



meta(φ)

Sprechen, Verstehen, Verändern

J.S. Mills Diskussionsfreiheit

• Marco R. Schori

Der Schutz fundamentaler Rechte und die Demokratie am Arbeitsplatz

• Simon Richmond

The Problem of the Barber-Shop Paradox by Lewis Carroll

• Vera Moser

Die Rolle der Selbstbestimmung in der Rechtfertigung passiver und aktiver Sterbehilfe

• Andreas Müller

Pornography's Authority to Subordinate Women

• Xenia Schmidli

Reassessing Credibility Excesses in Epistemic Injustice

• Dominik Fitze

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

Timo Junger

Lektorat

Xenia Schmidli, Raffael Grolimund, Luisa Pacozzi

Produktion und Gestaltung

Joannes B. Campell

Coverbild: Timo Junger

ISSN

Gedruckte Ausg.: meta(φ) ISSN:2297-9948

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

Text-Einreichungen

Einreichungen zur Veröffentlichung sind erwünscht an metaphi.philo@lists.unibe.ch. 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 Redaktions-Mitglied. In diesem Magazin veröffentlichte Texte gelten als wissenschaftliche Publikation und sind somit zitierfähig.

Erste Worte

Liebe Leserin, lieber Leser

Vor einigen Wochen sah sich das Redaktionsteam des *meta(φ)* erneut mit der Aufgabe konfrontiert, einen Titel für das neue Heft zu finden. Nach angeregter Diskussion einigten wir uns schliesslich auf eben diejenige Überschrift, die nun das Cover dieser bereits sechsten Ausgabe ziert: Sprechen, Verstehen, Verändern. Indem wir diese drei Verben auswählten und schliesslich ins Dokument einfügten, haben wir durch Sprache Einfluss auf die Welt genommen. Dieser Sprechakt, wie John L. Austin unser Vorgehen bezeichnen würde, zeigt—wenngleich in bescheidenem Kontext—welche Macht von Sprache ausgehen kann. Passend dazu vereint das vorliegende Heft sechs wundervolle Texte, die allesamt auf unterschiedliche Weise auf die Phänomene der Sprache, des Verstehens oder des Veränderns Bezug nehmen.

Den Auftakt im ersten Teil namens *Miteinander Sprechen* macht Marco R. Schoris Text zu J.S. Mills Diskussionsfreiheit, in welchem unter anderem argumentiert wird, dass auf Basis von Mills Schrift *On Liberty* eine progressive Diskussionskultur gerechtfertigt werden kann. Anschliessend legt Simon Richmond seine philosophische Analyse des Arbeitsplatzes dar und argumentiert, dass demokratische Mitspracherechte, welche die Funktion haben, fundamentale Rechte zu schützen, auch am Arbeitsplatz gelten sollten. Im Abschnitt *Verstehen und Bestimmen* folgt dann zunächst Vera Mosers Auseinandersetzung mit Lewis Carrolls Barber-Shop Paradox, in welcher untersucht wird, wie Konditionale, also sogenannte „wenn-dann Sätze“, zu verstehen sind. Eine

zentrale These wird sein, dass formale Logik nicht zwingend mit menschlichen Intuitionen in Einklang stehen muss. Im Fachwort dieser Ausgabe geht Andreas Müller der Frage nach, welche Rolle der Selbstbestimmung in verschiedenen Formen der Sterbehilfe zukommt. Untersucht wird dabei insbesondere, ob und wie angesichts der breiten Akzeptanz von passiver Sterbehilfe gegen aktive Sterbehilfe argumentiert werden kann. Der dritte Teil trägt den Titel *Sprache und Autorität* und vereint Aufsätze von Xenia Schmidli und Dominik Fitze. Im ersten Text geht die Autorin dem Phänomen der Pornographie auf den Grund und argumentiert, dieser komme in unserer Gesellschaft die Autorität zu, Frauen zu unterdrücken. Wichtige argumentative Ressource ist dabei die Sprechakttheorie von John L. Austin. Abschliessend setzt sich Dominik Fitze mit Miranda Frickers Idee der epistemischen Ungerechtigkeit auseinander und legt sein Augenmerk insbesondere auf sogenannte „credibility excesses“, also Situationen, in denen einer Person unfairerweise nicht zu wenig, sondern zu viel Glaubwürdigkeit zugeschrieben wird.

Zum Schluss wünschen wir Ihnen eine inspirierende und unterhaltsame Lektüre und hoffen, Sie erhalten einen erhellenden Einblick in die Ideen und Argumente, welche am Institut für Philosophie der Universität Bern entstehen und ausgearbeitet werden.

Herzlichst, Timo Junger
im Namen des Redaktionsteams



Xenia Schmidli



Timo Junger



Joannes B. Campell



Luisa Pacozzi



Raffael Grolimund

“It was for too long the assumption of philosophers that the business of a ‘statement’ can only be to ‘describe’ some state of affairs, or to ‘state some fact’, which it must do either truly or falsely.”

John L. Austin
How to do Things with Words, 1962

“The most striking aspect of linguistic competence is what we may call the ‘creativity of language’, that is, the speaker’s ability to produce new sentences, sentences that are immediately understood by other speakers although they bear no physical resemblance to sentences which are ‘familiar’.”

Noam Chomsky
Chomsky: Selected Readings, 1971

Inhaltsverzeichnis

Miteinander Sprechen

- J.S. Mills Diskussionsfreiheit
• Marco R. Schori 6
- Der Schutz fundamentaler Rechte und die Demokratie am Arbeitsplatz
Richard Arnesons Theorie demokratischer Rechte angewandt auf die Arbeitswelt • Simon Richmond 15

Verstehen und Bestimmen

- The Problem of the Barber-Shop Paradox by Lewis Carroll
A Clash between Intuition and Rational Thinking • Vera Moser 24
- Die Rolle der Selbstbestimmung in der Rechtfertigung passiver und aktiver Sterbehilfe
• Andreas Müller 35

Sprache und Autorität

- Pornography's Authority to Subordinate Women
• Xenia Schmidli 48
- Reassessing Credibility Excesses in Epistemic Injustice
• Dominik Fitze 62

J.S. Mills Diskussionsfreiheit

Einleitung

John Stuart Mill, bekannt für seine utilitaristischen und liberalen Schriften, behandelt in *On Liberty* (1859, 217) die „bürgerliche oder soziale Freiheit“. Ein beträchtlicher Teil dieser Schrift widmet sich der Erörterung dessen, was er die *Freiheit des Denkens und der Diskussion* nennt. Im gleichnamigen Kapitel verteidigt er hauptsächlich eine *Meinungsäusserungsfreiheit*, die noch heute Gegenstand philosophischer Diskussionen ist. Nach Daniel Jacobson (2016, 440–41) gibt es drei verschiedene Interpretationen dieser Freiheit, welche sich alle auf die Schrift von Mill berufen: Die pragmatische Interpretation vertrete die Ansicht, dass Mill die Regulierung von *Meinungsäusserungen* zwar prinzipiell in der Zuständigkeit der Gesellschaft sehe, solche Regulierungen aber niemals optimal seien und Mill daher keine Eingriffe erlaube. Die qualifizierte Interpretation gehe hingegen davon aus, dass Mill gar keine *Meinungsäusserungsfreiheit* vertrete, sondern eine *Freiheit der nicht-schädigenden oder harmlosen Meinungen*. Jacobson behauptet, entgegen dieser beiden Interpretationen, dass Mill für eine *uneingeschränkte Meinungsäusserungsfreiheit* argumentiert habe. Die pragmatische Interpretation gehe fehl darin zu glauben, dass Mill nicht für ein Recht argumentiert habe, welches prinzipiell vor Eingriffen geschützt sei, und die qualifizierte Interpretation verkenne die Signifikanz und die Rolle des Getreidehändler-Beispiels (Jacobson 2016, 440–41). Jacobson zieht den Schluss, dass die *Meinungsäusserungsfreiheit* von Mill auch schädigende Äusserungen schützt—beispielsweise beleidigende oder abwertende Aussagen—und diese daher toleriert werden müssten, sofern sie nicht performative Sprechakte seien, die nicht mehr immun vor Eingriffen seien (Jacobson 2016, 452). Diese Interpretation geht auf ein Verständnis von Mills Liberalismus zurück, der aus einem Schädigungsprinzip und einer Freiheitssphäre besteht. Das Schädigungsprinzip

konkretisiert, welche Handlungen gerechtfertigt eingeschränkt werden dürfen und die Freiheitssphäre beschreibt Rechte, welche vom Schädigungsprinzip ausgenommen sind (Jacobson 2016, 446).

Im vorliegenden Essay wird die These vertreten, dass auf der Basis von *Mills Argumentation für die Meinungsäusserungsfreiheit, entgegen Jacobsons Behauptung, für eine progressive Diskussionskultur argumentiert werden kann*. Dabei soll nicht unterstellt werden, dass Mill selbst für eine solche Diskussionskultur argumentiert habe. Er hat jedoch die argumentativen Grundlagen für ein solches Argument gelegt. Unter einer „progressiven Diskussionskultur“ wird im Folgenden verstanden, dass eine Gesellschaft weder rechtlich noch moralisch dazu verpflichtet ist, beleidigenden und abwertenden Äusserungen aufgrund von nicht wählbaren Eigenschaften—beispielsweise Geschlecht, sexuelle Orientierung, Herkunft, etc.—eine (öffentliche) Plattform zu gewähren.

Um das Argument aufzubauen, wird in einem ersten Schritt die Interpretation des Mill'schen Liberalismus (1) und dessen zentralen Prinzipien, das Schädigungsprinzip (1.1) und die Freiheitssphäre (1.2), nach Jacobson anhand des Artikels „Mill on Liberty, Speech, and the Free Society“ (2000) und dem Beitrag „Mill on Freedom of Speech“ in *A Companion to Mill* (2016) aufgezeigt. Dies bietet die Grundlage für die *uneingeschränkte Meinungsäusserungsfreiheit* nach Jacobson, welche daraufhin besprochen wird (1.3). Der nächste Abschnitt geht auf einige Aspekte des Arguments von Mills *On Liberty* ein, welche eine andere Interpretation der *Meinungsäusserungsfreiheit* zulassen (2.1). Dazu gehört erstens die Unterscheidung zwischen einer blossen *Ausdrucksfreiheit* und einer *Meinungsäusserungsfreiheit* (2.1.1). Zweitens wird auf die Rolle der *Diskussionsfreiheit* in den epistemischen Prozessen einer Gesellschaft verwiesen (2.1.2). Der dritte Aspekt

beschäftigt sich mit dem zentralen Anspruch Mills, ein Freiheitskonzept zu entwerfen, welches dem Menschen als entwicklungsfähigem Wesen, als denkendes Individuum, gerecht wird (2.1.3). Diese Blickpunkte, gemeinsam mit dem Mill'schen Liberalismus nach Jacobson, bieten die Grundlage eines Argumentes für eine progressive Diskussionskultur (2.2). Das Fazit rekapituliert die Ergebnisse und gibt einen kurzen Ausblick.

1. Mills Liberalismus nach Jacobson

Jacobson argumentiert für eine Interpretation der Mill'schen *Meinungsäusserungsfreiheit*, die jegliche Beeinträchtigung dieser Freiheit einer Akteurin als Verletzung ihrer Freiheitsrechte versteht (Jacobson 2000, 278). Diese Interpretation beruht auf Jacobsons Verständnis des Mill'schen Liberalismus, welches durch folgende drei Thesen zusammengefasst werden kann (vgl. Jacobson 2000, 277):

1. Mill befürwortet eine Form des Liberalismus, die Personen gewisse fundamentale Rechte zuspricht, ein Recht auf freie Meinungsäusserung miteingeschlossen.
2. Mill akzeptiert zwar, dass Sprechakte schädigend sein können, aber seine Verteidigung der *Meinungsäusserungsfreiheit* beschränkt sich nicht auf harmlose Sprechakte.
3. Die Freiheitssphäre folgt nicht aus dem Schädigungsprinzip und ist nicht auf harmlose Handlungen beschränkt.

Mills Liberalismus könne somit nicht auf das Schädigungsprinzip reduziert werden, sondern es beinhalte sogar eine strikte Ablehnung paternalistischer Eingriffe sowie eine substantielle Freiheitssphäre, welche von der Gesellschaft nicht beeinträchtigt werden darf. Zudem sei das Schädigungsprinzip auf Handlungen beschränkt, die nicht von der Freiheitssphäre geschützt sind und die Freiheitssphäre umfasst nicht nur Handlungen, die für andere nicht-schädigend sind (Jacobson 2000, 278). Die Ablehnung paternalistischer Eingriffe könne aus dem Schädigungsprinzip gewonnen werden (Jacobson 2000, 292–93). Da die Äusserung von Meinungen und Gesinnungen von Mill zu den Rechten gezählt wird, welche die Freiheitssphäre konstituieren, schliesst Jacobson, dass Mill eine *uneingeschränkte Meinungsäusserungsfreiheit* vertreten habe. In einzelnen Fällen können Sprechakte zwar unterbunden werden, dies aber nicht aufgrund der

Schädigung Dritter, wie dies das Schädigungsprinzip vermuten lassen würde (Jacobson 2000, 278), sondern weil sie in gewissen Kontexten eine Verpflichtung—beispielsweise gegenüber der Familie oder eines Gläubigers—verletzen würden (Jacobson 2016, 447).

Diese Interpretation des Mill'schen Liberalismus bietet sowohl die Grundlage für die *uneingeschränkte Meinungsäusserungsfreiheit* nach Jacobson wie auch für die *Diskussionsfreiheit*, wie sie im vorliegenden Essay vertreten wird. Um die Kritik an Jacobsons Interpretation nachvollziehen zu können, wird im Folgenden der Inhalt und die argumentativen Rollen des Schädigungsprinzips und der Freiheitssphäre bei Jacobson aufgezeigt. Danach wird auf die *uneingeschränkte Meinungsäusserungsfreiheit* nach Jacobson eingegangen.

1.1 Inhalt und Rolle des Schädigungsprinzips

Mill präsentiert das Schädigungsprinzip in zwei unterschiedlichen Formulierungen, welche direkt aufeinander folgen:

„Einerseits sei Selbstschutz der einzig legitime Zweck für eine Person oder eine Gesellschaft, in die Handlungsfreiheit einer Person einzugreifen. Andererseits sei der einzige Grund, weshalb gerechtfertigt Macht über ein Mitglied der Gesellschaft ausgeübt werden kann, gegen dessen Willen, die Schädigung anderer zu verhindern.“¹ (Mill 1859, 223)

Diese Formulierungen des Schädigungsprinzips sind problematisch, da die beiden aufeinanderfolgenden Aussagen nicht als synonym verstanden werden können: Nicht jede Ausübung von Macht ist gleichzusetzen mit einer Beeinträchtigung der Freiheit einer Person. Jacobson schlägt daher eine Formulierung vor, welche er für weniger problematisch hält und dennoch beiden Aussagen gerecht werde:

„Der einzig gute Grund für eine Einmischung anderer ist es, die Schädigung Dritter zu verhindern.“² (Jacobson 2000, 290)

1 Übersetzungen von Textpassagen aus *On Liberty* orientieren sich an der Übersetzung von Bruno Lemke in der Reclam Ausgabe *On Liberty: Über Freiheit: Englisch/Deutsch* (2009), verweisen jedoch auf die Textstelle im englischen Original nach den *Collected Works of John Stuart Mill*, herausgegeben von John M. Robson, in 33 Bänden.

2 Eigene Übersetzungen, welche nahe am englischen Original sind, werden mit Anführungs- und Schlusszeichen gekennzeichnet.

