuterung* zu verstehen. Dies eröffnete aber die Möglichkeit, dass das Verhältnis zwischen den Termini *mereologisch* verstanden werden kann. Ein weiteres Problem bleibt, nämlich wie die modale respektive apodiktische Syllogistik zu verstehen ist. Das *dictum de omni et nullo* kann zwar modal interpretiert werden, ohne dass gleich die assertorische und modale Syllogistik zusammenfallen, ob jedoch die Probleme der apodiktischen Syllogistik damit gelöst werden können (beispielsweise weswegen Barbara NXN gültig ist), muss noch weiter untersucht werden.

Literatur

- Barnes, Jonathan. 2007. *Truth, etc: six lectures on ancient logic*. PhD thesis. Oxford: Clarendon Press.
- Łukasiewicz, Jan. 1957. *Aristotle's syllogistic: from the standpoint of modern formal logic*. 2nd ed. Oxford: Clarendon press.
- Malink, Marko. 2013. *Aristotle's modal syllogistic*. Cambridge, MA and London: Harvard University Press.
- Morison, Benjamin. 2015. "What is a Perfect Syllogism?" In *Oxford Studies in Ancient Philosophy. Volume XLVIII*, edited by Brad Inwood. 107–66. Oxford: Clarendon Press.
- Nortmann, Ulrich and Theodor Ebert. 2007. *Aristoteles: Analytica Priora. Buch I*. Lizenzausgabe. Werke in deutscher Übersetzung / Aristoteles. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Patzig, Günther. 1963. *Die aristotelische Syllogistik: logisch-philologische Untersuchungen über das Buch A der Ersten Analytiken. 2., verbesserte Auflage*. Abhandlungen der Akademie der Wissenschaften in Göttingen. Philologisch-Historische Klasse. Folge 3. Göttingen: Vandenhoeck und Ruprecht.
- Ross, William David. 1949. *Aristotle's prior and posterior analytics*. Oxford: Clarendon Press.
- Vlasits, Justin. 2019. "Mereology in Aristotle's Assertoric Syllogistic." *History and Philosophy of Logic* 40 (1): 1–11.

Marco R. Schori studiert Wissenschaftsphilosophie und Klassische Philologie mit Schwerpunkt Griechisch im Master an der Universität Bern. Er interessiert sich besonders für antike Philosophie, Geschichte und Philosophie der Wissenschaften und der Logik.

Anhang – Griechischer Text

Vorgehen im 1. Kapitel der *Analytica Priora* (*An. Pr.* 1. 1, 24^a 11-15):

εἴτα διορίσαι τί |
ἔστι πρότασις καὶ τί ὅρος καὶ τί συλλογισμός, καὶ ποῖος
| τέλειος καὶ ποῖος ἀτελής, μετὰ δὲ ταῦτα τί τὸ ἐν ὅλῳ
εἶ- | ναὶ ἢ μὴ εἶναι τότε τῷδε, καὶ τί λέγομεν τὸ κατὰ
παντὸς | ἢ μηδενὸς κατηγορεῖσθαι.

„Als nächstes müssen wir uns darüber verständigen, was eine (sylogistische) Aussage, was ein Terminus und was ein Syllogismus ist sowie welcher Syllogismus vollkommen und welcher unvollkommen ist; danach darüber, was es heisst, dass etwas in (etwas als) einem Ganzen ist oder nicht ist, und was wir mit dem Ausgesagtwerden-von-jedem | oder -von-keinem meinen.“ (Nortmann und Ebert 2007, 15)

Quantität von Aussagen (*An. Pr.* 1. 1, 24^a 18-22):

λέγω δὲ καθόλου μὲν τὸ παντὶ ἢ μηδενὶ ὑπάρχειν, ἐν μέρει
| δὲ τὸ τινὶ ἢ μὴ τινὶ ἢ μὴ παντὶ ὑπάρχειν, ἀδιόριστον δὲ
τὸ | ὑπάρχειν ἢ μὴ ὑπάρχειν ἄνευ τοῦ καθόλου ἢ κατὰ
μέρος, οἷον | τὸ τῶν ἐναντίων εἶναι τὴν αὐτὴν ἐπιστήμην
ἢ τὸ τὴν ἡδονὴν μὴ εἶ- | ναὶ ἀγαθόν.

„Allgemein‘ nenne ich (eine Aussage), die ein Jedem-Zukommen oder Keinem Zukommen (aussagt), ‚partikulär‘ (eine Aussage), die ein Irgendeinem-Zukommen, Irgendeinem-nicht-Zukommen oder Nicht-jedem-Zu-kommen aussagt, ‚unbestimmt‘ (eine Aussage), die | ein Zukommen oder Nicht-Zukommen aussagt ohne Angabe über den allgemeinen oder partikulären Charakter der Aussage, z. B. ‚Gegensätze fallen in dieselbe Wissenschaft‘, ‚Vergnügen ist kein Gut‘.“ (Nortmann und Ebert 2007, 15)

Definition von Syllogismus (*An. Pr.* 1. 1, 24^b 18-20):

συλλογισμός δὲ ἐστὶ λόγος ἐν |
ᾧ τεθέντων τινῶν ἕτερον τι τῶν κειμένων ἐξ ἀνάγκης
συμβαί- | νει τῷ ταῦτα εἶναι.

Definition von *vollständigen* Syllogismen (*An. Pr.* 1. 1, 24^b 22-26):

τέλειον μὲν οὖν
καλῶ συλλογισμὸν τὸν μηδενὸς ἄλλου προσδεόμενον
παρὰ τὰ | εἰλημμένα πρὸς τὸ φανῆναι τὸ ἀναγκαῖον,
ἀτελῆ δὲ τὸν προσ- | δεόμενον ἢ ἐνὸς ἢ πλειόνων, ἃ
ἔστι μὲν ἀναγκαῖα διὰ τῶν | ὑποκειμένων ὄρων, οὐ μὴν
εἰληπται διὰ προτάσεων.

Erläuterung von ‚Von-jedem-Ausgesagtwerden‘ (*An. Pr.* 1. 1, 24^b 26-28):

τὸ δὲ ἐν
ὅλῳ εἶναι ἕτερον ἐτέρῳ καὶ τὸ κατὰ παντὸς κατηγορεῖσθαι
| θατέρου θάτερον ταῦτόν ἐστιν.

dictum de omni et nullo (*An. Pr.* 1. 1, 24^b 28-30):

λέγομεν δὲ τὸ κατὰ παντὸς
κατηγορεῖσθαι ὅταν μηδὲν ἢ λαβεῖν [τῶν] [τοῦ
ὑποκειμένου] | καθ’ οὗ θάτερον οὐ λεχθήσεται· καὶ τὸ
κατὰ μηδενὸς ὡσαύτως.

Erste Figur, *Barbara XXX* und *Celarent XXX* (*An. Pr.* 1. 4, 25^b 32-35):

Ἵταν οὖν ὅροι τρεῖς οὕτως ἔχωσι πρὸς ἀλλήλους ὥστε
τὸν | ἔσχατον ἐν ὅλῳ εἶναι τῷ μέσῳ καὶ τὸν μέσον ἐν
ὅλῳ τῷ πρώτῳ | ἢ εἶναι ἢ μὴ | εἶναι, ἀνάγκη τῶν ἄκρων
εἶναι συλλογισμὸν | τέλειον.

Barbara XXX (*An. Pr.* 1. 4, 25^b 37-40):

ὡς γὰρ τὸ A κατὰ παν-
τὸς τοῦ B καὶ τὸ B κατὰ παντὸς τοῦ Γ, ἀνάγκη τὸ A κατὰ
| παντὸς τοῦ Γ κατηγορεῖσθαι· πρότερον γὰρ τὸ εἰρηται
πῶς τὸ | κατὰ παντὸς λέγομεν.

Darii XXX (*An. Pr.* 1. 4, 26^a 23-25):

ὑπαρχέτω γὰρ τὸ μὲν A παντὶ τῷ B, τὸ δὲ B τινὶ τῷ Γ.
οὐκοῦν εἰ ἔστι παντὸς κατηγορεῖσθαι τὸ ἐν ἀρχῇ
λεχθέν, ἀνάγκη | τὸ A τινὶ τῷ Γ ὑπάρχειν.

Beweis für *Camestres XXX* (*An. Pr.* 1. 5, 27^a 9-14):

πάλιν εἰ τὸ M τῷ μὲν N παντὶ τῷ Δ ἢ μηδενὶ,
οὐδὲ τὸ Ξ τῷ N οὐδενὶ ὑπάρξει (εἰ γὰρ τὸ M οὐδενὶ τῷ Ξ,
οὐδὲ | τὸ Ξ οὐδενὶ τῷ M· τὸ δὲ γε M παντὶ τῷ N ὑπῆρχεν·
τὸ ἄρα | Ξ οὐδενὶ τῷ N ὑπάρξει· γεγένηται γὰρ πάλιν
τὸ πρῶτον | σχῆμα)· ἐπεὶ δὲ ἀντιστρέφει τὸ στερητικόν,
οὐδὲ τὸ N οὐδενὶ τῷ | Ξ ὑπάρξει, ὥστ’ ἔσται ὁ αὐτὸς
συλλογισμός.

„Umgekehrt: Wenn das M jedem N, aber keinem X (zukommt), | wird auch das X keinem N zukommen. – Denn wenn das M keinem X, so auch das X keinem M; das M aber kam jedem N zu; also wird das X keinem N zukommen; es ist nämlich wieder die erste Figur zustande gekommen. – Da aber die verneinende (Aussage) konvertierbar ist, wird auch das N keinem X zukommen, so daß sich derselbe Syllogismus ergibt.“ (Nortmann und Ebert 2007, 22)

Felapton XXX (*An. Pr.* 1. 6, 28^a 26-28):

καὶ ἂν τὸ μὲν P παντὶ τῷ Σ, τὸ δὲ Π μηδενὶ
ὑπάρχη, ἔσται συλλογισμός ὅτι τὸ Π τινὶ τῷ Ρ οὐχ ὑπάρ-
| ξει ἐξ ἀνάγκης·

„Auch wenn das R jedem S, das P aber keinem zukommt, ergibt sich ein Syllogismus, da notwendigerweise das P irgendeinem R nicht zukommen wird.“ (Nortmann und Ebert 2007, 25–26)

Reduktionsverhältnisse der assertorischen Modi

(*An. Pr.* 1, 7, 29^b 15-19):

ὥστ' ἐπεὶ οἱ μὲν ἐν τῷ μέσῳ σχήματι συλλογισμοὶ πάντες ἀνάγονται εἰς τοὺς ἐν τῷ πρώτῳ καθόλου | συλλογισμοὺς, οἱ δὲ κατὰ μέρος ἐν τῷ πρώτῳ εἰς τοὺς ἐν | τῷ μέσῳ, φανερόν ὅτι καὶ οἱ κατὰ μέρος ἀναχθήσονται εἰς | τοὺς ἐν τῷ πρώτῳ σχήματι καθόλου συλλογισμούς.

„Da also die Syllogismen in der mittleren Figur sich alle auf die allgemeinen Syllogismen in der ersten reduzieren lassen, die partikulären (Syllogismen) in der ersten aber auf die in der mittleren, so ist klar, daß auch die partikulären auf die allgemeinen Syllogismen in der ersten Figur reduziert werden.“ (Nortmann und Ebert 2007, 29)

e- und *a*-Konversion von *N*-Prämissen (*An. Pr.* 1, 8, 30^a 2-3):

τό τε γὰρ στερητικὸν ὡσαύτως ἀντιστρέφει, καὶ τὸ ἐν | ὅλῳ εἶναι καὶ τὸ κατὰ παντὸς ὁμοίως ἀποδῶσομεν.

„Die verneinende (Aussage) ist nämlich ebenso konvertierbar (wie die assertorische), und das ‚in (etwas als) einem Ganzen sein‘ wie auch das ‚von jedem (ausgesagt werden)‘ haben wir analog zu verstehen.“ (Nortmann und Ebert 2007, 30)

Barbara NXN (*An. Pr.* 1, 9, 30^a 15-23):

Συμβαίνει δὲ ποτε καὶ τῆς ἐτέρας προτάσεως ἀναγκαίας οὐσης ἀναγκαῖον γίνεσθαι τὸν συλλογισμόν, πλὴν οὐχ | ὀποτέρας ἔτυχεν, ἀλλὰ τῆς πρὸς τὸ μείζον ἄκρον, οἷον εἰ τὸ | μὲν *A* τῷ *B* ἐξ ἀνάγκης εἴληπται ὑπάρχον ἢ μὴ ὑπάρχον, | τὸ δὲ *B* τῷ *Γ* ὑπάρχον μόνον· οὕτως γὰρ εἰλημμένων τῶν | προτάσεων ἐξ ἀνάγκης τὸ *A* τῷ *Γ* ὑπάρξει ἢ οὐχ ὑπάρξει. | ἐπεὶ γὰρ παντὶ τῷ *B* ἐξ ἀνάγκης ὑπάρχει ἢ οὐχ ὑπάρχει | τὸ *A*, τὸ δὲ *Γ* τῶν *B* ἐστὶ, φανερόν ὅτι καὶ τῷ *Γ* ἐξ ἀνάγκης ἔσται θάτερον τούτων.

„Es kommt aber unter bestimmten Bedingungen vor, daß der Syllogismus auch dann notwendig ausfällt, wenn (nur) eine Prämisse von beiden notwendig ist – allerdings ist es nicht gleichgültig, welche das ist; sondern (es muss) die mit dem grösseren Aussenterminus (sein), wie wenn etwa vorausgesetzt ist, das *A* komme dem *B* mit Notwendigkeit zu – oder nicht zu –, das *B* aber komme dem *C* bloß zu. Sind nämlich die | Prämissen auf diese Weise angenommen, so wird mit Notwendigkeit das *A* dem *C* zukommen oder nicht zukommen. Denn da jedem *B* das *A* mit Notwendigkeit zukommt oder nicht zukommt, das *C* aber einen Teil der *B*-(Dinge) darstellt, **so ist klar**, dass auch für das *C* das eine oder das andere davon mit Notwendigkeit gelten wird.“ (Nortmann und Ebert 2007, 41, Hervorhebung M.R.S.)

aoi-QQQ (*An. Pr.* 1, 14, 33^a 27-34):

ἐὰν δὲ στε-

ρητικὴ ληφθῆ ἢ ἐν μέρει πρότασις, ἢ δὲ καθόλου καταφατικὴ, | τῇ δὲ θέσει ὁμοίως ἔχουσιν (οἷον τὸ μὲν *A* παντὶ τῷ *B* ἐνδέ- | χεται, τὸ δὲ *B* τινὶ τῷ *Γ* ἐνδέχεται μὴ ὑπάρχειν), διὰ μὲν | τῶν εἰλημμένων προτάσεων οὐ γίνεται φανερός συλλογισμός, | ἀντιστραφείσης δὲ τῆς ἐν μέρει καὶ τεθέντος τοῦ *B* τινὶ τῷ *Γ* | ἐνδέχεσθαι ὑπάρχειν τὸ αὐτὸ ἔσται συμπέρασμα ὃ καὶ πρό- | τερον, καθάπερ ἐν τοῖς ἐξ ἄρχης.

„Falls dagegen die partikuläre Prämisse als verneinend angenommen wird und die allgemeine als bejahend und ihre Stellung analog ist – (was soviel heißt) wie, das *A* (kommt) jedem *B* möglicherweise (zu), | das *B* kommt irgendeinem *C* möglicherweise nicht zu –, so kommt aufgrund der angenommenen Prämissen kein **evidenter Syllogismus** zustande; wird aber die partikuläre (Prämisse bezüglich des Möglichen) umgeformt und vorausgesetzt, daß das *B* irgendeinem *C* möglicherweise zukommt, so ergibt sich dieselbe Konklusion, die (sich) auch vorhin (ergab) – ebenso wie bei den zu Beginn (behandelten Syllogismen).“ (Nortmann und Ebert 2007, 39–40, Hervorhebung M.R.S.)

Hinweis auf modales (mögliches) ‚Von-jedem-Ausgesagtwerden‘ (*An. Pr.* 1, 14, 32^b 40-33^a 1):

τοῦτο δὲ φανερόν ἐκ τοῦ ὀρισμοῦ· τὸ γὰρ | ἐνδέχεσθαι παντὶ ὑπάρχειν οὕτως ἐλέγομεν.

„Das ist klar aufgrund der Definition. Das | Möglicherweise-jedem-Zukommen haben wir nämlich so erklärt.“ (Nortmann und Ebert 2007, 38)

Hinweis auf modales (mögliches) ‚Von-jedem-Ausgesagtwerden‘ (*An. Pr.* 1, 14, 33^b 24-25):

τοῦτο δὲ φανερόν ἐκ τοῦ | ὀρισμοῦ τοῦ ἐνδέχεσθαι.

„Das ist klar aufgrund der | Definition des ‚möglicherweise‘.“ (Nortmann und Ebert 2007, 39)

Barbara QXQ (*An. Pr.* 1, 15, 33^b 33-36):

ἐνδεχέσθω γὰρ τὸ *A* παντὶ τῷ *B*, τὸ δὲ *B* παντὶ τῷ *Γ* κείσθω ὑπάρχειν. ἐπεὶ οὖν ὑπὸ τὸ *B* ἐστὶ τὸ | *Γ*, τῷ δὲ *B* παντὶ ἐνδέχεται τὸ *A*, φανερόν ὅτι καὶ τῷ *Γ* | παντὶ ἐνδέχεται. γίνεται δὴ τέλειος συλλογισμός.

„Es soll nämlich das *A* jedem *B* möglicherweise (zukommen) und von *B* festgesetzt sein, daß es jedem *C* zukomme. Weil nun unter dem *B* das | *C* steht, jedem *B* aber das *A* möglicherweise (zukommt), **ist offensichtlich**, daß (es) auch jedem *C* möglicherweise (zukommt). Es kommt also ein vollkommener Syllogismus zustande.“ (Nortmann und Ebert 2007, 41, Hervorhebung M.R.S.)

Apriorität und Notwendigkeit

Das kontingente Apriori gemäss Saul Kripke

1. Einleitung

Die Philosophie ist keine empirische Wissenschaft. Viele der Überlegungen, die sie anstellt, basieren nicht direkt auf Erfahrung und weisen dementsprechend apriorische Elemente auf. Weiter besteht Interesse an einer gewissen Allgemeingültigkeit philosophischer Aussagen. Infolge dieser Betrachtungen wird der metaphysische Status von apriorischen Propositionen relevant. Wegweisend ist hier die Frage, ob diese Propositionen kontingent oder notwendig sind. Insbesondere für die Erkenntnistheorie ist es zentral, auf solche grundlegenden Problematiken Antworten zu finden und sich zu überlegen, inwiefern wir ohne Erfahrung überhaupt etwas wissen können und ob es sich dabei um Sachverhalte handelt, die auch anders sein könnten. Historisch gesehen ist die Verbindung der beiden Begriffe „Apriorität“ und „Notwendigkeit“ sehr eng gezogen. Zum Beispiel sind gemäss einigen Forschenden¹ die beiden Begriffe bei Kant äquivalent (oder zumindest koextensiv).

Die grobe Idee hinter dieser scheinbaren Übereinstimmung ist folgende: Falls eine Proposition unabhängig von der Erfahrung gewusst werden kann, muss sie für alle verschiedenen möglichen Erfahrungen gültig sein. In diesem Zusammenhang auftauchende Unterscheidungen wie „analytisch“/„synthetisch“, „a priori“/„a posteriori“ oder „notwendig“/„kontingent“ sowie die Frage, ob diese Begriffe überhaupt klar und in verständlicher Weise voneinander abzugrenzen sind, wurden in der Vergangenheit kontrovers diskutiert.²

¹ So schreibt Jonathan Bennett bezüglich Kants Begrifflichkeit: „necessity and universality are entailed by apriority as well as entailing it“ (1966, 9). Anthony Quinton vertritt eine ähnliche Position: „‘A priori’ means [...], following Kant, ‘necessary’“ (1964, 32). Ausführungen dazu finden sich bei Kant in verschiedenen Werken, insbesondere in der *Grundlegung zur Metaphysik der Sitten* (Kant [1785] 2016, 86), oder der *Kritik der praktischen Vernunft* (Kant [1788] 2003, 39).

² Beispielsweise bedeutet „Notwendigkeit“ gemäss Rudolf Carnap

Saul Kripke greift in *Naming and Necessity* die Frage nach dem Status von a priori Propositionen auf und zeigt, entgegen der oben erwähnten Idee, dass es mindestens eine Kategorie von a priori Wahrheiten gibt, die nicht notwendig sind (Kripke 1980). Das Ziel des vorliegenden Beitrags besteht in einer kritischen Auseinandersetzung mit Kripkes Überlegung. Dazu gilt es zunächst, auf die Konzepte der Notwendigkeit und der Apriorität etwas genauer einzugehen. Weiter folgt eine Einführung in Kripkes Theorie der Eigennamen. Darauf aufbauend werde ich sein bekanntes Beispiel vom Urmeter in Paris rekonstruieren sowie einige Prämissen seines Beispiels diskutieren und diese auf ihre Plausibilität prüfen. Abschliessend wird Kripkes Ansatz als Ganzes betrachtet, um dessen Auswirkungen auf den metaphysischen Status von a priori Propositionen zu untersuchen.

Vorab gilt es somit, sich mit den zentralen Begrifflichkeiten zu beschäftigen.

2. Notwendigkeit und Apriorität

Zunächst sollen die zwei zentralen Konzepte des vorliegenden Beitrags eingehender betrachtet und insbesondere Kripkes Darstellung davon kurz erläutert werden. Es handelt sich um Notwendigkeit und Apriorität.

2.1 Notwendigkeit

Dem philosophischen Begriff der Notwendigkeit kann mittels modaler Fragestellungen angenähert werden. Eine Proposition ist nicht lediglich falsch oder wahr, sie ist dies auch auf eine bestimmte Art und Weise. Mit

(1947) so viel wie „logische Wahrheit“ oder Analytizität“. Richard Swingburne (1975) argumentiert dafür, dass alle analytischen Propositionen a priori sind. Dazu ist auch die Debatte bei Quine (1951) interessant. Er vertritt die Position, wonach der Begriff der Analytizität unbegründet ist. Gemäss Quine folgt daraus ebenfalls, dass modale Begriffe wie „notwendigerweise“ nicht intelligibel sind.

anderen Worten: Eine Proposition kann beispielsweise notwendigerweise wahr, möglicherweise wahr oder kontingenterweise wahr sein. Es handelt sich also nicht nur um eine Auseinandersetzung damit, wie die Welt ist, sondern auch damit, wie sie hätte sein können. Kripke umschreibt diesen Aspekt folgendermassen: „We ask whether something might have been true or might have been false“ (1980, 36). Beispielsweise könnte gefragt werden: „Hätte Kripke auch Geschichte statt Mathematik studieren können?“ Weiter könnte man fragen: „Wäre es möglich gewesen, dass Kripke Präsident von Amerika geworden wäre?“

Für die semantische Interpretation modaler Sätze ist insbesondere die Theorie der sogenannten möglichen Welten zentral.³ Dabei handelt es sich gewissermassen um grösstmöglich spezifizierte Weisen, wie die Welt hätte sein können.⁴ Innerhalb dieser Praxis ist eine Proposition genau dann notwendig, wenn sie in allen möglichen Welten wahr ist. Analog dazu ist sie genau dann möglich, wenn sie in mindestens einer Welt wahr ist und kontingent, wenn sie möglich, aber nicht notwendig ist. Es handelt sich bei diesen modalen Überlegungen also um metaphysische Fragen. Ein wesentlicher Punkt ist hierbei folgender: „This [die Frage, ob etwas möglich ist] in and of itself has nothing to do with anyone’s knowledge of anything“ (Kripke 1980, 36).

Kripke hebt hervor, dass der metaphysische Status einer Proposition nicht davon abhängt, ob sie von jemandem gewusst wird oder überhaupt gewusst werden kann. Er illustriert dies anhand des Beispiels der Vermutung von Goldbach. Sie besagt, dass jede gerade Zahl, grösser als zwei, die Summe aus zwei Primzahlen ist. Die Vermutung ist bis heute unbewiesen. Falls sie sich als falsch herausstellt, wird sie dies – Kripke geht von einer „classical view of mathematics“ (1980, 36) aus – notwendigerweise sein, falls sie sich als wahr herausstellt, wird sie dies ebenfalls notwendigerweise sein. Aus einer epistemischen Perspektive ist zurzeit nicht klar, welche der beiden Möglichkeiten der Fall ist. Unabhängig davon gehört der Wahrheitswert, welcher der Vermutung zukommt, aber notwendigerweise zu ihr (Kripke 1980, 36).

2.2 Apriorität

Apriorität ist ein Konzept aus der Epistemologie. Es handelt davon, wie wir etwas wissen können. Was unter „a priori Wissen (oder Begründung)“ genau zu verstehen ist, ist umstritten.⁵ Kripke weist auf die Vagheit des Begriffs hin (1980, 36). Eine ausführliche Diskussion hierüber würde den Rahmen dieses Beitrags indes sprengen. Weil er dessen ungeachtet zentral für diese Untersuchung ist, werde ich einige Überlegungen dazu anführen. Bruce Russell schlägt für a priori Rechtfertigungen folgende Definition vor:

“Roughly speaking, a priori justification provides reasons for thinking a proposition is true that comes from merely understanding, or thinking about, that proposition.” (Russell 2020, Ziff. 1.)

Der wesentliche Punkt von Apriorität ist, wie etwas gewusst oder begründet werden kann. Ich lege Russells Definition folgendermassen aus: Die Frage nach der Apriorität einer Proposition wird darauf zurückgeführt, ob die Gründe, die für ihre Wahrheit sprechen, a priori sind. Es gibt bestimmte Rechtfertigungsgründe, die als a priori gelten. Diese können sich einerseits aus dem Verständnis der Proposition ergeben. Das bedeutet nach meinem Dafürhalten, dass allein aufgrund der Bedeutung des Satzes ersichtlich ist, dass seine Wahrheitsbedingungen erfüllt sind. Wer den Satz versteht, weiss, dass er wahr ist. Andererseits kann sich die Einsicht in die Wahrheit der Proposition auch aus dem Nachdenken über die verstandene Proposition ergeben. Diese beiden Arten, eine Proposition berechtigterweise für wahr zu halten, sind a priori. Beide sind unabhängig von jeglicher empirischen Evidenz.

Kripke weist darauf hin, dass die traditionelle Charakterisierung von Kants Verständnis des Begriffs folgende war: A priori Wahrheiten sind jene, die unabhängig von jeglicher Erfahrung gewusst werden können. Hierbei meint Kripke, dass das Wissen darum in einem gewissen Sinn möglich ist, jedoch nicht, dass der entsprechende Inhalt tatsächlich jemand weiss (Kripke 1980, 34). Weiter ist nach Kripke nicht klar, für wen das Wissen möglich ist. In Anbetracht dieser und anderer Schwie-

3 Die Semantik der möglichen Welten geht zurück bis auf Carnap (1947). Eine grundlegende Darstellung findet sich beispielsweise bei Jaakko Hintikka (1961). Weiter sind auch die Ausführungen von David Lewis (1986) zentral.

4 Dabei geht es nicht nur um Möglichkeit innerhalb unserer (i.e. der aktuellen) Welt, sondern um (logische, metaphysische, nomologische und andere Arten von) Möglichkeit im Allgemeinen.