Das Schädigungsprinzip müsse als notwendige, jedoch nicht als hinreichende Bedingung für eine gerechtfertigte Einmischung der Gesellschaft oder einzelner in die Handlungsfreiheit einer Person verstanden werden (Jacobson 2000, 290). Dies sei sowohl mit den oben wiedergegebenen Aussagen von Mill vereinbar, als auch mit der Aussage, dass „[...] die Schädigung, oder die Wahrscheinlichkeit einer Schädigung, der Interessen anderer die Einmischung der Gesellschaft rechtfertigen könne, dies jedoch nicht in jedem Fall ein solches Eingreifen rechtfertigen würde“ (Mill 1859, 292). Das heisst, auch wenn Meinungsäusserungen nach Mill prinzipiell einschränkbar wären, was sowohl von Jacobson als auch in diesem Essay bestritten wird, würde der alleinige Umstand, dass die Äusserung schädlich ist, nicht ausreichen, um sie zu unterbinden.

Ein weiterer Aspekt des Schädigungsprinzips ist laut Jacobson, dass es einer Beschränkung unterworfen ist. Eine Einmischung der Gesellschaft sei demnach nur dann gerechtfertigt, wenn die schädigende Handlung nicht Bestandteil der Freiheitsrechte ist (Jacobson 2000, 291). Jegliche Einschränkung ist nach Mill ein Übel, auch wenn sie manchmal das kleinere Übel ist (Jacobson 2000, 293). Besonders deutlich wird diese Ansicht Mills in der Besprechung des Freihandels: „Das Prinzip der individuellen Freiheit ist in dieser Lehre nicht vorhanden und daher auch in den meisten Fragen nicht zu finden, welche deren Grenzen betreffen [...]“ (Mill 1859, 293). Jacobson schliesst aus dieser Aussage, dass einerseits nicht jede Ausübung von gesellschaftlicher Macht über eine Person eine Verletzung ihrer Freiheit darstellt und dass es andererseits einen grundlegenden Einwand gibt als Unzweckmässigkeit, wenn Einschränkungen die Freiheit einer Person begrenzen (Jacobson 2000, 293).

Das Schädigungsprinzip bietet somit die Grundlage für legitime Eingriffe der Gesellschaft oder einzelner in den Handlungsspielraum einer Person, sofern die betreffende Handlung nicht Bestandteil der Freiheitsrechte dieser Person ist. Die Schädlichkeit einer Handlung allein bietet jedoch keinen hinreichenden Grund für Eingriffe. Beide Bedingungen sind somit notwendig und ergeben einen *guten Grund*, sind jedoch nicht hinreichend (Jacobson 2000, 290–91).

1.2 Inhalt und Rolle der Freiheitssphäre

„Self-regarding action“ ist Mills Kunstbegriff für eine Handlung (im weitesten Sinne), welche in die Freiheitssphäre fällt (Jacobson 2000, 298). Zur Freiheitssphäre gehört erstens „[...] das innere Feld des Bewusstseins

und dies fordert Gewissensfreiheit im weitesten Sinne [...]“ (Mill 1859, 225). Auch zur ersten Freiheit zählt Mill die „Freiheit des Denkens und Fühlens [und die] unbedingte Unabhängigkeit der Meinung und Gesinnung in allen Fragen“ (Mill 1859, 225). Auch die Freiheit, diese Meinung und Gesinnung öffentlich zu vertreten zählt er zur ersten Freiheit. Zweitens gehört hierhin die Freiheit „des Geschmacks und der Studien“ (Mill 1859, 226). Dies dient dazu, einen Lebensplan entwerfen zu können, der dem Charakter der Person entspricht, sofern dies anderen keinen Schaden zufügt. Drittens folgt hierauf die Freiheit, dies alles auch mit anderen zu tun: Die Freiheit des Zusammenschlusses. Keine Gesellschaft könne frei genannt werden, in welcher diese Freiheiten nicht absolut und unqualifiziert garantiert würden. (Mill 1859, 225–26)

Laut Jacobson (2000, 299) vereint nicht ein gemeinsames, beschreibendes Merkmal die drei Bereiche der Freiheitssphäre, sondern eine gemeinsame Rechtfertigung. Diese ist:

„Um Autonomie erlangen zu können, müssten wir sowohl in der Lage sein, verschiedene Lebensentwürfe zu vergleichen, zu kritisieren und auszuprobieren als auch selbst Kritik gegenüber offen sein.“ (Jacobson 2000, 299)

Die Verteidigung von absoluten Freiheitsrechten mit dem Utilitarismus von Mill zu vereinbaren, scheine auf den ersten Blick nicht einfach zu sein, könne jedoch in vielen Schriften von Mill beobachtet werden (Jacobson 2016, 448–49). In „Utilitarianism“ (Mill 1861, 250) beschreibt Mill die Rechte einer Person als etwas, worauf sie einen gerechtfertigten Anspruch hat, dass deren Besitz und deren Schutz durch die Gesellschaft gewährleistet wird. In *On Liberty* begründet er Rechte in ihrem Nutzen, möchte diesen „Nutzen [aber] im weitesten Sinne“ verstanden wissen (Mill 1859, 224). Dies ist Jacobson (2016, 449) zufolge so zu verstehen, dass die Freiheitsrechte einer Person Nutzenabwägungen in Einzelfällen immer überwiegen, da sie aus moralischen Regeln erwachsen, welche in erster Linie die Grundlagen des menschlichen Wohlergehens betreffen. Daher konstituierten sie—im Gegensatz zu anderen Verhaltensregeln—absolute Verpflichtungen.

Dass in der Freiheitssphäre auch Handlungen mit eingeschlossen sind, die andere schädigen können, ist hiermit wohl klar ersichtlich. Mill (1859, 281) macht am Beispiel der Trunkenheit deutlich, dass eine Handlung, die zwar ein Laster ist, aber in die Freiheitssphäre und damit zu den *self-regarding action* gehört und immun

gegenüber rechtlichen und moralischen Sanktionen ist, aus der Sphäre herausfallen kann. Wenn eine Person aufgrund ihrer Trunkenheit eine Verpflichtung vernachlässigt, beispielsweise gegenüber der Familie oder einem Gläubiger, könne sie getadelt oder—je nach Schwere des Vergehens—rechtlich belangt werden. Grund für den Tadel oder die rechtliche Belangung ist dann jedoch laut Jacobson (2000, 300) nicht die Trunkenheit, sondern die Verletzung einer Verpflichtung gegenüber einer anderen Person.

1.3 Uneingeschränkte Meinungsäusserungsfreiheit

Das Schädigungsprinzip und die Freiheitssphäre konstituieren gemeinsam den Liberalismus von Mills *On Liberty*. Diese Überlegungen bilden auch die Grundlage für die Position von Jacobson (2016, 440–41), dass Mill für eine *uneingeschränkte Meinungsäusserungsfreiheit* argumentiert. Dazu muss er aufzeigen, dass Mill die Freiheit, eine Meinung zu äussern, als Freiheitsrecht in der Freiheitssphäre versteht.

Dass Mill die *Meinungsäusserung* dem ersten Bereich der Freiheitssphäre—der Gewissensfreiheit im weitesten Sinne—zuordnet, „[...] obwohl sie unter ein anderes Prinzip zu fallen scheint [...]“, findet sich in der Beschreibung der Freiheitssphäre (Mill 1859, 225). Die ambige Anmerkung, dass sie nur darunter zu fallen *scheint*, kann man nach Jacobson (2000, 283) auf zwei Arten verstehen: Entweder werde behauptet, dass sie nur scheinbar andere betreffe, oder dass sie, obwohl sie andere betreffe, trotzdem unter dasselbe Prinzip falle. Beide Lesarten ordnen *Meinungsäusserungen* klar der Freiheitssphäre zu (Jacobson 2000, 283). Zudem lasse die Aussage, dass die *Meinungsäusserungsfreiheit* praktisch untrennbar von der Freiheit der Gedanken sei, vermuten, dass Mill *Diskussion* als eine Art des Denkens verstehe und Äusserungen von Meinungen näher an das Denken als an das Handeln rücke (Jacobson 2000, 284). Demnach wäre das Sprechen nicht als praktische Art, seine Gedanken auszudrücken, zu verstehen, sondern als Medium des Denkens selbst. Ein weiterer Hinweis, dass Mill den Meinungs Austausch als ein Medium des Denkens versteht, finde sich in den epistemischen Argumenten des zweiten Kapitels von *On Liberty*, in welchen die Rechtfertigung von Wissen als grundlegender sozialer Prozess charakterisiert wird, wonach die Entstehung von Wissen nur durch den diskursiven Austausch mit anderen Meinungen möglich sei (Jacobson 2016, 444). Um eine Überzeugung gerechtfertigt zu glauben, müsse sie gegen die besten entgegenstehenden Überzeugungen bestand haben (Jacobson 2016, 444).

Wenn es nun korrekt ist, dass Mill die *Meinungsäusserungsfreiheit* der Freiheitssphäre zuordnet, muss Jacobson das Getreidehändler-Beispiel einordnen und erklären können. Mill zeigt am Getreidehändler-Beispiel auf, dass gewisse Sprechakte unterbunden werden dürfen: Wenn jemand während einer Hungersnot vor dem Haus eines Getreidehändlers einem aufgebrachten Mob zuruft, dass alle Getreidehändler die Armen aushungern würden, dürfe dies unterbunden werden. Würde diese Meinung anders vorgebracht, beispielsweise in einem Zeitungsartikel, hingegen nicht (Mill 1859, 260). Jacobson geht dabei von folgendem Zitat aus, welches aufzeigt, wann eine *Meinungsäusserung* ihre Immunität verliert:

„Niemand behauptet, dass Handlungen ebenso frei sein dürften wie Meinungen. Im Gegenteil: selbst Gedanken verlieren ihre Strafflosigkeit, wenn die Umstände, unter denen sie ausgesprochen werden, von der Art sind, dass ihr Ausdruck eine direkte Anstiftung zu irgendeiner Schandtät bildet.“ (Mill 1859, 260)

Jacobson (2000, 286–87) schliesst aus dieser Textstelle, dass performative, illokutionäre³ Sprechakte, wie beispielsweise kommissive Sprechakte⁴, nichtsdestotrotz *self-regarding* sind, da sie entweder in den Bereich der Freiheit des Geschmacks und der Studien oder der Freiheit des Zusammenschlusses fallen. Anders verhalte es sich mit assertiven⁵ und expressiven⁶ Sprechakten. Diese würden in den Bereich der Gewissensfreiheit im weitesten Sinne fallen. Das Getreidehändler-Beispiel diene dazu, aufzuzeigen, in welchen Kontexten genau das Kundtun einer Meinung als Handlung in substantiellerem Sinne, als *direkte Anstiftung*, zu verstehen sei. Es werde damit hinreichend klar gemacht, dass nicht der propositionale Gehalt des Sprechaktes der Grund für die

3 Ein illokutionärer Sprechakt ist nach John L. Austin (1962) die kleinste, vollständige, sprachliche Einheit. Searle (1999) unterscheidet fünf unterschiedliche, illokutionäre Sprechakte voneinander, wovon die drei für diese Diskussion relevanten hier angesprochen werden.

4 Nach der Sprechakttheorie von Searle ist ein kommissiver Sprechakt eine Verpflichtung der Sprecherin Handlungen auszuführen, welche den propositionalen Gehalt des Sprechaktes wiedergeben. Dies können Versprechen, Gelübde, Zusicherungen, Verträge und Verbürgungen sein. (Searle 1999, 149)

5 Ein assertiver Sprechakt dient dazu, den Hörer von der Wahrheit der geäusserten Proposition zu überzeugen. Die Proposition wird als Tatsache, welche die Welt wiedergibt, vorgebracht. Beispiele sind Feststellungen, Beschreibungen, Klassifikationen und Erklärungen. (Searle 1999, 148)

6 Ein expressiver Sprechakt verbürgt die Ernsthaftigkeit des Sprechaktes. Beispiele sind Entschuldigungen, Danksagungen, Gratulationen, Begrüssungen und Beileidsbekundungen. (Searle 1999, 149)

Einschränkung der Handlungsfreiheit der Sprecherin sei, sondern der Kontext, in welchem er geäußert werde. Dieselbe Meinung, veröffentlicht in einer Zeitung, dürfte nicht zensiert werden. (Jacobson 2000, 286–87)

Mill verstehe unter Meinungen, dass verschiedene Personen die gleiche Meinung vertreten würden, wenn sie eine Proposition glauben. Unterschiedliche Äußerungen derselben Meinung könnten aber verschiedene Handlungen konstituieren, je nach Kontext der Äußerung. Diese Auffassung mache implizit eine Typenidentifikation von Äußerungsakten anhand der Meinung oder Gesinnung, die geäußert wurde, was wiederum festlege, welche Eingriffe in die *Meinungsäußerungsfreiheit* als Einschränkung der Freiheit gelten würden. Die *uneingeschränkte Meinungsäußerungsfreiheit*, für die Mill argumentiere, sei die Freiheit, jegliche faktische oder normative Meinung äußern zu können, wobei Meinungen durch ihren Inhalt individuiert würden. Dies habe zur Folge, dass alle expressiven Sprechakte immun vor Eingriffen der Gesellschaft seien, was jedoch nicht für die restlichen vier illokutionären Sprechakte gelte. (Jacobson 2000, 285–90)

2. Diskussionsfreiheit

Im folgenden Abschnitt wird ein Argument für eine *progressive Diskussionskultur* auf der Grundlage des Mill'schen Liberalismus nach Jacobson formuliert. Dazu wird in einem ersten Schritt die Analyse der Mill'schen *Meinungsäußerungsfreiheit* von Jacobson kritisiert und ergänzt. Danach wird anhand der angepassten Analyse aufgezeigt, wie auf dieser Grundlage für eine *progressive Diskussionskultur* argumentiert werden kann. Es wird jedoch nicht behauptet, dass Mill selbst dieses Argument vorgebracht habe, sondern vielmehr, dass er die Grundlage dafür bereitet hat. Einigkeit zwischen der Position dieses Essays und Jacobsons Position besteht in der Analyse des Mill'schen Liberalismus. Die Ansichten unterscheiden sich jedoch in Bezug auf den Inhalt und die Rolle der Rechtfertigung der *Meinungsäußerungsfreiheit* als Teil des ersten Freiheitsbereichs. Im Folgenden wird nun darauf eingegangen, wie die Analyse ergänzt werden muss, um ein vollständiges Bild der *Meinungsäußerungsfreiheit* nach Mill zu erhalten.

2.1 Einwände gegen Jacobson

2.1.1 Der Inhalt: Ausdrucksfreiheit vs. Meinungsäußerungsfreiheit

Mill bespricht die *Freiheit, Meinungen zu äußern*, in einem begrenzten Kontext: Dabei werden *Meinungen*, im Gegensatz zum heutigen subjektivistischen Ver-

ständnis, als propositionale Überzeugungen behandelt, die nicht nur einzelnen Personen zugänglich sind und dazu dienen, falsche Vorstellungen als solche zu identifizieren. Er „[...] stellt [...] im zweiten Kapitel die Freiheit, seine Meinung zu äußern [...] stets in den Zusammenhang der Wahrheit und Falschheit propositionaler Überzeugungen“ (Niesen 2015, 38). Dies spreche gegen eine bloße Ausdrucksfreiheit ohne Berücksichtigung der Wahrheitsfindung. Mill habe die theoretischen und sprachlichen Grundlagen zur Verfügung gehabt, um für eine weiter greifende Ausdrucksfreiheit zu argumentieren. Er begrenze die Diskussion der *Meinungsäußerungsfreiheit* jedoch auf den Kontext der *Freiheit des Denkens und der Diskussion* und charakterisiere sie somit als Mittel epistemischer Überlegungen (Niesen 2015, 38). Dies kann anhand des Verhältnisses von *Meinungen* und *Tatsachen* aufgezeigt werden: Meinungen seien demnach Tatsacheninterpretationen und machten Tatsachen dem Geist erst zugänglich. Auch dürften *wahre Meinungen* nicht mit Tatsachen gleichgesetzt werden (Niesen 2015, 38–39 und Mill 1859, 234). Eine Besprechung von falschen Tatsachenbehauptungen, sowohl wissentlich als auch unwissentlich geäußerte, lässt sich in *On Liberty* nicht finden. Dass Mill seine Ansicht aus „Law of Libel and Liberty of the Press“ (vgl. Mill 1825), dass falsche Tatsachenbehauptungen prinzipiell eingeschränkt werden dürfen, ändern müsste, sei aus den in *On Liberty* vorgebrachten Argumenten nicht zu erschliessen (Niesen 2015, 43–44). Nicht nur das Verhältnis von Tatsachen und Meinungen, sondern auch der Zusammenhang zwischen der *Freiheit des Denkens und der Diskussion* und der *Meinungsäußerungsfreiheit*, gibt einen Eindruck davon, dass es bei Mill nicht um eine subjektivistische Ausdrucksfreiheit geht. Dieser Zusammenhang lasse sich mit der Wahrheitsfindung einerseits und der „argumentativen Suche nach wohlverstandenen Überzeugungen und vernünftigen Handlungsplänen“ andererseits erklären (Niesen 2015, 40). Dass die *Meinungsäußerungsfreiheit* der Wahrheitsfindung diene, erkennt man an der Art und Weise, wie Mill begründet, dass sie zu den Freiheitsrechten einer Person gehört. Dass sie andererseits der individuellen und gesellschaftlichen Deliberation diene, begründet sich in der Annahme, dass die Rechtfertigung von Meinungen ein primär soziales Unterfangen sei. Dieser zweite Aspekt wird im nächsten Unterkapitel 2.1.2 besprochen.

Mill (1859, 229–43) beginnt die Rechtfertigung, dass die *Meinungsäußerungsfreiheit* zu den Freiheitsrechten einer Person gehört, mit der Besprechung wahrer Meinungen. Eine wahre Meinung müsse aus zwei

Gründen den Schutz vor gesellschaftlichen Eingriffen genießen: Einerseits würde mit der Unterdrückung einer (potentiell) wahren Meinung ein möglicherweise entscheidender Beitrag zur Diskussion vorenthalten werden (Mill 1859, 229). Andererseits müsste man sich damit Unfehlbarkeit an. Sollte die Unterbindung einer (potentiell) wahren Meinung aus der Überlegung im Sinne des Wohlergehens aller Gesellschaftsmitglieder geschehen, würde dies einer Beanspruchung von Unfehlbarkeit in der Beurteilung der Schädlichkeit oder Nützlichkeit einer Meinung gleichkommen (Mill 1859, 233). Zudem sei die Wahrheit von der Nützlichkeit einer Meinung nicht sinnvoll zu trennen (Niesen 2015, 37). Der Schutz von *wahren Meinungen* scheint dabei den Zugang der Person zu wahren Überzeugungen zu sichern. Im Zentrum steht weniger der Sprecher, sondern eher die Hörerin. Dies legt zumindest der Fokus auf die Vorenthaltung eines wichtigen Diskussionsbeitrags und die Einschätzung der Nützlichkeit einer (potentiell) wahren *Meinung* nahe.

Falsche Meinungen werden von Mill (1859, 243–52) unter den gleichen absoluten Schutz gestellt, wie wahre Meinungen. Einerseits könnten sie einen wahren Teil enthalten, andererseits aber würden sie dazu beitragen, dass wahre Meinungen nicht zu dogmatischen Glaubenssätzen verkommen. Einzig durch das Aufeinandertreffen von unterschiedlichen, teilweise wahren Meinungen könne sichergestellt werden, dass die Wahrheit in einer Diskussion ans Licht komme (Mill 1859, 258). Mit dieser Verteidigung komme auch zum Ausdruck, dass für Mill die bloße Wahrheit einer Meinung nicht ausreiche, um ihren vollen Nutzen zu entfalten. Eine wahre Meinung zu haben sei nur nützlich, wenn sie durch deliberative Prozesse zustande komme (Niesen 2015, 40). Es ist sowohl im Interesse der Sprecherin als auch des Hörers, dass falsche und teilweise wahre Meinungen in einer Diskussion vorgebracht werden können.

Hiermit konnte nun aufgezeigt werden, dass Mill mit dem Begriff „Meinung“ keine *expressiven Sprechakte* gemeint haben kann, wie dies Jacobson (2000, 287) vorschlägt. Die Rechtfertigung der *Meinungsäußerungsfreiheit* orientiert sich an Propositionen, welche für andere zugänglich sind. Die Wahrheit des Inhalts eines expressiven Sprechaktes wird immer angenommen und kann daher nicht als Interpretation einer Tatsachenbehauptung verstanden werden. Es ist also anzunehmen, dass lediglich *assertive* Sprechakte von der Rechtfertigung der *Meinungsäußerungsfreiheit* von Mill dem ersten Bereich der Freiheitssphäre zugeordnet werden können. Die Abwesenheit der Be-

sprechung von *expressiven* Sprechakten in Mills Verteidigung der *Meinungsäußerungsfreiheit* kann besser dadurch erklärt werden, dass sie zum zweiten Bereich der Freiheitssphäre gehören.

2.1.2 Das Wissen: ein sozialer Prozess

James Zappen (1993) argumentiert dafür, die beiden Werke, *A System of Logic* (Mill 1843a und Mill 1843b) und *On Liberty*, als das logische, resp. rhetorische Gegenüber des jeweils anderen anzusehen. Er zeigt auf, dass Mill für die Probleme seiner induktiven Logik nicht nur in *A System of Logic*, sondern auch in *On Liberty* Lösungen präsentiert (Zappen 1993, 191). Diese enge Verbindung zwischen den beiden Schriften ermöglicht einen weiteren Einblick in die Rolle der *Meinungsäußerungsfreiheit* für die *Freiheit des Denkens und der Diskussion*. Die Rechtfertigung von Wissen und die Entwicklung eines freien Denkprozesses sind demnach inhärent soziale Prozesse.

Mill versucht in *A System of Logic* (1843a, 184) aufzuzeigen, dass deduktive Logiken ein ihnen innewohnendes epistemisches Problem haben. Deduktive Schlüsse würden demnach immer einen Zirkelschluss aufweisen. Mit universellen Aussagen könne man streng genommen partikuläre Aussagen nicht beweisen, da nur solche Aussagen abgeleitet werden könnten, die bereits von der universellen Aussage als vorausgesetzt angenommen würden (Mill 1843a, 184). Mill sucht daher den Ausweg in einer induktiven Logik, die die Grundlage für menschliche Urteile bilde. Er glaubt, dass eine Beweisführung auf zwei Prozessen beruht: Der erste, die Induktion, diene zur Gewinnung einer universellen Proposition aus partikulären Aussagen. Der zweite, die Deduktion, biete die Interpretation der universellen Proposition (Mill 1843a, 203). Dies stellt ihn nun vor ein Problem: Wenn es wahr ist, dass von einer Partikularaussage auf eine Universalaussage geschlossen wird, dann kann nicht von der Sicherheit oder Gültigkeit der daraus gewonnenen Universalaussage ausgegangen werden. Die aus dieser Aussage gewonnenen, partikulären Aussagen können aufgrund der Unsicherheit der universellen Aussage nicht mit Sicherheit abgeleitet werden (Mill 1843a, 306–7). Wenn er für dieses Problem keine Lösung findet, ist sein epistemisches Unterfangen zum Scheitern verurteilt.

Das erste Problem löst er in *A System of Logic*, indem er die Verallgemeinerung durch die Erfahrung ersetzt. Sowohl die Erfahrung der Person selbst als auch die kumulative Erfahrung der gesamten Menschheit dient somit als Korrektiv für Universalaussagen (Mill 1843a, 319). Das zweite Problem, die Ableitung neuer partiku-

lärer Aussagen, bleibt jedoch weiterhin bestehen. Die Lösung dafür wird laut Zappen (1993, 195) in *On Liberty* vorgebracht. Mill halte hier fest, dass die Erfahrung (ob kumulativ oder nicht) alleine nicht ausreiche, um Universalaussagen zu stützen, weshalb sie durch eine weitere Vorrichtung ergänzt werden müsste (Zappen 1993, 196). Zu diesem Zweck biete er eine Rhetorik der öffentlichen Diskussion, welche sicherstellen soll, dass Argumentationen und Beweisführungen korrekt seien. Die ständige und andauernde Korrektur soll mithilfe der Sokratischen Fragemethode geschehen. Diese auf Gesprächen basierende Methode müsse in allen Wissensbereichen (ausser in der Mathematik) als Standard gelehrt und angewandt werden (Zappen 1993, 195–96). Voraussetzung für das Gelingen sei erstens der Zugang zu entgegengesetzten Meinungen, zweitens eine Verteidigung der eigenen Meinung gegen konkurrierende und drittens ein Profitieren durch die Korrektur bzw. eine Vervollständigung der eigenen Meinung dank dieser Auseinandersetzung (Cherwitz und Hikins 1979, 16).

In diesen drei Voraussetzungen erkennt man den Nutzen von falschen und nur teilweise wahren Meinungen für Sprecher und Hörerinnen: Durch ständige Untersuchung und Hinterfragung von Meinungen werden einer Person die Begründungen für wahre Meinungen präsent und teilweise wahre Meinungen können angepasst und verbessert werden. Dieser Prozess ist inhärent sozial und bedingt daher auch, dass sich möglichst viele Personen daran beteiligen können, um einen möglichst grossen Pool an Meinungen und Erfahrungen zu erhalten.

Dies untermauert noch einmal die Ansicht, dass es sich bei den Sprechakten, die zum ersten Bereich der Freiheitssphäre gezählt werden, um assertive Sprechakte handeln muss. Das Generieren von Wissen findet in der öffentlichen und privaten Diskussion statt und muss daher mittels Meinungen geschehen, deren Wahrheitsgehalt sinnvoll diskutiert werden kann. Zudem wird durch das Korrektiv der kumulativen und eigenen Erfahrung und durch die Sokratische Methode ein Hinweis auf die Rolle der Meinungsäusserungsfreiheit gegeben. Auch Jacobson (2016, 449–52) geht auf diese ein. Er bespricht sie aus der Perspektive von Hörern, Sprecherinnen und im Hinblick auf Mills Demokratieverständnis. Durch das Auslassen der *Wahrheitsfunktionalität*⁷ von Meinungen ist die Analyse jedoch unvollständig.

2.1.3 Die Rolle: Die denkende Person

Seana Shiffrin (2014) merkt in ihrem Werk *Speech matters: on lying, morality, and the law* an, dass Mills Theorie der *Meinungsäusserungsfreiheit* meist zu jenen gezählt wird, die auf die Wahrheitsfindung fokussieren. Sie weist jedoch auch Ähnlichkeiten zu Theorien auf, welche die Person als denkendes Wesen ins Zentrum stellen (Shiffrin 2014, 83, Anmerkung 4). Denkerbasierte Theorien begründen die *Meinungsäusserungsfreiheit* auf den Interessen und Bedürfnissen der denkenden Person. Den Ausgangspunkt bilden bei Shiffrin (2014, 86) die Annahmen, dass Menschen individuelle Akteure mit signifikanten (jedoch wesentlich unvollkommenen) rationalen, emotionalen sowie moralischen Fähigkeiten und Empfindungsvermögen sowie Wahrnehmungsfähigkeiten sind, die sich überschneiden und einander beeinflussen. Der Besitz und die Ausübung dieser Fähigkeiten konstituieren den Kern dessen, was Menschen an sich selbst wertschätzen. Und schliesslich konstituiere das besondere Verhältnis, die Entwicklung und die Ausübung dieser Fähigkeiten gemeinsam mit den physischen Attributen den Menschen als das distinkte Individuum, das es in seinem geteilten historischen und sozialen Milieu ist (Shiffrin 2014, 86). Um den notwendigen Rahmen zu untersuchen, welcher diese Fähigkeiten und deren sinnvolle Ausübung garantiert, listet Shiffrin acht Interessen auf (Shiffrin 2014, 86–88). Diese können auch in Mills *On Liberty* wiedergefunden werden. Aus Platzgründen werden hier nur jene zwei Interessen besprochen, welche für das Argument von Bedeutung sind. (Shiffrin 2014, 82–88)

Das erste Interesse besteht in einer „entwickelten Fähigkeit für praktisches und theoretisches Denken. Eine Denkerin hat ein grundlegendes Interesse daran, ihre mentalen Fähigkeiten dazu zu entwickeln, aufnahmefähig, verständnisvoll und ansprechbar gegenüber Begründungen und Fakten praktischer und theoretischer Gedanken zu sein, das heisst, sich des Wahren, Falschen und Unbekannten bewusst zu sein und adäquat reagieren zu können“ (Shiffrin 2014, 86).

Dieses Interesse ist in der Verteidigung der *Meinungsäusserungsfreiheit* von Mill wiedererkennbar: Mill fordert sie, um genau diese Fähigkeiten ausbilden und am Leben erhalten zu können (Mill 1859, 243). Zudem spiegelt sich dieses Interesse in der Forderung, dass die Sokratische Methode als rhetorische Vorrichtung zur Rechtfertigung von Wissen in allen wissenschaftlichen und moralischen Belangen (ausser in der Mathematik) installiert werden soll (Mill 1859, 252).

Ein weiteres zentrales Interesse eines Denkers ist eine „angebrachte Anerkennung und Behandlung. Jeder

⁷ Wahrheitsfunktionalität wird in diesem Kontext nicht im technisch-logischen Gebrauch verwendet, sondern es wird darunter verstanden, dass der propositionale Gehalt von Äusserungen sinnvoll als wahr oder falsch bezeichnet werden kann.

Denker hat ein grundlegendes Interesse von anderen Akteuren als die Person anerkannt zu werden, die sie ist, und dass sie von anderen moralisch gut behandelt wird.“ (Shiffrin 2014, 88)

Mill (1974, 260–75) bespricht im dritten Kapitel, *Über Individualität als eins der Elemente der Wohlfahrt*, den zweiten Bereich der Freiheitssphäre, welcher die Freiheit „des Geschmacks und der Studien“ betrifft. Diese Freiheit diene dazu, einen Lebensplan entwerfen zu können, der dem Charakter der Person entspricht, sofern dies anderen keinen Schaden zufüge. Behinderungen durch Dritte seien demnach nicht gestattet, auch wenn sie die Lebensweise für „verrückt, verderbt oder falsch halten“ (Mill 1974, 226). Es sei notwendig, dass Menschen „[...] nach ihrer Meinung [...] handeln und sie im Leben umsetzen, ohne von ihren Mitmenschen durch physischen oder moralischen Zwang daran verhindert zu werden [...]“ (Mill 1974, 260).

Beleidigende Sprechakte können weder durch das erste noch das zweite Interesse als schützenswert eingestuft werden. Beleidigende und verletzende Äusserungen tragen nicht zur Entwicklung intellektueller Fähigkeiten bei. Dem anderen Interesse wirken sie entgegen. Durch solche Äusserungen wird die Fähigkeit anderer unterminiert, einen selbstbestimmten und freien Lebensplan zu entwerfen.

2.2 Progressive Diskussionskultur

Die drei oben angesprochenen Aspekte der Verteidigung der *Meinungsäusserungsfreiheit* von Mill bilden gemeinsam mit der Interpretation des Mill'schen Liberalismus von Jacobson die Grundlage eines Argumentes für eine *progressive Diskussionskultur*. Entgegen Jacobsons Auffassung wird aufgezeigt, wie anhand von Mills Argumenten beleidigende und abwertende Äusserungen nicht als schützenswerte *Meinungen* betrachtet werden müssen und die Gesellschaft daher weder rechtlich noch moralisch in der Pflicht steht, solchen Äusserungen eine (öffentliche) Plattform zu bieten.