5 Der Standardansatz besteht darin, eine negative Definition anzugeben: Eine Proposition ist a priori wahr, wenn sie unabhängig von der Erfahrung gewusst werden kann (dazu siehe Bruce Russell 2020). Hier stellen sich die Fragen, in welchem Sinn „unabhängig“ verstanden wird und was genau unter den Begriff der Erfahrung fällt. Es gibt aber auch Forschende, wie z.B. Laurence Bonjour (1998), die für eine positive Definition argumentieren: A priori Wissen bezieht sich auf eine Art rationaler Einsicht, mittels der die Wahrheit einer Proposition erfasst wird.

rigkeiten schlägt er vor, sich stattdessen auf die Frage zu konzentrieren, ob eine bestimmte Person etwas a priori weiss oder aufgrund von a priori Evidenz davon überzeugt ist (Kripke 1980, 34–35). Kripke spezifiziert jedoch nicht weiter, was alles als a priori Evidenz gilt.

2.3 Verknüpfung von Apriorität und Notwendigkeit

Kripke zeigt auf, dass eine grundlegende kategoriale Differenz zwischen den beiden Begriffen besteht. Während es sich bei der Notwendigkeit um ein metaphysisches Konzept handelt, ist „a priori“ ein epistemischer Begriff.

Eine enge Verbindung der beiden Begriffe ist, wie in der Einleitung bereits erwähnt, indes historisch stark verwurzelt. Kripke weist darauf hin, dass „notwendig“ und „a priori“ trotz der kategorialen Differenz oftmals so gebraucht werden, als würden sie dasselbe bedeuten (1980, 34). Eine Äquivalenz würde einerseits implizieren, dass jede notwendige Proposition auch a priori gewusst werden kann. Andererseits müsste aus der Apriorität einer Aussage ihre Notwendigkeit folgen. Der zweite Fall ist für vorliegenden Beitrag interessant. Dies würde bedeuten, dass aufgrund dessen, dass eine Proposition a priori eingesehen werden kann, auf ihre Notwendigkeit geschlossen werden darf. Wie Kripke zeigt, muss diese Beziehung indes begründet werden: „It’s certainly a philosophical thesis, and not a matter of obvious definitional equivalence [...] that everything a priori is necessary“ (1980, 36).

Eine mögliche, intuitive Rechtfertigung für diese These lautet folgendermassen: Wenn etwas a priori gewusst werden kann, bedeutet das, dass die dafürsprechende Evidenz in einem gewissen Sinn unabhängig von der Erfahrung sein muss. Dem liegt nahe, dass die wissende Person nicht in die Welt zu schauen braucht, um zu diesem Wissen zu gelangen. Somit muss die a priori eingesehene Proposition in allen Welten der Fall sein. Denn wie ist es möglich, etwas über eine ganz bestimmte Welt zu wissen, ohne diese Welt zu betrachten?

Kripke ist sich diesem Gedankengang bewusst, schreibt jedoch:

“This depends on the thesis that there can’t be a way of knowing about the actual world without looking that wouldn’t be a way of knowing the same thing about every possible world.” (1980, 36)

Die angesprochene These weist er zurück, indem er aufzeigt, wie es möglich ist, ohne jegliche Erfahrung etwas über eine spezifische Welt zu wissen, was nicht für alle anderen Welten gilt.

3. Theoretische Grundlagen bei Kripke

Wie angedeutet, führt Kripke ein Argument in Form eines Gegenbeispiels an, das zeigen soll, wie a priori Wissen über eine bestimmte Welt möglich ist, ohne dass die dahinterliegende Proposition notwendig ist. Um sein Argument zu untersuchen, ist es zunächst von zentraler Bedeutung, einige seiner Grundkonzepte zu erläutern. Dieses Kapitel beschäftigt sich einleitend mit dem Konzept des starren Designators und erläutert Kripkes These, nach der dieses Konzept auf Eigennamen zutrifft. Zudem gilt es, den Unterschied zwischen einer Definition, die den Bezug eines Eigennamens festlegt, sowie einer, welche seine Bedeutung gibt, zu erläutern. Weiter wird aufgezeigt, wie Kripke in diesem Zusammenhang über mögliche Welten nachdenkt.

3.1 Eigennamen als starre Designatoren

Kripkes Ansatz baut auf seinem Verständnis von Sprache, insbesondere von Eigennamen auf. Beim sogenannten „rigid designator“ (starrer Designator) handelt es sich um eine Eigenschaft, die referierenden Ausdrücken zukommen kann (Kripke 1980, 48). Hierbei ist es zentral, auf welche Art und Weise der Ausdruck ein Objekt bezeichnet. Kripke definiert sein Konzept des starren Designators folgendermassen: „Let’s call something a rigid designator if in every possible world it designates the same object“ (1980, 48).

Es handelt sich um einen Ausdruck, der in jeder möglichen Welt dasselbe Objekt bezeichnet. Das bedeutet gemäss Kripke keinesfalls, dass dieses in jeder möglichen Welt existieren muss, also notwendigerweise existiert. Seine Idee ist vielmehr, dass der starre Designator das Objekt in jeder möglichen Welt bezeichnet, in der es existiert (Kripke 1980, 48). Die Erklärung des Konzepts setzt noch nicht voraus, dass tatsächlich sprachliche Ausdrücke existieren, die diesen Bedingungen entsprechen. Zunächst ist lediglich relevant, dass klar wird, was es für einen solchen bedeuten würde, etwas starr zu designieren.

Weiter stellt Kripke die These auf, dass Eigennamen starre Designatoren sind (1980, 49). Er wendet sich damit explizit gegen die zur Zeit seines Schaffens konventionelle These der „description theory of proper names“ (Kripke 1980, 5), verstanden als Theorie der Bedeutung von Eigennamen. Später im Text referiert er darauf

als „the Frege-Russell view“⁶ (1980, 27). Ich habe ihn so verstanden, dass er dieser Sicht zuschreibt, sie besage, dass es für einen Eigennamen eine Beschreibung gibt, die seinen kognitiven Gehalt konstituiert und seinen Bezug festlegt. Kripke sagt dazu: “[A] proper name, properly used, simply was a definite description abbreviated or disguised” (1980, 27).

Somit liefert die Beschreibung eines Eigennamens dessen Bedeutung. „Aristoteles“ könnte zum Beispiel „der letzte grosse Philosoph der Antike“ bedeuten. An dieser Stelle soll nicht näher auf die genauere Darstellung, Probleme oder Weiterentwicklungen dieser Theorie eingegangen werden. Kripke kontrastiert diese Vorstellung der Bedeutung von Eigennamen mit seiner eigenen Theorie dazu. Er unterscheidet vorab zwischen zwei unterschiedlichen Arten, eine Beschreibung im Zusammenhang mit einem Eigennamen aufzufassen. Sie kann die Bedeutung des Namens geben oder lediglich seinen Bezug festlegen. Kripke weist Ersteres zurück und geht davon aus, dass nur Letzteres der Fall ist (beziehungsweise der Fall sein kann!), womit er der „Frege-Russell view“ widerspricht (Kripke 1980, 57). Wenn „Der letzte grosse Philosoph der Antike“ in der aktuellen Welt genau auf Aristoteles zutrifft und eine grosse Anzahl Sprechende diese Beschreibung Aristoteles zuschreibt, folgt daraus nicht, dass diese Beschreibung äquivalent mit „Aristoteles“ ist, was auch nicht für andere korrekte, individualisierende Beschreibungen von Aristoteles der Fall ist. Eine solche Beschreibung kann lediglich dafür benutzt werden, in unserer Welt festzulegen, welches Objekt mit „Aristoteles“ gemeint ist. Gewissermassen pickt sie das richtige Individuum heraus. Sobald sein Bezug festgelegt ist, bezieht sich „Aristoteles“ immer auf Aristoteles, gleichgültig in welcher möglichen Welt (Kripke 1980, 57).

Kripke zeichnet diesen Unterschied explizit in seiner Einleitung zu *Naming and Necessity*. Die Idee ist folgende: Gegeben sei der Satz: „Aristoteles mochte Katzen“. Was sind nun dessen Wahrheitsbedingungen in der aktuellen Welt? Nach Kripke ist die Proposition genau dann wahr, wenn Aristoteles Katzen mochte (1980, 6). Nach Bertrand Russells Ansatz müsste die Aussage laut Kripke so analysiert werden, dass sie genau dann wahr ist, wenn der letzte grosse Philosoph der Antike Katzen mochte. Das ist, wenn das Objekt, das durch die Beschreibung „der letzte grosse Philosoph der Antike“ herausgepickt wird, tatsächlich Katzen mochte. Kripke hält fest, dass die Wahrheitsbedingungen beider Interpretationsansätze in der aktuellen Welt extensional überein-

stimmen. Er hebt jedoch hervor, dass das angemessene Verständnis einer Aussage auch voraussetzt, dass deren Wahrheitsbedingungen für kontrafaktische Situationen angegeben werden können. Diesbezüglich differiert Kripkes Ansatz von der Bertrand Russell zugeschriebenen Theorie. Während sich „Aristoteles“, verstanden als starrer Designator, in jeder kontrafaktischen Welt per Definition auf Aristoteles bezieht, ist dies für „Der letzte grosse Philosoph der Antike“ keineswegs der Fall (Kripke 1980, 6–7). Stellen wir uns eine Welt vor, in der Aristoteles Katzen mochte, aber nicht der letzte grosse Philosoph der Antike war. Nehmen wir behelfsmässig weiter an, dass der letzte grosse Philosoph der Antike ein gewisser *X* war und Katzen überhaupt nicht mochte. Nach Kripkes Verständnis wäre die untersuchte Proposition also wahr. Nach Bertrand Russell würde sich „Aristoteles“ in dieser Welt aber auf *X* beziehen und nicht auf Aristoteles. Somit würde die Proposition eine Aussage über *X* machen und in unserem Beispiel falsch sein. Das Russellsche Verständnis ist gemäss Kripke, der dafür plädiert, dass wir eine direkte Intuition der Rigidität von Eigennamen haben, kontraintuitiv (1980, 14).

3.2 Taufe und mögliche Welten

Wie kommt „Aristoteles“ jedoch zu seinem Bezug? In seiner zweiten Lesung in *Naming and Necessity* schreibt Kripke von einem „initial baptism“ (1980, 97). Es handelt sich dabei um eine Art Taufe, die abläuft, wenn der Bezug des Ausdrucks zu Beginn festgelegt wird. Das kann zum Beispiel über eine Beschreibung geschehen, die ein einzelnes Individuum herauspickt. Eine andere Möglichkeit sieht Kripke in der „ostension“ (1980, 97). Dies ergibt sich dadurch, dass der Festleger des Bezugs gleichsam auf etwas zeigt und es damit tauft. Kripke weist aber darauf hin, dass er mit dieser Darstellung nicht intendiert, notwendige und hinreichende Bedingungen für die Referenz von Ausdrücken aufzustellen (1980, 94–96). Weiter führt er aus, wie diese Bezugnahme von Person zu Person in der Sprache weitergegeben wird. Ich werde darauf aber nicht näher eingehen, weil es für die Fragestellung des vorliegenden Beitrags nicht relevant ist.

Ein weiterer wesentlicher und stark mit den Erläuterungen über Eigennamen verknüpfter Aspekt betrifft die Art und Weise, wie Kripke mögliche Welten versteht. Er vertritt die Position, dass mögliche Welten gedanklich sehr oft falsch benutzt werden. Um dies zu illustrieren, bedient er sich des klassischen Problems der Identität über mögliche Welten. Hier geht es darum, bei modalen Aussagen, die Eigennamen enthalten, zu bestimmen, welches Objekt in einer möglichen Welt durch den je-

⁶ Kripke bezieht sich dabei auf Bertrand Russell (1905) und Gottlob Frege (1892).

weiligen Eigennamen bezeichnet wird. Nehmen wir, um dieser Bestimmung nachzugehen, analog zu Kripke die Frage, ob Richard Nixon die Wahl im Jahr 1968 auch hätte verlieren können. (Kripke 1980, 40–47)

Zunächst gilt es, aufzuzeigen, wie der nach Kripkes Dafürhalten verkehrte Ansatz ungefähr funktioniert. Eine mögliche Welt ist durch die deskriptiven Bedingungen gegeben, die wir damit verbinden. Die Schwierigkeit dabei ist, dass vielleicht niemand in dieser Welt qualitativ identisch mit Nixon ist. Um die Frage für diese Welt zu prüfen, müssen wir aber wissen, wer darin Nixon ist. Es können lediglich Eigenschaften beobachtet werden, aber nicht, dass jemand Nixon ist. Irgendwie müssen im Vorherein qualitative Kriterien festgelegt werden, die es erlauben, für jede Welt zu sagen, wer Nixon ist. Einen Ansatz, dies zu lösen, den Kripke auch erwähnt, bietet beispielsweise die „counterpart theory“ von David Lewis (Kripke 1980, 45). Lewis argumentiert dafür, dass Ähnlichkeiten zwischen verschiedenen möglichen Welten ein Gegenstück des vom Eigennamen bezeichneten Objekts determinieren, welches dann für die Evaluation der Aussage relevant ist (vgl. Lewis 1986).

Gemäss Kripke ist jedoch dieser gesamte Ansatz, über mögliche Welten nachzudenken, verkehrt: “Possible worlds’ are stipulated, not discovered by powerful telescopes” (1980, 45). Mögliche Welten sind also nicht so aufzufassen, als wären sie weit entfernte Entitäten, die irgendwie zu entdecken wären. Wir legen fest, dass es sich um Nixon handelt und fragen, was in einer kontrafaktischen Situation – oder in einer anderen möglichen Welt – mit ihm passiert wäre. Somit ist es Teil der Beschreibung der stipulierten kontrafaktischen Welt, dass sie Nixon enthält. Wir können auf ihn zeigen und fragen, was mit ihm geschehen wäre, wenn bestimmte Ereignisse anders gewesen wären (Kripke 1980, 42–45). In diesem Fall ist gemäss Kripke die Frage nach der Identität hinsichtlich möglicher Welten kein Problem, da die Identifizierung von Dingen in der Welt, die mit Eigennamen bezeichnet werden, überhaupt nicht erforderlich ist. Zusammenfassend lässt sich sagen, dass Kripke die Reihenfolge umdreht:

“[W]e begin with the objects, which we have, and identify, in the actual world. We can then ask whether certain things might have been true of the objects” (1980, 53).

Gleichsam picken wir Nixon in der aktualen Welt heraus und setzen ihn in eine stipulierte mögliche Welt. Nun gilt es, sich Kripkes bereits mehrmals erwähntem Beispiel zu widmen.

4. Kripkes Argument

Anhand eines Beispiels zeigt Kripke auf, inwiefern es möglich ist, dass eine Proposition a priori wahr ist, zugleich aber nicht notwendig (vgl. 1980, 54–56). In vorliegendem Kapitel werde ich dieses Beispiel mittels teilweiser Formalisierung rekonstruieren. Anschliessend sollen einigen kritische Prämissen diskutiert werden, bevor die aus dem Beispiel gewonnene Erkenntnis verallgemeinert werden kann.

4.1 Rekonstruktion

Zunächst gilt es, die von Kripke zurückgewiesene These klar zu formulieren:

T: Wenn eine Proposition a priori wahr ist, dann ist sie notwendigerweise wahr.

Kripke bedient sich des Beispiels vom Urmeter in Paris. Es handelt sich dabei um einen spezifischen Stock, der benutzt wurde, um das metrische System zu standardisieren. Die in diesem Kontext relevante Proposition, welche laut Kripke kontingent sowie a priori ist, lautet folgendermassen:

P: Die Länge des Stocks S zum Zeitpunkt t_0 ist ein Meter.

Diese Identitätsaussage ist nach Kripke nicht so zu verstehen, dass „die Länge des Stocks S zum Zeitpunkt t_0 “ und „ein Meter“ synonym sind. Die Proposition bezieht sich vielmehr darauf, dass das, worauf die beiden Ausdrücke referieren, identisch ist (Kripke 1980, 56). Beide beziehen sich auf dieselbe Länge.

P1: „Die Länge des Stocks S zum Zeitpunkt t_0 “ legt den Bezug von „ein Meter“ fest.

Diese Prämisse ergibt sich aus dem Rahmen des Beispiels. Zudem ist relevant, dass Kripke dieses Festlegen des metrischen Systems nicht als definitionsgebend versteht. Es handelt sich lediglich darum, den Ausdruck „ein Meter“ zu initialisieren – zu taufen – und dadurch seinen Bezug, also eine bestimmte Länge, festzulegen. Wie oben erläutert wurde, besteht ein gewichtiger Unterschied zwischen diesen beiden Arten: Das Determinieren des Bezugs eines Ausdrucks mittels einer Beschreibung hat nämlich nicht zur Folge, dass die Beschreibung des Ausdrucks auch dessen Bedeutung ist.

P2: P1 \rightarrow P

Wenn es der Fall ist, das P1 wahr ist, ist auch P wahr.

K1: P ist wahr. (aus P1 und P2)

Da Kripke im Rahmen des Beispiels von P1 ausgeht, ist P ebenfalls wahr. Dies erlaubt ihm jedoch noch nicht, eine Aussage über den epistemischen Status von P zu machen:

P3: P1 ist a priori wahr.

P3 bedeutet, dass a priori gewusst werden kann, wie der Bezug eines Begriffs festgelegt ist. Nach meinem Dafürhalten muss Kripke diese These aufstellen, um sinnvollerweise sagen zu können, dass P a priori ist. Diese Prämisse erachte ich jedoch als unklar und werde dies später diskutieren.

P4: Wenn B aus A folgt und A a priori wahr ist, ist auch B a priori wahr.

P4 ist nötig, um sicherzustellen, dass der epistemische Status von P1 bei der Implikation von P2 gewissermaßen ebenfalls mitvererbt werden kann. Dies scheint meines Erachtens vorauszusetzen, dass die Implikationsbeziehung zwischen P1 und P ebenfalls a priori eingesehen werden können muss.

K2: P ist a priori wahr. (aus P1, P2, P3 und P4)

Da P1 a priori wahr ist und P daraus folgt, kann dank P4 darauf geschlossen werden, dass die Proposition P a priori wahr ist. Wer also weiss, wie der Bezug des Begriffs „ein Meter“ festgelegt wurde, braucht nicht in die Welt zu schauen und den Stock S zum Zeitpunkt t_0 zu messen, sondern kann seine Länge einfach folgern. Somit steht zunächst der epistemische Status der zu untersuchenden Proposition fest. Weiter gilt es, deren metaphysischen Status zu prüfen.

P5: „Ein Meter“ ist ein starrer Designator.

Kripke ist überzeugt, dass „ein Meter“ als starrer Designator verstanden werden sollte. Das bedeutet in diesem Fall, dass er eine bestimmte Länge bezeichnet, die bei seiner Taufe mittels Beschreibung festgelegt wurde. Dieser Vorgang betrifft unsere Welt, also die Welt, welche die Sprache bestimmt, in der „ein Meter“ ein Ausdruck ist. Ungeachtet der Welt, auf die „ein Meter“ nun angewendet wird, wird damit immer diese bestimmte Länge bezeichnet.

P6: „Die Länge des Stocks S zum Zeitpunkt t_0 “ ist nicht ein starrer Designator.

Dieser Ausdruck ist eine Beschreibung einer Eigenschaft eines bestimmten Objekts zu einem spezifischen Zeitpunkt. Er fixiert somit nicht eine bestimmte Länge in jeder möglichen Welt, sondern immer jene, die in der betrachteten Welt die Eigenschaft aufweist, die Länge des Stocks S zum Zeitpunkt t_0 zu sein. Dies schliesst nicht aus, dass dies in jeder Welt dieselbe Länge ist. Es zeigt lediglich, dass der Ausdruck „die Länge des Stocks S zum Zeitpunkt t_0 “ nicht zwingend in jeder Welt dieselbe Länge bezeichnet.

P7: „Die Länge des Stocks S zum Zeitpunkt t_0 “ ist keine essentielle Eigenschaft des Stocks S.

Diese Prämisse dient in meiner Rekonstruktion dazu, auszuschliessen, dass „die Länge des Stocks S zum Zeitpunkt t_0 “ in jeder Welt dieselbe Länge bezeichnet. Das ist wesentlich, um sicherzustellen, dass die Länge des Stocks S eine kontingente Eigenschaft ist und somit nicht notwendigerweise einem Meter entspricht. Dies scheint einleuchtend: Abzustreiten, dass Stöcke eine unveränderliche Länge haben, wäre absurd. Zudem könnte dieser Stock aus verschiedensten möglichen Gründen etwas grösser oder kleiner sein, ohne plötzlich ein anderer Stock zu sein. Kripke spricht davon, dass beispielsweise mechanische Spannung auf den Stock hätte einwirken können; oder eine andersartige Belastung, die seine Länge verändert hätte (1980, 55).

K3: Es gibt eine mögliche Welt w und eine Länge m , für die gilt: „Die Länge des Stocks S zum Zeitpunkt t_0 ist m “ und „Ein Meter ist nicht gleich m “. (aus P5, P6 und P7)

Diese Konklusion ist das zweite Kernstück in Kripkes Argument. „Die Länge des Stocks S zum Zeitpunkt t_0 “ ist kein starrer Designator und es gibt möglicherweise Längen des Stocks S, die von einem Meter abweichen.

K4: Es gibt eine mögliche Welt w , für die gilt: P ist falsch (aus K3; Reformulierung)

Diese Reformulierung zeigt klarer, dass P möglicherweise falsch sein könnte.

P8: Eine Proposition ist notwendigerweise wahr, genau dann, wenn sie in allen möglichen Welten wahr ist.

K5: P ist nicht notwendigerweise wahr. (aus K4 und P8)

K7: Es gibt eine Proposition P, die a priori wahr ist, aber nicht notwendigerweise wahr ist. (aus K2 und K5)

Diese Konklusion fasst die beiden Kernaussagen von Kripkes Argument zusammen. Ihr zufolge ist die zu untersuchende Proposition a priori einzusehen, wahr und kontingent. K7 ist die Negation von T und steht damit im Widerspruch mit T. Somit kann T zurückgewiesen werden.

Das Argument ist in dieser Reformulierung meines Erachtens gültig. Ich werde nun einige Prämissen, die nach meinem Dafürhalten genauerer Betrachtung bedürfen, diskutieren.

4.2 Plausibilität der Prämissen

Vorweg einige Worte zur Prämisse P7. Hier ist anzumerken, dass deren Zurückweisung zwar dieses spezifische Beispiel entkräften würde, aber analoge Beispiele, die sich auf eine nicht essenzielle Eigenschaft beziehen, nicht. Höchstens könnte behauptet werden, dass das Festlegen des Bezugs eines Ausdrucks nur mittels Beschreibung, die eine essenzielle Eigenschaft eines Objekts bezeichnet, möglich ist. Dies ist aber eine sehr delikate These: So müsste nämlich begründet werden, warum beispielsweise eine Beschreibung, die exakt ein Individuum herauspicks, sich dafür jedoch nicht auf seine essenziellen Eigenschaften beruft, nicht referenzbestimmend sein kann. Kripke weist selbst darauf hin, dass die Eigenschaften, mittels derer der Bezug in der aktuellen Welt festgelegt wird, nicht essenzielle Eigenschaften sein müssen (1980, 53).

Weiter möchte ich einige Anmerkungen zur Konklusion K3 anbringen. Es könnte dagegen eingewendet werden, dass in einer möglichen Welt, in welcher der Stock S eine andere Länge hat, „ein Meter“ auch etwas anderes bedeuten würde. Die Länge, mittels welcher der Bezug des Ausdrucks festgelegt wurde, wäre eine andere. Dem gilt es, das Konzept vom starren Designator entgegenzuhalten. Es ist charakteristisch dafür, dass die Festlegung des Bezugs in der aktuellen Welt passiert. Um den Wahrheitswert von P in einer anderen Welt zu bestimmen, wird der Ausdruck „ein Meter“ aus unserer Sprache (und damit mit der Bedeutung und dem Bezug unserer Sprache) gewissermassen mitgenommen. Der Einwand kann aber wie folgt erweitert werden: Was, wenn die kontrafaktische Aussage sich nicht darauf bezieht, was in einer möglichen kontrafaktischen Welt passieren würde, gegeben die aktuelle Welt, sondern

darauf, dass unsere Welt möglicherweise anders sein könnte? Das ist, wenn sich die Möglichkeitsaussage auf die Welt bezieht, in welcher der Bezug der Ausdrücke festgelegt wird. Einen möglichen Interpretationsansatz, der dieser Unterscheidung Rechnung trägt, findet sich beispielsweise bei David Chalmers (2006). In der vorliegenden Arbeit werde ich auf diesen Aspekt jedoch nicht weiter eingehen. Nach meinem Dafürhalten hält es Kripke für adäquat, unsere Möglichkeiten so auszuliegen, dass wir uns mit unseren Ausdrücken auf unsere Sprache (und somit auf die Festlegungen der Bezüge in unserer Welt) beziehen, wenn wir modale Aussagen treffen und uns fragen, ob ein bestimmter Sachverhalt möglich ist oder möglich gewesen wäre.

Zur Prämisse P8 gilt es hervorzuheben, dass ich sie eingefügt habe, weil es nicht bereits im Begriff der Notwendigkeit liegt, diesen auf semantischer Ebene mittels möglicher Welten zu interpretieren. Ich habe das Konzept der Notwendigkeit hier auf die gegebene Weise ausgelegt, da Kripke selbst in *Naming and Necessity* spezifisch darauf hinweist, wie er mögliche Welten versteht. Zudem referiert der Begriff des starren Designators in seiner Erklärung auf die möglichen Welten. In der Apologie des starren Designators verweist er auf „transworld identity“ (Kripke 1980, 46).

Die Prämisse P5 besagt, dass „ein Meter“ ein starrer Designator ist. Diese Prämisse könnte nach meinem Dafürhalten dann zurückgewiesen werden, wenn Kripkes Namenstheorie ebenfalls für falsch gehalten wird. Letztere hat durchaus einige Nachteile, wie beispielsweise, dass sie keine klare Interpretation von singulären Existenzaussagen bietet (Kripke 1980, 59). Andererseits scheint P5 insbesondere der Intuition in Bezug auf Eigennamen adäquat Rechnung zu tragen. Ich werde an dieser Stelle aber nicht weiter auf Kripkes Namenstheorie eingehen.

Die am meisten Unklarheiten hervorrufende Prämisse ist meines Erachtens P3. Dabei geht es um das Verständnis von „a priori“. Wie eingangs in der Definition von Bruce Russell eingeführt, ist eine Proposition dann a priori wahr, wenn ihre Wahrheit aus ihrem Verständnis oder dem Nachdenken über sie generiert werden kann. Die Proposition P bezieht sich darauf, wie der Bezug von „ein Meter“ festgelegt wurde. Die Frage ist, ob das Wissen, wie der Bezug festgelegt wurde, unter die Kategorie „a priori“ fällt. Zunächst scheint mir die Wahrheit dieser Proposition nicht aus ihrer Bedeutung zu folgen (was auch immer unter der Bedeutung einer Proposition verstanden wird). Kripke insistiert, dass das Festlegen des Bezugs nicht die Bedeutung des Ausdrucks ist. Wenn die Beschrei-

bung, mittels derer „ein Meter“ getauft wurde, nicht Teil seiner Bedeutung ist, sehe ich nicht, wie es aus dem reinen Verständnis der Proposition „Die Länge des Stocks S zum Zeitpunkt t_0 legt den Bezug von ‘ein Meter’ fest“ folgen soll, dass sie wahr ist. Weiter scheint auch Nachdenken nicht weiterzuhelfen. Gegeben, dass jemand eine Proposition dann verstanden hat, wenn die Person ihre Wahrheitsbedingungen eingesehen hat, setzt zudem nach meinem Dafürhalten nicht notwendigerweise das Wissen um die Taufe voraus. Dies ist konform mit Kripkes Ansicht über seinen eigenen Vorschlag: “[I]t takes the notion of intending to use the same reference as a given” (1980, 79).