Wie oben besprochen wird für dieses Argument vom Mill'schen Liberalismus nach Jacobson ausgegangen. Das heisst, dass Mill für eine Freiheitslehre argumentiert, welche grundlegend aus zwei Prinzipien besteht: Dem Schädigungsprinzip und der Freiheitssphäre. Das Schädigungsprinzip besagt, dass „der einzig gute Grund für eine Einmischung anderer ist, die Schädigung Dritter zu verhindern“ (Jacobson 2000, 290). Es ist dabei noch der Einschränkung unterworfen, dass nur Handlungen davon betroffen sein können, die nicht auf Freiheitsrechten der Freiheitssphäre basieren. Die Freiheitssphäre umfasst drei Bereiche: Das innere Feld

des Bewusstseins, die Freiheit des Geschmacks und der Studien sowie die Freiheit des Zusammenschlusses. Handlungen, welche auf diesen drei Freiheitsbereichen beruhen, dürfen nur eingeschränkt werden, wenn durch deren Ausübung eine Verpflichtung gegenüber einer anderen Person verletzt wird. Bis hierhin kann Jacobson noch zugestimmt werden.

Aufgrund der Tatsache, dass Mill die Zuordnung der *Meinungsäusserungsfreiheit* zum ersten Bereich anhand der Wahrheitsfunktionalität von *Meinungen* vornimmt, können einzig *assertive* Sprechakte als dieser Gruppe zugehörig verstanden werden. Dies legt auch deren Rolle in der Entwicklung einer freien Bürgerin nahe. Die restlichen Sprechakte werden somit dem zweiten und dritten Bereich zugeordnet, welche explizit auf nicht-schädigende Handlungen begrenzt werden. Assertive Sprechakte können, wie Jacobson (2000, 286–87) korrekt interpretiert, aufgrund des Kontexts der Äusserung unterbunden werden. Dies kann geschehen, wenn der Kontext der Äusserung eine Verletzung der Pflicht gegenüber einer anderen Person darstellt (Mill 1859, 281). Zudem legt Mill diesen Überlegungen die Interessen einer Person (verstanden als Denkerin) zugrunde, das heisst, einer Person mit rationalen, emotionalen, moralischen Fähigkeiten und Empfindungsvermögen sowie Wahrnehmungsfähigkeiten (Shiffrin 2014, 83, Anmerkung 4). Dies führt dazu, dass der Ausschluss aus epistemischen Prozessen als sowohl moralisches wie auch epistemisches Übel verstanden werden muss: Beleidigende und abwertende Äusserungen aufgrund von Identitätsmerkmalen unterminieren die Fähigkeit und Möglichkeit an epistemischen Prozessen teilnehmen zu können. Daraus wird ersichtlich, dass eine Gesellschaft nicht in der Pflicht ist, solchen Äusserungen eine (öffentliche) Plattform zu gewähren. Denn die Unterminierung der Fähigkeit und Möglichkeit der Teilnahme an epistemischen Prozessen konstituiert hinreichend eine Verletzung der Pflichten gegenüber anderen Denkern.

Fazit

Ziel dieses Essays war es, aufzuzeigen, dass auf der Grundlage des Mill'schen Argumentes für die *Meinungsäusserungsfreiheit* für eine *progressive Diskussionskultur* argumentiert werden kann. Dazu wurden in einem ersten Schritt die drei grundlegenden Thesen des Mill'schen Liberalismus nach Jacobson (2000) erläutert (1). Um diese nachvollziehen zu können, wurden daraufhin die zentralen Prinzipien dieser Position beschrieben. Das Schädigungsprinzip (1.1) beschreibt demnach, dass ein Eingriff in die Handlungsfreiheit

einer Person durch die Gesellschaft nur dann legitim sei, wenn dadurch die Schädigung Dritter verhindert werden könne (Jacobson 2000, 290). Davon ausgenommen seien Handlungen, welche auf Freiheitsrechten basieren. Die Freiheitssphäre (1.2) umfasst drei Bereiche, welche vor einer Einschränkung durch die Gesellschaft geschützt sind. Alle drei werden auf die gleiche Art und Weise gerechtfertigt und beinhalten auch schädigende Handlungen (im weitesten Sinn). Da die *Freiheit, Meinungen zu äussern*, zur ersten Freiheitssphäre gehört, ist für Jacobson die Annahme bestätigt, dass Mill für eine *uneingeschränkte Meinungsäusserungsfreiheit* (1.3) argumentiert habe. Daraufhin wurde diese Analyse von Jacobson kritisiert (2) und ergänzt. Demnach wird von Jacobson fälschlicherweise behauptet, dass nur *expressive Sprechakte* immun vor Eingriffen seien, obwohl die Verteidigung der *Meinungsäusserungsfreiheit* von Mill keinen solchen Schluss zulässt, da einzig von *Meinungen* im Sinne von *assertiven Sprechakten* die Rede sein kann (2.1.1 und 2.1.2). Der letzte Aspekt ergänzt die Analyse um die Erkenntnis, dass Mill eine denkerbasierte *Meinungsäusserungsfreiheit* vorgebracht hat, welche auch eine angebrachte Anerkennung und Behandlung der Person als *denkendes* Wesen beinhaltet (2.1.3). Dies lässt den Schluss zu, dass beleidigenden und abwertenden Äusserungen keine (öffentliche) Plattform gewährt werden muss, da dies die Beteiligung von Personen bei epistemischen Prozessen unterminiert (2.2).

Die Argumentation von Mill ist überzeugend und nachvollziehbar. Es ist daher auch verständlich, dass sie noch heute Gegenstand philosophischer und politischer Diskussionen ist. Für künftige Projekte könnte beispielsweise auch untersucht werden, ob auf der Basis dieser Argumente auch dafür argumentiert werden kann, dass eine Gesellschaft geradezu moralisch dazu verpflichtet ist, beleidigenden und abwertenden Äusserungen keine (öffentliche) Plattform zu gewähren.

Literatur

- Austin, John Langshaw. 1962. *How to Do Things with Words*. Herausgegeben von James Opie Urmson und Marina Sbisa (1975). Cambridge (MA), Harvard University Press.
- Cherwitz, Richard A., und James W. Hinkins. 1979. "John Stuart Mill's on liberty: Implications for the epistemology of the new rhetoric." *Quarterly Journal of Speech* 65 (1), 12–24.
- Jacobson, Daniel. 2000. "Mill on Liberty, Speech, and the Free Society." *Philosophy & Public Affairs* 29 (3), 276–309.
- Jacobson, Daniel. 2016. "Mill on Freedom of Speech." In *A Companion to Mill*, herausgegeben von Christopher Macleod und Dale E. Miller, 440–453. Hoboken: Wiley.
- Macleod, Christopher und Dale E. Miller. 2016. *A Companion to Mill*. Hoboken: Wiley.
- Mill, John Stuart. 1825. "Law of Libel and Liberty of the Press." In *The Collected Works of John Stuart Mill Volume XXI*, herausgegeben von John M. Robson (1984). Toronto and Buffalo, University of Toronto Press.
- Mill, John Stuart. 1843a. "A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation. Part I." In *The Collected Works of John Stuart Mill Volume VII*, herausgegeben von John M. Robson (1973). Toronto and Buffalo: University of Toronto Press.
- Mill, John Stuart. 1843b. "A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation. Part II." In *The Collected Works of John Stuart Mill Volume VIII*, herausgegeben von John M. Robson (1974). Toronto and Buffalo: University of Toronto Press.
- Mill, John Stuart. 1859. "On Liberty." In *The Collected Works of John Stuart Mill Volume XVIII*, herausgegeben von John M. Robson (1977). Toronto and Buffalo: University of Toronto Press.
- Mill, John Stuart. 1859. *On Liberty: Über Freiheit: Englisch/Deutsch*. Herausgegeben von Bernd Gräfrath (2009). Übersetzt von Bruno Lemke. Stuttgart: Reclam.
- Mill, John Stuart. 1861. "Utilitarianism." In *The Collected Works of John Stuart Mill Volume X*, herausgegeben von John M. Robson (1969). Toronto and Buffalo: University of Toronto Press.
- Niesen, Peter. 2015. "Über Die Freiheit Des Denkens Und Der Diskussion." In *John Stuart Mill: Über Die Freiheit*. Herausgegeben von Michael Schefczyk und Thomas Schramme, 33–54. Berlin: De Gruyter.
- Searle, John Rogers. 1999. *Mind, language and society: doing philosophy in the real world*. London: Weidenfeld; Nicolson.
- Shiffrin, Seana Valentine. 2014. *Speech matters: on lying, morality, and the law*. Princeton, New Jersey: Princeton University Press.
- Zappen, James P. 1993. "The Logic and Rhetoric of John Stuart Mill Author." *Philosophy & Rhetoric* 26 (3): 191–200.

Marco R. Schori studiert Wissenschaftsphilosophie und Klassische Philologie mit Schwerpunkt Griechisch im 1. Semester des Masters. Er interessiert sich besonders für antike Philosophie, Geschichte und Philosophie der Wissenschaften, der Logik und der Mathematik.

Der Schutz fundamentaler Rechte und die Demokratie am Arbeitsplatz

Richard Arnesons Theorie demokratischer Rechte angewandt auf die Arbeitswelt

Einleitung

Am Arbeitsplatz gelten Regeln. Üblicherweise werden sie von Vorgesetzten festgelegt, von Arbeitnehmer*innen eingehalten und Regelbrüchen folgen Sanktionen. Ähnlich verhält es sich mit strategischen Entscheidungen in Unternehmen: Arbeitnehmer*innen führen aus, an der Entscheidung beteiligt sind sie nicht. Zwar haben die meisten Angestellten in irgendeiner Form ihren Zuständigkeitsbereich und da eine gewisse Gestaltungsfreiheit; grundlegend darüber mitbestimmen, welche Regeln gelten, wer die unternehmerischen Entscheidungen trifft und zu welchen Konditionen gearbeitet wird, können aber die wenigsten. Über Arbeitnehmer*innen wird entschieden und wer über sie entscheidet, liegt ausserhalb ihrer Kontrolle. So beschrieben, gleichen Arbeitsplätze kleinen Diktaturen und man mag versucht sein, die Abwesenheit demokratischer Strukturen in unserer Arbeitswelt zu kritisieren. Doch haben Arbeitnehmer*innen ein moralisches Recht auf demokratische Mitbestimmung am Arbeitsplatz? Nach einigen Jahrzehnten des Schattendaseins erhält diese Frage über die akademische Debatte hinaus nun wieder mehr Aufmerksamkeit (Frega, Herzog und Neuhäuser 2019, 1–2).

Eine Strategie, um für ein Recht auf demokratische Mitbestimmung am Arbeitsplatz zu argumentieren, besteht darin, Argumente für ein Recht auf Mitbestimmung in staatlichen Angelegenheiten auf die Arbeitswelt zu übertragen. Hierfür wird zunächst die Annahme getroffen, dass Staaten demokratisch organisiert sein sollten. Typischerweise wird diese Annahme an näher zu bestimmende charakteristische Eigenschaften von Staaten geknüpft, etwa die Aufgabe von Staaten, öffentliche

Regeln festzulegen, die im gemeinsamen Interesse der Bürger*innen sein sollten. Anschliessend wird zu zeigen versucht, dass dieselben oder—in relevanter Weise—ähnliche Eigenschaften auch Unternehmen auszeichnen. Daraus wird dann geschlossen, dass die Gründe, die für ein Recht auf demokratische Mitsprache im Staat sprechen, auch eine demokratische Organisation von Unternehmen verlangen (Cohen 1989, 27).

Angelehnt an dieses Schema soll in dieser Arbeit versucht werden, die Begründung eines Rechts auf Demokratie, die Richard Arneson in „Democratic Rights at National and Workplace Levels“ (1993) gibt, auf die Arbeitswelt zu übertragen. Arneson vertritt darin die These, demokratische Mitspracherechte hätten die Funktion, andere, fundamentalere Rechte zu schützen, und seien gerechtfertigt, insofern sie dieser Funktion gerecht würden (Arneson 1993, 118). Eine kritische Diskussion dieses Ansatzes hat im Umfang dieser Arbeit keinen Platz und deshalb wird er als theoretische Rahmenannahme vorausgesetzt. Das Ziel ist nicht, die These der Schutzfunktion demokratischer Rechte zu verteidigen, sondern abzutasten, inwieweit sie sich auf die Frage der Demokratie am Arbeitsplatz anwenden lässt. Nun hängt aber die Stärke des Analogieschlusses vom Recht auf Demokratie im Staat auf ein Recht auf Demokratie am Arbeitsplatz eng mit der Plausibilität der vorausgehenden Argumentation für die Demokratie im Staat zusammen. Da diese hier nicht evaluiert wird, ist die Tragweite dieser Arbeit entsprechend begrenzt.

In einem ersten Schritt wird die demokratietheoretische Position Arnesons knapp umrissen und auf die Arbeitswelt angewandt. Dabei soll gezeigt werden,

dass die These der Schutzfunktion demokratischer Rechte zur Unterstützung eines Rechts auf Demokratie am Arbeitsplatz herangezogen werden kann. Die Analogie zwischen Staaten und Unternehmen ist aus verschiedenen Gründen umstritten und Argumente, die sich auf sie berufen, wurden von einigen Philosoph*innen—darunter Arneson selbst—kritisiert (Frega, Herzog und Neuhäuser 2019, 2–3). In einem zweiten Schritt wird deshalb Arnesons zentraler Einwand gegen den Analogieschluss dargestellt und vor dem Hintergrund seiner demokratietheoretischen Position diskutiert, um den Einwand zu entkräften. Schliesslich, um die Tragfähigkeit der Überlegungen zu testen, werden mögliche weitere Einwände geprüft.¹

1. Arnesons Theorie demokratischer Rechte und die Demokratie am Arbeitsplatz

Um über einen Analogieschluss von einem Recht auf Demokratie im politischen Bereich zu einem solchen im wirtschaftlichen Bereich zu gelangen, muss, wie gesagt, zuerst für das erste Recht argumentiert werden. Ein Recht auf Demokratie in staatlichen Angelegenheiten lässt sich argumentativ auf zwei grundlegend verschiedenen Wegen erreichen. Einerseits kann man die demokratischen *Verfahren* betrachten und argumentieren, diese seien gerecht, beispielsweise weil ideale demokratische Prozesse allen Menschen dieselbe Mitsprache gewähren und dadurch ein Prinzip der Gleichheit verwirklichen würden. Diese Perspektive legt den Fokus auf die Struktur der Entscheidungsverfahren und sucht Gerechtigkeit in Eigenschaften der Verfahren selbst. Andererseits kann man sich den *Ergebnissen* zuwenden, die demokratische Verfahren erzielen.

Arneson beschreitet diesen zweiten Weg. Seine Position baut auf die normative Grundannahme, niemand habe ein Recht auf eine soziale Position, die die Ausübung von Macht über andere Menschen ohne deren Einwilligung beinhalte—es sei denn, die Machtausübung gewährleiste den Schutz der fundamentalen Rechte der betroffenen Menschen. Nun sei die Teilnahme an demokratischen Entscheidungen eine Form von Machtausübung über andere und ohne die Annahme, demokratische Verfahren würden dem Schutz fundamentaler Rechte dienen, sei ein Recht auf demokratische Mitsprache nicht plausibel. Ein Recht auf Demokratie, das auf verfahrensinternen Gerech-

tigkeitskriterien statt auf den erzielten Resultaten der Verfahren beruht, schliesst Arneson aus. Politische Prozesse seien anhand der Ergebnisse zu bewerten, die man von ihnen vernünftigerweise erwarten könne; der primäre Standard, um die Ergebnisse zu beurteilen, sei der Schutz fundamentaler Rechte (Arneson 1993, 118–25).