Ich möchte nicht genauer darauf eingehen, wie sich Kripke dies konkret vorstellt. Es ist jedoch vernünftigerweise anzunehmen, dass die zusätzliche Information darüber, wie „ein Meter“ getauft wurde, erforderlich ist, um die Wahrheit der Proposition „Die Länge des Stocks S zum Zeitpunkt t_0 legt den Bezug von ‘ein Meter’ fest“ einzusehen. Die wesentliche Frage ist, ob diese zusätzliche Information als a priori Evidenz aufgefasst werden sollte.

Folgende Überlegung ist möglich: Um den Bezug des Begriffs „ein Meter“ festzulegen, muss in die Welt geschaut werden und ein spezifischer, konkreter Stock in der Welt bestimmt werden. Je nach Auslegung des Begriffs „a priori“ wäre das eine Begründung dafür, dass hier ein empirisches Element mitspielt. Denn das Festlegen des Bezugs passiert nicht unabhängig von jeglicher bestimmten Welt. Was ist von diesem Einwand zu halten? Zunächst scheint es mir bezeichnend, dass Kripke in seinem eingangs erwähnten Statement über seine Auffassung von Apriorität nicht auf die Unabhängigkeit von möglichen Welten referiert. Er betont vielmehr, dass die Rechtfertigungsgründe, die für die Wahrheit einer Proposition angeführt werden, auf a priori Evidenz basieren.

Zwecks der Bestimmung dieser Schwierigkeit ist für Kripke, wie ich ihn verstanden habe, die Frage relevant, ob jemand durch das Festlegen des Bezugs eines Ausdrucks etwas über die Welt dazulernt.

“But merely by fixing a system of measurement, has he thereby learned some (contingent) information about the world, some new fact that he did not know before?” (Kripke 1980, 63)

Dieser Aspekt scheint mir relevant. Wenn es der Fall wäre, dass durch das Festlegen des Bezugs eine neue Information über die Welt erlernt würde, würde dies

dafür sprechen, dass die Festlegung des metrischen Systems in gewisser Weise empirischen Gehalt hat und das Wissen darum nicht rein a priori ist. Dies ist meines Erachtens eher unplausibel. In einem gewissen Sinn lernt eine Person, die erfährt, wie der Bezug eines Begriffs festgelegt wurde, etwas – sie erhält eine zusätzliche Information. Dabei scheint es sich aber eher um sprachliches Wissen zu handeln: Die Person lernt etwas über den Begriff und seinen Bezug. Wenn eine Person festlegt, dass der Bezug von „ein Meter“ die Länge des besagten Stocks ist, sieht sie aufgrund dessen gewissermaßen die Wahrheit der Proposition „die Länge des Stocks S zum Zeitpunkt t_0 ist ein Meter“ ein. Dieses neue Wissen darüber, dass ein bestimmter kontingenter Satz über die Welt wahr ist, wird jedoch nicht durch Erfahrung gewonnen, sondern durch die Festlegung. Die Stipulation schafft eine a priori Wahrheit. In diesem Sinne hat die Identitätsaussage P meines Erachtens keinen kognitiven Gehalt, der über die Festlegung der Begriffe hinaus geht.

Insofern ist es vernünftig, das Wissen darüber, wie ein Bezug eines Begriffs festgelegt wird, als a priori einsehbar zu qualifizieren. Trotzdem ist meines Erachtens die Grenze zwischen Apriorität und Aposteriorität in diesem Argument nicht sehr explizit gezogen, was wohl auch daran liegt, dass Kripke ein Beispiel verwendet, um seine Gedanken zu veranschaulichen. Seine Ideen sind zwar intuitiv einsichtig, es ist aber schwierig, zu sehen, woran das liegt, und sie konzeptuell festzumachen.

4.3 Verallgemeinerung des Beispiels

Nachdem nun einige Prämissen eingehender betrachtet wurden, gilt es, das beispielhafte Argument zu verallgemeinern. Welche Eigenschaften muss eine Aussage aufweisen, um analog zum vorliegenden Exempel als a priori und kontingent eingestuft zu werden?

Es ist für Kripkes Beispiel charakteristisch, dass es sich bei der Proposition P um eine Identitätsaussage handelt. P postuliert die Identität zwischen den Bezugsobjekten zweier Ausdrücke, in diesem Fall zwischen den Bezugsobjekten von „die Länge des Stocks S zum Zeitpunkt t_0 “ und „ein Meter“. Von diesen beiden Ausdrücken ist einer ein starrer Designator – nämlich „ein Meter“ – und der andere nicht. Zudem ist es für den nicht-starren Designator zentral, dass er nicht angesichts metaphysischer Gründe in jeder möglichen Welt dasselbe Objekt designiert. Dieser Umstand ist jedoch nicht hinreichend für Kripkes Beispiel. Wäre es hinreichend, müsste auch der Satz „die Länge des Stocks X zum Zeitpunkt t_0 ist ein Meter“ – gegeben, dass der Stock X vom Stock S verschieden ist – falls er wahr

ist, ebenfalls für das Argument funktionieren. Dies ist aber nicht der Fall. Um die Apriorität der in Frage kommenden Proposition zu garantieren, muss weiter der nicht-starre Designator so sein, dass er den Bezug des starren Designators festlegt. Somit entsteht die apriorische Wahrheit mittels dieser Festlegung.

Die hinreichenden Kriterien, die eine Proposition aufweisen muss, damit sie wie das behandelte Beispiel als a priori und kontingent eingestuft werden kann, sind damit folgende:

- Die Proposition muss die Identität der Bezugsubjekte zweier Ausdrücke aussagen.
- Einer der Ausdrücke muss ein starrer Designator sein.
- Einer der Ausdrücke darf kein starrer Designator sein.
- Der nicht-starre Designator darf nicht in jeder möglichen Welt dasselbe Objekt bezeichnen.
- Der nicht-starre Designator muss den Bezug des starren Designators festlegen.

Diese Kriterien schliessen nicht aus, dass nur Propositionen, die sie erfüllen, a priori und kontingent sind.⁷ Sie sind hinreichend dafür, aber nicht notwendig.

5. Diskussion

Nachdem die wichtigsten Punkte von Kripkes Argumentation besprochen wurden, werden nun in diesem Abschnitt – unter der Voraussetzung, dass die Argumentation schlüssig ist – deren Auswirkungen diskutiert. Dabei wird ferner darauf eingegangen, wie die These, wonach aus Apriorität Notwendigkeit geschlossen werden kann, gegen Einwände verteidigt werden könnte.⁸

Kripkes Beispiel zeigt durch seine Art und Weise, über Eigennamen nachzudenken, Folgendes auf: Es ist möglich, etwas über die aktuelle Welt zu wissen, ohne in diese Welt zu schauen und ohne, dass dies deshalb auch für andere mögliche Welten gelten muss. Zudem trifft seine Theorie eine zentrale Unterscheidung: Wissen über eine spezifische Welt ohne empirische

Erkenntnis dieser Welt ist nicht dasselbe wie Wissen über eine spezifische Welt, unabhängig davon, um welche Welt es sich handelt.

So stellt Kripke fest, dass es falsch ist, ohne zusätzliches philosophisches Argument davon auszugehen, dass die Apriorität einer Proposition deren Notwendigkeit impliziert. Meines Erachtens wäre es aber nicht korrekt, ihm zu unterstellen, dass er die Position vertritt, wonach diese beiden Konzepte in jedem Fall unabhängig sind. Er weist lediglich auf den Unterschied zwischen metaphysischer und epistemischer Ebene hin und zeigt auf, dass es mindestens eine Kategorie von Aussagen gibt, die a priori und kontingent sind.

In einem gewissen Sinn erscheint Kripkes Beispiel banal und nicht wirklich das zu treffen, was die Intuition, wonach a priori Propositionen notwendig sind, meint. Seine Beispiele von Propositionen, die zugleich a priori und kontingent sind, sind im Prinzip einfach per Stipulation geschaffen. Er äussert dies selbst an einer Stelle und schlägt vor:

“Maybe the thesis about a prioricity implying necessity can be modified. It does appear to state some insight which might be important, and true, about epistemology.” (1980, 63)

Kripkes Gegenbeispiel zu T betrifft eine mitunter ziemlich klar abgrenzbare Menge von Aussagen. Es sind dies solche, welche die Festlegung des Bezugs eines Ausdrucks in direkter Weise betreffen, genauer jene, welche die oben erwähnten Kriterien erfüllen. Wäre es möglich, T so zu modifizieren, dass sie der intuitiven Einsicht Rechnung trägt und gleichzeitig diese Konterbeispiele umgeht? Aufgrund Kripkes Überlegungen ist es schwierig, lediglich die Kategorie von Sätzen, welche die Kriterien erfüllen, auszuschliessen und sich ansonsten auf die Intuition zu verlassen, wonach die Implikation gilt. Es ist seiner Ansicht nach ein Argument erforderlich, dass die Brücke von der Apriorität zur Notwendigkeit schlägt. Kripke selbst betont in einer Fussnote, dass er sich an keiner Reformulierung versucht und nicht sieht, wie eine solche aussehen müsste, falls sie nicht zur Folge haben soll, dass T trivial wird (1980, 63).

Ich möchte im Folgenden kurz drei Ansätze andeuten, mittels derer die These T, wonach die Apriorität einer Aussage ihre Notwendigkeit impliziert, vielleicht erfolgreich angepasst werden könnte. Kripkes Gegenbeispiel widerlegt die Annahme, wonach es unmöglich sei, etwas a priori zu wissen, was für eine bestimmte Welt wahr ist, aber nicht für alle anderen möglichen Welten. T basiert jedoch, wie oben aufgezeigt, auf

⁷ Andere Vorschläge für kontingente a priori Propositionen stammen beispielsweise von John Turri (2011) oder Gareth Evans (1979).

⁸ Im Folgenden werde ich auf diese These mit dem Ausdruck „T“ referieren.

ebendieser Annahme. Zunächst könnte versucht werden, T ohne Rückgriff auf Letztere zu begründen, was sich als sehr schwieriges Unterfangen herausstellen könnte, da nach meinem Dafürhalten gerade darin deren intuitive Anziehungskraft liegt.

Eine andere Idee besteht darin, den Begriff der Apriorität für T so auszulegen, dass er diese stipulierten a priori Wahrheiten nicht miteinschließt. Gleichzeitig darf er aber nicht zu nahe an der Notwendigkeit sein, ansonsten ist T nicht sehr aussagekräftig. Beispielsweise zu sagen, dass a priori etwas ist, das unabhängig von und für alle Welten gewusst werden kann, würde in einem gewissen Sinn die Notwendigkeit bereits miteinschließen, sofern sie als „wahr in allen möglichen Welten“ definiert wird. Wenn der Begriff der Apriorität genug unterschiedlich zur Notwendigkeit ist, wird ein Argument, das die Lücke zwischen Apriorität und Notwendigkeit schließt, erforderlich. Diese Grenze so zu ziehen, dass sie nicht willkürlich ist und T gleichzeitig intelligibel und aussagekräftig bleibt, scheint mir sehr schwierig.

T könnte auch etwas abgeschwächt werden, indem ihr Anwendungsbereich eingeschränkt wird. Hier würde es darum gehen, zu zeigen, dass es zwar durchaus einen Weg gibt, etwas a priori über eine Welt zu wissen, was nicht auch für alle anderen gilt, aber, dass dies insbesondere (oder sogar nur) über Stipulation möglich ist. Dies würde vielleicht den Weg öffnen, T für bestimmte Arten von Propositionen und Bedingungen trotzdem zu halten. Das scheint mir aber ebenfalls ein beschwerliches Unterfangen zu sein. Zumal die Brücke zwischen epistemischem und metaphysischem Status trotzdem irgendwie gerechtfertigt werden muss.

Kripkes Argument hat zur Folge, dass der Schluss von der Apriorität einer Aussage auf ihre Notwendigkeit oder eine ähnlich strikte Verbindung schwierig haltbar ist. Ungeachtet dessen gesteht Kripke der These T eine gewisse epistemische Einsicht zu. Ich möchte im Folgenden einige Überlegungen anführen, die für deren Plausibilität sprechen.

Zunächst ist es vernünftig, davon auszugehen, dass a priori wahre Propositionen zumindest möglicherweise notwendig sind und dass deren Notwendigkeit ebenfalls a priori eingesehen werden kann. Der apriorische Zugang ist gewissermaßen für diese Art von Wissen prädestiniert. Unsere Erfahrung und damit unser Wahrnehmungswissen sind in vielerlei Hinsicht kontingent und auf Tatsachen in der aktuellen Welt gestützt. Sie hängen davon ab, was wir wahrnehmen, wie wir es wahrnehmen und wie die aktuelle Welt zufällig beschaffen ist. Apriorisches Wissen hängt gerade nicht von diesen Faktoren ab. Überlegungen, die auf welt-unabhängiger

Evidenz beruhen, sind im Vorteil, um allgemeine Propositionen über alle logisch möglichen Welten einzusehen. Insofern scheint es möglich – vorausgesetzt, dass es überhaupt Einsicht in Notwendigkeit gibt – a priori notwendig wahre Propositionen einzusehen.

Weiter kann für die Apologie der Plausibilität einer starken Verbindung von Apriorität hin zur Notwendigkeit auch die inverse Relation betrachtet werden. An die meisten unserer kontingenten Wahrheiten gelangen wir nur über empirische Kanäle. Wir sind in Kontakt mit der Welt, haben Sinneswahrnehmungen und betreiben empirische Wissenschaft. Ich werde an dieser Stelle nicht auf potenzielle Gegenbeispiele (wie jene, die Kripke selbst anführt (1980, 128–129)) und -argumente eingehen. Nichtsdestotrotz scheint mir die Evidenz auch hier dafür zu sprechen, dass kontingente Wahrheiten in den allermeisten Fällen a posteriori sind.

6. Fazit

Zusammenfassend gilt es festzuhalten, dass Kripke einen klaren Unterschied zwischen dem metaphysischen Begriff der Notwendigkeit und dem epistemischen Begriff der Apriorität aufzeigt. Die beiden Konzepte sind für Kripke somit keinesfalls äquivalent. Somit fordert er eine Begründung für die These, wonach von der Apriorität einer Proposition auf ihre Notwendigkeit geschlossen werden kann. Dementsprechend zeigt er anhand eines Beispiels auf, dass es mindestens eine Kategorie von Aussagen gibt, die zwar a priori eingesehen werden können, aber lediglich kontingent sind.

Es bleibt jedoch offen, was die Folgen von Kripkes Arbeit für den metaphysischen Status von a priori Propositionen im Allgemeinen sind. Es wird sich zeigen, ob die These, dass aus Apriorität Notwendigkeit folgt, trotz Kripkes Gegenbeispiels dahingehend modifiziert werden kann, dass die These haltbar ist und ihre epistemische Einsicht zum Tragen kommt.

Literatur

- Bennett, Jonathan. 1966. *Kant's Analytic*. London: Cambridge University Press.
- Bonjour, Laurence. 1998. *In Defense of Pure Reason: A Rationalist Account of A Priori Justification*. Cambridge: Cambridge University Press.
- Carnap, Rudolf. 1947. *Meaning and Necessity*. Chicago: The University of Chicago Press.
- Chalmers, David. 2006. "Two-dimensional semantics." In *The Oxford Handbook to the Philosophy of Language*, Edited by Ernest Lepore and Barry C. Smith, Oxford: Oxford University Press.
- Evans, Gareth. 1979. "Reference and Contingency." *Monist* 62 (2). 161–189.
- Frege, Gottlob. 1892. "Über Sinn und Bedeutung." In *Zeitschrift für Philosophie und philosophische Kritik*. Bd. 100. 25–50.
- Hintikka, Jaakko. 1961. "Modality and Quantification." *Theoria* 27. 119–28.
- Kant, Immanuel. (1785) 2016. *Grundlegung zur Metaphysik der Sitten*. Hamburg: Felix Meiner Verlag.
- ——— (1788) 2003. *Kritik der praktischen Vernunft*. Hamburg: Felix Meiner Verlag.
- Kripke, Saul. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Lewis, David. 1986. *On the plurality of worlds*. Oxford: Blackwell.
- Quine, Willard Van Orman. 1951. "Two Dogmas of Empiricism." *Philosophical Review* 60 (1). 20–43.
- Quinton, Anthony. 1964. "The 'A Priori' and the Analytic." *Proceedings of the Aristotelian Society* 64. 31–54.
- Russell, Bertrand. 1905. "On Denoting." *Mind* 14 (56). 479–493.
- Russell, Bruce. 2020. "A Priori Justification and Knowledge." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. <<https://plato.stanford.edu/archives/sum2020/entries/apriori/>>.
- Swingburne, Richard. 1975. "Analyticity, Necessity and Apriority." *Mind* 84 (334). 225–243.
- Turri, John. 2011. "Contingent A Priori Knowledge." *Philosophy and Phenomenological Research* 83 (2). 327–44.

Rafaëla Schinner studiert Philosophie im Bachelor an der Universität Fribourg. Sie interessiert sich im Besonderen für politische Philosophie, Philosophie der Ökonomie sowie Wahrheitstheorien. Weiter beschäftigen sie aktuell Fragen zu normativer Rationalität sowie zur Beziehung zwischen Rationalität und Freiheit.

Wie man spricht

Über das unterdrückende Potenzial legitimer Sprache

1. Einleitung

Für Pierre Bourdieu dient Sprache nicht nur der Kommunikation, sondern ist zugleich Schauplatz alltäglicher sozialer Machtkämpfe. In diesem Sinne kritisiert er im Text *Zur Ökonomie des sprachlichen Tauschs* (1982) das gemeinhin verbreitete Verständnis der Sprache als neutrales, universelles Gut, an dem alle Sprecher:innen einer Sprachgemeinschaft prinzipiell gleichermaßen teilhaben können. Er zeigt auf, wie *legitime Sprache* produziert und reproduziert wird, wie sie soziale Unterschiede widerspiegelt und dabei weitgehende Auswirkungen auf das gesellschaftliche sowie individuelle Leben hat. (Bourdieu 2017, 107–178)

Einige dieser Auswirkungen schränken die Entwicklungs- und Entfaltungsmöglichkeiten der Sprecher:innen ein. Damit scheinen sie, mit Iris Marion Young, auch *unterdrückend* zu wirken. Im Text *Five Faces of Oppression* (1990) argumentiert Young für ein breites, strukturelles Verständnis von *Unterdrückung*, die sich – oft unerkannt und im sozialen Alltag eingebettet – in fünf Facetten¹ zeigt: Als *Ausbeutung*, *Marginalisierung*, *Machtlosigkeit*, *Kulturimperialismus* und *Gewalt* (1990, 39–65).

Ein unterdrückendes Potenzial legitimer Sprache wird in Bourdieus Analyse kaum explizit genannt, schwingt aber mit. Benannt werden nicht-legitime Sprachpraxen etwa als „unterdrückte [...] Sprachkompetenzen“ (Bourdieu 2017, 124–125) oder „unterdrückte [...] Sprachformen“ (Bourdieu 2017, 139). Genauer beleuchtet werden diese Beobachtungen in Bourdieus Text aber nicht. Vorliegend wird deshalb versucht, legitime Sprache als Unterdrückung nach Young zu beschreiben. Verwendet werden dazu die zwei Facetten: *Kulturimperialismus* und *Gewalt*. Dies hat zwei Grün-

de: Erstens sind Ausbeutung, Marginalisierung und Machtlosigkeit gemäss Young stark mit Erwerbsarbeit verbunden (1990, 58). Eine Anwendung dieser Facetten jenseits der Arbeitswelt bedingt daher zunächst eine Erweiterung von Youngs Konzept auf andere Lebensbereiche, was in diesem Essay nicht geleistet werden kann. Zweitens ist Unterdrückung als Kulturimperialismus eine Erfahrung, die viele unterdrückten Gruppen machen (Young 1990, 64). Das macht dieses Gesicht – und das damit eng verschränkte Gesicht der Gewalt (Young 1990, 63) – zu sinnvollen Kategorien, um das unterdrückende Potenzial legitimer Sprache anzugehen.

Zwei Aspekte stehen im Zentrum der folgenden Analyse: Zunächst soll herausgearbeitet werden, inwiefern legitime Sprache kulturimperialistisch und gewaltvoll sein kann. Weiter wird aufzuzeigen versucht, inwiefern Youngs Gedanken einige Folgen legitimer Sprache genauer beleuchten und so für die Lektüre von Bourdieus Text gewinnbringend sein können. Die Leitfrage lautet dabei: *Inwiefern kann Bourdieus legitime Sprache als Kulturimperialismus und Gewalt nach Young verstanden werden?*

Dazu werden zuerst Bourdieus Konzept der *legitimen Sprache* skizziert und einige Auswirkungen aufgezeigt, die sich für die Sprecher:innen ergeben (Kapitel 2). Danach wird Youngs Verständnis von *Unterdrückung* dargestellt und die Gesichter *Kulturimperialismus* und *Gewalt* genauer beschrieben (Kapitel 3). Darauf aufbauend soll gezeigt werden, wie legitime Sprache als *Kulturimperialismus* und *Gewalt* unterdrückend wirken kann und inwiefern eine Betrachtung legitimer Sprache mithilfe von Youngs Ansatz einige Auswirkungen legitimer Sprache verdeutlicht (Kapitel 4).

¹ *Face of Oppression* wird in dieser Arbeit mit *Variante*, *Facette*, *Form* und *Gesicht der Unterdrückung* übersetzt.

2. Legitime Sprache nach Bourdieu

Bourdieu denkt Sprache immer einhergehend mit Macht und Herrschaft. Er argumentiert für eine Analyse von Sprache, die ihre soziale Bedingungen in den Blick nimmt (Bourdieu 2017, 107). Vor diesem Hintergrund zeigt er auf, wie *offizielle Sprache* durch politisches Kalkül geschaffen und als *legitime Sprache* zur Norm gekrönt wird (2017, 120–125). Diese Produktion und die damit verbundene Reproduktion legitimer Sprache werden in diesem Kapitel nachgezeichnet, um dann einige ihrer Auswirkungen für Sprecher:innen anderer, nicht-legitimer Sprachen aufzuzeigen.

2.1 Produktion und Reproduktion legitimer Sprache

Als politisches Instrument dient die Schaffung einer offiziellen Sprache während sozialer Kämpfe der (Re-)Formierung von Denkstrukturen: Die neue soziale Welt wird sag- und denkbar, indem Interessen, Ideen und Vorstellungen der herrschenden Schicht mit neuem politischem Vokabular, neuen Metaphern und Verweissystemen vermittelt werden (Bourdieu 2017, 122). Das jeweilige Bildungssystem spielt hier eine fundamentale Rolle: In der Schule wird gelehrt und gelernt, *wie man was sagt* (Bourdieu 2017, 123). Einmal etabliert, wird legitime Sprache auf vielfältige Weise reproduziert: Das *Zusammenspiel von Bildung und Arbeitsmarkt* etwa setzt (oft unbewusste) Anreize für ihren Gebrauch (Bourdieu 2017, 124–125). Der Zugang zu begehrten Posten bedingt oft legitime Sprachkompetenz, was durch entsprechende Schulabschlüsse vermeintlich bescheinigt wird (vgl. Bourdieu 2017, 133). Folglich weisen Eltern ihre Kinder an, *richtig* zu sprechen oder drängen sie zum Besuch weiterführender Schulen – im Wissen, dass die Zukunft ihrer Kinder unter anderem von der Fähigkeit, sich richtig ausdrücken zu können, abhängt (Bourdieu 2017, 124–125). Auch unterschiedliche Varianten von *Objektivierung und Kodifizierung* – beispielsweise Bücher oder Gesetzestexte – wirken bei der Festsetzung oder Eingrenzung der Sprache mit, die gesellschaftlich zur Verfügung steht (Bourdieu 2017, 136).

Geschaffen wird so eine *von der legitimen Sprache durchzogene soziale Welt*, in der die vorherrschende Macht und Autorität der Sprache verschleiert werden: Eine Sprache, die ursprünglich eine unter vielen war, wird zur *legitimen Sprache*, die mit einer *legitimen Sprachpraxis* einhergeht. Alle anderen Sprachen und Sprachpraxen werden nun an dieser einen gemessen; *falsche* Aussprache, *andere* Redewendungen oder *abweichende* Dialekte werden, jeweils in Relation zur geltenden Norm, hörbar. (Bourdieu 2017, 130)

Zentral ist nun, dass sich nach Bourdieu sprachliche Unterschiede in soziale Unterschiede rückübersetzen lassen: Vokabular, Aussprache, Grammatik oder Gesten geben Aufschluss über soziale Herkunft und Position (2017, 131). Erklärt wird dies mithilfe des sogenannten *Habituskonzeptes*. Der Habitus kann grob als eine Art Netz von dauerhaften Veranlagungen (Dispositionen) verstanden werden, das durch gesellschaftlich geprägte Erfahrungen entsteht, welche das Individuum inkorporiert (Rehbein und Saalman 2009, 110–118). Er leitet Wahrnehmung, Handeln und Denken eines Individuums, u.a. durch die Tendenz, so zu handeln, wie man es gelernt hat (Rehbein und Saalman 2009, 110–118). *Sprechen* ist als körperliche Technik Teil des Habitus und beinhaltet neben der Aussprache auch beispielsweise Körperhaltung oder Mundbewegungen (Bourdieu 2017, 173). *Sprechen* ist damit die Aneignung eines *Sprachstils*, einer spezifischen Art und Weise des Kommunizierens und nicht nur die Produktion einer sprachlichen Äusserung (vgl. Bourdieu 2017, 147). Als Teil des Habitus ist dieser Sprachstil erfahrungsbedingt und somit auch gesellschaftlich geprägt (vgl. Bourdieu 2017, 109). Zudem sind die Sprachstile auf eine Weise geordnet, welche die Hierarchie der jeweiligen sozialen Gruppen widerspiegelt (Bourdieu 2017, 131). *Wie* jemand spricht, verrät damit einiges über die soziale Position der Sprecher:innen und die gesellschaftlichen Hierarchien:

„[...] in jeder Interaktion (und damit bei jedem Diskurs) ist über die verwendeten Sprachen, über ihre Sprecher und über die sozialen Gruppen – definiert über den Besitz der entsprechenden Sprachkompetenz – die ganze Sozialstruktur präsent.“ (Bourdieu 2017, 148)

Basierend auf dem Habitus bilden Individuen dann sprachliche Produkte, die sie auf einem bestimmten sprachlichen Markt anbieten (Bourdieu 2017, 108). Der Wert, den Sprache und sprachliche Produkte auf diesem Markt haben, liegt in der damit einhergehenden sozialen Anerkennung (Bourdieu 2017, 136). In einer von legitimer Sprache durchzogenen Welt bedeutet das, dass der eigene Sprachstil immer in Relation zur herrschenden Norm erfahren wird (vgl. Bourdieu 2017, 109).

Die alltägliche Anerkennung der legitimen Sprache – und damit auch die Anerkennung der Differenz des eigenen Sprachstils zur Norm – geschieht unbewusst und ohne offensichtlichen Zwang, jedoch mit schwerwiegenden Konsequenzen (vgl. Bourdieu

2017, 126). Da der sprachliche Habitus immer durch soziale Erfahrungen geprägt und auch vererbt wird, *reproduziert legitime Sprache bzw. die Anerkennung der legitimen Sprache auch die mit ihr einhergehenden sozialen Unterschiede* (Bourdieu 2017, 135). Differenzen im Sprachstil erlauben soziale Distinktion, die mit entsprechenden Konnotationen einhergehen – immer in Relation zur Norm (vgl. Bourdieu 2017, 109). Hier wird die Macht, die von legitimer Sprache ausgeht, besonders deutlich sichtbar: Legitime Sprachkompetenz wird zur statusabhängigen Fähigkeit (vgl. Bourdieu 2017, 151). Sie wird linguistisches Kapital:

„Mit anderen Worten, die Diskurse sind nicht nur (oder nur ausnahmsweise) Zeichen, die dechiffriert und verstanden werden sollen; sie sind auch Zeichen des Reichtums, zu taxieren und zu bewerten, und Zeichen der Autorität, denen geglaubt und gehorcht werden soll.“ (Bourdieu 2017, 147)

Das hat konkrete, fundamentale Auswirkungen für Menschen, denen legitime Sprachkompetenz fehlt, wie im folgenden Abschnitt ersichtlich wird.