Um von diesen Überlegungen ausgehend für ein Recht auf Demokratie zu argumentieren, muss Arneson zeigen, dass demokratische Systeme tatsächlich besser als alternative Staatsformen diesen Schutz fundamentaler Rechte gewährleisten. Arneson rechtfertigt diese empirische These mit dem Hinweis auf das Problem der Selektion von Personen für politische Machtpositionen. Er argumentiert, es gebe kein nicht-demokratisches Verfahren, das zuverlässig die Auswahl kompetenter Personen gewährleiste, die ihre Macht nutzen würden, um die Rechte aller zu sichern. Historische Erfahrungen würden zeigen, dass gerade autokratische Selektionsprozesse in dieser Hinsicht ohne Aussicht auf Erfolg seien. Garantieren, dass die fundamentalen Rechte gesichert seien, könne allerdings kein Selektionsverfahren und eine Ideallösung sei entsprechend nicht verfügbar. Doch vorausgesetzt, jemand müsse eine Regierung wählen, sei die Frage, ob diese politische Macht demokratisch unter allen oder nur unter einem Teil der Bürger*innen verteilt sein sollte. Vor dem Hintergrund der Abwesenheit einer Ideallösung hält er die Demokratie für die zweitbeste Lösung (Arneson 1993, 122–23).

Politische Wahlen sind diesem Verständnis nach also ein effektives Mittel, Grundrechte zu schützen. Dies setzt voraus, dass eine Mehrheit der Beteiligten eine allfällige Beschneidung mindestens *ihrer* Rechte verhindern würde. Was aber ist in Fällen, in denen Mehrheitsentscheidungen oder Entscheidungen demokratisch gewählter Körperschaften den Schutz fundamentaler Rechte mancher Menschen bedrohen? Dann, so Arneson, seien die Grenzen demokratischer Rechte erreicht und deren Einschränkung—beispielsweise durch ein Verfassungsgericht—zulässig.²

Hier ist anzumerken: Zu diesem letzten Schluss gelangen kann Arneson, weil er demokratische Rechte ausschliesslich instrumentell versteht und keine verfahrensinternen Gerechtigkeitskriterien zulässt.

¹ Ein Einwand, der diskutiert wird, ist lose angelehnt an Argumente von Daniel Jacob und Christian Neuhäuser (Jacob und Neuhäuser 2018, 927–44).

² Arneson vertritt nicht die These, der Schutz fundamentaler Rechte sei der einzige zulässige Grund, die Gestaltungsfreiheit gewählter Regierungen durch Einrichtungen wie ein Verfassungsgericht zurückzubinden. Er lässt auch die Möglichkeit für anderweitig motivierte Einschränkungen offen, etwa um Mehrheitsinteressen zu schützen, solange der Schutz fundamentaler Rechte gewährt bleibt (Arneson 1993, 124–25).

Folgt man seiner Verteidigung der Demokratie, ist leicht einzusehen, weshalb demokratische Rechte eingeschränkt werden dürfen, wenn sie jene primäre Schutzfunktion nicht erfüllen, die sie überhaupt erst wertvoll macht. Ansätze, die demokratische Verfahren an sich als wertvoll beschreiben, würden mehr argumentative Arbeit verlangen, um zum selben Schluss zu gelangen. Arneson weist die Vorstellung verfahrensinthetischer Gerechtigkeit allerdings nicht nur im Kontext der demokratiethoretischen Debatte zurück—er erklärt sie grundsätzlich für unzulänglich. Gerechte Verfahren seien nicht als Verfahren gerecht und könnten ausschliesslich aufgrund ihrer Ergebnisse moralisch beurteilt werden. Rechte, die Individuen in Verfahren hätten, seien bloss instrumenteller Natur und dazu da, moralisch wünschenswerte Ergebnisse zu erzielen (Arneson 1993, 120–29).

Nutzen wir diese Position als Schablone und legen sie auf den Untersuchungsgegenstand der Arbeitswelt, müssen wir daher nicht fragen, ob die Entscheidungsprozesse am Arbeitsplatz selbst fair sind oder inwiefern intrinsische Beteiligungsrechte gewährleistet oder verletzt werden. Stattdessen müssen wir nach den Ergebnissen fragen, die durch die Entscheidungsverfahren erzielt werden. Und damit sind wir angelangt bei der Frage: Inwiefern spricht Arnesons Ansatz für ein Recht auf Demokratie am Arbeitsplatz?

Wie im Staat gibt es am Arbeitsplatz Positionen mit Macht über andere Menschen. Personen in diesen Machtpositionen treffen Entscheidungen für andere—manchmal gegen den Willen von Arbeitnehmer*innen und oft ohne deren explizite Zustimmung. Und wie im Staat gibt es am Arbeitsplatz Selektionsverfahren für die Besetzung dieser Positionen. Diese müssen, da wir ohne verfahrensinthetisches Gerechtigkeitskriterium operieren, an ihren Ergebnissen gemessen werden. Weil jene, die in den Selektionsverfahren für Führungspositionen in Unternehmen mitentscheiden, in einer Position der Macht über andere sind, scheint der Schutz fundamentaler Rechte auch hier als ein Massstab der Beurteilung der Ergebnisse angemessen. Nun gibt es—wenn es kein nicht-demokratisches Verfahren gibt, das die Auswahl geeigneter Personen für staatliche Machtpositionen gewährleistet—keinen Grund, anzunehmen, die Selektion von Personen in unternehmensinterne Machtpositionen, die Untergebenen gegenüber nicht demokratisch rechenschaftspflichtig sind, sichere den Schutz der Rechte dieser Untergebenen zuverlässig. Auf den ersten Blick scheint damit auch die Demokratie am Arbeitsplatz als zweitbeste Lösung betrachtet werden zu können.

Diese Argumentation geht davon aus, dass es fundamentale Rechte gibt, die am Arbeitsplatz bedroht sein können. Auch Arnesons Verteidigung der Demokratie setzt die Existenz schützenswerter Grundrechte voraus. Im Einzelnen schlägt er unter anderem individuelle Freiheitsrechte und egalitäre Rechte auf materielle Ressourcen als fundamentale Rechte vor, lässt aber gleichzeitig den Katalog an Grundrechten offen für Ergänzungen: Unser epistemischer Zugang zu fundamentalen Rechten sei beschränkt und es sei davon auszugehen, dass Menschen künftig andere inhaltliche Vorstellungen fundamentaler Rechte haben würden. Die grundsätzliche Existenz solcher Rechte sei aber breit akzeptiert und seine Argumentation würde sich auch mit einem alternativen Set an Rechten nachvollziehen lassen (Arneson 1993, 118–19). Einigermassen unabhängig davon, welches Set an Rechten letztlich postuliert wird, scheint klar, dass es in der Macht von Staaten liegt, diese Rechte zu missachten. Für Arnesons Argumentation ist das Fehlen eines klar abgeschlossenen Rechtekatalogs daher unproblematisch. Ob dasselbe für Unternehmen gilt, ist dagegen weniger klar. Deshalb muss kurz etwas dazu gesagt werden, welche Rechte denn konkret durch ein Recht auf Demokratie am Arbeitsplatz zu schützen sind.

Verschiedene Kandidaten für Grundrechte, die im Zusammenhang mit der Arbeitswelt geschützt werden müssen, wie beispielsweise Rechte auf materielle Ressourcen, sind denkbar. Ein Recht, das für die Zwecke dieser Arbeit postuliert werden soll, ist das Recht auf körperliche und psychische Unversehrtheit. Weshalb wird hier gerade mit diesem Recht gearbeitet? Einerseits deshalb, weil physische und psychische Gesundheit eine Voraussetzung für den vollen Genuss einer Reihe anderer Rechte ist, und andererseits auch aus der Hoffnung heraus, dass ein solches Recht einigermaßen unstrittig sein dürfte. Was genau körperliche und psychische Unversehrtheit bedeutet, ist vage, und die Grenzen im Umgang mit Menschen, die durch ein Recht auf Unversehrtheit etabliert würden, sind möglicherweise von Person zu Person unterschiedlich und nicht objektivierbar.³ Auch ist nicht gänzlich klar, ob umfangreiche Ansprüche auf positive Leistungen aus

3 Dies bedroht die Argumentation dieser Arbeit jedenfalls nicht; denn würde sich tatsächlich herausstellen, dass abschliessende objektive Kriterien nicht verfügbar sind, könnte dies das von der Schutzfunktion demokratischer Rechte ausgehende Argument für die Demokratie am Arbeitsplatz zusätzlich stärken. Sieht man nämlich die Deutungshoheit darüber, wann genau am Arbeitsplatz eine Verletzung des Rechts auf körperliche und psychische Unversehrtheit vorliegt, in Teilen bei den betroffenen Personen selbst, liefert dies einen weiteren Grund, demokratische Kontrolle durch die betroffenen Personen hierarchischen Kontrollmechanismen vorzuziehen.

einem solchen Recht resultieren würden. Ich denke aber, es ist einsehbar, dass in der Arbeitswelt die physische und körperliche Unversehrtheit manchmal bedroht sein kann, und so gehe ich davon aus, dass die Argumentation dieser Arbeit nachvollziehbar ist, auch wenn unterschiedliche Konzeptionen von körperlicher und psychischer Unversehrtheit vorhanden sein mögen.

Nun kann man sich vielfältige Vorgänge am Arbeitsplatz denken, die das Recht auf körperliche und psychische Unversehrtheit bedrohen können: Etwa harte körperliche Arbeit, Kontakt mit giftigen Substanzen, Nacharbeit, Crunch-Time, Überstunden, Stress oder Mobbing. Hinzu kommt die Bedeutung der Arbeit für die Freiheit, ein gesundes und nicht von Existenzsorgen geplagtes Leben zu führen. Die Hoffnung ist nicht unbegründet, dass demokratische Kontrolle durch Arbeitnehmer*innen ein effektives Mittel wäre, ihren Ansprüchen mehr Durchschlagskraft zu verleihen. Immerhin würde eine demokratische Rechenschaftspflicht der Vorgesetzten gegenüber den Angestellten die Machtverhältnisse am Arbeitsplatz in einer Art und Weise umdrehen, die Vorgesetzte stärker zwingen würde, die Anliegen ihrer Untergebenen zu berücksichtigen. Zumindest gibt es, solange wir davon ausgehen, dass demokratische Kontrolle im Staat ein effektives Mittel ist, die Rechte der Bürger*innen zu schützen, keine Gründe zur Annahme, dass die Demokratie am Arbeitsplatz nicht eine ähnliche Funktion erfüllen könnte. Doch welche konkrete Form von Demokratie am Arbeitsplatz verlangt diese Argumentation? Da sich die Argumentation innerhalb einer ergebnisorientierten Perspektive bewegt, ist dies weitgehend eine empirische Frage, wobei die Möglichkeit besteht, dass für unterschiedliche Formen von Unternehmen unterschiedliche Formen der demokratischen Einbindung der Arbeitnehmer*innen geeignet sind. Schliesslich kann—genau wie die Behauptung eines Rechts auf Demokratie noch nicht allzu viel darüber sagt, welche konkreten Mitspracherechte die Bürger*innen eines Staats haben sollten—die Forderung nach einer Demokratisierung der Arbeitswelt verschiedenartig interpretiert werden. Die Frage, welche konkreten Mitspracherechte ein Recht auf Demokratie am Arbeitsplatz verlangt, wäre alleine Stoff für eine ausführliche Diskussion.

Deshalb soll hier unter „Demokratie am Arbeitsplatz“ grob die Übergabe von Entscheidungsmacht an die Arbeitnehmer*innen eines Unternehmens oder an Personen, die durch die Arbeitnehmer*innen gewählt sind, verstanden werden. Eine demokratische

Rechenschaftspflicht der Vorgesetzten gegenüber der Angestellten bildet die Mindestforderung und analog zu Arnesons Argumentation ist die Möglichkeit zugelassen, demokratische Rechte am Arbeitsplatz einzuschränken, wenn Grundrechte oder andere fundamentale Interessen von Menschen—nicht nur innerhalb des Unternehmens—bedroht sind.⁴

Die eben entwickelte Argumentation baut darauf, dass keine überzeugenden Gründe vorliegen, weshalb der Schutz fundamentaler Rechte die demokratische Kontrolle staatlicher Macht verlangt, nicht aber die Demokratisierung der Machtverhältnisse innerhalb des Unternehmens. Daher werden im nächsten Abschnitt Einwände gegen die Analogie zwischen Staaten und Unternehmen geprüft.

2. Die Freiwilligkeit des Angestelltenverhältnisses

Arneson weist die Analogie zwischen Staaten und Unternehmen zurück. Seine Erklärung, weshalb sein Argument der Schutzfunktion demokratischer Rechte kein moralisches Recht auf demokratische Mitsprache am Arbeitsplatz impliziere, baut darauf, dass sich Angestellte aus freien Stücken einem Unternehmen anschliessen und das Unternehmen ebenso freiwillig wieder verlassen können. Anders als die Mitgliedschaft in einem Unternehmen, betont Arneson, komme die Zugehörigkeit zu einem Staat normalerweise nicht freiwillig zustande. Menschen würden auf dem Territorium eines Staates geboren, da aufwachsen, Beziehungen eingehen und sich an lokale Gegebenheiten anpassen. In ein anderes Land zu ziehen, würde ihre Lebensperspektiven massiv verändern. Verglichen damit, sei der Wechsel des Arbeitsplatzes meistens deutlich weniger einschneidend und etwas, das zu einer gewöhnlichen Biographie gehöre (Arneson 1993, 138–41).

Darin liege der entscheidende Unterschied zwischen Staaten und Unternehmen: Im Bereich der Arbeit gebe es valide Exitoptionen und die Exitkosten seien vergleichsweise gering.⁵ Während sich also Bürger*innen den Konsequenzen staatlichen Handelns nur durch die drastische Massnahme des Auswanderns entziehen könnten, hätten Arbeitnehmer*innen die Freiheit,

4 Dies könnte etwa durch Regulierungen auf Staatsebene oder durch unternehmensinterne Gremien analog zu Verfassungsgerichten geschehen.

5 Die Annahme, Arbeitnehmer*innen hätten stets valide Exitoptionen mit geringen Exitkosten wurde verschiedentlich als unrealistisch kritisiert, da viele Arbeitnehmer*innen regional und branchenspezifisch gebunden sind. Ausserdem liegen die Exitoptionen meist in einer erneuten Anstellung in einem nicht-demokratisch strukturierten Unternehmen (Frega, Herzog und Neuhäuser 2019, 5).

angesichts unerträglicher Arbeitsbedingungen das Unternehmen zu verlassen. So könnten Arbeitnehmer*innen auch ohne demokratische Mitsprache vermeiden, ungewünschte Auswirkungen unternehmerischer Entscheidungen ertragen zu müssen.⁶ Würden Arbeitnehmer*innen trotz Optionen im Unternehmen bleiben, sei dies als Hinweis darauf zu werten, dass die Arbeitssituation erträglich sei und keine grundlegenden Rechte verletzt würden (Arneson 1993, 139–40).

In diesem Verständnis sind Exitoptionen und demokratische Rechte theoretisch austauschbare Mittel, um denselben Zweck—den Schutz fundamentaler Rechte—zu erreichen. Damit sind demokratische Rechte gradueller Natur und abhängig von den jeweiligen Exitoptionen und Exitkosten einer spezifischen Machtkonstellation. Arneson verdeutlicht dies anhand zweier einander gegenübergestellter Beispiele: Er skizziert zwei fiktive aneinander angrenzende Kleinstaaten, die sich sprachlich, wirtschaftlich und kulturell ähnlich sind und die grenzüberschreitende Migration nicht einschränken. Der einzige substantielle Unterschied zwischen den beiden Staaten ist, dass einer demokratisch verfasst ist und der andere nicht. Weil in einem solchen Szenario die Bewohner*innen des nicht-demokratischen Staates eine valide Exitoption mit relativ geringen Exitkosten hätten, meint Arneson, wäre der verletzte Anspruch auf Demokratie entsprechend gering (Arneson 1993, 140).

Dieses Szenario kontrastiert er mit einem zweiten fiktiven Beispiel. Er schildert die Situation von Minenarbeiter*innen in einer isolierten Region, die sich kulturell vom restlichen Land deutlich unterscheidet. Andere Arbeit als die Anstellung in den Minen gibt es in der Region nicht. Wollten die Minenarbeiter*innen die Stelle wechseln, müssten sie in eine andere Landesregion ziehen, womit die Exitkosten damit vergleichbar wären, auszuwandern. Weil in diesem Fall alle Exitoptionen mit exorbitanten Kosten verbunden wären, würde Arneson zufolge der Analogieschluss vom Staat auf die Unternehmen eher überzeugen und die Minenarbeiter*innen hätten einen massgeblichen Anspruch auf Demokratie am Arbeitsplatz (Arneson 1993, 139–40).

Generell dürften aber für Angestellte in Unternehmen mehr zumutbare Exitoptionen zu geringeren Kosten

verfügbar sein als für Bürger*innen im Staat, obschon dies zweifellos von den politischen und wirtschaftlichen Rahmenbedingungen abhängt.⁷ Doch wie plausibel ist der behauptete proportionale Zusammenhang zwischen demokratischen Rechten und Exitoptionen unter der Voraussetzung der Schutzfunktion demokratischer Rechte? Da verschiedene Machtsituationen unterschiedliche Gefahrenpotentiale für die Verletzung von Grundrechten bergen, sind graduelle Unterschiede im moralischen Anspruch auf demokratische Mitsprache einigermassen plausibel. Auch der Gedanke, dass Exitoptionen Betroffenen einen Ausweg aus unerträglichen Verhältnissen geben, leuchtet ein. Es scheint aber, wenn die Position der Schutzfunktion demokratischer Rechte ernst genommen wird, machen Exitoptionen demokratische Rechte nicht überflüssig. Entscheidend für die Notwendigkeit, Grundrechte zu schützen, ist der Umstand des Ausgeliefert-Seins der Träger*innen dieser Rechte. Exitoptionen können einen Beitrag zum Schutz der Grundrechte leisten, reichen aber für sich genommen nicht aus: Einerseits bieten Exitoptionen keinen ausreichenden Schutz. Die Regierung des nicht-demokratischen Staates aus Arnesons Szenario könnte beispielsweise Gesetze verabschieden, die Grundrechte verletzen, und diese Gesetze umsetzen, bevor Bürger*innen ihre Auswanderung organisieren können. Und sie könnte beschliessen, Grenzkontrollen einzuführen, um die Auswanderung von Bürger*innen zu unterbinden.

Andererseits mögen sich manche Menschen aus diversen Gründen entscheiden, trotz der Verletzungen ihrer Rechte im Land zu bleiben. Auch diese Menschen haben einen moralischen Anspruch auf Berücksichtigung ihrer fundamentalen Rechte. Werden diese Rechte verletzt, ist das—unabhängig von den Gründen für die Entscheidung zu bleiben—ein moralischer Verlust. Entsprechend sollten die Rechte aller geschützt werden, auch die Rechte jener, die freiwillig bleiben. Gemäss Arneson ist der Schutz der Grundrechte vor dem Handeln von Körperschaften, in deren Machtbereich sich Menschen befinden, dasjenige, was demokratische Mitsprache wertvoll macht. Und die blosser Möglichkeit, sich dem Machtbereich dieser Körperschaften zu entziehen, sollte daran nichts ändern.

Wichtig ist hier, dass Arnesons Position keine verfahrensinterne Gerechtigkeit zulässt. Den Bürger*innen von Arnesons nicht-demokratischem Kleinstaat ein Recht auf Demokratie mit der Begründung abzuspre-

6 Einen ähnlichen Einwand äussern Jacob und Neuhäuser, sie kommen allerdings nicht ganz zum selben Schluss wie Arneson. Jacob und Neuhäuser fordern—anders als Arneson—staatliche Eingriffe, um Arbeitnehmer*innen echte Exitoptionen zu garantieren, beispielsweise durch ein bedingungsloses Grundeinkommen (Jacob und Neuhäuser 2018, 938–41). In dieser Arbeit wird Arnesons Version des Einwands diskutiert. Mit einzelnen Anpassungen sollten die unten folgenden Ausführungen aber auch auf andere Versionen anwendbar sein.

7 So könnte möglicherweise der Anspruch auf Demokratie, den Angestellte gewisser monopolistischer Firmen behaupten können, stärker sein als jener von Bürger*innen von Kleinststaaten.

chen, sie hätten die Möglichkeit, das Land zu verlassen, obwohl die Einführung demokratischer Strukturen den Schutz ihrer fundamentalen Rechte verbessern würde, lässt sich aber alleine mit ergebnisorientierten Kriterien nicht rechtfertigen. Entscheidend ist also nicht, ob Menschen Exitoptionen haben, sondern ob Demokratie ein effektives Mittel ist, um den Schutz ihrer fundamentalen Rechte zu verbessern. Solange dies der Fall ist, haben auch Menschen, die allen Widrigkeiten zum Trotz ihre Exitoptionen nicht ergreifen, weiterhin einen gewichtigen Anspruch auf demokratische Mitsprache.

Im Staat ist die Gefahr der Verletzung fundamentaler Rechte durch Exitoptionen alleine nicht gebannt, und dasselbe gilt für den Arbeitsplatz. Hohe Exitkosten helfen zwar, zu verstehen, weshalb Angestellte manchmal Ungerechtigkeiten am Arbeitsplatz akzeptieren. Geringe Exitkosten alleine reichen aber nicht, um den Schutz der Grundrechte als gewährleistet anzusehen. Selbst wenn wir, wie Arneson, davon ausgehen, dass für Arbeitnehmer*innen unter gewöhnlichen Bedingungen zumutbare Alternativen zum aktuellen Unternehmen verfügbar sind, können wir eine Entscheidung, das Unternehmen nicht zu verlassen, nicht als Nachweis dafür werten, dass keine fundamentalen Rechte verletzt werden. Die Gründe von Arbeitnehmer*innen, trotz Verletzung ihrer Rechte im Unternehmen zu bleiben, können vielfältig und individuell sein.⁸ Und wie die Existenz von Exitoptionen sind sie—unter der Voraussetzung der Schutzfunktion demokratischer Rechte—für die Frage nach einem Recht auf Demokratie am Arbeitsplatz irrelevant. Denn die Argumentation hier ist nicht, dass Menschen vor dem geschützt werden sollten, was hinter ihren Gründen steckt, ein Unternehmen nicht zu verlassen. Selbst wenn keine nachvollziehbaren Gründe vorliegen, sollte dies nicht als Rechtfertigung dienen, ein Machtssystem nicht zu ändern, wenn die Systemänderung zu weniger Verletzungen von Grundrechten führen würde. Die Argumentation konzentriert sich deshalb bloss auf die Frage, ob die Einführung demokratischer Verfahren in ein bestimmtes Entscheidungssystem, in diesem Fall den Arbeitsplatz, den Schutz der fundamentalen Rechte jener, über die entschieden wird, verbessern würde.

Der Aspekt der Freiwilligkeit böte ferner Stoff für einen möglichen zweiten Einwand: Statt wie Arneson

auf ein allfälliges Ende, könnte man auch auf den Beginn einer Anstellung fokussieren. Freiwillig ist an der Mitgliedschaft in Unternehmen ja nicht nur, weiterhin bei einem Unternehmen angestellt zu bleiben, dem man bereits angehört. Freiwillig ist auch der ursprüngliche Beitritt. Wie oben dargestellt, eröffnet Arneson seine Argumentation zugunsten eines Rechts auf Demokratie mit der Position, Machtausübung *ohne Einwilligung* der Betroffenen sei nur dann gerechtfertigt, wenn sie dem Schutz fundamentaler Rechte diene. Wer sich von einem Unternehmen anstellen lasse, könnte man nun argumentieren, willige damit ein, in die Machtstrukturen eines Unternehmens eingegliedert zu werden. Die Zugehörigkeit zu einem Staat dagegen erlange man meist durch Geburt und damit ohne vorhergehende Einwilligung und daher impliziere Arnesons Argument für ein Recht auf Demokratie kein Recht auf Demokratie am Arbeitsplatz.

Eine mögliche Antwort auf diesen Einwand findet sich in Arnesons Verteidigung der These, es gebe keine verfahrensinterne Gerechtigkeit. Darin diskutiert er die Rolle, die Zustimmung als Gerechtigkeitskriterium für Verfahren spielt. Würde von den Ergebnissen des Verfahrens unabhängig die freiwillige Zustimmung der Beteiligten ein Verfahren gerecht machen, spräche dies gegen seine These. Er schlägt daher vor, die Bedeutung von Zustimmung ergebnisorientiert zu verstehen. Die Forderung, niemand dürfe gegen den eigenen Willen und ohne Einwilligung gezwungen werden, Vereinigungen einzugehen, sei ein Mittel, um sicherzustellen, dass die Interessen aller angemessen berücksichtigt würden. Wenn aber die Bedingung der Zustimmung aller Beteiligten diesen Zweck nicht erfülle, dann seien gegebenenfalls paternalistische Massnahmen angebracht (Arneson 1993, 128–29).

Dieser Gedanke liefert die Antwort auf den Einwand, wer freiwillig einem Unternehmen beitrete, stimme den unternehmensinternen Machtstrukturen zu, womit diese gerecht seien: Wenn am Arbeitsplatz das Risiko besteht, dass fundamentale Rechte von Arbeitnehmer*innen verletzt werden, und wenn die Entscheidungsverfahren im Unternehmen so angepasst werden können, dass die Rechte der Arbeitnehmer*innen besser geschützt sind, dann müssen die Verfahren auch dann geändert werden, wenn zuvor alle Beteiligten ebendieses Verfahren implizit oder explizit zugestimmt haben.⁹

⁸ Ausserdem sei hier darauf hingewiesen, dass es eine Reihe von Exitkosten gibt, die auch in für Arbeitnehmer*innen idealen wirtschaftlichen Bedingungen weiterhin bestehen. Dazu gehören beispielsweise der Wechsel des sozialen Umfelds und das verlassen gewohnter Arbeitsabläufe.

⁹ Dies gilt auch dann, wenn die Interessen mancher Beteiligter durch die Einführung demokratischer Strukturen frustriert werden und einige deshalb die Demokratie am Arbeitsplatz ablehnen, solange der Schutz fundamentaler Rechte verbessert wird.

Wenn also der Schutz fundamentaler Rechte die primäre Funktion des Rechts auf Demokratie ist und die Einführung von Demokratie am Arbeitsplatz diese Funktion ebenfalls erfüllt, dann wird ein Recht auf demokratische Mitsprache im Unternehmen weder durch offenstehende Exitoptionen noch durch die initiale Einwilligung der Angestellten überflüssig gemacht.

3. Unterschiede zwischen der Autorität von Staaten und Unternehmen

Weiter oben wurde darauf hingewiesen, dass Arnesons nicht-demokratischer Kleinstaat Gesetze erlassen und umsetzen könnte, die fundamentale Rechte seiner Bürger*innen verletzen würden, bevor diese die Gelegenheit hätten, in den demokratischen Nachbarstaat auszuwandern. Mit diesem Hinweis sollte gezeigt werden, dass Ausstiegsoptionen aus Hierarchien unter Umständen nicht ausreichen, um die Grundrechte der Menschen in diesen Hierarchien zu schützen. Dagegen könnte man einwenden, der Staat könne wohl solche Gesetze erlassen, Unternehmen hingegen könnten dies nicht. Staaten haben umfassendere Kompetenzen als Unternehmen: Sie können die grenzüberschreitende Migration regulieren und haben ein Gewaltmonopol. Ausserdem können Staaten Regeln setzen, denen sich die Unternehmen auf ihrem Territorium fügen müssen, und sie haben die Möglichkeit, das Einhalten dieser Regeln zu kontrollieren und durchsetzen.

Ein Argument, das in diese Richtung zielt, formulieren Daniel Jacob und Christian Neuhäuser in „Workplace Democracy, Market Competition and Republican Self-Respect“ (2018). Sie kritisieren den Analogieschluss mit dem Einwand, die Autorität von Staaten sei abschliessend. Zwar seien auch Staaten in ihrem Handeln durch internationales Recht eingeschränkt, doch hätten sie alleine die Autorität, die Regeln festzulegen, die auf ihrem Staatsgebiet gelten würden. Unternehmen hingegen seien der Autorität von Staaten unterworfen und durch deren Gesetze reguliert. Solange sich Unternehmen innerhalb dieser staatlichen Regulierungen bewegen würden, dürften sie legitimerweise Autorität über Angestellte ausüben. Insofern gebe es einen relevanten Unterschied zwischen den Gesetzen im Staat und den Regeln im Unternehmen (Jacob und Neuhäuser 2018, 932).

Darin steckt die Vorstellung, die moralische Legitimität der staatlichen Autorität übertrage sich auf die Autorität der Unternehmen innerhalb des Staates (Jacob

und Neuhäuser 2018, 932). Diese Vorstellung wurde hier, da diese Arbeit eingangs die ergebnisorientierte Perspektive Arnesons als theoretische Grundannahme voraussetzte, von vornherein ausgeschlossen. Wird der Schutz fundamentaler Rechte als Standard akzeptiert, kann nicht angenommen werden, die Autorität von Unternehmen würde ihre moralische Legitimität von jener der Autorität von Staaten beziehen, ohne auf die Frage einzugehen, ob der geforderte Schutz effektiv gewährleistet ist.

Ein Grund, weshalb der beschriebene Unterschied der Autorität von Staaten und Unternehmen einem Analogieschluss ausgehend von der demokratietheoretischen Position Arnesons dennoch im Weg stehen könnte, liegt darin, dass Staaten Gesetze verabschieden können, die auf den Schutz der moralischen Ansprüche von Arbeitnehmer*innen abzielen. Jacob und Neuhäuser schlagen beispielsweise vor, eine allgemeine Beschreibung des Aufgabenbereichs von Angestellten in Arbeitsverträgen zu verlangen und den Anliegen von Arbeitnehmer*innen durch Betriebsräte Gehör zu verschaffen (Jacob und Neuhäuser 2018, 935–36). Tatsächlich stehen Staaten viele Massnahmen offen, um ein mögliches Recht auf körperliche und psychische Unversehrtheit zu schützen, ohne Unternehmen zu verpflichten, Vorgesetzte durch Angestellte wählen zu lassen oder in demokratischer Weise die Kontrolle über zentrale unternehmerische Entscheidungen an die Arbeitnehmer*innen zu übergeben. Dazu gehören etwa gesetzliche Sicherheitsvorschriften für die Arbeit mit gefährlichen Maschinen, die Regulierung der Arbeits- und Ruhezeiten und Lohnanspruch bei krankheitsbedingter Abwesenheit.

Hierzu ist zunächst festzuhalten, dass auch zwischen einem Staat und den Gemeinden dieses Staates ein ähnlicher Unterschied besteht. Staaten können auf demokratischem Weg Gesetze erlassen, welche die Reichweite der Autorität der Gemeinden eingrenzen und die Rechte der Gemeindebürger*innen vor dem Zugriff der Gemeindepolitik schützen. Ausserdem könnte man eine demokratische supranationale Organisation schaffen, deren Autorität wiederum in einem ähnlichen Verhältnis zu einzelnen Staaten stünde, wie die Autorität der Staaten zu ihren Gemeinden. Diese supranationale Organisation könnte den Mitgliedsstaaten Vorschriften machen, um die Rechte der Bürger*innen dieser Staaten zu schützen. Dennoch würde damit die Demokratie in den Mitgliedstaaten nicht überflüssig, genauso wenig wie die demokratische Organisation der Gemeinden demokratischer Staaten überflüssig ist.