2.2 Auswirkungen legitimer Sprache

Eine erste Auswirkung legitimer Sprache ist, dass Sprecher:innen nicht-legitimer Sprachen der *Zugang zu bestimmten sozialen Sphären erschwert oder sogar verwehrt wird* (vgl. Bourdieu 2017, 132–133). Gerade prestigeträchtige und ertragsreiche Positionen scheinen nur mit ausreichend legitimer Sprachkompetenz zugänglich. Wird legitime Sprache erst (zu) spät erlernt, bedeutet dies oft einen entscheidenden Nachteil im Kampf um materielle Sicherheit (Bourdieu 2017, 142). Damit ist legitime Sprache ein *Element von Chancenungleichheit*. Das kann für Sprecher:innen ohne legitime Sprachkompetenz aber nicht nur aus materieller oder finanzieller Sicht fatal sein, sondern auch hinsichtlich einer potenziellen Änderung der herrschenden Verhältnisse: Da die Möglichkeit, einen Posten mit sprachlicher Definitionsmacht – etwa als Lehrer:in oder Jurist:in – zu erhalten, von legitimer Sprachkompetenz abhängt, reproduzieren sich die bestehenden Verhältnisse erneut.

Des Weiteren werden Sprecher:innen nicht-legitimer Sprachen oft *nicht gehört* (vgl. Bourdieu 2017, 131–132). Das kann einerseits tatsächlich an ihrem Unvermögen liegen, offizielle sprachliche Äusserungen zu produzieren. Andererseits kann es dadurch bedingt sein, dass

Beiträge, die in nicht-legitimer Weise geäußert werden, nicht als entsprechende Wünsche, Interessen oder Beschwerden erkannt werden. Diese Auswirkungen sind sowohl aus individueller Perspektive als auch auf gesellschaftlicher Ebene höchst problematisch – beispielsweise für eine angestrebte politisch ausgeglichene Repräsentation in Demokratien.

Ferner leben Sprecher:innen nicht-legitimer Sprachpraxen ein Leben, in dem sie dem offiziellen Sprachgesetz *immer unterworfen* scheinen. Ihre Sprachpraxen werden *systematisch abgewertet*, sie werden etwa als „Mundart“ oder „vulgärer Slang“ bezeichnet (Bourdieu 2017, 130). Sie werden ihrer eigenen Sprache *beraubt*, da die legitime Sprache diejenige ist, die gesprochen werden muss, um gehört zu werden (Bourdieu 2017, 131–132). Dadurch werden die herrschenden Strukturen wiederum bestätigt: In der Anerkennung der Differenz des eigenen Sprachstils zur legitimen Sprache und im Versuch, sie zu sprechen, wird ihre Autorität erneut anerkannt (Bourdieu 2017, 143). Mit der Aneignung eines Sprachstils geht zudem die Entwicklung eines Sinns für den sozialen Wert von Sprache einher (Bourdieu 2017, 168). Dieser erlaubt die Vorwegnahme des potenziellen Profits sprachlicher Äusserungen auf einem bestimmten sprachlichen Markt, indem etwa Sanktionen antizipiert werden (Bourdieu 2017, 161). Sprecher:innen passen ihre Sprache deshalb je nach sozialer Situation den herrschenden Konditionen an und *zensieren sich damit selbst* (vgl. Bourdieu 2017, 162).

Alle diese Auswirkungen schränken Sprecher:innen nicht-legitimer Sprachen ein, sei es mit Blick auf berufliche Laufbahnen, politische Teilhabe oder persönlicher Entfaltungsmöglichkeiten. Damit liegt, mit Iris Marion Young, der Verdacht nahe, dass sie auch *unterdrückend* wirken können (vgl. Young 1990, 40). Um dieser Vermutung nachzugehen und die unterdrückenden Momente legitimer Sprache freilegen zu können, wird im nächsten Kapitel Youngs Konzept der *Unterdrückung* vorgestellt.

3. Unterdrückung nach Young

Gemeinhin wird Unterdrückung als bewusste Tyrannei der herrschenden Gruppe über andere Teile der Gesellschaft verstanden. Damit einher geht oft die Überzeugung, dass Unterdrückung in westlichen, liberalen Gesellschaften nicht existiere: Unterdrückt wird bei den *Anderen*. (Young 1990, 40–41)

Young kritisiert diese Auffassung scharf. Sie entwirft stattdessen ein differenziertes Bild von Unterdrückung, die diffus und subtil das soziale Leben in allen Gesell-

schaften durchzieht. Nach Young zeigt sich Unterdrückung in fünf Formen: als *Ausbeutung*, *Marginalisierung*, *Machtlosigkeit*, *Kulturimperialismus* und *Gewalt* (Young 1990, 40). In allen Varianten hat Unterdrückung eine einschränkende, beschneidende Auswirkung auf die Unterdrückten und ihr Leben:

“In the most general sense, all oppressed people suffer some inhibition of their ability to develop and exercise their capacities and express their needs, thoughts, and feelings.”
(Young 1990, 40)

Abgesehen davon sind konkrete Unterdrückungserfahrungen und -auswirkungen aber vielfältig, vielschichtig, fluide und oft ineinander verwoben (Young 1990, 64). Nachfolgend werden einige allgemeine Merkmale von Unterdrückung sowie die beiden konkreten Aspekte *Kulturimperialismus* und *Gewalt* genauer beschrieben.

3.1 Merkmale

Unterdrückung ist bei Young ein zentrales *Element sozialer Ungerechtigkeit*, die jenseits materieller Ungleichheit *kulturelle und alltägliche Aspekte* betrifft. Sie denkt Unterdrückung als *Zustand sozialer Gruppen*. Der Begriff der sozialen Gruppe wird dabei nicht als essentialistisch oder deterministisch, sondern als sozial konstruiert verstanden: Eine Gruppe definiert sich in Abgrenzung zu einer anderen und deren kulturellen Praxen. Dabei ist ein Individuum oft Teil verschiedener sozialer Gruppen, wovon einige unterdrückt sein können und andere nicht. Obwohl Gruppenzugehörigkeit oft von aussen definiert wird, ist sie für die Mitglieder selbst häufig konstitutiver Teil der Identität: “[...] one finds oneself as a member of a group, which one experiences as always already having been” (Young 1990, 46). (Young 1990, 39–47)

Wichtig ist, dass Unterdrückung *keine aktiv unterdrückende Gruppe benötigt*. Bewusste Herrschaft ist damit keine notwendige Bedingung von Unterdrückung (kann aber durchaus damit einhergehen). Unterdrückung hat stattdessen primär *strukturelle Ursachen*: Sie entsteht oft durch das unhinterfragte, *unbewusste Akzeptieren* gesellschaftlicher Normen oder Stereotypen, welche die Basis für institutionelle Regeln darstellen. Durch diverse ökonomische, politische und kulturelle Institutionen mitsamt ihren gesellschaftlichen Konsequenzen wird Unterdrückung so tagtäglich (re-)produziert. Eingebettet in unseren sozialen Alltag geschieht sie oft unbewusst und wirkt *subtil*. Von der Unterdrückung einer Gruppe *profitiert immer eine andere, privilegierte Gruppe*. Letztere ist am Erhalt ihrer Privilegien interessiert und stützt deshalb (oft unbewusst) das geltende System und damit die Unterdrückung anderer Gruppen. Besonders deutlich zeigt sich das youngsche Verständnis von Unterdrückung in Form des Kulturimperialismus. (Young 1990, 38–42)

legierte Gruppe. Letztere ist am Erhalt ihrer Privilegien interessiert und stützt deshalb (oft unbewusst) das geltende System und damit die Unterdrückung anderer Gruppen. Besonders deutlich zeigt sich das youngsche Verständnis von Unterdrückung in Form des Kulturimperialismus. (Young 1990, 38–42)

3.2 Kulturimperialismus

Young definiert Kulturimperialismus als *Universalisierung der Erfahrungen und Kultur einer dominanten, d.h. zahlenmässig grossen, Gruppe* und die damit verbundene Krönung dieser Kultur zur Norm. Durch privilegierten oder sogar exklusiven Zugang der dominanten Gruppe zu Interpretationsmitteln der sozialen Welt und gesellschaftlicher Kommunikation projiziert die dominante Gruppe ihre Werte, Erfahrungen, Ziele und Perspektiven (unbewusst) auf Mitglieder anderer Gruppen. Dass sie so ihre eigene kulturelle Perspektive als *einzig* Perspektive verkennt, hat für anderen sozialen Gruppen einschneidende Folgen. Deren Sicht- und Lebensweise werden auf paradoxe Weise doppelt unterdrückt. Einerseits werden sie mindestens als von der Norm abweichend, andererseits als unterlegen oder sogar als kriminell markiert und abgewertet. Das Dominante ist das *Normale*, das Nicht-Dominante das *Andere*. Die dabei entstehenden Stereotypen verzahnen sich mit bereits vorhandenen Vorurteilen und tragen zur Reproduktion der Unterdrückung bei. Demgegenüber werden nicht-dominante kulturelle Aspekte unsichtbar: Indem die dominante Kultur ihre eigenen Elemente universalisiert, lässt sie höchstens in einer stereotypisierten oder marginalisierten Form Platz für anderes. (Young 1990, 58–61)

Kulturimperialismus beschneidet damit die Entwicklung und Entfaltung der Mitglieder nicht-dominanter Gruppen massiv:

“While the subject desires recognition as human, capable of activity, full of hope and possibility, she receives from the dominant culture only the judgment that she is different, marked, or inferior.” (Young 1990, 60)

Dabei generiert die Zugehörigkeit zur Gruppe der Unterdrückten eine tatsächliche Andersartigkeit, nämlich *spezifische Erfahrungen und Perspektiven des Unterdrücktseins*, welche von der dominanten Kultur wiederum nicht oder kaum wahrgenommen werden. Die Mitglieder der dominanten Gruppe sind in ihrer Rolle als Profiteur:innen mindestens unbewusst am Erhalt des Systems interessiert. Auf Herausforderun-

gen des universellen Anspruchs der dominanten Kultur durch nicht-dominante Gruppen reagiert sie aber nicht nur mit gesellschaftlicher Unterjochung oder Stereotypisierung der *Anderen*, sondern auch mit Gewalt. (Young 1990, 58–61)

3.3 Gewalt

Gewalt gegen eine soziale Gruppe äussert sich in vielfältiger Art und Weise. Was Gewalt für Young zu einem Gesicht der Unterdrückung macht, ist weniger der unmittelbare Akt oder das Ausmass ihrer Brutalität als der soziale Kontext, der sie umgibt: Gewalt gegen Mitglieder einer unterdrückten Gruppe wird dadurch charakterisiert, dass sie sozial möglich, toleriert und sogar akzeptiert ist. Es handelt sich um eine soziale Praxis. Direkte Gewaltanwendung ist dabei gar nicht notwendig:

“The oppression of violence consists not only in direct victimization, but in the daily knowledge shared by all members of oppressed groups that they are *liable* to violation, solely on account of their group identity.”
(Young 1990, 62)

Das Wissen, dass man selbst oder Mitglieder der eigenen sozialen Gruppe jederzeit – und sozial akzeptiert – Gewalt fürchten müssen, beschneidet Freiheit und Würde und wirkt damit unterdrückend. Young betont hier vor allem den unterdrückenden Aspekt dieses Wissens an sich. Unterdrückung durch Gewalt scheint sich dadurch aber auch indirekt im Alltag zu zeigen, beispielweise durch Verhaltensänderungen oder Selbstbeschränkungen. Young hat somit ein breites Verständnis von Gewalt: Auch Bedrohung, Spott und Belästigung sind Varianten unterdrückender Gewalt, sobald sie dazu dienen, Menschen aufgrund ihrer Gruppenzugehörigkeit zu erniedrigen, zu demütigen oder zu stigmatisieren. (Young 1990, 61–62)

Während Kulturimperialismus einen unterdrückten sozialen Kontext (re-)produziert, scheint dieser mittels Gewalt verteidigt und ausgeweitet werden zu können. Gleichzeitig schafft die Herrschaft der dominanten Kultur eine Wirklichkeit, in der Gewalt sozial akzeptiert wird, wodurch sie strukturell wird (vgl. Young 1990, 62). Beide Facetten sind also miteinander verknüpft (vgl. Young 1990, 62). Sie schränken auf ihre eigene Weise die Möglichkeit zur freien Entfaltung der Unterdrückten ein und verhindern, dass ihre Bedürfnisse, Perspektiven und Gefühle gesellschaftlich wahr- und ernstgenommen werden.

4. Legitime Sprache und Unterdrückung

In diesem Kapitel wird nun, basierend auf den vorhergehenden Kapiteln, das unterdrückende Potenzial legitimer Sprache freigelegt. Danach werden die Erkenntnisse aus dieser Analyse genauer diskutiert, um festzuhalten, inwiefern der Blick auf die unterdrückenden Aspekte legitimer Sprache für die Lektüre von Bourdieus Text gewinnbringend sein kann.

4.1 Legitime Sprache als Kulturimperialismus

Legitime Sprache ist mindestens dreifach unter der Perspektive eines unterdrückenden Kulturimperialismus denkbar: Sie ist erstens selbst *Element der dominanten Kultur*, zweitens ein *Ort der indirekten Reproduktion von Unterdrückung in anderen kulturellen Bereichen* und drittens ein *direktes Mittel zur Verbreitung der dominanten Kultur*. Dabei sind diese Momente weder trennscharf abgrenzbar noch unabhängig voneinander. Sie müssen eher als Blickrichtungen auf das unterdrückende Potenzial legitimer Sprache verstanden werden, die je nach Fokus unterschiedliche Aspekte beleuchten.

Betrachtet man legitime Sprache als *Element der herrschenden Kultur*, so zeigt sie sämtliche Merkmale kulturimperialistischer Unterdrückung: Durch ihre breite Anerkennung als Mass aller Sprachen – individuell inkorporiert und im sozialen Alltag eingebettet – wird ihre Autorität normalisiert und verschleiert. Kaschiert als *die* Sprache, dient sie den Interessen derjenigen sozialen Gruppe von Menschen, die sie sprechen und die von ihr (unabsichtlich) profitieren. Für Sprecher:innen mit nicht-legitimer Sprachkompetenz hat dies die von Young beschriebene doppelte paradoxe Unterdrückung zur Folge (vgl. 1990, 59): So werden einerseits alle anderen Sprachen und Sprachpraxen systematisch abgewertet und ihre Sprecher:innen stigmatisiert. Jeweils dem Vergleich zur offiziellen Sprache ausgesetzt, werden Dialekte als *Abweichungen* und kulturelle Metaphern, Ausdrücke oder spezifische Aussprachen als sprachliche *Fehler* aufgefasst. Sprecher:innen werden korrigiert und damit negativ markiert. Andererseits wird nicht-legitime Sprache jenseits dieser Entwertung unhörbar: Wer seine Anliegen auf sprachlich nicht-legitime Art vorbringt, hat gesellschaftlich, politisch oder rechtlich grundsätzlich schlechtere Karten: Die auf diese Weise vertretenen Interessen, Wünsche oder Präferenzen werden als weniger wichtig aufgefasst oder bleiben tatsächlich unverstanden und ungehört.

Dies ist unterdrückend, da die individuellen Entfaltungs- und Erfahrungsmöglichkeiten qua Gruppenzugehörigkeit klar beschnitten werden (vgl. Young 1990,

40). Erstens schränkt die durch legitime Sprache geschaffene soziale Welt das Denk- und Sagbare massiv ein. Da legitime Sprache bestimmt, was gesagt werden kann, darf und soll, werden gruppenspezifische Erfahrungen, Wünsche und Bedürfnisse – beispielsweise die Formulierung homosexuellen Begehrens oder die Thematisierung rassistischer Erfahrungen – unsagbar, ungehört und unerhört. Das kann sowohl für soziale Gruppen wie auch für die einzelnen Gruppenmitglieder fatal sein: *Es fehlen einem die Worte, fundamentale Elemente der eigenen Identität zu fassen, auszudrücken und von anderen wahrgenommen zu werden.*

Zweitens ist legitime Sprache ein *strukturelles Element von Chancenungleichheit*, da legitime Sprachkompetenz oft eine notwendige Bedingung für den Zugang zu bestimmten sozialen Sphären und Posten ist (vgl. Bourdieu 2017, 133). Dabei scheint es ebenso paradox wie ungerecht, dass es sich um ebendiese Sphären und Posten mit sprachlicher Definitionsmacht, politischem Einfluss und materieller Sicherheit handelt. Legitime Sprache muss aber nicht zwingend als direktes, bewusstes Ausschlusskriterium wirken: Sprecher:innen nicht-legitimer Sprachen scheinen vom legitimen Diskurs eher gar nicht adressiert und auch nicht gehört zu werden. Damit einher geht auch die fehlende gesellschaftliche Repräsentation von Sprecher:innen nicht-legitimer Sprachen, welche die Situation verschärft: Wer keinen Zugang hat oder ungehört bleibt, kann in den bestehenden Strukturen schlecht für die eigenen Interessen kämpfen. Dies ist insbesondere mit Blick auf diejenigen spezifischen Interessen, Wünsche und Erfahrungen störend, die gerade durch den Prozess des Ausgeschlossen- und Ungehörtseins generiert werden.

Mit Blick auf die Rolle legitimer Sprache als Spiegel sozialer Differenzen geht von ihr die Gefahr der *indirekten Reproduktion von Unterdrückung in anderen Bereichen* aus. (Legitime) Sprache wird hier zum Medium, das Informationen über soziale Herkunft transportiert. Wenn ein bestimmter Sprachstil Rückschluss auf Herkunft oder Identität erlaubt, so wird Sprechen zur Gefahr für Betroffene mit Blick auf Unterdrückung, basierend auf der jeweiligen Herkunft oder Identität: Denn durch Sprache werden bestehende gesellschaftliche Strukturen affirmiert. Zentral ist hier, dass diese Reproduktion meist unbewusst, subtil und unabsichtlich geschieht, etwa durch die angenommene vermeintliche Deckung von sprachlichen Fähigkeiten und fachlicher Kompetenz: Einer Person, die die offizielle Sprache beherrscht und sie auf legitime Art und Weise spricht, wird zugleich fachliche Kompetenz zugeschrieben (Bourdieu 2017, 151–152). In der Folge wird

Unterdrückung in anderen Bereichen, vermittelt über die Sprache, erneuert oder sogar verstärkt.

Als dritte Variante ist Sprache in der Form einer *Gehilfin des Kulturimperialismus* als Unterdrückung fassbar: Legitime Sprache kann als *Instrument* verwendet werden, bzw. wirkt als Trägerin, um die *dominante Kultur zu verbreiten und zu etablieren*. Sie bestimmt, wer und was gehört wird, wer was sagen darf. Insbesondere im Bereich des Rechtsprechens erreicht kulturimperialistische Unterdrückung via legitime Sprache eine ungeheure Tiefe:

„Das rechte, das formal richtige Sprechen erhebt von daher – und durchaus mit nennenswerten Aussichten auf Erfolg – den Anspruch, Recht zu sprechen, das heißt zu sagen, was sein soll.“ (Bourdieu 2017, 113)

In der kodifizierten und objektivierten Form von Gesetzestexten (und auch anderen offiziellen Dokumenten) verleiht legitime Sprache der dominanten Kultur mit ihren Ideen, Konzepten und Vorstellungen eine kollektiv anerkannte, rechtliche Existenz, Basis und Legitimation (vgl. Bourdieu 2017, 114). Diese Macht legitimer Sprache, bzw. derer, die sie beherrschen, scheint fundamental: Sie durchzieht und bestimmt sowohl das individuelle als auch das gesellschaftliche Leben. Nicht umsonst bezeichnet Bourdieu Sprache als „*Medium par excellence* des Traums von der absoluten Macht“ (2017, 114). Wer über die legitime Sprache herrscht und sie definieren kann, bestimmt über die gesellschaftliche Realität – bewusst oder unbewusst, absichtlich oder unabsichtlich.

Legitime Sprache kann also auf vielfältige Weise kulturimperialistisch unterdrückend verstanden werden. Sie erfüllt alle von Young genannten Kriterien dieses Gesichts der Unterdrückung: Legitime Sprache wird normalisiert, universalisiert und unbewusst im sozialen Alltag eingebettet. Sie markiert in ihrer Rolle als Mass aller Sprachen die Sprachpraxen anderer sozialer Gruppen als abweichend und verhindert gleichzeitig, dass deren Sprecher:innen gehört werden. Während diese Tatsache Sprecher:innen nicht-legitimer Sprachen in ihrem Leben massiv einschränkt, profitiert eine andere Gruppe, nämlich diejenige, die legitime Sprachkompetenz vorweisen kann. Das geschieht meist unbewusst und unbemerkt: Lehrpersonen, Beamt:innen oder Autor:innen tragen somit durch ihr alltägliches Schreiben und Sprechen – meist ohne böse Absicht – zur Reproduktion bestehender Ungleichheiten bei. Betont werden muss, dass die Autorität legiti-

mer Sprache nicht eigenständig wirkt, sondern immer durch ihre alltägliche Reproduktion innerhalb eines sozialen Kontexts bedingt wird:

„Die legitime Sprache verfügt aus sich selbst heraus ebenso wenig über die Macht, ihr eigenes zeitliches Fortbestehen zu sichern, wie über die Macht, ihre räumliche Verbreitung zu bestimmen. Nur jene *creatio continua*, die sich in den ständigen Auseinandersetzungen zwischen den verschiedenen Autoritäten vollzieht, die im Felde der spezialisierten Produktion um das Monopol und die Durchsetzung der legitimen Ausdrucksweise konkurrieren, kann die Permanenz der legitimen Sprache und ihres Wertes, das heisst der Anerkennung, die sie finde, gewährleisten.“ (Bourdieu 2017, 136)

Legitime Sprache ist aber nicht nur kulturimperialistisch, sondern scheint auch als Gewalt fassbar, wie im Folgenden gezeigt wird.

4.2 Legitime Sprache als Gewalt

Als Form von Gewalt scheint legitime Sprache mindestens auf zwei Arten denkbar: Einerseits kann legitime Sprache als *direktes und indirektes Mittel* verstanden werden, um kulturimperialistische Vorstellungen durchzusetzen und zu verteidigen. Hier wird klar, wie eng Kulturimperialismus und Gewalt miteinander verbunden sind. Andererseits scheinen bestimmte Aspekte der *legitimen Sprache an sich* bzw. des sprachlichen Marktes, den sie strukturiert und dominiert, als Gewalt fassbar.

Wird legitime Sprache als Gehilfin des Kulturimperialismus verstanden, scheint es eine Grenze zu geben, wo sie zur Gewalt wird – vor allem mit Blick auf ihre konkreten Folgen für nicht-dominante Gruppen. Legitime Sprache ist hier nicht mehr nur Trägerin der dominanten Kultur, sondern ein Mittel zu ihrer *Verteidigung*. Als Reaktion auf explizite und implizite Herausforderungen anderer sozialer Gruppen wird Sprache bewusst und unbewusst von der profitierenden Gruppe genutzt, um die herrschenden Verhältnisse zu erhalten. Das geschieht etwa dann, wenn politische Parteien die rechtliche Verankerung der Ehe als Gemeinschaft zwischen Mann und Frau zementieren wollen, um Homosexuellen die Ehe zu verunmöglichen, da deren Lebensweise nicht der dominanten Kultur entspricht. Sprache wird so zur *unterdrückenden Gewalt*, da sie auf Ausschluss, Demütigung und Stig-

matisierung von Menschen ausschliesslich aufgrund ihrer sozialen Gruppenzugehörigkeit zielt (vgl. Young 2017, 62). Besonders sichtbar wird das, wenn etwa *People of Colour* bedroht, Frauen sprachlich belästigt oder Menschen mit Behinderung verspottet werden, weil sie dunkelhäutig, weiblich oder beeinträchtigt sind. Gerade hier muss der herrschende soziale Kontext miteinbezogen werden: Legitime Sprache wirkt dann als unterdrückende Gewalt, wenn sie im Dienst und auf Grundlage kulturimperialistischer Herrschaft steht und akzeptiert wird. Somit wird sie zur *systematischen und strukturellen Gewalt*.

Oft zeigt sich das gewaltvolle Potenzial legitimer Sprache aber subtiler als durch direkte und offen geäußerte Worte. So kann eine dominante Kultur ihre Werte und Vorstellungen mittels legitimer Sprache verteidigen, indem beispielsweise Formulare ein bestimmtes Geschlechtersystem voraussetzen, bei dem nicht-binäre Geschlechtsvarianten unzulässig sind. Entscheidend, um solche im sozialen Alltag eingebetteten sprachlichen Äusserungen als Gewalt charakterisieren zu können, scheint der Fokus auf den Auswirkungen für die unterdrückten Gruppen zu sein. Tagtäglich wird nicht-dominanten Gruppen auf sprachliche Art und Weise ihr Anderssein vermittelt. Dies kann mindestens psychisch verletzend wirken. Mit Bourdieus Habituskonzept scheint es aber auch möglich, diese verletzenden Folgen als physische Gewalterfahrung zu beschreiben: Wie oben ausgeführt, basiert der Habitus auf gesellschaftlich geprägten Dispositionen, die individuell inkorporiert werden. Die Folgen müssen *körperlich* gedacht werden. Beleidigungen, Spott, Belästigung und Drohungen prägen, wie wir denken, leben, uns bewegen oder sprechen. Legitime Sprache hat damit Auswirkungen, die physisch verletzen können, etwa indem einem Worte fehlen oder gestottert oder gebückt gesprochen wird (vgl. Bourdieu 2017, 173–174).

Legitime Sprache kann auf diese Weise auch als *Gewalt an sich* verstanden werden, indem sie den möglichen Lebens- und Denkraum einschränkt. Erstens kann die eigene, freie Entfaltung erschwert oder verhindert werden, da Gefühle, Erfahrungen oder Ideen sprachlich nicht fassbar oder negativ konnotiert sind. *Es fehlen einem die Worte*, zentrale Aspekte des eigenen Lebens zu formulieren. Insbesondere konkrete Unterdrückungserfahrungen bleiben in den Worten der legitimen Sprache unausgesprochen, da sie dort nicht existieren bzw. ungehört bleiben. Legitime Sprache schafft, zweitens, einen sozialen Raum mit potenzieller Sanktionsgefahr im Falle sprachlicher Abweichung. Ganz ähnlich, wie beispielsweise manche Frauen keine

kurzen Kleider anziehen oder Menschen mit Behinderung manche Ausbildungen von vornherein ausschließen, zensieren sich Sprecher:innen nicht-legitimer Sprachen oder passen sich an, um sich selbst nicht in Gefahr zu bringen. Sowohl das *Wissen* um die potenzielle Gefahr als auch die *daraus resultierenden Verhaltensänderungen* sind Zeichen unterdrückender Gewalt. Bourdieu selbst spricht in diesem Zusammenhang von *lautlos-unsichtbarer Gewalt*:

„Am allerdeutlichsten wird sie [die Gewalt legitimer Sprache] jedoch in all den Korrekturen, die die Beherrschten punktuell oder auf Dauer im aussichtslosen Bemühen um Korrektheit bewusst oder unbewusst an den stigmatisierten Merkmalen ihrer Aussprache, ihres Wortschatzes [...] und ihrer Syntax vornehmen; oder in der Verwirrung, die sie ‚ganz aus der Fassung bringt‘ und ihnen ‚die Sprache verschlägt‘, als hätte man ihnen plötzlich ihre eigene Sprache weggenommen.“ (Bourdieu 2017, 128)

Gewalt in Form von legitimer Sprache wirkt ohne direkte:n Angreifer:in hiermit unsichtbar. Sie ist eingebettet in unsere sozialen Strukturen, normalisiert in der Anerkennung der Norm und im Versuch, bestehende Differenzen zu verstecken.