Dies hat nicht damit zu tun, dass die übergeordneten demokratischen Ebenen nicht durch die Regulierung der Autorität der tieferen Ebenen massgeblich am Schutz der Rechte der Menschen teilhätten. Vielmehr als die bloss Beschränkung der Autorität einer Körperschaft durch eine umfassendere Autorität ist entscheidend, ob die Autorität dieser Körperschaft die Fähigkeit miteinschliesst, Grundrechte zu verletzen. Solange nämlich in einem Entscheidungssystem Grundrechte verletzt werden können, sind die Rechte der Beteiligten zu schützen. Dies dürfte auf dem Weg der demokratischen Kontrolle des Entscheidungssystems durch die Beteiligten zuverlässiger gelingen, als wenn ausschliesslich der Umweg über eine übergeordnete demokratische Körperschaft offensteht. Denn die Verletzung von Grundrechten durch lokale Autoritäten kann spezifisch sein und ausserhalb der Aufmerksamkeit der nationalen oder internationalen Öffentlichkeit geschehen.

Nicht anders verhält sich die Sache bei Unternehmen. Auch hier können die Verletzungen fundamentaler Rechte vielfältig, branchen- und unternehmensspezifisch und ohne öffentliche Aufmerksamkeit sein und sich angesichts staatlicher Regulierungen wandeln. Die Arbeitnehmer*innen selbst sind direkter informiert über Vorgänge im Unternehmen als die Masse der Wähler*innen im ganzen Staat oder eine vergleichsweise weit entfernte Regierung. Ausserdem ist zu beachten, dass die Arbeitnehmer*innen eines einzelnen Unternehmens oder einer einzelnen Branche im demokratischen Prozess des Staates eine Minderheitsposition innehaben. Und nicht alle Bevölkerungsteile haben dieselben materiellen Interessen, sodass in den demokratischen Prozess Interessen miteinfließen, die dem Schutz der Rechte mancher Arbeitnehmer*innen entgegenstehen können. Dementsprechend scheint wahrscheinlich, dass zusätzlich zu staatlichen Regulierungen die Einführung von Demokratie am Arbeitsplatz ein wirksames Mittel wäre, den Schutz fundamentaler Rechte auszubauen.

Die zentrale—letztlich empirische—Frage, zu der wir an dieser Stelle wieder zurückkehren, bleibt damit: Wie lassen sich die fundamentalen Rechte der Arbeitnehmer*innen am besten schützen? Die staatliche Regulierung unternehmerischen Handelns und die Verfügbarkeit attraktiver Exitoptionen können sicherlich einen Beitrag leisten. Ich denke aber, dass die zusätzliche Einführung demokratischer Mitspracherechte für Arbeitnehmer*innen den Schutz weiter ausbauen würde. Mindestens solange wir davon ausgehen, dass im Staat die Demokratie ein effektives Mittel für den Schutz fundamentaler Rechte ist, spricht wenig dafür, dass dies im Unternehmen anders ist.

Fazit

Diese Arbeit war ein Versuch, Arnesons Theorie demokratischer Rechte auf die Arbeitswelt zu übertragen. Wird seine These der Schutzfunktion demokratischer Rechte akzeptiert, dann stellt sich die Frage, ob—alles andere bleibe gleich—die Einführung demokratischer Mitsprache für die Arbeitnehmer*innen den Schutz fundamentaler Rechte am Arbeitsplatz verbessern würde. Wenn die Übergabe demokratischer Kontrolle an die Bürger*innen eines Staates als adäquates Mittel zum Schutz ihrer fundamentalen Rechte betrachtet wird, ist schwer einzusehen, weshalb die Demokratisierung eines Unternehmens für die Arbeitnehmer*innen nicht dasselbe leisten sollte. Arnesons Standard für die Evaluation von Entscheidungsverfahren spricht deshalb nicht nur für ein Recht auf Demokratie im Staat, sondern auch für ein Recht auf Demokratie am Arbeitsplatz.

Zwar gibt es zwischen Staaten und Unternehmen Unterschiede: Die Autorität von Staaten ist umfassender als jene von Unternehmen. Und Arbeitnehmer*innen schliessen sich, anders als die Bürger*innen eines Staates, ihrem Unternehmen freiwillig an. Ausserdem stehen den Angestellten eines Unternehmens üblicherweise mehr Exitoptionen zu geringeren Exitkosten offen als den Bürger*innen eines Staates. Doch solange die Demokratisierung von Unternehmen den Schutz der Grundrechte der Arbeitnehmer*innen ausbaut, bleibt ein starker Grund für ein Recht auf Demokratie am Arbeitsplatz bestehen—unabhängig davon, ob Staaten gleichzeitig bemüht sind, Unternehmen zu regulieren, oder ob Arbeitnehmer*innen ein Unternehmen theoretisch verlassen könnten, um Rechtsverletzungen zu vermeiden.

Wo liegen Schwächen dieser Argumentation? Einerseits lassen sich die demokratietheoretischen Grundannahmen dieser Arbeit—Arnesons These der Schutzfunktion demokratischer Rechte und seine Zurückweisung verfahrensinterner Gerechtigkeitskriterien—kritisieren. Möglicherweise würde allerdings auch eine abgeschwächte Form von Arnesons Position, die im Schutz von Grundrechten *eine* von vielen Funktionen demokratischer Rechte sehen würde, als Fundament für ein Recht auf Demokratie am Arbeitsplatz reichen. Andererseits kann die Argumentation innerhalb des theoretischen Rahmens kritisiert werden, indem nach fundamentalen Rechten gefragt wird, die durch eine Demokratiepflcht für Unternehmen verletzt würden. Anbieten würden sich dazu beispielsweise Eigentumsrechte, deren Verhältnis zur Demokratie am Arbeitsplatz mindestens unklar ist.

Ferner wurde, aufgrund des begrenzten Rahmens dieser Arbeit, für einige zentrale Annahmen nur spärlich argumentiert. Dazu gehört etwa die Behauptung eines Rechts auf körperliche und psychische Unversehrtheit, das am Arbeitsplatz bedroht sein kann. Auch der konkrete Zusammenhang zwischen fundamentalen Rechten und der Demokratie am Arbeitsplatz wäre genauer auszuarbeiten. Ob die Demokratie am Arbeitsplatz effektiver als alternative Einrichtungen den Schutz fundamentaler Rechte gewährleistet, ist letztlich eine empirische Frage, die im Rahmen dieser Arbeit nicht zu klären ist. Trotzdem konnte in dieser Arbeit gezeigt werden, dass sich auf der Basis von Arnesons Theorie demokratischer Rechte ein Argument für ein Recht auf Demokratie am Arbeitsplatz formulieren lässt.

Literatur

- Arneson, Richard. 1993. „Democratic Rights at National and Workplace Levels.” In *The Idea of Democracy*, herausgegeben von David Copp, Jean Hampton, und John Roemer, 118–138. Cambridge: Cambridge University Press.
- Cohen, Joshua. 1989. „The Economic Basis of Deliberative Democracy.” *Social Philosophy and Policy* 6 (2): 25–50.
- Frega, Roberto, Lisa Herzog, und Christian Neuhäuser. 2019. „Workplace Democracy. The Recent Debate.” *Philosophy Compass* 14 (4): 1–11.
- Jacob, Daniel, und Christian Neuhäuser. 2018. „Workplace Democracy, Market Competition and Republican Self-Respect.” *Ethical Theory and Moral Practice* 21 (4): 927–944.

Simon Richmond (26) hat im Bachelor Philosophie mit Nebenfach Geschichte studiert. Er interessiert sich für Fragen nach Demokratie und Gerechtigkeit, für politische Philosophie im Allgemeinen und für Metaethik.

The Problem of the Barber-Shop Paradox by Lewis Carroll

A Clash between Intuition and Rational Thinking

Introduction

The Barber-Shop Paradox, pointed out by Lewis Carroll at the end of the 19th century, is a problem in the field of logic. By stating this paradox, Carroll indicates some issues concerning the understanding of if-then clauses. For example, one question raised by Carroll in connection with the Barber-Shop Paradox is whether we can accept the sentence “if p , then q ” while also asserting “if p , then not q ”. Although truth-functional logic delivers answers to this and to other related questions, the Barber-Shop Paradox is still being discussed in philosophy. The aim of my essay is to show why the Barber-Shop Paradox is still problematic, even though it has been solved by means of logic. For that purpose, I will show that the crux of the problem lies in two different kinds of understanding if-then clauses. Take for example the sentence “If my sister drinks a beer, I drink a beer”. While sitting in a pub with a couple of friends, we understand this sentence without thinking much. But we can also analyse it in a philosophy course and ask ourselves: “What exactly does this sentence mean? May I drink a beer even if my sister does not drink one? Do I have to drink a beer if she does drink one? Is this sentence a true sentence? What does it even mean for such a sentence to be true? May I truthfully utter the sentence ‘I drink no beer if my sister does drink one’ as well?”

In order to answer such questions about if-then clauses, we may take those clauses out of context and analyse their truth conditions by formalising them, or we may leave them in their settings and give answers which depend on our world. On the basis of these two different ways of handling if-then clauses, I will pursue the aim that I have characterised above. For that pur-

pose, I will start embedding the Barber-Shop Paradox into the history of logic, especially by considering the understanding of if-then clauses at the end of the 19th century. Thereafter, I will reconstruct the Barber-Shop Paradox and indicate a dispute between Carroll and John Cook Wilson (a colleague of Carroll) with the intent to show different opinions about if-then clauses. Further, I identify three central contentious points in that dispute and explicate the way of solving them by means of truth-functional logic. Although the solution I am going to present shows that the Barber-Shop Paradox does not have to be understood as a paradox, it does not solve the problem entirely. For that reason, I elaborate two different kinds of understanding if-then clauses by means of the three identified contentious points. While one of these understandings is only formal and does not depend on the world we live in, the other one is an everyday approach which depends on our world. Revealing these two kinds of understandings shall then help to solve the Barber-Shop Paradox, and with it the contentious points of that problem, in a satisfying way. However, the solution I am going to present is of a different kind than the ones already given by means of logic. For my approach does not intend to solve the Barber-Shop Paradox logically, but rather it works out a conflict between intuition and rational thinking regarding the problem of if-then clauses pointed out by Carroll.

1. The Paradox

The Barber-Shop Paradox is dealing with the truth of propositions that function as an antecedent. The following shall show the main issue Carroll intends to il-

illustrate with that paradox (this list refers to Carroll 1977, 451):

1. Let p , q , and r be propositions.
2. Let us accept the sentence “if p is true, then q is true”.
3. Let us also accept the sentence “if r is true, then if p is true, q is false”

Given 1.–3., one question raised by the Barber-Shop Paradox is whether it is possible that the proposition r is true (Carroll 1977, 451). Whereas representatives of a truth-functional logic, which is the dominant approach in contemporary logic, answer this question with “yes”, some other people hold the view that r cannot be true. In this section, I will describe the Barber-Shop Paradox without evaluating it yet. For this purpose, I will first put this paradox in the historical context of logic in order to demonstrate why, at the end of the 19th century, it has not yet been as clear how to solve problems concerning if-then clauses in a logical way as it is today. Second, I will give a more detailed reconstruction of the Barber-Shop Paradox and establish the conflict between Carroll and Cook Wilson. Within my reconstruction, I will focus on three questions about if-then clauses which are intensely discussed in the context of this paradox.

1.1 Embedded in the History of Logic

The oldest sources of logical approaches in the West are found in ancient Greece. Probably the most important ideas regarding the development of western logic arise from Aristotle (384–322 BCE) (Klement 2019, sect. 2; Smith 2019, before Introduction). However, with respect to if-then clauses, the sources of Aristotle’s logic do not deliver much, although he (or his pupil Theophrastus) held the view that a doctrine concerning if-then clauses should be elaborated (Klement 2019, sect. 2). More about if-then clauses can be found in texts regarding the approaches of Diodorus Cronus and his pupil Philo (late 3rd century BCE) (Klement 2019, sect. 2; Sedley 2018, Introduction). Philo considered that an if-then clause is only false if the antecedent is true while the consequent is false, otherwise it is true. This notion accords with the truth-functional logic that is dominant today. But Diodorus Cronus begged to differ. Although he accepted that an if-then clause is false in the mentioned case, he believed that it is also false if it is (or was) not possible that the antecedent is false and the consequent is true (Bobzien 2019, sect. 3).

After the time of the ancient Greek philosophers, western logic had not developed much for a couple of centuries (Klement 2019, sect. 2). Only at the mid 19th century, philosophers engaged in logical studies to an increasing degree, and the so called “symbolic logic” or “algebraic logic” came up. This approach handles logical thoughts not with natural language but with symbols and is closer to mathematics than Aristotle’s work. This kind of logic, especially the work of George Boole, influenced the now dominant truth-functional logic (Klement 2019, sect. 2). But it were mostly the texts of Gottlob Frege that established the truth-functional logic and with it the view that an if-then clause with a false antecedent is always true, regardless of the truth-value of the consequent (Edgington 2014, sect. 2.1). Although Frege published the relevant texts at the end of the 19th century, hence at the same time Carroll pointed out the Barber-Shop Paradox, they were not well-received. Only during the early 20th century, when Bertrand Russell, Ludwig Wittgenstein and other philosophers dealt with Frege’s notions, the thoughts of Frege gained more recognition (Zalta 2019, sect. 1). So, the understanding of if-then clauses in the truth-functional way we basically owe Frege, Russell, Wittgenstein and other philosophers who published at the early 20th century. Thus, the truth-functional understanding of if-then clauses was not familiar to Carroll when he came up with the Barber-Shop Paradox.

Hence, the issues concerning if-then clauses in the context of the Barber-Shop Paradox were pointed out at a time when logical approaches came up again after a long period of silence, but before the truth-functional logic, which delivers a clear answer to those issues, has been prevailed. So at the time Carroll published the Barber-Shop Paradox, there was no established theory of how if-then clauses should be understood in a logical manner. Only a discussion about the understanding of if-then clauses, which has its roots in the ancient Greece, had already been there. Moreover, there is a sign that Carroll engaged both in Aristotle’s logic and the upcoming more mathematical logic, when we consider that Carroll’s book “Symbolic Logic” is dedicated to the memory of Aristotle (Carroll 1977, Dedication) but named after the then upcoming approach. All this shows that Carroll wrote his work at a time in which the dominant logic approaches were changing, but the new form of truth-functional logic had not yet been enforced. All in all, it should now be intelligible why, at the end of the 19th century when Carroll pointed out the Barber-Shop Paradox, it had not yet been as clear how if-then clauses are to be understood by means of logic as it is today.

1.2 Reconstruction

As my remarks about the historical context of the Barber-Shop Paradox show, there were, from Carroll's point of view, some concerns about the understanding of if-then clauses. In order to develop a better understanding, Carroll discussed the problems he identified with other logicians, especially with his colleague John Cook Wilson (Carroll 1977, 444–445). Carroll's aim was that logicians came to an agreement about how to deal with if-then clauses (Carroll 1977, 442). This is the reason why he put forward the Barber-Shop Paradox.

All in all, there are no less than eight versions of this paradox (Carroll 1977, 449). Most of them refer to a discussion between the two logicians Carroll and Cook Wilson. There were some letter exchanges about the understanding of if-then clauses, but the most familiar version was published in *Mind* in July 1894 (Carroll 1977, 445). There, Carroll tells a story about a child and two men on their way to a barbershop:

Sometimes, when the two men, Joe and Jim are with the child, who is the narrator of this story, they talk about logical puzzles. On the way to a barbershop, Uncle Jim mentions that he would like the barber Carr to shave him. However, the barbershop they are heading to is run by three barbers, Allen, Brown, and Carr. Since not all of them are always in the shop, it might be that Carr is out, and Jim could not get a shave from him. This is the reason why Joe was motivated to prove logically that Carr is there. For this purpose, he uses two premisses and one assumption:

- (1) If Carr is out, then if Allen is out, Brown is in.
- (2) If Allen is out, then Brown is out.
- (A) Carr is out.