4.3 Legitime Sprache durch die Brille der Unterdrückung

Die Auswirkungen von Bourdieus legitimer Sprache können also in vielfältiger Weise als Form struktureller Unterdrückung nach Young verstanden werden. Youngs genaue Beschreibung der kulturimperialistischen Herrschaft einer dominanten sozialen Gruppe schafft einen theoretischen Rahmen, um legitime Sprache im sozialen Kontext zu verorten: Die (Re-)Produktion legitimer Sprache und legitimer Sprachpraxis spielt eine fundamentale Rolle bei der Produktion, Reproduktion, Verbreitung und Verteidigung der dominanten Kultur: sowohl als eigenes kulturelles Element als auch als Spiegel sozialer Differenzen und sprachliches Instrument der profitierenden Gruppe bzw. Träger der herrschenden Normen. Dabei kann sie als unterdrückende Gewalt verstanden werden, wenn man den Fokus auf die konkreten physischen und psychischen Auswirkungen für Unterdrückte legt und sie als Verteidigungsmittel der profitierenden Gruppe gegen implizite und explizite Herausforderungen nicht-dominanter Gruppen denkt.

Diese Rollen und Auswirkungen legitimer Sprache bei gesellschaftlicher Unterdrückung kommen in Bourdieus Analyse zu kurz. Er scheint deren (Re-)Produktion und den sprachlichen Markt primär mit dem Ziel zu analysieren, die „Illusion des Sprachkommunismus“ (Bourdieu 2017, 115) aufzudecken und die scheinbare Natürlichkeit der Verhältnisse als geschaffenes System der Herrschenden zu entlarven. Bei den Auswirkungen auf der individuellen Ebene fokussiert er auf die einverlebte Erfahrung legitimer Sprache als Teil des (Klassen-)Habitus. Unterdrückung scheint dabei zwar immer mitgedacht zu werden, wird aber nicht konkret ausformuliert, da andere Aspekte im Zentrum der Analyse stehen.

Mit Youngs Ansatz hingegen wird fassbar, inwiefern sich legitime Sprache und das dazugehörige Herrschaftssystem im sozialen Alltag anfühlen: eher als subtiles, oft unsichtbares Gefühl anders zu sein, nicht dazugehören, nicht gehört zu werden, nicht gemeint zu sein, nicht angesprochen zu werden, sich anpassen zu müssen oder ständig potenzieller Gefahr ausgesetzt zu sein. Das von Bourdieu oft betonte unbewusste Moment von Herrschaft lässt sich mithilfe des Konzepts der Unterdrückung präziser beschreiben und erklären, da Unterdrückung keine direkt unterdrückende, sondern lediglich eine unbewusst profitierende Gruppe bedingt. Durch die „youngsche Brille der Unterdrückung“ betrachtet, werden die Auswirkungen legitimer Sprache damit genauer beleuchtet, wobei sich je nach Blickrichtung ein anderes Gesicht zeigt.

5. Fazit

Die Leitfrage dieses Textes lautet: *Inwiefern kann Bourdieus legitime Sprache als Kulturimperialismus und Gewalt nach Young verstanden werden?* Nach der Analyse wird klar, dass Youngs Konzept der Unterdrückung sinnvoll und gewinnbringend eingesetzt werden kann, um einige Auswirkungen legitimer Sprache für Sprecher:innen nicht-legitimer Sprachen genauer zu charakterisieren. Legitime Sprache ist in verschiedenen Ausprägungen ein fundamentaler Teil unterdrückender kulturimperialistischer sozialer Systeme und eine Form von unterdrückender Gewalt mit psychischen und physischen Auswirkungen.

Während Bourdieu versucht, die Bedingungen der Aneignung legitimer Sprachkompetenz zu beschreiben und die Entstehung des sprachlichen Marktes skizziert, auf dem definiert wird, was als legitim gilt und was nicht, konzentriert sich Young eher auf die Auswirkungen solcher Prozesse. Durch Youngs Ansatz wird Bour-

dieus Analyse um den Aspekt einer facettenreichen und vielschichtigen Unterdrückung erweitert.

In der (Re-)Produktion legitimer Sprache wird aber auch die Verschränkung von Kulturimperialismus und Gewalt besonders deutlich: Sprachliche Gewalt, ob bewusst oder unbewusst, direkt oder indirekt, wird ermöglicht und bedingt durch das Interesse der profitierenden Gruppe, die soziale und sprachliche Definitivsmacht zu erhalten, zu verbreiten und zu verteidigen.

Mit Bourdieu und Young lässt sich erkennen, dass unser sozialer, politischer, rechtlicher und ökonomischer Alltag von legitimer Sprache durchzogen ist und Sprecher:innen nicht-legitimer Sprachen dadurch unterdrückt werden. Das ist sowohl störend als auch ungerecht. Offen bleibt, wie aus diesem selbst reproduzierenden System und den damit einhergehenden unterdrückenden Auswirkungen ausgebrochen werden kann. Dennoch leisten beide Autor:innen einen sehr wichtigen Beitrag zur Problemlösung, indem sie unsichtbare gesellschaftliche Strukturen sichtbar machen und deren scheinbare Neutralität und Auswirkungen analysieren und kritisieren: Sprache ist kein neutrales Gut und Unterdrückung bedeutet nicht bewusste Tyrannei. Stattdessen sind wir durch unseren Sprachgebrauch – meist unbewusst, unfreiwillig und in unterschiedlichen Rollen – an der Reproduktion unterdrückender legitimer Sprache beteiligt.

Literatur

- Bourdieu, Pierre. 2017. „Zur Ökonomie des sprachlichen Tauschs (1982).“ In Schultheis, Franz und Stephan Egger (Hg.): *Pierre Bourdieu. Sprache. Schriften zur Kulturosoziologie I*, 107–178. Berlin: Suhrkamp.
- Rehbein, Boike und Gernot Saalmann. 2009. „Habitus.“ In Fröhlich, Gerhard und Boike Rehbein (Hg.): *Bourdieu Handbuch. Leben – Werk – Wirkung*, 110–118. Stuttgart: Metzler.
- Young, Iris Marion. 1990. „Five Faces of Oppression.“ In Young, Iris Marion (Hg.): *Justice and the Politics of Difference*, 39–65. Princeton (New Jersey): Princeton University Press.

Luisa Paozzi studiert im Master *Political, Legal, and Economic Philosophy in Bern*. Sie interessiert sich für die verschiedenen Schnittstellen von Recht, Politik und Philosophie.

What is Wrong about Crimes?

On the Distinction between *mala prohibita* and *mala in se* in Criminal Law

1. Introduction

Murder is a crime. Likewise, it is a crime to drive on the left side of the road. These two acts can both be criminalised by a criminal justice system, both are equally binding since for both one must suffer the consequences in case of violation. But there seems to be an important difference on a theoretical level: murder is considered wrong not only in legal terms, but also because it is *inherently* wrong or evil to kill someone. Killing is a *malum in se* (Duff 2009, 89). On the contrary, driving on the left side of the road does not seem to be wrong *per se* or inherently so. It is regarded as wrong only because it has been established by law that one must drive on the right side. In other words: it is wrong because it is forbidden, it is a *malum prohibitum* (Duff 2009, 89). In theory, criminal law offences are distinguished according to whether the behaviour in question is to be considered wrong, inherently and independently of the law, or whether it concerns rather normal behaviour which is forbidden by law and thereby criminalised. Broadly speaking, this is the distinction between *mala prohibita* and *mala in se* (Duff 2009, 89). But what theoretical framework are we talking about?

On a theoretical level, we might ask if there is a grand unifying theory of criminal law (Duff 2009, 6). Ideally such a theory would define the most fundamental principles of criminal law, independent of different criminal law systems, and thus explain how criminal law works, what it punishes, how it punishes, and why it punishes. One of the fundamental questions underlying this debate is: what are we criminally responsible for? This question aims at the object of criminal responsibility and ultimately at what is and what is not criminalised by the state. In this framework, the distinction between *mala in se* and *mala prohibita* becomes essential. On the one hand, the law

seems to criminalise behaviour that is considered inherently wrong, and on the other hand, there are numerous conducts which are made wrong by the law. The distinction is therefore initially of a theoretical nature, and it is unclear what practical implications are raised. But even in theory, the distinction is controversial because it is unclear and imprecise. This is not only in terms of content and definitions, but also in terms of practical relevance and its added value for understanding criminal law. (Duff 2009, 6–10)

The aim of this paper is to work out whether the distinction between *mala prohibita* and *mala in se* makes sense by trying to determine what this distinction should be. The source text is chapter four “Criminally responsible for what? Crimes as wrongs” in the monograph by R. A. Duff: *Answering for crime—responsibility and liability in the criminal law* (Duff 2009, 79–93). In the first part of this thesis, Duff’s approach is being reconstructed. Subsequently, the distinction between *mala prohibita* and *mala in se* will be analysed in more detail according to their definitions. In a third part, the distinction will be critically reflected and discussed in terms of its practical relevance. In this paper, I will argue that there is no meaningful distinction between *mala prohibita* and *mala in se*, because there seems to be no coherent conception of *mala prohibita*. I will further argue that the distinction is of neither theoretical nor practical importance.

2. What are we responsible for? Duff on *mala prohibita* and *mala in se*

The following section constitutes a reconstruction of Duff’s approach to *mala prohibita* and *mala in se*. However, before Duff’s conception of the distinction can be traced in more detail, it seems necessary to explain the larger framework of the discussion. For, in

order to understand a term or a distinction, it is indispensable to consider the context in which the term or the distinction is used. The distinction between *mala prohibita* and *mala in se* is a distinction in relation to something, and only by looking at this something, we may understand what the distinction is.

2.1 The object of criminal responsibility: crimes as wrongful actions

What are we criminally responsible for? This question is the subject of chapter four in *Answering for Crime* (see Duff 2009, 79). At the beginning of this chapter, Duff summarises what has been established in the previous explanations: the conditions for criminal responsibility (Duff 2009, 79). On the one hand, we are only criminally responsible for what is within our control (*control requirement*) and, on the other hand, we have to be aware of what we are doing (*epistemic condition*). An essential aspect of these two conditions is that they only circumscribe responsibility retrospectively (Duff 2009, 79). However, there is also the possibility of circumscribing responsibility prospectively. We are responsible as citizens to our fellow citizens through the particular position we hold in society and the roles and duties that come with it. Duff's central thesis is that as citizens we are responsible for wrongful actions (Duff 2009, 80). Why we are responsible for actions rather than thoughts, choices, or even our character, is explained elsewhere (Duff 2009, 95–121) and, in view of the subject matter of this paper, cannot be discussed further. Instead, the focus of this paper is on the criterion of wrongfulness. If criminal responsibility consists of wrongful actions, then the question is what “wrong” means. Intuitively, it could be said that crimes should be moral wrongs (Duff 2009, 81). Why moral? It seems to be an important aspect of criminal law that it *condemns* our behaviour. Taxes, for example, are there, among other things, to guide our behaviour, criminal law, however, also wants to *condemn* our wrong behaviour (Duff 2009, 81). When our behaviour is condemned, it is not only expressed that a legal rule has been broken, but a stamp is also put on the behaviour that clearly declares it to be wrong. Duff also discusses the objection that criminal law is merely a catalogue of rules that leaves us free to break the rule or not (Duff 2009, 81). However, in his view criminal law cannot be just a collection of value-free rules, for there are certain reasons for criminalising some behaviour. These reasons, according to Duff, consist in moral wrongfulness (Duff 2009, 82). At this point, it is interesting that we have moved from

the question of what we are criminally responsible for to the question of what behaviour is criminalised by the state. The focus has thus moved away from the object of criminal responsibility to the object of criminalisation. In order to understand what kind of behaviour is criminalised by the state, it is first necessary to clarify the aspect of wrongfulness in crimes.

2.2 The role of ‘wrong’ in crimes

Two distinctions are fundamental to Duff's approach: first, the distinction between *object* and *condition* of criminal responsibility, and second, the distinction between *harmful* and *wrongful*. The first distinction runs through Duff's entire theory. The question is whether wrongfulness is the object or a condition of criminal responsibility. For example, it would be a condition of my criminal responsibility that I consciously performed the act in question and was not, for example, sleepwalking. However, consciously performing the act is not the object of criminal responsibility. I am not responsible for the fact I was not sleepwalking while acting, but for what I did. With regard to the second distinction, it can be stated that Duff distinguishes whether a behaviour is harmful to the environment or wrongful in a moral sense. (Duff 2009, 82–89)

Examining these two aspects together, the central question for Duff arises: are we criminally responsible for the harm or wrong that we do; or are we responsible for something else, on condition that our conduct was harmful or wrongful (Duff 2009, 83)? The question here is whether harm or wrongfulness are the primary reason for criminalising a behaviour, or whether they are only a condition to enable criminalisation, with the primary cause for punishment found elsewhere. One way to answer this question is to take the view that wrongfulness should be the primary reason for criminalisation: this position is called *legal moralism* (Duff 2009, 84). Implicit in this position is the assumption that the purpose of criminal law is retributive justice. In this respect, all those should be punished who are also morally guilty because they have committed a morally reprehensible act. Wrongfulness would thus be the intentional object of criminal responsibility. Legal moralism thus denotes a position whose advocates assume that wrongfulness is a good reason for criminalisation (Duff 2009, 84).

There are some problems with this account. If only morally wrong things are punished, then a rather large part of the existing criminal law ceases to be applicable. The reason for this is not that all offences covered by criminal law are of a moral nature, or at least, do not

seem to be of a moral nature. Further, the question arises whether such a position is tenable under the limitations of a liberal constitutional state. How is it possible that in a pluralistic society diverse values are combined under a morally coherent value structure? How authoritarian and binding should regulations be for someone who does not share the same values? The latter question essentially depends on what demands are made on moral regulations and how specific they are. For example, the law may criminalise unjust enrichment, leaving room for interpretation as to the meaning of “unjust” and the values that go with it. How to deal with this conceptual vagueness in practice is another question. However, by means of abstract and general formulations, law can encompass different values and summarise minimal basic requirements and values that different cultures share and thus set basic requirements for our social coexistence.

The situation is different with the liberal demand that behaviour should not be criminalised only in order that it can be punished retributively. This thesis is also known as the Walker-principle. In its complete formulation it reads:

“Prohibitions should not be included in the criminal law for the sole purpose of ensuring that breaches of them are visited with retributive punishment” (Duff 2002, 97).

This principle thus sets a minimum requirement for criminalisation. It does not say what is to be criminalised, but it creates a barrier to criminalisation by stipulating what at least may not be criminalised, namely behaviour that is not intrinsically wrong, but is only forbidden for the reason that it can be punished. Consequently, in the latter case, this would mean that the only thing wrong with the criminalised act is the fact that it has been criminalised. The discussion of this principle leads to the distinction between *mala prohibita* and *mala in se* (Duff 2009, 89). While *mala in se* denote crimes that are inherently wrong and therefore punishable, *mala prohibita* are characterised by the fact that they are wrong only due to the fact that they were declared wrong (Duff 2009, 89). In this respect, the Walker-principle seems to allude primarily to the questionable and unjustified regulation of *mala prohibita*. Duff, however, does not only reject the principle for *mala in se*, but he also finds it untenable in relation to *mala prohibita* (Duff 2002, 98). The following section intends to trace why he holds this view and how he justifies it.

2.3 Defending legal moralism: rejecting the Walker-principle

Duff rejects the Walker-principle both in relation to *mala in se* and in relation to *mala prohibita* (Duff 2002, 97). It is questionable how he justifies this. With regard to the Walker-principle, there is an important distinction that Duff introduces, and which is important to his argument because it reflects his understanding of retributivism. Specifically, it concerns the question of what exactly Nigel Walker means when he says “[...] for the sole purpose of ensuring that breaches of them are visited with retributive punishment” (Walker 1969, 26; see Duff 2002, 97). According to Duff, there is a difference in the question of whether someone is punished because their actions deserve to be met with the consequence threatened by the law, or whether the behaviour is criminalised for the only reason that it can be punished (Duff 2002, 106). Nigel Walker’s principle is hardly intended to deny that the penalties threatened by the law must consequently also be borne. Instead, the Walker-principle corresponds to the latter: we should not criminalise conduct for the sole reason of punishing it.

So why does Duff reject the Walker-principle in relation to *mala in se*? Duff gives a rather superficial definition of *mala in se*. He says that *mala in se* refers to crimes that are “wrong in themselves” or “inherently evil” (Duff 2009, 85). We must first again become aware of the problem Duff is trying to solve here. As mentioned above, advocates of legal moralism—that Duff is trying to defend here—would assume that moral wrongfulness is a good reason to criminalise behaviour. This would also mean that they would advocate criminalising a behaviour simply because it is morally wrong, even if it has no bad effects on the environment, which means that it is not *harmful*. From the perspective of virtue ethics, it would be the task of the law to educate the citizen to be a morally good person, and in this respect, the fear of the punishment should force the citizen to behave morally good and condemn their moral wrongfulness (virtue jurisprudence) (Duff 2009, 84). The Walker-principle, however, states that no behaviour should be criminalised for the mere reason that it can be punished. How can this difference between the two positions be resolved?

With regard to *mala in se*, Duff provides two arguments against the Walker-principle. Firstly, he argues that it is wrong to take the view that criminal law only imposes prohibitions (Duff 2009, 85). Secondly, the very nature of *mala in se* crimes requires to have reasons beyond the legal regulation not to do a certain

act. How are we supposed to understand this? We do not need to elaborate these two arguments separately, for they are essentially interdependent. What is important is to understand what Walker's understanding of law is. By thinking that law consists only of prohibitions, he assumes that law is the instrument of a sovereign who arbitrarily punishes certain behaviour. The fact that the act is criminalised is supposed to give us a reason to refrain from committing the act. Especially with regard to *mala in se*, it is not the case that we do not commit certain acts just because they are forbidden. In most cases, we have reasons beyond prohibition why we do not commit a certain act like, for instance, killing. One reason could be that we consider it as morally wrong to end someone's life, because everyone has a right to life. My reason for not killing someone is not simply because it is forbidden by law. Rather, I have other reasons, and it is precisely for these reasons that the ban seems understandable to me. On the other hand, in the case of offences which are said to be wrong only because they are forbidden, it is usually the case that we cannot understand the prohibition, and therefore the inhibition threshold to commit the offence is also much lower. In most cases, the law does not consist of unfounded and arbitrary prohibitions, but, since it addresses us as citizens and not only as compliant subjects, it provides us with comprehensible reasons why we should refrain from committing an act. Law succeeds in this by not simply prohibiting, but by conveying values. Those who commit an offence should realise that they are not simply violating a legal prohibition, but that they are violating community interests and thereby committing a *public wrong* (Duff 2009, 89). What does this mean for the Walker-principle? Regarding *mala in se*, it is obviously not tenable in the form presented by Duff. The acts in question are not only criminalized in order to punish their realisation. Rather, the punishment has a declaratory function by making it clear that committing such an act was a public moral wrongdoing.

The situation with *mala prohibita* is more difficult. By definition, *mala prohibita* are precisely conducts which only receive their wrongfulness through legal prohibition. Why should the Walker-principle also be rejected in relation to *mala prohibita*?

If *mala prohibita* are characterised by their wrongfulness arising only from legal regulation, this means *e contrario* that the conduct was not wrongful prior to its regulation. But what does "prior to" mean? Duff introduces a distinction to characterise this prelegal situation (Duff 2002, 104). In the process of criminal

legislation, a distinction must be made between legal regulation and criminalisation. According to Duff, criminal legislation functions in such a way that, in a first step, it is considered whether a certain conduct should be legally regulated at all (legal regulation). Only in a second step it is decided what kind of legal regulation should be used to guide our behaviour. A criminal law definition does not appear to be mandatory in order to guide behaviour. A tax regulation would also be possible. Hence, Duff's approach is to argue that the conducts in question need not have been wrongful before legal regulation, but that they were wrongful before criminalization (Duff 2002, 98). Therefore, the relevant behaviour is not wrongful before legal regulation but becomes wrongful with it, which in turn gives reason to criminalise the act. *Prima facie*, this differentiation seems somewhat artificial, and it is, on the one hand, difficult to understand what is meant by it, and on the other hand, it is questionable whether this reasoning is satisfying. For *mala in se* provide reasons beyond the law not to perform a certain act. According to Duff, however, these reasons are supposed to be found in legal regulations which antecede criminalisation. He accounts for this by grounding the intrinsic badness of the act that legal moralism demands in the "logical space" (Duff 2009, 90) between legal regulation and criminalisation. It is questionable whether this is convincing. The Walker-principle requires that behaviour is not criminalised for the sole reason of punishing it. However, this is exactly what Duff describes with his approach and then uses it in reverse to reject the principle. *Mala prohibita* are characterised by the fact that their badness is predetermined by law and cannot be justified in any other respect. Whether the badness is regulated before or after legal regulation or before or after criminalisation does not change the fact that the badness is founded only and exclusively in the law. Duff himself describes the aim of the distinction as follows:

"My aim here has been only to show that, if we distinguish the question of whether to regulate from that of how the regulations should be enforced, and define *mala prohibita* as offences consisting in conduct that is not wrongful prior to its legal regulation, we can see how legal moralists who make wrongfulness a criterion of criminalization can still justify some *mala prohibita*. What is crucial for them is that the conduct is wrong-

ful prior to or independently of its criminalisation and, as we have just seen, that can be true of conduct that is not wrongful prior to or independently of its legal regulation.” (Duff 2009, 92)

It seems that, with this approach, Duff is trying to rob *mala prohibita* of their character by artificially trying to justify their badness by distinguishing regulation from criminalisation, without giving any reason for why we should do that. Duff concludes that the Walker-principle must be rejected for both *mala in se* and *mala prohibita*. However, for this, he delivers not a convincing argument.

So far, the framework in which Duff discusses the distinction between *mala in se* and *mala prohibita* has been clarified. In general, it is a question of what we are criminally responsible for, namely—according to Duff—for wrongful actions (Duff 2009, 80). Further, it was stated that wrongfulness is not only a condition of criminal responsibility, but its object. This view corresponds to the position of legal moralism. Duff defends legal moralism for both *mala in se* and *mala prohibita* by rejecting the Walker-principle in both respects. Thus, he uses the distinction between *mala in se* and *mala prohibita* in his treatise on criminal responsibility. However, he uses the distinction instrumentally and does not give a definition of these two concepts. He only refers to the intuitive differentiation that *mala in se* are inherently wrong, while *mala prohibita* are wrong due to prohibition. With this, he does not address the ambiguities of the concepts. The following section will address the ambiguities of Duff's argumentation by looking at the conceptual definitions of the distinction between *mala prohibita* and *mala in se*.

3. *Mala prohibita* and *mala in se*: an empty distinction?

In *Answering for Crime*, Duff uses the distinction between *mala in se* and *mala prohibita* without explaining the content of the concepts in detail or properly distinguishing between them. It seems that Duff uses a superficial distinction which, in his opinion, needs no further elaboration (Duff 2009, 89). Instead, he states that while the traditional approach to explaining *mala prohibita* is “misleading” (Duff 2009, 89), its definition is “a real one” (Duff 2009, 89). However, he refrains from explaining in which respect the definition is misleading, but nevertheless genuine. In another essay, he also does not address the substantive distinction between *mala in se* and *mala prohibita*, but claims “I

assume that the distinction [...] is still clear and substantial enough to be a useful one” (Duff 2002, 98). The following section attempts to conceptualise the distinction between *mala prohibita* and *mala in se*. Thus, the terms are to be defined. For that, I will, in a first step, look at the historical background of the distinction and, in a second step, explain various attempts at defining the difficulties arising from the concepts.

3.1 Historical background

In 1496, the concepts *mala prohibita* and *mala in se* seem to have appeared for the first time (Wolfe 1981, 131). Also, the Columbia Law Review Association mentions this year as the time of the first appearance (1930, 74). However, there are authors who argue that the distinction is much older and already appears in antiquity in the work of Aulus Gellius (Evans 1914, 397). In 1496, Chief Justice Fineux used the distinction to analyse the power of the Crown. Basically, the point was that the Crown's prerogative to pardon someone should not apply to *mala in se* offences (Wolfe 1981, 132). The distinction thus served to identify the particularly serious offences among many and to establish that no excuses would apply to these offences. The distinction was therefore originally practical and not theoretical. Further, it does not refer to criminal responsibility. It is also interesting to note that the distinction was originally limited to marking the severity of the offence. It is precisely this aspect, however, that Duff later wants to exclude from the distinction (Duff 2009, 90).

3.2 Some basic definitions

Today the concepts *mala prohibita* and *mala in se* differ from their original usage. But how are they to be understood in modern times? A basic definition has already been given in the first part of this paper: Duff speaks of *mala in se* as crimes that are “wrong in themselves” or “inherently evil” (Duff 2009, 89). Black Law's Dictionary of 1979 speaks of it as a wrong and adds “acts morally wrong” (Wolfe 1981, 132). According to this dictionary *mala prohibita* are offenses “which are made criminal by statute but which, of themselves, are not criminal” (Wolfe 1981, 132). Ballentine's Law Dictionary speaks of *mala prohibita* as “wrongs only forbidden by positive law”. *Prima facie*, this seems plausible; Wolfe writes that the distinction has a “deceptive simplicity” (Wolfe 1981, 132). Probably for this reason, the distinction is often used without going into what it consists of, what it is supposed to serve and what added value it creates. If one tries to apply this purely theoretical distinction to categories, the dis-

inction suddenly becomes unclear and leads to what Wolfe calls a “frustrating ambiguity” (Wolfe 1981, 132).

There are many aspects of the distinction that are questionable. In my opinion, however, the major difficulties arise from the criteria *inherently* and *prelegal*, on which I would like to focus below.

Mala in se are bad in themselves, regardless of their regulation in positive law. *Mala prohibita* are not bad in themselves and only become so through their statutory regulation. In their prelegal state, they are not. From this, we gain the criterion of statutory regulation for the distinction. We can therefore ask whether the origin of wrongfulness arises from the statute or has already existed beforehand (Wolfe 1981, 132–133). It is thus noteworthy that a characteristic regarding regulations is at the centre of the distinction and not a substantive content-dependent one. This is interesting for the reason that, intuitively, *mala in se* is often understood to mean exclusively offences against life and limb, while *mala prohibita* forms the less important remainder. According to this basic definition, however, such an anticipation of content is not justified. What is decisive for the distinction is apparently not which *malum* is at issue, but when positive law declares it as such. Is it not ironic that a distinction that is supposed to tell us which behaviour is to be criminalised does not provide us with any substantive criteria? However, it is not excluded that a substantive criterion can be given. For Wolfe distinguishes whether the nature of the statutory regulation is creative or declaratory in nature (Wolfe 1981, 133). If the statutory regulation is only declaratory in nature, then it must be assumed that the wrongfulness of the act already existed before the legal regulation and insofar it is a *malum in se* and not a *malum prohibitum*. *Mala prohibita* require that their wrongfulness is created by law. The question here seems to be how wrongfulness is to be created this way. In order to assess this question, it seems necessary to look at the content of wrongfulness in more detail. In what respect must something be wrongful in order to be considered as a *malum in se*? Often, *mens rea* is mentioned as a criterion (Wolfe 1981, 133). Offences that are considered as a *malum in se*, are often characterised by an evil intent (Wolfe 1981, 133–134). However plausible this may seem at first glance, the consequence would be absurd. Wolfe correctly explains that it would be strange to frame *mala prohibita* by not requiring evil intent to be present. (Wolfe 1981, 133–134). One could also dispense with the imprecise concept of *mens rea* and state that everything that is common law must be considered *malum in se* (Wolfe 1981, 134). This approach

seems particularly ineffective because it circumvents the very character of *mala in se* that establishes that these crimes are *malum in se* regardless of the law. Common law certainly contains provisions that are considered *malum in se*, but these are certainly not *malum in se* because they belong to the common law.

It is difficult to pinpoint the concepts of *mala in se* and *mala prohibita* to a specific wrong in terms of content. This is, not least of all, because *mala in se* appear in so many different ways: they are by no means only offences against life and limb, but also offences against property, assets, even misdemeanours can be included. On the basis of these inconsistencies, Wolfe concludes that “[the distinction] is not an indispensable element of substantive or procedural criminal law” (Wolfe 1981, 141).

3.3 Critical review

The imprecision and incoherence associated with *mala in se* and *mala prohibita* are not satisfactory. Due to the diverse use of the terms and the manifold application of the distinction in different contexts, it is not only unclear what the distinction is, but also what the distinction is about. Due to this fact, the relevant question, in my opinion, is not what *mala prohibita* are, but whether *mala prohibita* exist at all.

In this regard, the criteria *inherently* and *prelegal* need to be revisited. As mentioned earlier in this paper, acts are not arbitrarily regulated by law and made punishable. As soon as it is considered whether behaviour should be criminalised, there are already various reasons suggesting the discussion of potential criminalisation. Duff himself says that, even in the case of *mala prohibita*, the law should give citizens reasons to understand the regulation (Duff 2002, 103). But could this not thereby make the regulated conducts *mala in se* again? As soon as the law gives us further reasons to understand the regulation, we may see the purpose of the regulation in the protection of these reasons and interests. The wrongfulness then arises in relation to the vulnerable interests and not purely from the fact that the conduct is regulated.

It would of course be possible to argue that the norms which regulate *mala prohibita* may in some circumstances protect interests that are attacked by *mala in se* offences, but that, nevertheless, the act, which constitutes a *malum prohibitum*, is not inherently wrong. For example, the right-hand lane rule protects legitimate interests such as road safety. Driving in the left lane is a morally neutral act until the law defines it otherwise, and this would be the difference to *mala in se*. But we should not argue this way, because of the reason that

we could argue the very same way in the reverse direction: even *malum in se* only get their characterisation when all circumstances are taken into account. For example, it is a *malum in se* if I cheat on my partner. However, the fact that I would be dating someone else is not in itself a *malum in se* act. What makes it so are the circumstances: for example the fact that we are in a monogamous relationship and that my partner relies on my fidelity.

The concept of *mala prohibita* seems empty of content since the law always protects certain interests with a regulation. As Duff rightly says, it is about the declaration of public wrongs. The fact that there should be no *mala prohibita* seems unsatisfactory. Perhaps, however, it is only a matter of revising the concept of *mala in se* as well: we should understand the term more generally as a violation of community interests. Driving on the left can also be seen as a *malum in se*, because it endangers road safety and thus puts fellow citizens in danger. Every regulation, as banal as it may seem *prima facie*, protects a certain interest. In saying this, I am not only saying that *mala prohibita* is an empty concept, but also that its emptiness is a good thing. For this concept would allow a totalitarian state to arbitrarily punish behaviour without giving comprehensible reasons for complying. If someone tells us that they do not commit murder merely because it is forbidden, they have not understood something essential in life.

The added value of the distinction, which is questionable from a practical point of view, has now also been challenged on a theoretical level. The distinction between *mala prohibita* and *mala in se* does not seem to make sense either in terms of content or concept. Instead, *mala in se* should be understood generally as offences that violate common interests. In the liberal view, however, this conception could lead to considerable difficulties: where are the limits of community interests? Could the regulation of conducts based on this conception result in over-criminalisation? However, these questions are not the subject of this paper and cannot be dealt with within its limited scope. It can only be stated that it still seems better to base wrongfulness on the violation of public interests than on the differentiation of regulation and criminalisation, without explaining the reasons beyond that.

4. Conclusion

The subject of this paper was the distinction between *mala prohibita* and *mala in se*. Its focus was the question of what the distinction consists of, with a view to what it is used for. This paper was prompted by a

chapter from Duff's monograph *Answering for Crime*. Firstly, it has been clarified in which framework Duff uses the distinction in order to gain insight into the discussion for which the distinction should be relevant. Secondly, it has been shown that Duff uses the distinction in his defence of legal moralism against the Walker-principle. This discussion is in turn embedded in the question of what we are criminally responsible for. It has also been shown that the latter question cannot be discussed independently of the question of what behaviour is criminalised. The reconstructed context has shown that Duff uses the distinction mainly instrumentally and does not address substantive inconsistencies.

For this reason, the second part of this paper is devoted to the various definitions that are central to *mala in se* and *mala prohibita*. Although the distinction seems attractive because of the elegance of its simplicity, several problems have emerged in relation to its categorical application. Ultimately, these differences in content have led to the fact that it is neither clear what the distinction is to be used for at all nor what its practical relevance is.

Further, I argued that the more important question is not what is meant by *mala prohibita*, but whether such exist at all. It has been shown how the criteria *inherently* and *prelegal* are problematic. This has led to the consequence that we should assume no *mala prohibita*, since already in the prelegal state considerations of why a behaviour should be criminalised are possible. This means that certain interests which indicate some actions to be inherently wrong, are already there in this state. However, by focusing on these interests, if we take the definition of *mala prohibita* in terms of content, it seems even good that there are no *mala prohibita* in this sense. For if a behaviour proves its wrongfulness only by being forbidden, it would allow authoritarian states to justify arbitrary rules.

So what does this mean for Duff's chapter in *Answering for Crime*? The superficiality with which he treats (or does not treat) the distinction in terms of content may seem disappointing. To be fair, however, the distinction and its definitional consequences are not the focus of the examined chapter. The distinction serves Duff to defend legal moralism against the Walker-principle. Thus, his aim is framed within a larger framework, but in one sense it is also much narrower since he does not claim to have a conclusive and correct conception of the concepts *mala in se* and *mala prohibita*. Nevertheless, Duff does not conclude with a clear understanding of what has been said on

these concepts, but rather, a certain vagueness resonates in his conclusion when he says: “On one view, the criminal law is a moralised practice, which focuses on moral wrongs; on the other, it is a technique of harm-prevention which is constrained by the demands of justice or fairness” (Duff2009, 93).

Of course, due to its limited scope, this paper has only covered a minimal part of the debate around *mala prohibita* and *mala in se*. Many questions and aspects had to be left open for future clarification, such as the question of how a pluralistic community of values should deal with moral wrongs or the question of whether the definition of public wrongs as the protection of community interests cannot lead to over-criminalisation. Moreover, the criterion of community interests only serves as a minimal positive criterion (or negative criterion in the sense of a barrier) for determining behaviour that should be criminalised by the state. This does not exclude the possibility that other criteria must be considered in order to determine this.

References

- Columbia Law Review Association. 1930. “The Distinction between ‘Mala prohibita’ and ‘Mala in Se’ in Criminal Law.” *Columbia Law Review* 30 (1): 74–86.
- Duff, R. A. 2002. “Crime, Prohibition and Punishment.” *Journal of Applied Philosophy* 19 (2): 97–108.
- ———. 2009. *Answering for Crime, Responsibility and Liability in the Criminal Law*. Oxford and Portland: Hart Publishing.
- Evans, Alvin E. 1914. “Aulus Gellius on Mala prohibita vs. Mala in se.” *The Classical Journal* 9 (9): 396–398.
- Walker, Nigel. 1969. *Sentencing in a Rational Society*. London: Allen Lane.
- Wolfe, Nancy Travis. 1981. “Mala in Se, A Disappearing Doctrine?” *Criminology* 19 (1): 131–143.

Céline Hoog (25) studiert im Master Rechtswissenschaften an der Universität Bern. Sie hat den Bachelor in Philosophie absolviert und ihre Interessen gelten der antiken Philosophie, der praktischen Philosophie und insbesondere der Rechtsphilosophie.

How Should We Program Autonomous Driving Vehicles Under Moral Disagreement?

1. Introduction

Consider a type of scenario where harm in an accident of an autonomous driving vehicle (ADV) is unavoidable, and where the available options are different distributions of harm to different objects or persons. How should ADVs be programmed to allocate the harm? Several scenarios of this type have recently been put forward in the philosophical literature (e.g. Lin 2015; Gerdes and Thornton 2015; Bonnefon et al. 2016). This issue has been called “the moral design problem” (Keeling 2018, 413). In his 2017 paper, Santoni de Sio proposed a legal-philosophical approach to find a solution to the moral design problem, because neither traditional nor experimental ethics could lead us to a consensus (Santoni de Sio 2017). In a response, Geoff Keeling suggested the Restricted Pareto Principle (RPP) as an interpretation of the doctrine of necessity. In his paper, he asserts that the RPP delivers a partial solution to the moral design problem (Keeling 2018). In this paper, I argue that an alternative solution to Keeling’s Restricted Pareto Principle should be found, because the principle is too limited in terms of the variety of cases it applies to, because it is ethically unsupported, because it is hardly extendable despite its minimalistic nature, and because its conditions are insufficient even within the cases where it applies. I propose that the traditional ethical approach should not be rejected because of the disagreements between the standard ethical theories. To cope with their incompatibility, we should instead make use of a bargaining game theoretic approach to find the set of principles all ethical theories disagree with *the least*. First, I will introduce the legal-philosophical approach and the doctrine of necessity as discussed in Santoni de Sio’s paper. I will then discuss Keeling’s *Restricted Pareto Principle* and the problems I see with it, before presenting my take on a game theoretic approach for applied ethics, which promises a pragmatic consensus between the disagreeing ethical theories.

2. The legal-philosophical approach

Various distinct approaches to solving the moral design problem have been proposed. Three of them have turned out to be particularly popular: First, it has been proposed that ethical programming of ADVs in such scenarios is an issue of public acceptance of new technologies, and that therefore, the right programming of ADVs for such scenarios is to be found by experimental ethics. More concretely, laypeople’s feelings, opinions, and intuitions on the moral status of different programming options should be evaluated in empirical studies and implemented in the ADVs (e.g. Bonnefon et al. 2016). Second, some philosophers have argued for a traditional ethical approach (e.g. Gerdes and Thornton, 2017). They argue that the views and arguments of contemporary ethicists should be determining for the programming of ADVs (Gerdes and Thornton, 2016). Third and most recently, a legal-philosophical approach, i.e. extracting judgments for the solution of the moral design problem from past verdicts of judges, has been suggested (Santoni de Sio, 2017; Keeling, 2018).

The former two have been accused of being problematic for fairly obvious reasons: Both philosophers and laypeople disagree about what is morally prohibited, permissible or obligatory in such scenarios (Santoni de Sio 2017, 412). The scenarios strongly resemble a classic: the trolley case (cf. Foot 1967; Thomson 1985). In both the trolley case and scenarios which give rise to the moral design problem, a decision about the allocation of harm among different objects or persons has to be made. Philosophers as well as laypeople have—even today—not found common ground on what ought to be done in the trolley case. In the relevant scenarios for the moral design problem, the same disagreements arise, and in light of this, it seems that neither the judgments of laypeople nor those of philosophers can offer an unambiguous solution for the moral design problem (Santoni de Sio 2017, 412). Because of this, Santoni de

Sio has suggested a legal-philosophical approach (2017, 412). He argues that we ought to look at how the law has previously regulated similarly difficult choices in other emergency scenarios. He argues that the legal doctrine of necessity would be particularly auspicious, as it regulates cases in which human agents have to intentionally cause damages to life and property in order to avoid other damages and losses, when avoiding all harm is deemed impossible (Santoni de Sio 2017, 412).

3. The doctrine of necessity

Santoni de Sio suggests a legal-philosophical approach. He holds that we ought to examine how the law has previously regulated similarly difficult choices in other scenarios and proposes to consider the legal doctrine of necessity (2017, 412–413.). According to Santoni de Sio, the doctrine of necessity “concerns the legitimate exceptions to the compliance to a legal prohibitive norm. Behaviours that are prima facie prohibited by criminal law may be permitted under exceptional circumstances” (Santoni de Sio 2017, 413). To find out when the defense of necessity is granted, he presents two cases:

Dudley & Stephens: The seamen Tom Dudley and Edwin Stephens were shipwrecked along with two other men. Thinking that they would all die very soon from starvation, Dudley and Stephens killed the cabin boy and ate him. They were then rescued and returned home, where they were convicted of murder. (R v Dudley and Stephens 1884)

Gracie & Rosie: Gracie and Rosie were conjoined twins. The doctors could either perform the operation and separate the twins, killing Rosie in the process, and giving Gracie a chance of survival of 90%, or not perform the operation, which would most likely lead to the death of both Gracie and Rosie. Brooke LJ permitted the defence of necessity as a legal justification for the doctors’ performing the operation and killing Rosie in the process. (Re A 2001)

Santoni de Sio then argues that recently, less restrictive¹ approaches to the interpretation of the doctrine of necessity occurred in English law—hence the different judgments in the two cases.

Geoff Keeling holds against Santoni de Sio and argues

¹ Note that *Gracie & Rosie* was much more recent than *Dudley & Stephens*.

that the interpretation of the doctrine of necessity has not changed. He proposes a different interpretation, which he calls the *Restricted Pareto Principle*.

4. The Restricted Pareto Principle

Keeling argues that the legal-philosophical approach could bring philosophers to a partial consensus on the moral design problem (Keeling 2018, 420), and that what he calls the “Restricted Pareto Principle” (RPP) as a reading of the doctrine of necessity would be that partial consensus (Keeling 2018, 422):

Restricted Pareto Principle (RPP): In situations where all following conditions are met, ceteris paribus, bringing about the Pareto efficient allocation of harm is justified (Keeling 2018, 420):

- i) Harm to at least one person is unavoidable.
- ii) A choice about how to allocate harm between different persons is required.
- iii) A unique Pareto efficient allocation of harm across different persons exists.

Keeling clarifies that “an allocation of harm is Pareto efficient if, and only if, there exists no alternative allocation of harm in which all affected parties are at least as well-off, and some affected party is strictly better off” (Keeling 2018, 421). Moreover, Keeling holds that there is a *unique* Pareto efficient allocation of harm when there is exactly *one* Pareto efficient allocation of harm, and that RPP is a sufficient condition on justified harm in the cases it applies to (Keeling 2018, 421).

Conditions i) and ii) are applicable to both the case of *Dudley & Stephens* and the case of *Gracie & Rosie*. In *Dudley & Stephens*, harm to at least one person was considered unavoidable by the judge: He accepted the fact that Dudley and Stephens would have died, had they not eaten the cabin boy. Therefore, a choice between killing one to save three and letting two die had to be made—i.e. a choice on how to allocate the harm between different persons. Similarly, harm to at least one person was unavoidable in the case of the conjoined twins *Gracie & Rosie*: If separated, Gracie would survive, and Rosie would die. If not separated, both would die. Therefore, a choice had to be made on how to allocate harm between Gracie and Rosie.

However, the two cases differ in their fulfilment of condition iii): In *Dudley & Stephens*, there was no *unique* Pareto efficient allocation of harm, but *four*: Instead of killing the cabin boy, the sailors could have killed Dudley,

or Stephens, or the fourth, unnamed sailor (Keeling 2018, 422). In contrast, separating Gracie from Rosie and—in doing so—killing Rosie, was the unique Pareto efficient allocation of harm, since Rosie would have died, even if the twins had not been separated (Re A 2001). With this in mind, RPP seems to account for the judgments in the two legal cases and appears to be a strong reading of the doctrine of necessity.

5. The Restricted Pareto Principle in light of ethical theories

With RPP, it seems that Keeling has put forward a plausible reading of the doctrine of necessity. He then argues that the by RPP provided account for the conditions under which necessity provides a justification for harm can be used to formulate a partial answer to the moral design problem in the following way:

*Restricted Pareto Principle** (RPP*): In collisions, where all following conditions are met, *ceteris paribus*, programming a driverless car to bring about the Pareto efficient allocation of harm is justified (Keeling 2018, 422):

- i) Harm to at least one person is unavoidable.
- ii) A choice about how to allocate harm between different persons is required.
- iii) A unique Pareto efficient allocation of harm across different persons exists.

In the fifth section of his paper, Keeling argues that proponents of utilitarianism², contractualism³, and deontology, all have reason to accept RPP* where it applies.

In particular, Keeling argues that deontologists have reason to accept RPP* in scenarios where it applies, because it accounts for the difference between killing and letting die, following Frances Kamm's definition of killing. She holds that "[i]n killing, we cause someone to lose a life that he would have had independently of our efforts at that time" (Kamm 2007, 18). According to Keeling, killing (according to the aforementioned definition) cannot occur in any cases where RPP* applies (Keeling 2018, 425).⁴

2 If an outcome Pareto-dominates another, the dominating outcome has strictly greater utility than the dominated outcome (Coleman 1980, 515). In scenarios where RPP* applies, RPP* only justifies bringing about an outcome if that outcome is utility maximising (Keeling 2018, 423).

3 Keeling discusses Scanlon's account of contractualism (Scanlon 1998). In scenarios where RPP* applies, no individual is rationally able to strictly prefer any alternative outcome than the one RPP* suggests, because iii) makes sure that the Pareto-efficient outcome is *unique* (Keeling 2018, 423).

4 Whenever RPP* applies there must be a *unique* Pareto efficient

6. Why Keeling's Restricted Pareto Principle* fails

There are several problems with Keeling's solution to the moral design problem. I will offer four accounts to reject RPP*. I argue that RPP* is too limited in terms of what variety of cases it applies to, that it is ethically unsupported, that it is hardly extendable despite its minimalistic nature, and that RPP*'s conditions are insufficient even in the cases where it applies.

6.1 The limitation of the RPP*

The RPP* only addresses a very limited number of problems that would occur in traffic. Of course, no single philosophical paper is under the obligation to solve all problems at once, nor to only solve fundamental problems. Nevertheless, it occurs to me that RPP* applies to an extraordinarily small number of cases, and that the ones where it does apply are neither very controversial nor likely to occur in traffic. A good abductive quality for a view is that it is *decisive* or *almost decisive*, i.e. that the view delivers a verdict in all or most of the cases with which we are concerned. However, RPP* is not even *almost decisive*.

RPP* does not deliver a verdict on Trolley-style cases, because in such cases, no *unique* Pareto efficient allocation of harm exists. In the classic Trolley case "Bystander at the switch" (Thomson 1985, 1397), the bystander can either let a trolley run over and kill 5 people or divert the trolley to the sidetrack by throwing a switch, thus killing the one person on the sidetrack. In such a case, there is no *unique* Pareto efficient allocation of harm. If the switch is not thrown (one person is left to live and five left to die), then no change could lead to the living of the five, without sacrificing the life of the one. Hence, not throwing the switch is Pareto efficient. If the switch is thrown (one person is sacrificed for the lives of the five), then no change could lead to the living of the one, without sacrificing the lives of the five. Hence, throwing the switch is also Pareto efficient, and hence, there are *two* Pareto efficient allocations of harm.

Provided that a scenario in traffic is a variation of the trolley case, RPP* will therefore not put forward any verdict. For example, in any scenario where an ADV can swerve off to save the life of at least one, but in doing so, would end the life of at least one other, RPP* does not apply. This leaves us with a very limited number of cases, like the one provided in Keeling's paper.

option, which means that all the persons that die (or get harmed) in the option RPP* suggests would have suffered at least equally bad harm. In light of Kamm's definition of "killing", this means that RPP* never suggests options such that people are killed, but only options such that people are left to die.

However, for all such cases I can think of, it appears to me that the decision is either trivial, or the programming of the car should clearly entail avoiding such cases to happen in the first place, or both. In most realistic cases, the decision appears to be trivial. If, for example, a car's brakes fail, and it can either hit the wall straight ahead, thereby killing only the driver, or hit the wall on the left, thereby hitting a pedestrian and killing both the pedestrian and the driver, it appears to me that programming the car so that it does not run over the pedestrian is uncontested as the morally preferable thing to do.

6.2 Ethical grounding of the RPP*

The RPP (and therefore the RPP*) is derived from analysis of past judgements. However, those judgements could very well have been faulty all along. In what regard they could be faulty is, of course, dependent on the chosen approach.

Taking an experimental ethics approach, one might argue that laypeople may disagree with the judgements. I have neither conducted any experiments on laypeople's opinions on those cases, nor can I cite any authors who have. The relevance of this objection simply lies in the fact that *we do not know yet* whether those judgements are—from the perspective of the experimental ethicists—faulty, or not.

Taking the traditional philosophical approach, one might object that RPP* cannot be supported by ethical deliberations. The immediate response would certainly be that RPP* seems to be supported by all three most dominant ethical theories. I see two problems with this defence of the principle: First, deontology seems to be quite strongly misrepresented in the paper, i.e. Kamm's definition of killing is very controversial and certainly an unusual deontological one. To give an illustration: According to Kamm's definition, it would not be considered killing if someone was to drive up to a car, which is rapidly rolling towards a wall and which does not have working brakes, and then to shoot the passenger, thereby causing his death.⁵ A consequence of this account of "killing" is that it would result in the strange conception that many shootings were *coincidentally* not killings, merely because the victims would have died anyway in that moment, and even though all gunmen had no idea about the upcoming "natural" deaths of the victims.

Most importantly, however, I argue that there are cases where RPP* applies, but where the outcome is rejected by all three theories (see 4.3). This raises strong doubts as to whether RPP* can be supported by tradi-

tional ethical deliberations—contrary to Keeling's claims.

The third approach one might choose is the legal-philosophical approach. This approach seems to be more relevant for evaluating the validity of the judgments in *Gracie & Rosie* and *Dudley & Stephens* than the experimental and traditional approach: Both Keeling and Santoni de Sio have decided to make use of the legal-philosophical approach *because* the previous two approaches did not offer decisive solutions. It is therefore to some extent obvious that the legal-philosophical approach leads to solutions which do not seem supportable by traditional ethics.

It seems to me that the legal-philosophical approach *could* offer answers to the question, but it has not done so *yet*. Like in the experimental approach mentioned above, the validity of the judgments in the two cases would have to be investigated by other judges. It could be that one or both judgments were far off what other judges would have adjudicated, but it could also be the case that all or most judges would have come to the same verdict. One way or the other, as long as this is not empirically investigated, the legal-philosophical approach does neither provide additional support nor any additional trouble for the validity of the judgments used by Keeling, and thus for RPP*.

6.3 RPP*'s insufficiency

Although RPP* appears—not least because of its minimalistic nature—to be very strong at first, I believe that it is not even sufficient in all the cases where it applies. In some instances, it would be obscure to grant someone the defence of necessity. Consider a variation of *Dudley & Stephens*:

Atlas & Eddie: Two sailors, Atlas and Eddie, are cast away on a raft. They do not have access to any food or water and are sitting face to face with each other, wondering when one of them will make the first move to kill the other. Both know that they will starve to death, if they do not eat anything. Suddenly, Atlas, the larger and stronger sailor, stands up, attacks, and kills Eddie.

Surely, as in *Dudley & Stephens*, the judge would not grant Atlas *defence of necessity*. This coincides with Keeling's reading of the defence of necessity: In this situation, RPP does not apply, because there are multiple Pareto efficient allocations of harm. So far, so good. With this in mind, consider the following addition to *Atlas & Eddie*:

⁵ Assuming that, without the intervention, the car would have hit the wall what would have caused the passenger's death.

What Atlas did not know was that Eddie was a quadriplegic. He could never have killed or eaten Atlas.

Atlas' lack of knowledge about Eddie's inability to kill him should not change the fact that the judge does not grant Atlas the *defence of necessity*, for if he would, then the consequences for Atlas' actions would be (certainly from Atlas' perspective) random. It seems to me that granting Atlas a defence of necessity because Eddie's condition (which Atlas did not know) would be accompanied by accepting and introducing to law an entire new set of moral luck (Williams 1981).

The problem for RPP as a reading of the defence of necessity is that if we consider Eddie's inability to kill Atlas, a *unique* Pareto efficient option would have indeed existed in *Atlas & Eddie*. Because Eddie could have neither killed nor eaten Atlas, Eddie would have died either way, but Atlas would have died only if he had not killed and eaten Eddie. Hence, RPP applies, but the defence of necessity should not be granted.

Now, one might object that Atlas should have asked Eddie about his condition in order to avoid moral luck, but that just seems to come with accepting that knowledge about the condition is a necessary, additional condition for the defence of necessity to apply—and therefore that *as it is*, RPP is insufficient.

6.4 RPP* – what now?

A fourth issue I see is that there is no obvious way to extend the remit of RPP* without losing its agreeableness with the three moral theories. This has also been pointed out by Keeling in one of his later works (Keeling 2020, 89). In section 5.1, we have already established that RPP* is applicable only to a very limited set of cases. For the vast majority of the relevant scenarios, solutions still have to be found. If RPP* were a starting point for the legal-philosophical approach, which could then be extended and could provide solutions to more scenarios, then RPP*'s limitation would not be too alarming. But howsoever we fiddle around with RPP*, the justifications for RPP* seem to decrease with every potential change. If, for example, we weaken RPP* to “an act is justified if Pareto efficient”, we lose the utilitarian and the deontologist: In the Trolley case, both *throwing the switch* and *not throwing the switch* are Pareto efficient, while the preferences of the two groups of ethicists diverge. This divisiveness in judgment translates to many relevant dilemmas in ADV programming, like swerving off the road to save two bikers, but killing one pedestrian instead.

In response, one might bring up that those delibera-

tions merely demonstrate the limits of the compatibility of deontology, utilitarianism, and contractualism: It appears that RPP* is simply the best principle we can come up with that is still compatible with all three theories. Keeling himself has later suggested that because of this, we ought to reject the legal-philosophical approach.⁶ While I do agree that the RPP* might not be extendable without losing proponents of one of those three ethical theories, I must stress that the reason Keeling turned his head to the legal-philosophical approach in the first place was that there is no grand unity to be found in the three ethical theories. The legal-philosophical approach might therefore still be highly fruitful in finding principles that are efficient, applicable, and that each of the ethical theories *disagree with the least*.⁷ But to find those principles, I instead propose a bargaining game theoretic approach.

7. Gradual disagreement – towards a bargaining approach

Keeling and Santoni de Sio have introduced the legal-philosophical approach to find solutions on how ADVs should be programmed to allocate harm in scenarios where harm is unavoidable and a decision on how to allocate it is required. With Keeling's RPP*, it appears that the limit of the compatibility of utilitarianism, contractualism, and deontology is found. I do not think that this is sufficient reason to discard the legal-philosophical approach, and I invite the legal philosophers, who are far more competent than myself on this topic, to investigate it further. However, neither do I believe that the diverging opinions of the traditional ethical approach is sufficient reason to discard it.

One thing that strikes me as crucial is that agreement and disagreement, even in ethical theories, is not, and never has been, a binary thing. Rather, it is gradual. Recent literature has repeatedly accepted or declined the traditional approach (and occasionally even the legal-philosophical approach) on the grounds that there are disagreements between proponents of different ethical theories (e.g. Bonnefon et al. 2016; Gerdes and Thornton 2016; Keeling 2018). Philosophers have known for a while now that there are problems in unifying the prominent ethical theories in a way they can all fully

⁶ He argues that “there is [...] no obvious means to relax the principle whilst remaining agreeable from the point of view of the three moral theories that I considered” (Keeling 2020, 90).

⁷ If policy makers wanted to go for maximum argumentative consistency, then choosing one consistent ethical theory would do the trick. But in that case, two of the three prominent ethical theories (maybe even *all* three) would run the high risk of being ignored altogether.

agree. The question now is whether we—because of this—discard all ethical theories for policy making, as has been proposed. The reason I hold the gradual nature of agreement to be so crucial is that I believe it to be the key in unifying ethical theories for policy making. Consider a life and death scenario involving an ADV and three persons to distribute the harm:

An ADV is driving down a narrow road. To its right is a kerb with one person taking a walk when, suddenly, two bikers join the street from the left out of their garage. The program of the car immediately assesses that—given the speed of the car and the amount of road left to break—the impact would most likely result in death for both bikers. However, the car could still swerve to the right and hit the stroller, who has an equally high probability to die from that accident. Either way, the passenger of the ADV bears no risk of injury, and no participants in this scenario have obviously violated any traffic laws.⁸

At the bottom of it, we are dealing with a *one versus two lives* (1v2) scenario, where RPP* does not apply, because there are two Pareto efficient solutions. The utilitarian disagrees with the option of hitting the two bikers and agrees with the option of hitting the pedestrian. The opposite is the case for the deontologist, who disagrees with the killing of the pedestrian, but agrees with driving into the bikers. Depending on the specific contractualist, each of the options can be valid. But what happens if we raise the numbers?

If we are looking at a *one versus ten lives* (1v10) scenario, the judgments do not change, but the strength of disagreement does. In the 1v10 scenario, the utilitarian disagrees even more with killing the larger group, than he does in the 1v2 scenario. In fact, the utilitarian disagrees exactly five times as much: He assesses five 1v2 cases, where the two always die, as equally bad as one 1v10 case, where the ten die. This is, however, not true for the deontologist: The deontologist judges the hitting of the larger group in the 1v2 case as equally bad or only

slightly better than in the 1v10 case. As the utilitarian cares much more about the varying number of deaths than the deontologist and the contractualist, it appears to me that proponents of the three theories could easily “trade off” claims on judgments of certain cases.

So, what if utilitarians were to grant the deontologists full right of judgment on all the 1v2 cases? For the utilitarians, this does not seem to be a terrible outcome (although certainly not the preferred one), as the difference of number of lives is only *one* each time. And what if, in exchange, the deontologist grants the utilitarians the full right of judgment on all the, let us say, *1v5-or-higher* cases? Would it not be preferable for both, if they got to make judgments on all the cases that are most important to them, in exchange for the judgment on the cases *other* ethicists care most about? Maybe, the alternative is risking not being able to have a say, *at all*. The goal here is to find a set of judgments which is *as acceptable as possible* to proponents of *all* relevant theories; among the available options, this should be the preferred one for each theory, since fully implementing either *one single theory* or *all theories simultaneously* is not an option.

8. A pragmatic solution

Given the considerations outlined in section 7, I therefore suggest a very pragmatic line of action:

- A. Identify the most relevant scenarios, in which ADVs must allocate harm.
- B. Evaluate how proponents of the relevant theories would decide in those scenarios.
- C. Evaluate how much moral weight the relevant theories attribute to those scenarios.
- D. Subject the theories (with their respective evaluations and judgments) to bargaining game theory.

Of course, the outlined steps demand further elucidations. In A, the word “relevant” might raise some questions. There are two main properties which make a scenario relevant in the pertinent sense. First, there is the *frequency* with which a certain scenario occurs.⁹ Second,

⁸ The obvious misconduct in traffic severely complicates such thought experiments. One might believe, for example, that all and only the persons violating traffic laws should bear all the risk and harm. This would create multiple problems, like for example: What amount of consideration for traffic law violations should be given to children in such situations? Or: To what extent will ADVs be able to immediately judge who has and has not violated traffic laws in a given scenario? Respective investigations will have to be made, but those are beyond the scope of this paper. However, what I argue should be compatible with anything that might come up in such a debate, and maybe even be a tool for finding relevant solutions.

⁹ Unfortunately, the *introduction* of ADVs will have an impact on the frequency. But for the time being, it appears that our best bet is looking at current frequencies of scenarios, and maybe (although with this, I would be hesitant) slightly manipulate the numbers according to what kind of cases will certainly stop oc-

there is the *moral weight* different ethical theories attribute to certain scenarios: There are scenarios which may be extremely infrequent but are of crucial importance for one of the relevant theories.¹⁰

Steps B and C seem to be clearer, although there are different ways the judgments could be evaluated. In my estimation, the best tool for those steps is experimental philosophy—although for this particular issue, done on experts instead of lay people. Of course, there is the possibility of a considerable bias within the field of ethics, which might distort the validity of the results. However, I do not believe that suspicion of a bias is decisive; there are plenty of methods to counter that. Conclusively, I suggest that the evaluation of B and C take place simultaneously and with the same persons.

For this paper, I will not go into technical details on how the game of step D could be constructed, although conceptually, it is quite straightforward: The aim is to find the set of judgments all relevant theories disagree with the least, assuming that “not making a deal” is not an option. The set of judgments is derived from step A, and proponents of each theory cede judgments on cases to proponents of another theory, so that the agreement of proponents of each theory, represented by the expectancy value, is maximized. The aim is to find the Nash-equilibrium (Nash 1950) for the judgments on the programming of ADVs.

I am fully aware that such methods are somewhat atypical in ethics, though I am not exactly sure why. Certainly, in this case, I believe that bargaining is necessary. For as long as the great moral theories stand in explicit non-compoundable contradiction, policy makers, lawyers, and even governments are in the right to say that they cannot gain anything applicable from the traditional ethical approach for the programming of ADVs. In this case, the ability to influence policy making should trump the ability to stay fully consistent, for the latter results in total deprivation of influence by ethicists.

9. Conclusion

Santoni de Sio and Keeling tried to extract principles that are acceptable to proponents of the three main ethical theories via the legal-philosophical approach. Keeling suggests the Restricted Pareto Principle*, which—as we have established—is problematic for several reasons. RPP* applies only to a very limited number of cases, is

curing, and start occurring, once ADVs have been introduced.

¹⁰ It might be, for example, extremely unlikely that an ADV is driving towards a group of 50 people, which its sensors can only detect too late. Nonetheless, such cases should be included: Because of the moral weight they have for certain moral theories, they play a crucial role in the eventual negotiation.

ethically barely supported, is hardly extendable despite its minimalistic nature, and its conditions are insufficient even within some cases where it applies. I argued that we have to recognise that there is simply no broadly applicable principle proponents of all ethical theories do agree on. Therefore, I have suggested that we try to find a set of principles everyone disagrees with the least. To find these principles, I proposed we make use of experimental philosophy applied on traditional ethics and subject the results to bargaining game theory. This way, traditional ethicists have the chance to propose a set of principles to governments and policy makers, which is strictly preferable for all ethical theories over being a realm of discussion from which policy makers cannot draw anything because of its incapability of unification.

References

- Bonnefon, Jean-François, Azim Shariff, and Iyad Rahwan. 2016. “The social dilemma of autonomous vehicles.” *Science* 352 (6293): 1573–1576.
- Coleman, Jules. 1980. “Efficiency, utility, and wealth maximization.” *Hofstra Law Review* 8 (3): 509–551.
- Foot, Philippa. 1967. “The problem of abortion and the doctrine of double effect.” *Oxford Review* 5: 5–15.
- Gerdes, Christian and Sarah Thornton. 2016. “Implementable ethics for autonomous vehicles.” In *Autonomous Driving: Technical, Legal and Social Aspects*, edited by Markus Maurer, J. Christian Gerdes, Barbara Lenz, and Hermann Winner, 87–102. Berlin and Heidelberg: Springer.
- Kamm, Frances. 2007. *Intricate ethics: rights, responsibilities, and permissible harm*. Oxford: Oxford University Press.
- Keeling, Geoff. 2018. “Legal Necessity, Pareto Efficiency & Justified Killing in Autonomous Vehicle Collisions.” *Ethical Theory and Moral Practice* 21: 413–427.
- ———. 2020. *The Ethics of Automated Vehicles*. Dissertation, University of Bristol.
- Lin, Patrick. 2015. “Why ethics matters for autonomous cars.” In *Autonomes Fahren*, edited by Markus Maurer, J. Christian Gerdes, Barbara Lenz, and Hermann Winner, 69–85. Berlin and Heidelberg: Springer.
- Nash, John Forbes. 1950. *Non-cooperative games*. Dissertation, Princeton University.
- R v Dudley and Stephens. 1884. QBD 273.
- Re A (Conjoined twins). 2001. 2 WLR 480 (CA).
- Santoni de Sio, Filippo. 2017. “Killing by autonomous vehicles and the legal doctrine of necessity.” *Ethical Theory and Moral Practice* 20 (2): 411–429.
- Scanlon, Thomas Michael. 1998. *What we owe to each other*. Cambridge: Harvard University Press.
- Thomson, Judith Jarvis. 1985. “The trolley problem.” *The Yale Law Journal* 94 (6): 1395–1415.
- Williams, Bernard. 1981. *Moral Luck*. Cambridge: Cambridge University Press.

Patrick Meinrad Müller (23) hat Philosophie in Bern studiert. Er interessiert sich besonders für die Lösung philosophischer Probleme durch Erkenntnisse in den Naturwissenschaften und beschäftigt sich zurzeit vor allem mit dem Zusammenhang zwischen Ontologie und Wahrnehmung aus neurowissenschaftlicher Perspektive.

Wie entsteht wissenschaftlich fundierte Politik?

Dieser Beitrag erschien erstmals im August 2021 in: VSH-Bulletin (47) 2: 15–22. Wir danken der Vereinigung der Schweizerischen Hochschuldozierenden für die Erlaubnis, ihn im meta(φ) erneut abdrucken zu dürfen.¹

1. Einleitung

Die COVID-19-Pandemie hat einmal mehr vor Augen geführt, wie wichtig wissenschaftliche Expertise für die Politik ist. Über das neue Corona-Virus war zunächst nur sehr wenig bekannt, und so tappte die Politik im Frühjahr 2020 im Halbdunklen, als sie versuchte, der Pandemie Einhalt zu gebieten. Aber schon bald wurden wissenschaftliche Studien zu den Verbreitungswegen des Virus publiziert; Virologinnen und Epidemiologen erteilten Ratschläge an die Politik; und es wurden wissenschaftliche Taskforces gebildet. Dabei ging es immer darum, die politischen Reaktionen auf das Virus auf eine wissenschaftliche Grundlage zu stellen.

Doch ist es in der Schweiz und anderswo gelungen, die vorhandene wissenschaftliche Expertise angemessen in die Politik einfließen zu lassen? Diese Frage wird recht unterschiedlich beantwortet. Einige kritische Stimmen machten geltend, Forschende hätten in der Pandemie zu viel Einfluss gewonnen. „Die Politik hat abgedankt, Virologen regieren die Welt“, schrieb etwa Reinhard K. Sprenger Ende März 2020.² Andere hingegen beklagten, die Politik habe wissenschaftliche Ergebnisse nicht ausreichend berücksichtigt und sich über wissenschaftliche Empfehlungen hinweggesetzt. In der Schweiz kritisierten insbesondere einige Mitglieder der Taskforce wichtige Entscheidungen des Bundesrats und beschwerten sich, die Politik habe die gesammelte wissenschaftliche Expertise nicht ernst genug genommen.³

Hat die Forschung nun zu viel oder zu wenig Einfluss auf die politischen Entscheidungen gehabt? Oder wurde ihre Expertise wenigstens aufs Ganze gesehen richtig berücksichtigt? Diese Fragen lassen sich nur beantworten, wenn zuvor grundlegender geklärt wird, wie wissenschaftliche Expertise die Politik beeinflussen *sollte*. An welchem *Ideal* können wir uns orientieren, wenn es darum geht, politische Entscheidungen auf wissenschaftlichem Wissen zu fundieren?

Um diese Frage soll es in diesem Artikel gehen. Ich möchte ein bestimmtes Ideal für das Zusammenspiel von Politik und Wissenschaft diskutieren. Genauer geht es dabei bloss um die Naturwissenschaften und Teile der Sozialwissenschaften. Andere Wissenschaften muss ich aus sachlichen Gründen, die noch deutlich werden, ausklammern. Das soll natürlich nicht heissen, dass die anderen Wissenschaften zweitrangig wären. Das Ideal, das ich untersuche, ist selbstverständlich nicht nur für die Pandemie relevant, sondern auch für andere Herausforderungen der Politik, etwa den Klimawandel. Dabei geht es nur um die Frage, wie wissenschaftliche Expertise in die Politik einfließen soll. Andere Fragen, welche das Verhältnis zwischen Politik und Wissenschaft betreffen, etwa wie wissenschaftliche Forschung durch den Staat gefördert werden sollte, klammere ich aus.

2. Politik und Wissenschaft: Eine funktionale Differenzierung

In Platons Idealstaat ist alles ganz einfach: Dort herrschen die Philosophen. Denn nur sie verfügen über die einschlägige Expertise. Für Platon besteht diese letztlich im Wissen darüber, was gut ist – in seiner Terminologie kennen die Philosophen die Idee des Guten. Politik wird also von den Experten betrieben.

1 Die Links wurden letztmalig im Juni 2021 geprüft.

2 <https://www.nzz.ch/feuilleton/coronavirus-virologen-regieren-die-politik-hat-abgedankt-ld.1549096>.

3 Siehe etwa <https://zofingertagblatt.ch/?id=293341&Wegen%20>

[%C2%ABpolitischem%20Korsett%C2%BB%3A%20Wissenschaftler%20tritt%20nach%20Lockerungsentscheid%20des%20Bundesrats%20aus%20Taskforce%20aus.](#)

Platons Ideal hat bekanntlich seine Attraktivität verloren. Eine besonders scharfe Kritik des Platonischen Idealstaates findet sich in Karl Poppers Klassiker „Die offene Gesellschaft und ihre Feinde“. Dort sucht Karl Popper nach den gedanklichen Grundlagen für die totalitären Systeme in der ersten Hälfte des 20. Jahrhunderts. Diese findet er auch bei Platon. Popper versucht zu zeigen, wie eine „utopische Sozialtechnik“ den griechischen Philosophen dazu führt, einen Idealstaat zu entwickeln, der totalitäre Züge annimmt. Dort dürfen die Herrschenden sogar das Volk belügen.⁴

Nicht jede Expertokratie muss totalitär werden. Aber aus der Sicht Poppers ist die Herrschaft durch Expertinnen und Experten grundsätzlich problematisch. Denn sie unterläuft eine fundamentale Unterscheidung zwischen Tatsachen auf der einen Seite und Zielen oder Werten auf der anderen. Die Unterscheidung lässt sich am besten auf der Basis von Urteilen erklären, die Menschen im Alltag fällen. Einige dieser Urteile versuchen die Welt so zu beschreiben, wie sie ist, etwa wenn jemand behauptet, Kinder seien durch das Virus weniger gefährdet als Menschen anderer Altersgruppen. Im besten Fall sind die Urteile wahr und entsprechen dann den Tatsachen. Werturteile hingegen (hier in einem sehr weiten Sinn verstanden) bewerten oder normieren Gegenstände und Handlungen, zum Beispiel, wenn jemand sagt, bestimmte Massnahmen seien geboten. Letztlich legen Werturteile Ziele nahe oder gar fest. Das Problem mit der Expertokratie ist dann, dass sie nicht bloss Tatsachen, sondern auch Werte und Ziele zum Gegenstand der Expertise erklärt. Damit wird den Menschen die Möglichkeit genommen, selbst über ihre Ziele zu entscheiden.

Die Dichotomie zwischen Tatsachen und Werten bildet eine grundlegende Voraussetzung für Poppers Kritik an Platon und das Ideal, das hier diskutiert werden soll. Die Dichotomie ist oft beschworen worden, etwa von David Hume⁵, wird aber gelegentlich in Abrede gestellt. So behauptet der Philosoph Hilary Putnam, Werte und Tatsachen seien unentwerrbar miteinander verschränkt.⁶ In der Tat trennen wir im Alltag nicht immer klar zwischen den beiden Arten von Urteilen. Wenn wir etwa eine Person als freundlich beurteilen, dann versuchen wir damit einerseits, gewisse Tatsachen zu beschreiben

(etwa, dass die Person im Gespräch auf andere zugeht). Andererseits stellen wir die Person damit aber auch als gut hin, wenigstens in einer bestimmten Hinsicht. Auch Begriffe wie jener der Krankheit verbinden Beschreibung und Wertung. Wenn wir aber wollen, dann können wir die beschreibenden und die wertenden Aspekte von Urteilen oft auseinanderdividieren. Um unser Ideal entwickeln zu können, müssen wir davon ausgehen, dass wir wenigstens einige reine Tatsachen denken können.

Die Dichotomie zwischen Tatsachen und Werten prägt nun auch ein einfaches Modell der rationalen Handlung.⁷ Demnach basiert eine rationale Handlung auf einem Ziel oder einem Wert auf der einen Seite und einer Überzeugung, wie das Ziel realisiert wird, auf der anderen. In diesem Sinne ist es rational, ein Lexikon zur Hand zu nehmen, wenn man wissen möchte, was das englische Wort „perpendicular“ heisst, und der Überzeugung ist, diese Information in einem Lexikon zu finden. Dabei ist die Überzeugung der Versuch, eine Tatsache zu erfassen. Idealerweise spiegelt die Überzeugung also eine Tatsache wider. Dann ist die ausgeführte Handlung erfolgreich. Ist die Überzeugung hingegen falsch, so verfehlt die Handlung in der Regel ihr Ziel, etwa wenn eine Person in einer Shakespeare-Gesamtausgabe nachsehen will, was das Wort „perpendicular“ bedeutet.

Die genannte Vorstellung von rationalem Handeln kann man nicht nur auf das Handeln individueller Personen beziehen, sondern auch auf jenes von Staaten und anderen politischen Akteuren, kurz auf politisches Handeln. Genau das tut das Ideal, um das es im Folgenden gehen soll. Es ordnet zusätzlich die unterschiedlichen Aspekte rationalen politischen Handelns unterschiedlichen gesellschaftlichen Subsystemen zu: Die Wissenschaft hat demnach die Aufgabe, richtige Überzeugungen über Tatsachen zu liefern, die beschreiben, wie Ziele erreicht werden. Die Politik hingegen soll die Ziele bestimmen, idealerweise auf der Basis einer öffentlichen Diskussion. Gleichzeitig ist es die Politik, welche die Entscheidung darüber trifft, wie die Ziele verfolgt werden. In der Praxis lassen sich die Bestimmungen der Ziele und die Entscheidung über konkreten Massnahmen oft gar nicht auseinanderhalten. Denn es ist eher eine Ausnahme, wenn die Politik zunächst Ziele (wie etwa Klimaziele) festlegt und erst später darüber entscheidet, wie sie umgesetzt werden. Öfter wird das genaue Ziel gleichzeitig mit den Massnahmen festgelegt.

Damit ist die Grundidee des Ideals, um das es gehen soll, deutlich: Wissenschaft und Politik erhalten ihre Aufgaben gemäss der Unterscheidung zwischen

4 Popper, K. R. (1945), *The Open Society and Its Enemies. Volume I: The Spell of Plato*, London: Routledge & Kegan Paul, hier zitiert nach der deutschen Übersetzung: ders. (1992), *Die offene Gesellschaft und ihre Feinde. Band I: Der Zauber Platons*, 7. Auflage Tübingen: J. C. B. Mohr (Paul Siebeck), S. 187.

5 Hume, D., *A Treatise of Human Nature*, III.1.i, etwa in der Ausgabe Hume, D. (1978), *A Treatise of Human Nature*. Edited by Selby-Bigge, L. A./Nidditch, P. H., 2. Auflage Oxford: Clarendon Press, S. 469.

6 Putnam, H. (2002), *The Collapse of the Fact/Value Dichotomy and Other Essays*, Cambridge (MA): Harvard University Press.

7 Etwa Davidson, D. (1963), „Actions, Reasons, and Causes“, *Journal for Philosophy* 60, 685–700.

Tatsachen und Werten. Allerdings ist das noch zu einfach gedacht. Das Ideal muss daher in drei Hinsichten qualifiziert werden.

Erstens erreichen erfolgreiche Handlungen nicht nur ihr Ziel, sondern haben auch Nebenfolgen. Kontaktbeschränkungen zu Zeiten von Corona führen nicht nur zu einer Abnahme der Ansteckungen, sondern auch zu psychischen Belastungen. Diese Nebenfolgen hat Popper hervorgehoben; er sah sogar eine wesentliche Aufgabe der Sozialwissenschaften darin, die Nebenfolgen politischer Massnahmen zu bestimmen.⁸ Die Nebenfolgen sind ein Problem, wenn sie unerwünscht sind. Das aber können sie nur sein, wenn andere Ziele oder Werte auf dem Spiel stehen. In der Tat hat ja ein Gemeinwesen (ebenso wie eine einzelne Person) viele Ziele und Werte, die in Konflikt miteinander geraten können. Ein adäquates Modell des Handelns muss das berücksichtigen. Daher muss unser Ideal wie folgt verfeinert werden: Die Abwägung zwischen konfligierenden Zielen und Werten wird der Politik zugeordnet. Das Wissen darüber zu gewinnen, welche Ziele sich realistischerweise gemeinsam verfolgen lassen, ist hingegen Aufgabe der Wissenschaften.

Zweitens lässt sich oft nicht genau bestimmen, wie Ziele erreicht und Werte realisiert werden können und welche Nebenfolgen eintreten. Ein halbwegs plausibles Modell des Handelns muss daher berücksichtigen, dass sich Handeln meist unter den Bedingungen von unvollständigem Wissen abspielt. Die Pandemie liefert dafür viele Beispiele. So war die Rolle von Kindern bei der Ausbreitung der Pandemie lange unklar. In der Entscheidungstheorie werden nun je nach Art und Umfang des vorliegenden Wissens idealtypisch drei Arten von Situationen unterschieden: In Situationen der Sicherheit sind alle relevanten Folgen der zu Gebote stehenden Handlungssituationen bekannt. In Entscheidungen mit Risiko ist hingegen manchmal nur bekannt, mit welcher Wahrscheinlichkeit eine Massnahme bestimmte Folgen zeitigt. In Situationen des Unwissens ist sogar nur noch bekannt, welche Konsequenzen möglich sind, wenn eine bestimmte Option gewählt wird.⁹

Wissenschaftliches Wissen kann nun nicht nur in der ersten Situation eine wichtige Rolle spielen. Vielmehr können die Wissenschaften auch in der zweiten Situation etwas beisteuern, nämlich die Wahrscheinlichkeiten. Die Wahrscheinlichkeit dafür, dass eine bestimmte Impfung erfolgreich ist, lässt sich etwa durch wissenschaftliche Forschung bestimmen.

In der dritten Situation wäre es Aufgabe der Wissenschaft herauszufinden, welche Folgen überhaupt möglich sind, wenn eine bestimmte Massnahme ergriffen wird.¹⁰ Forschung könnte etwa zeigen, dass bestimmte Lockerungen zu einer weiteren Infektionswelle führen könnten. In den Klimawissenschaften wurde herausgefunden, dass durch weitere Erwärmung sog. Kipppunkte überschritten werden können.

Insgesamt zeigt sich damit, dass Ergebnisse wissenschaftlicher Forschung auch dort einbezogen werden sollten, wo sie die Standards sicheren Wissens über die Folgen unterschreiten. Das untersuchte Ideal wird daher meist so verstanden, dass je nach Situationstyp die relevante Form von Wissen in die Entscheidung eingeht.

Eine dritte Komplikation ergibt sich aus dem Umstand, dass sich die Feststellung einer Tatsache oft in ein Werturteil übersetzen lässt. Dass ein exponentieller Anstieg von Infektionen unter bestimmten Bedingungen nur durch eine geringere Anzahl von Kontakten vermieden werden kann, ist eine Tatsache. Statt diese Tatsache direkt festzustellen wie eben geschehen, kann man auch sagen, wir sollten die Anzahl der Kontakte begrenzen, um einen exponentiellen Anstieg von Infektionen zu vermeiden. Damit wird zur Tatsache nichts hinzugefügt, dennoch haben wir es jetzt mit einem Werturteil zu tun. Dieses Werturteil ist allerdings hypothetisch, denn es steht unter der Bedingung einer Zielvorgabe. Ein unbedingtes Müssen ergibt sich daraus erst, wenn wir auch das Ziel verfolgen müssen. Das Ideal, das wir gerade entfalten, erlaubt daher, dass die Wissenschaften bedingte Werturteile fällen. Die wissenschaftliche Expertise kann deshalb in hypothetischen Werturteilen formuliert werden.

Insgesamt favorisiert unser Ideal also folgende Arbeitsaufteilung, um den Problemen einer Expertokratie zu entgehen: Entscheidungen über Ziele und ihre Gewichtung, aber auch ihre Umsetzung obliegen der Politik. Die Wissenschaft hat die Aufgabe, diese Entscheidungen zu informieren; das kann auch durch die Spezifikation von Wahrscheinlichkeiten oder blossen Möglichkeiten geschehen.

3. Die Folgen für die Wissenschaft: Wertfreiheit als Postulat

Was bedeutet das skizzierte Ideal nun für die Wissenschaften? Die wichtigste Folge ist sicher die Wertfreiheit der Wissenschaften. Darunter wird zunächst die Auffassung verstanden, dass jene wissenschaftlichen Resultate, die sich sprachlich formulieren lassen und damit in Zeit-

8 Popper, K. (1957), *The Poverty of Historicism*, London: Routledge, dt. Übersetzung in ders. (1965), *Das Elend des Historizismus*, Tübingen: J. C. B. Mohr.

9 Resnik, M. (1987), *Choices*, Minneapolis: University of Minnesota Press.

10 Resnik, D. B. (2003), "Is the Precautionary Principle Unscientific?", *Stud. Hist. Phil. Biol. & Biomed. Sci.* 34, 329–344.

schriften und Büchern publiziert werden können, wertfrei sind. Sie sind es insofern, als sie weder unbedingte Werturteile sind noch solche enthalten.

Ein prominenter Vertreter der Wertfreiheit ist Max Weber. Die Wertfreiheit im eben beschriebenen Sinn drückt er aus, wenn er schreibt, dass

„Tatsachenfeststellung, Feststellung mathematischer oder logischer Sachverhalte oder der inneren Struktur von Kulturgütern einerseits, und andererseits die Beantwortung der Frage nach dem Wert der Kultur und ihrer einzelnen Inhalte [...] ganz und gar *heterogene* Probleme sind.“¹¹

Oft wird die Wertfreiheit auch ins Praktische gewendet und tritt dann als Postulat auf. Weber fordert etwa in diesem Sinne, Politik (also das Vertreten politischer Ansichten) gehöre nicht in den Hörsaal. Genauer können wir zwischen zwei Arten von Forderungen an die Forschenden unterscheiden. Erstens sind diese gehalten, die eigene Forschung möglichst frei von Wertungen zu halten. Zweitens sind sie aufgerufen, in ihrer Rolle als Forschende in der Interaktion mit Politik und Öffentlichkeit keine unbedingten Werturteile zu vertreten.

Die erste Forderung der Wertfreiheit ist offensichtlich durch die Sorge motiviert, wissenschaftliche Ergebnisse könnten durch Wertungen beeinflusst werden und daher eben nicht mehr wertfrei sein. Es sind daher besondere Anstrengungen erforderlich, die den Einfluss von Wertungen auf den Forschungsprozess verhindern. Allerdings lassen sich viele Einflüsse von Wertungen auf die wissenschaftliche Forschung gar nicht sinnvoll ausschliessen.

Erstens werden Forschungsthemen und -fragen oft auf der Basis von Werten bestimmt. So entspricht es derzeit moralischen Werten, die eigene Forschungstätigkeit Corona und verwandten Themen zu widmen, weil damit ein Beitrag zur Bekämpfung der Pandemie und so letztlich zum Allgemeinwohl geleistet wird. Der Rekurs auf Werte ist hier legitim. Das Postulat einer wertfreien Wissenschaft sollte nicht ausschliessen, dass Wertungen die Wahl von Themen beeinflussen. Daher behaupten jene, welche die Wertfreiheit der Wissenschaften vertreten, oft, die Auswahl von Forschungsthemen und -fragen gehöre streng genommen nicht der wissenschaftlichen Forschung an und unterliege daher nicht dem Postulat der

Wertfreiheit, das für die Forschungstätigkeit selbst gelte.¹² Zweitens stützt sich auch die Wahl von Theorien und Modellen in der wissenschaftlichen Forschung sinnvollerweise auf Werte. So werden Theorien bevorzugt, wenn sie widerspruchsfrei und empirisch adäquat sind und gute Erklärungen liefern können. Die Widerspruchsfreiheit, die empirische Adäquatheit und die Erklärungskraft firmieren dabei als Werte oder Theorietugenden, d.h. als Eigenschaften, die Theorien gut machen. Solche Eigenschaften sollten offenbar in den Wissenschaften eine wichtige Rolle spielen. Das Postulat der wertfreien Wissenschaft ermöglicht das, indem es zwischen sog. epistemischen und anderen Werten unterscheidet. Epistemische Werte haben es mit den Zielen der Erkenntnisgewinnung zu tun und dürfen daher Entscheidungen in der wissenschaftlichen Forschung begründen. Die Wertfreiheit fordert dann nur, dass die Theoriewahl unabhängig von nicht-epistemischen, z.B. moralischen Werten bleibt. Eine Theorie darf nicht deshalb gewählt werden, weil sie das Selbstverständnis einer bestimmten Gruppe unterstützt. Vielmehr soll nur auf der Basis der vorhandenen Daten entschieden werden, ob eine Theorie akzeptiert wird.¹³

Drittens schleichen sich bei genauerer Betrachtung selbst bei der Überprüfung von Theorien und Hypothesen indirekt moralische Werte ein. Denn wie Richard Rudner argumentiert hat, lassen sich empirische Hypothesen nicht beweisen, sondern nur mehr oder weniger mit Belegen stützen. Das aber bedeutet, dass eine Schwelle definiert werden muss, welche die Belege überschreiten müssen, damit eine Hypothese akzeptiert wird. Wo diese Schwelle angesetzt wird, ist letztlich eine praktische Entscheidung, in die moralische Werte eingehen. So werden bei der Zulassung von Medikamenten sehr starke Belege für die Hypothese gefordert, dass die Nebenwirkungen gering sind. Der Grund liegt nur darin, dass so im Zweifel viel Schaden vermieden wird. Rudners Überlegungen zeigen zwar nicht, dass die Wahl von Theorien direkt von moralischen Werten beeinflusst werden darf. Aber ein indirekter Einfluss kann nicht ausgeschlossen werden.¹⁴

Insgesamt kann die Wertfreiheit der wissenschaftlichen Forschung also nur mit Einschränkungen gefordert werden. In der philosophischen Diskussion wurde aufgrund der Einschränkungen sogar oft argumentiert,

11 Weber, M. (1919), *Geistige Arbeit als Beruf. Vorträge vor dem Freistudentischen Bund. Erster Vortrag. Wissenschaft als Beruf*, München/Leipzig: Duncker und Humblot, S. 3–37, hier zitiert nach Weber, M. (1995), *Wissenschaft als Beruf*, Stuttgart: Reclam, S. 29–30.

12 Siehe etwa Rudner, R. (1953), “The Scientist Qua Scientist Makes Value Judgments”, *Philosophy of Science*, 20 (1):1–6, dt. Übersetzung in: Carrier, M. & Schurz, G. (Hrsg.) (2013), *Werte in den Wissenschaften. Neue Ansätze zum Werturteilsstreit*, Berlin: Suhrkamp, S. 108–117.

13 Siehe dazu etwa Lacey, H., (1999), *Is Science Value-Free? Values and Scientific Understanding*, London: Routledge.

14 Siehe dazu wieder Rudner (1953) und Douglas, H. (2000), “Inductive Risk and Values in Science”, *Philosophy of Science*, 67(4), 559–579.

das Postulat der Wertfreiheit sei obsolet.¹⁵ So macht die Standpunktepistemologie geltend, dass Menschen in ihren Erkenntnisbemühungen ihren Standpunkt und damit ihre Herkunft, Werte und Interessen nie völlig verleugnen können.¹⁶ Das ist gerade für einige Sozial- und viele Geisteswissenschaften ein Stück weit plausibel. Ohnehin entspricht die Vorstellung, dass diese Wissenschaften Tatsachenwissen generieren, nicht ihrem gängigen Selbstverständnis. Vielmehr ist ihr Ziel z.B., Interpretationen und Deutungen zu liefern, etwa von literarischen Texten oder Rechtscorpora. Gerade auch in der Ethik geht es fundamental darum, unbedingte Werturteile zu begründen. Daher ist das Ideal der wertfreien Wissenschaft für die genannten Wissenschaften wenig plausibel. Wir beschränken uns hier auf die Naturwissenschaften, für welche die Kritik der Standpunktepistemologie weniger einleuchtet. Für sie gilt, dass die genannten Einwände nicht den Kern der Wertfreiheit berühren, also der Vorstellung, dass wissenschaftliche Resultate keine Werturteile sind oder enthalten.

Aus dieser Auffassung wird nun oft eine zweite Art von Forderung abgeleitet. Diese ist für unsere Zwecke wichtiger, weil sie das Verhältnis der Wissenschaft zur Öffentlichkeit und Politik betrifft. Die Forderung lautet, dass Wissenschaftlerinnen und Wissenschaftler in dieser Rolle keine unbedingten Werturteile vertreten. Sie können dann zwar sagen, dass bestimmte Massnahmen notwendig sind, damit eine weitere Ausbreitung des Virus eingeschränkt wird. Sie dürfen aber nicht öffentlich oder im Dialog mit der Politik fordern, bestimmte Massnahmen sollten unbedingt eingesetzt werden.

Das Postulat der Wertfreiheit in diesem Sinne wird in der Praxis manchmal verletzt. So heisst es in einer Stellungnahme einer bekannten wissenschaftlichen Akademie aus dem Jahr 2020, es sei „aus wissenschaftlicher Sicht unbedingt notwendig, die weiterhin deutlich zu hohe Anzahl von Neuinfektionen durch einen harten Lockdown schnell und drastisch zu verringern“. Wissenschaftlich etablieren lässt sich aber allenfalls, dass ein Lockdown notwendig ist, um ein gewisses Ziel zu erreichen, etwa eine Überlastung der Krankenhäuser zu vermeiden.¹⁷ Die genannte Stellungnahme ist besonders problematisch, weil sie auf die negativen Nebenfolgen

eines Lockdowns gar nicht eingeht. Diese müssen aber bei der Bewertung eines Lockdowns mit einbezogen werden. Nur eine Abwägung zwischen den positiven und negativen Seiten eines Lockdowns kann die Frage beantworten, ob ein Lockdown notwendig ist. Aus der Sicht der Wertfreiheit ist es problematisch, dass die Stellungnahme der Akademie überhaupt nicht erkennen lässt, auf welcher Abwägung das unbedingte Plädoyer für einen Lockdown beruht.

Aber warum sollten Forschende in ihren Expertisen und Stellungnahmen auf unbedingte Werturteile verzichten? Wie genau ist die Wertfreiheit als Postulat hier begründet? Eine naheliegende Antwort lautet, dass Forschende ihre Kompetenzen überschreiten, wenn sie unbedingte Werturteile vertreten. Diese folgen nicht aus ihren wissenschaftlichen Ergebnissen, daher sollten sie auch nicht den Eindruck erwecken, unbedingte Wertungen wären wissenschaftlich.

Ausserdem kann das Vertreten unbedingter Werturteile das Vertrauen in die Wissenschaften unterminieren. Denn unbedingte Werturteile sind manchmal umstritten; im politischen Diskurs sind sie häufig konstitutiv für bestimmte politische Gruppierungen, die sich etwa durch ein Bekenntnis zur Freiheit oder zum Vorrang des Umweltschutzes identifizieren. Wenn Forschende unbedingte Werturteile vertreten, dann sind sie im politischen Diskurs nicht mehr neutral; sie können mit bestimmten politischen Richtungen assoziiert werden, ja als Partei wahrgenommen werden. Das kann dann dazu führen, dass auch wissenschaftliche Ergebnisse, die unabhängig von Wertungen bestehen, nicht mehr akzeptiert werden, weil unterstellt wird, sie seien Teil einer politischen Position. Das widerspricht aber dem Anspruch der Wissenschaften, Resultate zu erzielen, die von allen Menschen unabhängig von ihren Werten und politischen Einstellungen akzeptiert werden können. Wissenschaftliche Resultate, so heisst es oft, gelten objektiv; ihre Begründung muss im Prinzip von allen Menschen nachvollzogen werden können.

Es sprechen also gute Gründe dafür, dass Forschende in ihren Stellungnahmen auf unbedingte Werturteile verzichten. In der Praxis ist diese Forderung allerdings nur sehr schwierig umzusetzen.

Ein Hauptproblem liegt darin, dass Forschende auch andere Rollen einnehmen, etwa die von Bürgerinnen und Bürgern. In diesen Rollen kommen sie aber um unbedingte Werturteile nicht herum. Sie müssen sich zwischen unterschiedlichen politischen Positionen entscheiden, die unterschiedliche Werte priorisieren. Vielleicht engagieren sie sich sogar politisch und werden in der Politik aktiv. Nun ist die praktische For-

15 Douglas, H. (2007), "Rejecting the Ideal of Value-Free Science", in: Kincaid, H., Dupré, J. & Wylie, A. (Hrsg.) (2007), *Value-free Science? Ideals and Solutions*, New York: Oxford University Press, S. 120–142.

16 Siehe etwa Harding, S. (1992), "Rethinking Standpoint Epistemology: What is "Strong Objectivity?"", *The Centennial Review* 36/3, 437–470.

17 Leopoldina, 7. Ad-hoc-Stellungnahme vom 8.12.2020, <https://www.leopoldina.org/publikationen/detailansicht/publication/coronavirus-pandemie-die-feiertage-und-den-jahreswechsel-fuer-einen-harten-lockdown-nutzen-2020/>.

derung der Wertfreiheit nicht verletzt, sofern eine Person ihre Rollen klar trennt und als Forscherin oder Forscher keine unbedingten Werturteile vertritt. Allerdings ist eine solche Rollentrennung in der Praxis kaum durchzuhalten. Wenn sich eine Wissenschaftlerin in der Politik engagiert, dann wird sie von vielen oft immer noch als Wissenschaftlerin wahrgenommen. Im Prinzip könnte sie natürlich bei jeder Stellungnahme sagen, in welcher Rolle sie sich äussert, aber das wird schnell unpraktisch und verkompliziert die Stellungnahmen.

Damit hängt ein weiteres Problem zusammen. Selbst wenn ein Wissenschaftler stets penibel zwischen dem zu trennen versucht, was sich wissenschaftlich zu einer Frage sagen lässt, und dem, was er als Bürger sagen möchte, heisst das noch nicht, dass die Öffentlichkeit diese Trennung mitvollzieht. So werden Forschende oft direkt nach unbedingten Werturteilen gefragt: „Muss der Lockdown verlängert werden?“ oder „Sollten wir dem CO₂-Gesetz zustimmen?“ Es werden eindeutige Antworten auf solche Fragen erwartet, und wenn eine Wissenschaftlerin genau differenziert, geht das oft zu Lasten der Relevanz für die Meinungsbildung. Die Verbreitung von Äusserungen in den Medien und ihre Wahrnehmung durch die Öffentlichkeit gehorchen dann nochmal eigenen Gesetzen. So werden die Formulierungen von Forschenden schon für die Bildung von Überschriften oft von Journalistinnen vereinfacht. Wenn die Äusserungen dann durch andere Medien aufgegriffen werden, kommt es oft zu weiteren Verkürzungen, die der Veränderung von Aussagen beim Spiel „Stille Post“ ähneln. Bei einem politisch heiss diskutierten Thema wird eine Äusserung in der Öffentlichkeit schnell einem politischen Lager zugeordnet, auch wenn das gar nicht intendiert ist. So dürfte die wissenschaftlich fundierte Warnung von Kipppunkten des Klimas schnell als links oder grün wahrgenommen werden, weil der Klimawandel durch linke und grüne Parteien stärker thematisiert wird. Aber die Kipppunkte sind natürlich an sich weder links, Mitte oder rechts.

Angesichts der Schwierigkeiten, in der Praxis zwischen der wissenschaftlich begründeten Expertise von dem unbedingten Werturteil zu unterscheiden, mag man geneigt sein, das Postulat der wertfreien Wissenschaft ganz aufzugeben. Das wäre aber vorschnell. Postulate oder Ideale können auch dort eine Wirkung entfalten, wo sie sich nicht vollständig umsetzen lassen. In diesem Sinne kann das Postulat der Wertfreiheit durchaus als Orientierung fungieren.

4. Die Folgen für die Politik

Was bedeuten unsere Überlegungen nun aber für die Politik? Im Sinne der oben skizzierten Aufgabenteilung hat die Politik die Aufgabe, grundlegende Ziele und ihre Gewichte festzulegen, aber auch über konkretes Handeln zu entscheiden. Das Ideal rationalen Handelns impliziert dabei, dass die Entscheidungen der Politik auf dem einschlägigen Wissen beruhen sollten. Das bedeutet zum einen, dass die Bestimmung und Gewichtung der Ziele durch Wissen darüber fundiert sind, was sich realistisch erreichen lässt. So könnte es sein, dass die Ausbreitung von Covid-19 nicht vollständig eingedämmt werden kann. Wenn das so ist, dann ist das Ziel, die Neuansteckungen auf Null zu senken, unrealistisch und sollte nicht von der Politik proklamiert werden. Zum anderen sollte auch die Wahl von Mitteln durch einschlägiges Wissen fundiert sein. Wie bereits bemerkt, ist das nicht immer das Wissen, dass bestimmte Ziele durch gewisse Mittel erreicht werden. Manchmal gibt es nur Wissen über Wahrscheinlichkeiten oder Möglichkeiten. Aber auch dieses Wissen kann und sollte bei rationalen Entscheidungen mit einbezogen werden. (Hier und im Folgenden gehe ich davon aus, dass das einschlägige Wissen, das gebraucht wird, *wissenschaftliches* Wissen ist. Dieses ist besonders gut begründet, weil es sich Verfahren verdankt, die Irrtum besonders verlässlich ausschliessen.¹⁸)

Die Begründung dieser Forderungen ist ziemlich offensichtlich. Wer sein Handeln nicht auf der Basis von Wissen wählt, erreicht die gesetzten Ziele nicht oder doch weniger wahrscheinlich. Ähnliches gilt für einen Fall, in dem Ziele gewählt werden, die sich nicht gemeinsam erreichen lassen.

Diese einfache Begründung suggeriert vielleicht, dass die Forderungen, die sich aus unseren Überlegungen an die Politik ergeben, gar nicht verletzt werden. Wer will schon seine eigenen Ziele nicht erreichen? Zudem hat die Politik ja die Freiheit, über die Ziele und ihre Umsetzung zu entscheiden. Daher kann sie sozusagen an zwei Schrauben drehen. Sie hat daher einigen Spielraum, Ziele und Mittel so zu wählen, dass das wissenschaftliche Wissen berücksichtigt wird.

Allerdings ist die Sache nicht so einfach. Während der Pandemie gab es ja in der Tat immer wieder Versuche seitens der Politik, bestimmte wissenschaftliche Erkenntnisse in Abrede zu stellen und bei der Entscheidungsfindung nicht zu berücksichtigen. So haben etwa der frühere Präsident der USA und sein brasilianischer

18 Hoyningen-Huene, P. (2013), *Systematicity: The Nature of Science*, New York: Oxford University Press.

Kollege die Gefährlichkeit des Virus oft heruntergespielt und entsprechende Erkenntnisse auch nicht beachtet. Die Ursache dafür ist oft eine gewisse „Kurzsichtigkeit“. Politische Akteure orientieren sich eher an Gütern, die kurzfristig erreichbar sind, als an Gütern, die nur längerfristig realisiert werden. Wir ziehen den Spatz in der Hand der Taube auf dem Dach vor, wie es im Sprichwort heisst. Ein Lockdown hat kurzfristig viele negative Nebeneffekte; seine Wirkung stellt sich hingegen erst nach mehreren Wochen ein. Da mag es bequemer sein, zunächst einmal abzuwarten und auf einen Lockdown zu verzichten. Ähnliche Überlegungen gelten für die Mitigation des Klimawandels, wo die Folgen des eigenen Handelns noch weiter in der Zukunft liegen.¹⁹

Die Bevorzugung der naheliegenden Güter lässt sich dabei als Missachtung jener Erkenntnisse beschreiben, welche die fernerliegenden Güter betreffen. Man kann die Bevorzugung der naheliegenden Güter sogar dadurch rationalisieren, dass man behauptet, die Realisierung der fernerliegenden Güter sei unsicher oder gar nicht zu erwarten. So wird eine abwartende Haltung zum Klimawandel oft mit dem Hinweis gerechtfertigt, die Klimaveränderungen seien unsicher oder unabhängig von dem, was wir Menschen tun. Daran ist natürlich problematisch, dass wissenschaftliche Erkenntnisse ohne sachliche Gründe bezweifelt werden.

Eine weitere Ursache, warum wissenschaftliche Erkenntnisse seitens der Politik manchmal bezweifelt, vielleicht sogar geleugnet werden, hat es ebenfalls mit Zielkonflikten zu tun. Wie wir gerade gesehen haben, trägt ein Lockdown zur Erreichung gewisser Ziele bei, etwa des Schutzes vor COVID-19-Erkrankungen, verhindert aber das Erreichen anderer Ziele, etwa eines durchgehenden Unterrichtsangebots an Kinder. Aufgrund einer gewissen Priorisierung der Ziele mag nun ein bestimmter politischer Akteur dazu kommen, einen Lockdown abzulehnen. Um für diese Position zusätzliche Zustimmung zu gewinnen, lohnt es sich nun für den Akteur, die Gefahren der Erkrankung kleinzureden und die Probleme eines Lockdowns zu betonen. Denn so lassen sich leicht Mitstreiter gewinnen, die eigentlich den Schutz vor COVID-19 stärker priorisieren. Wenn man diese Akteure glauben machen kann, dass die Gefährdung durch die Infektion nicht besonders gross ist, dann sind sie eher gegen den Lockdown.

Das Vernachlässigen oder explizite Bezweifeln wissenschaftlicher Ergebnisse lässt sich dabei oft ein Stück weit

rationalisieren, und zwar gerade dann, wenn der Stand der wissenschaftlichen Forschung noch nicht besonders fortgeschritten ist. Das war zu Beginn der Corona-Pandemie der Fall, als ein neues Virus untersucht werden musste. Es gibt mindestens drei Gründe, warum sich in solchen Situationen, aber auch sonst wissenschaftliche Resultate mit recht einfachen Begründungen bezweifeln oder leugnen lassen (was aber nicht rational ist).

Erstens sind viele wissenschaftliche Ergebnisse mit Unsicherheiten behaftet. Gerade zu Beginn der Pandemie war die Datenlage oft verhältnismässig dünn, die Stichproben von Experimenten waren klein, so dass die Unsicherheiten hoch waren.²⁰ Weitere Unsicherheiten können entstehen, wenn die verwendeten Testverfahren nicht besonders gut sind. Auch die Interpretation der Daten kann angreifbar sein. Dann lassen sich die Ergebnisse einer Studie oft mit guten Gründen bezweifeln. Besonders einfach ist das, wenn Studienergebnisse als „erste“ oder „vorläufige“ Resultate deklariert werden.

Zweitens kommt es gerade bei der anfänglichen Erschliessung eines neuen Forschungsgebiets manchmal zu Ergebnissen, die einander widersprechen oder die wenigstens nicht zusammenzupassen scheinen. So gab es in der Corona-Pandemie Studien mit unterschiedlichen Ergebnissen zur Wirksamkeit von Lockdowns.²¹ Häufig kam es dann auch zu einander widersprechenden Expertenstatements in der Öffentlichkeit, oder Fachleute änderten ihre Meinung.²² In solchen Situation kann sich jede Person die Studienergebnisse aussuchen, die am besten zu ihrer Position passen, ohne dabei den wissenschaftlichen Diskussionsstand als ganzen zu berücksichtigen. Ausserdem ist es in solchen Situationen einfach, die Zuverlässigkeit wissenschaftlicher Forschung insgesamt zu bezweifeln.

Drittens ist die Relevanz vieler Studien für die Praxis nicht ganz klar, weil die Studien unter speziellen Bedingungen entstehen. Die Ergebnisse, die eine Laborstudie über die Ausbreitung des Virus über Aerosole erzielt, lassen sich nicht ohne Weiteres auf ein Bahnabteil oder ein Klassenzimmer übertragen.²³ Der Transfer in

19 Zur psychologischen Evidenz für die Bevorzugung kurzfristiger Ziele siehe etwa Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002), "Time discounting and time preference: A critical review", *Journal of Economic Literature* 40(2), 351–401.

20 Siehe etwa Lindner, A. K. et al. (2021), "Head-to-head comparison of SARS-CoV-2 antigen-detecting rapid test with professional-collected nasal versus nasopharyngeal swab", *European Respiratory Journal*, 2004430. Dort wurde die Sensitivität bestimmter Testverfahren untersucht. Da in den Daten nur 41 Personen krank waren, sind die Konfidenzintervalle sehr gross.

21 Siehe etwa <https://www.netdoktor.ch/coronavirus/studie-wirksamkeit-lockdown-10630439>.

22 Vgl. dazu <https://www.br.de/nachrichten/wissen/corona-studien-wenn-wissenschaft-widerspruechlich-ist,S8fX4QL>.

23 Siehe etwa https://www.medica.de/de/News/Archiv/Wie_breitet_sich_das_SARS-CoV-2-Virus_in_der_Raumluft_aus für ein Beispiel einer Laborstudie.

die Praxis ist nur mit einem Analogieschluss möglich, der mit Unsicherheiten behaftet ist. Daher ist es leicht, wenigstens die Relevanz wissenschaftlicher Ergebnisse für die Praxis zu bezweifeln.

Insgesamt ist es also gerade dann, wenn ein neues Phänomen untersucht wird, recht einfach, wissenschaftliche Resultate zu bezweifeln. Dabei kann begründete Kritik an einzelnen Studien durchaus legitim sein. Es ist aber nicht rational, die Zuverlässigkeit der Wissenschaften insgesamt in Zweifel zu ziehen, denn selbst schwache Nachweise sind besser als gar keine Nachweise. Ebenso wenig ist es rational, wissenschaftliche Resultate nur selektiv wahrzunehmen und sich die Resultate herauszupicken, die besser zur eigenen Position passen. Das Ideal einer Arbeitsteilung zwischen Politik und Wissenschaft fordert vielmehr, dass die Beurteilung von Tatsachen vollständig der Wissenschaft obliegt und nicht den Wünschen bestimmter politischer Akteure untergeordnet wird.

Dass wissenschaftliche Ergebnisse oft mit grossen Unsicherheiten behaftet sind, dass sie manchmal einander widersprechen und nicht direkt für die Praxis relevant sind, zeigt auch, dass die Berücksichtigung wissenschaftlichen Wissens viel schwieriger ist, als es die oben genannte idealtypische Unterscheidung von drei Entscheidungssituationen aus der Entscheidungstheorie suggeriert. In realistischen Situationen ist es nicht so, dass wir etwa beanspruchen können, die relevanten Wahrscheinlichkeiten zu wissen. Es mögen zwar Wahrscheinlichkeiten für die Folgen wichtiger Handlungsoptionen untersucht worden sein, aber dabei wird es zu unterschiedlichen Ergebnissen gekommen sein, die teilweise nicht direkt zur Praxis passen usw.

Besondere Vorsicht ist angesagt, wenn nur einige Aspekte eines Phänomens untersucht sind, andere aber gar nicht. In Zeiten der Pandemie wurden mithilfe von Modellen viele Vorhersagen für die Anzahl von Neuinfektionen für verschiedene Szenarien getroffen. Die prognostizierten Infektionszahlen wurden sinnfällig mit bunten Diagrammen veranschaulicht. Dabei zeigte sich oft, dass Kontaktbeschränkungen die Zahlen signifikant reduzieren würden. Viele Nebenfolgen eines Lockdowns waren dagegen nicht untersucht; wenigstens gab es dazu keine quantitativen Aussagen, die sich ähnlich sinnfällig darstellen liessen. In solchen Fällen ist es zwar naheliegend, aber letztlich nicht rational, nur die genau untersuchten Aspekte in die Untersuchung einzubeziehen. Das Ideal einer funktionalen Trennung von Wissenschaft und Politik fordert vielmehr, auch zu berücksichtigen, dass bestimmte Aspekte nicht untersucht wurden, und daher entsprechende Risiken und Unsicherheiten in Rechnung zu stellen. Es geht also nicht darum, nur die vorhandenen

Ergebnisse zu berücksichtigen - denn die Auswahl dessen, was untersucht wurde, mag von bestimmten Interessen getrieben sein. Vielmehr gilt es, den Wissensstand insgesamt zu berücksichtigen, mit seinen Lücken.

5. Fazit

In der Geschichte der westlichen Wissenschaften gibt es eine starke Tendenz, die wissenschaftliche Betrachtung von menschlichen Werten und Zielen zu emanzipieren.²⁴ Das führt zu einer Funktionstrennung zwischen Wissenschaft und Politik, die sich im Ideal rationalen politischen Handelns äussert, das wir diskutiert haben. Demnach haben die Wissenschaften die Aufgabe, Wissen über Tatsachen herauszufinden; die Politik soll auf dieser Basis Entscheidungen über Ziele, aber auch konkretere Massnahmen treffen. Diesem Ideal kann von beiden Seiten zuwidergehandelt werden. So treten manchmal Forschende in dieser Rolle mit unbedingten Werturteilen an die Öffentlichkeit und die Politik heran. Die Politik vernachlässigt manchmal wissenschaftliche Ergebnisse oder stellt sie sogar ohne gute Gründe in Abrede. All das zeigt aber nicht, dass das Ideal obsolet ist, sondern nur, dass es sich um ein Ideal handelt. Schwerer wiegt, dass das Ideal in der Praxis manchmal kaum erfüllt werden kann. Ausserdem beschreibt es das Zusammenwirken von wissenschaftlicher Expertise und politisch gesetzten Zielen konkret nur für sehr idealisierte Situationen, wie sie in der Entscheidungstheorie unterschieden werden. Schliesslich gilt es nicht für alle Wissenschaften, sondern nur für die Naturwissenschaften, wo der Anspruch, wertfreie Erkenntnisse zu gewinnen, plausibel ist. Dennoch kann das Ideal wenigstens dort, wo es um die naturwissenschaftliche Expertise geht, als grobe Orientierung dafür dienen, was wissenschaftlich fundierte Politik ist.

Claus Beisbart ist Extraordinarius für Wissenschaftsphilosophie am Institut für Philosophie der Universität Bern. In seiner Forschung interessiert er sich für die Wissenschaftsphilosophie (Erkenntnistheorie von Simulation und Modellierung; Philosophie der Kosmologie; Wahrscheinlichkeiten), für die Methode des Überlegungsgleichgewichts und für einige Fragen der Ethik. Claus Beisbart ist seit Kurzem auch Chair des Clusters „Ethics and Policy“ am neugegründeten *Multidisciplinary Center for Infectious Diseases* (MCID) und interessiert sich daher für philosophische Aspekte der Pandemie.

²⁴ Siehe dazu Daston, L. & Galison, P. (2007), *Objectivity*, 7. Auflage, New York: Zone Books, dt. Daston, L. & Galison, P. (2007), *Objektivität*, Suhrkamp: Frankfurt am Main.