For the sake of the argument, Joe asks Jim in a first step whether he accepts premiss (1). Jim answers: "Of course he [Brown] must [be in], or there'd be nobody to mind the shop" (Carroll 1894, 436). Based on a statement of Jim, Joe argues in a second step for premiss (2). It was one day ago, when Jim said to Joe that Allen does not leave the shop without Brown (because he was ill, and since then he is nervous about going out alone). So, due to his own statement, Uncle Jim grants premiss (2). In order to prove that Carr is in, Joe chooses the method *Reductio ad Absurdum*, hence he assumes the contrary of the statement he wants to prove. Because the two premisses and the assumption are inconsistent in Joe's opinion, the assumption has to be dropped (the premisses are granted and cannot be given up). But Jim does not see the inconsistency of the three statements. So, Joe

tries to persuade him. He claims that the two sentences "If Allen is out, then Brown is out" (which is premiss (2), therefore it is accepted by Jim) and "If Allen is out, then Brown is in" (which follows from the conjunction of the assumption and premiss (1)) cannot be true together. Therefore, the assumption is false, and Carr must be in the shop. But Jim does not accept the incompatibility of these two sentences. Joe then asks Jim:

"How can they [possibly be true together]? How can one and the same protasis¹ prove two contradictory apodoses²? You grant that the two apodoses 'Brown is *in*' and 'Brown is *out*,' are contradictory, I suppose?" (Carroll 1894, 437)

Although Jim grants that the two consequents "Brown is in" and "Brown is out" are contradictory, he still does not accept the contradiction between the two sentences "If Allen is out, then Brown is in" and "If Allen is out, then Brown is out". He replies:

"Why shouldn't the two Hypotheticals be true together? It seems to me that would simply prove '*Allen is in*' (...) why shouldn't we put it like this? If Allen is out, Brown is *out*. If Carr and Allen are *both* out, Brown is *in*. Which is absurd [because when Allen leaves the shop, he takes Brown with him]. Therefore Carr and Allen can't be *both* of them out. But so long as Allen is *in*, I don't see what's to hinder Carr from going *out*." (Carroll 1894, 437)

The conversation between Joe and Jim shows that they differ about the consistency of two sentences of the form "if p, then q" and "if p, then not q". In addition, there is another point shown by their conversation which is relevant for their dispute, and with it, for the discussion about the understanding of if-then clauses. As mentioned above, Jim accepts premiss (1) because it follows from the fact that at all times, there has to be someone in the shop. But premiss (1) is not the only thing that is inferable from this fact. There are other ways of interpreting this fact, as Jim points out. For example, it is possible that Carr and Allen are both out and Brown is in. However, since Jim and Joe both agree that Allen takes Brown with him when he leaves the shop, Carr and Allen cannot both be out because then nobody would mind the shop.

1 Greek word for "antecedent".

2 Greek word for "consequent".

In summary, it can be recorded that due to the two facts that (a) someone has to watch the shop, and that (b) Allen takes Brown with him when he leaves the shop, Joe and Jim both put forward an argument with different conclusions:

Joe's Argument

- | | |
|-----|---|
| (1) | If Carr is out, then if Allen is out, Brown is in. |
| (2) | If Allen is out, then Brown is out. |
| (A) | Carr is out. |
| (C) | If Allen is out, then Brown is out and if Allen is out, then Brown is in. |

Because Joe regards the conclusion as contradictory, he claims that the assumption has to be given up. This proves logically, as he believes, that Carr is in the shop.

Jim's Argument

- | | |
|-----|--|
| (1) | If Carr is out, then if Allen is out, Brown is in. |
| (2) | If Allen is out, then Brown is out. |
| (C) | Carr is in or Allen is in. |

As it can be seen, Jim builds an argument out of the two same facts ((a) and (b)) as Joe does, but he arrives at another conclusion. In reply to Jim's argument, Joe says that Jim is wrongly dividing the antecedent and the consequent of the clause "If Carr is out, then if Allen is out, Brown is in" (Carroll 1894, 437). Whether Jim actually does divide the relevant clause wrongly, will be discussed in section 3. At this point, I will only point out that the discussion between Jim and Joe indicates a dispute about the understanding of the sentence "If Carr is out, then if Allen is out, Brown is in".

As my remarks show so far, Joe and Jim disagree on two points. Firstly, Joe believes that two sentences of the form "if p , then q " and "if p , then not q " are contradictory, while Jim begs to differ. Secondly, they are not coming to an agreement about how to understand a sentence with the form "if p , then if q , then r ".

One further issue about if-then clauses, which Carroll discussed with other logicians, can be found in a note attached to the story about the three barbers. It is the question whether an if-then clause whose antecedent is false can be regarded as legitimate (Carroll 1894, 438). In the story itself, there is only a hint to this issue, as Uncle Joe says to Uncle Jim:

"You will also grant me that the truth of a Hypothetical—I mean its *validity* as a logical *sequence*—does not in the least depend on its *pro-tasis* being actually true, nor even on its being *possible*. The Hypothetical 'If you were to run from here to London in five minutes you would surprise people,' remains true as a *sequence*, whether you can do it or not." (Carroll 1894, 437)

Uncle Jim replies that he cannot do it, but he does not clearly answer Joe's question. Neither does he say whether he accepts Joe's suggestion, nor does he deny it. For the purpose of Joe's argument, this issue is relevant because otherwise Joe could not conscientiously use the sentence "If Carr is out, then if Allen is out, Brown is in" as a premiss. The antecedent is false, in Joe's opinion, but he needs the whole sentence to be true in order to build his argument.

In summary, the main three questions that are raised by the Barber-Shop Paradox are, as my reconstruction indicates and Carroll explicitly mentions (Carroll 1894, 437), the following:

1. Are two if-then clauses, of the forms "if p , then q " and "if p , then not q " compatible?
2. How do the following propositions differ (if they do)?
 - a) p , q , r cannot be all true at once
 - b) If p and r are true, q is not true
 - c) If r is true, then if p is true, q is not true
 - d) If p is true, then, if r is true, q is not true
3. Can an if-then clause, whose antecedent is false, be regarded as legitimate?

The first question results from the discussion between Joe and Jim about the consistency of the two sentences "If Allen is out, then Brown is out" and "If Allen is out, then Brown is in".

The second question is only partly found in the conversation between Joe and Jim. This can be shown by exchanging p with "Allen is out", q with "Brown is out", and r with "Carr is out". Putting it this way, Joe and Jim both accept (a), since they agree that not all of the three barbers can be out at once. But whereas Joe formulates his premiss (1) in a way that corresponds with (c), Jim phrases his premiss (1) in a way that corresponds with (b). So, while (a)–(c) appear in the conversation between Joe and Jim, there is no correspondence to (d) in the story about the three barbers. However, Carroll mentions (d) in the note attached to the story (Carroll 1894, 438).

Also, the third question can be found in the reconstructed story, namely when Joe asks Jim whether he grants that the validity of if-then clauses does not depend on its antecedent being actually true.

2. Solutions

In this section, I will discuss the aforementioned questions in order to show why the Barber-Shop Paradox is still problematic even though truth-functional logic delivers consistent answers to each of these questions. For that purpose, I will first give a brief overview of truth-functional logic, especially of the material implication, and show how this kind of logic answers the three identified questions. Since, as I will show in a second step, this way of solving the Barber-Shop Paradox does not solve the problem which Carroll had in mind entirely, I will eventually argue that we have to differentiate between two understandings of if-then clauses in order to understand and solve this problem in a satisfying way.

2.1 Solution by means of Truth-Functional Logic

In order to understand the solution to the Barber-Shop Paradox by means of truth-functional logic, I will start with a brief overview of truth-functionality. Most of what I am going to say in this context I consider to be familiar to the readers of this essay. But still, it is helpful to have that information in mind in order to comprehend in what content the Barber-Shop Paradox is solved with truth-functional logic and what part of the paradox remains problematic.

As mentioned above, truth-functional logic has been established only after Carroll had put forward his paradox. Moreover, this kind of logic was influenced by Boole, Frege, Russel, Wittgenstein and other logicians whose approaches were to a greater extent mathematical than the ones before. The idea behind truth-functionality is to work with operators in order to calculate truth-values. Each of these operators is used to calculate truth-values of propositions by a specific kind of rule. So, in order to calculate the truth-value of a given proposition (output), at least one truth value of a proposition (input) and one rule are required. In classical truth-functional logic, there are two principles which have their origins with Aristotle and are relevant for calculating the truth-values. The first one is known as “The Law of the Excluded Middle” (Klement 2019, sect. 2). It states that there are only two possible truth-values for a proposition: *true* and *false*. Take for example the proposition

“Socrates is a philosopher”. The Law of the Excluded Middle now says that this is the case (*true*) or it is not (*false*). There is no third possibility. The second principle is called “The Law of Contradiction” (Klement 2019, sect. 2). With this principle, it is claimed that the truth-values *true* and *false* are different in a way that they cannot be applied together. This means that the proposition mentioned above is either true or false, it cannot be both.

The former illustration shall now be enlightening in order to comprehend how if-then clauses have to be understood in a truth-functional way. The rule for such clauses is called “material implication” (Edgington 2014, sect. 2.1). It is best understood by means of a truth-table:

Material Implication

p	q	$p \rightarrow q$
<i>true</i>	<i>true</i>	<i>true</i>
<i>true</i>	<i>false</i>	<i>false</i>
<i>false</i>	<i>true</i>	<i>true</i>
<i>false</i>	<i>false</i>	<i>true</i>

As this table shows, an if-then clause, which is here indicated with the symbol “ $\rightarrow$ ” and named “material implication”, is only false when its antecedent is true whereas its consequent is false.

With the help of this illustration about the way of how if-then clauses are understood by means of truth-functional logic, I will now present the answers to the three questions Carroll has put forward in the context of the Barber-Shop Paradox.

Question 1

Probably the simplest way to answer the question “Are two if-then clauses, of the forms

- *if p, then q* ($p \rightarrow q$) and
- *if p, then not q* ($p \rightarrow \neg q$)

compatible?” by truth-functional logic is to compare the truth-tables of those two forms:

Truth-table of $p \rightarrow q$

p	q	$p \rightarrow q$
<i>true</i>	<i>true</i>	<i>true</i>
<i>true</i>	<i>false</i>	<i>false</i>
<i>false</i>	<i>true</i>	<i>true</i>
<i>false</i>	<i>false</i>	<i>true</i>

Truth-table of $p \rightarrow \neg q$

p	$\neg q$	$p \rightarrow \neg q$
true	false	false
true	true	true
false	false	true
false	true	true

In the second truth-table, two operators are found. One of them ($\rightarrow$) I have already explained before. The other one, the negation ($\neg$), is very simple. It results basically from the two principles mentioned above. Which means that the negation of q is always true when q is false, and vice-versa, for there are only these two truth-values and we may not attribute them together to one and the same proposition. So, the difference between $p \rightarrow q$ and $p \rightarrow \neg q$ just lies in the negation of q . Thus, $p \rightarrow \neg q$ is only false if p is true and $\neg q$ is false (what means that q is true).

Now, the two if-then clauses “if p , then q ” and “if p , then not q ” are compatible if it is possible that they are jointly true. With a look at the two truth-tables it is seen that this actually is possible, namely in the following two cases: If p and q are both false, or if p is false and q is true.

Hence, the application of material implication shows that whenever the antecedent of an if-then clause is false, the whole if-then clause is true (this corresponds with Philo’s point of view (see sect. 1.1)). So, the two sentences “If Allen is out, then Brown is in” and “If Allen is out, then Brown is out” are compatible by means of truth-functional logic, since it is possible that they are jointly true. Namely, if it is false that Allen is out.

Question 2

As shown in section 2, the second question which has to be answered by means of truth-functional logic is:

How do the following propositions differ (if they do)?

1. p , q and r cannot be all true at once
2. If p and r are true, q is not true
3. If r is true, then if p is true, q is not true
4. If p is true, then if r is true, q is not true

The answer to this question is rather simple. By applying the rules of truth-functional logic, these four sentences are equivalent, this means they have the same meaning which is, in truth-functional logic, the truth-value. This could be shown through truth-tables

in the same way as the former question was answered. However, regarding the purpose of my essay, it is not relevant how those rules work in detail. It is only important to notice that these four sentences have the same meaning in truth-functional logic and that the truth-values can be calculated by applying the relevant rules. For this reason, I will not bother the readers with a rather technical way of explicating that these sentences do mean the same in classical truth-functional logic. Still, if this kind of reasoning is considered, Jim does not divide the antecedent and the consequent of the clause “If Carr is out, then if Allen is out, Brown is in” wrongly as Joe accuses him of.

Question 3

The third and last question, which is connected to the Barber-Shop Paradox, is whether an if-then clause, whose antecedent is false, can be regarded as legitimate. As my explanation of the material implication shows, it is not only *possible* that such clauses are true, but they *are always* true within the rules of truth-functional logic. For they are true by definition of the material implication and the negation. However, this does not yet give an answer to the question above. One could even think that a sentence which is true by definition cannot be legitimate, because it does not deliver any insight, for it is vacuous. In fact, in mathematics and truth-functional logic one works with if-then clauses with a false antecedent. Thus, they are probably regarded as legitimate in these fields. Thus, the clause “If you were to run from here to London in five minutes, you would surprise people” is probably legitimate within truth-functional logic even though the addressee of that sentence does not actually run from there to London. However, a complete answer to this question depends on what is meant with “legitimate”. Since Carroll does not say whether the use of such sentences is sufficient for legitimacy, it might be that an additional virtue has to be on hand.

2.2 Unsolved Conflict

As it is seen through my explication given above, the three questions about if-then clauses raised by the Barber-Shop Paradox can be answered by means of truth-functional logic. Although the Barber-Shop Paradox is not paradoxical in a logical sense (since it contains no contradiction with regard to truth-functionality), not the entire conflict is solved. Even though Carroll was aware of the fact that his puzzle was not a real paradox, he was discussing the contentious points of the Barber-Shop Paradox with other

logicians. For he considered his paradox as “a very real difficulty in the Theory of Hypotheticals” (Carroll 1977, 442) because the usual intuition with regard to the questions about if-then clauses does not always match with the logical answer. One thing which is symptomatic of the latter is the following: With the material implication, which is normally taught to students of logic at the very beginning of their studies, most are struggling (Edgington 2014, sect. 2.1). This can be explained with the fact that material implication forces us to accept, for example, the sentence “If Zurich is the capital of Switzerland, then all dogs have three eyes”, or as we have seen “If Allen is out, then Brown is in and if Allen is out, then Browns is out”. However, sentences of this kind do not seem to be intuitively right.

In this subsection, I will point out the remaining problems of the Barber-Shop Paradox and argue that they result from two ways of understanding if-then clauses. For this purpose, I am again addressing the three questions one after the other, and I will show why the answers of truth-functional logic are not satisfactory.

Question 1

As we have seen, the two if-then clauses “If Allen is out, then Brown is in” and “If Allen is out, then Brown is out” are compatible in truth-functional logic because the falsity of the antecedent is sufficient for the truth of the whole conditional. This rule, however, which makes the two clauses compatible, is not uncontroversial. Although, within truth-functional logic, there is nothing to be said against this compatibility, it is only a stipulated one. Thus, the solution which is delivered by truth-functional logic to this contentious point does not really solve the conflict between the two opinions about the compatibility.

By differentiating an everyday understanding of if-then clauses from the material implication, I will now exemplify why two if-then clauses with the same antecedent but two contrary consequents do not seem compatible at first glance. Consider the following two sentences which are compatible within truth-functional logic: “If I touch this wire, I get an electric shock” and “If I touch this wire, I do not get an electric shock” (Edgington 2014, sect. 2.3). These sentences are only compatible if the consequents do not depend on the antecedent, in other words; if getting a shock is not a *consequence* (not an effect) of touching the wire. For if it were a consequence, the touching would cause the shock and we could not honestly say “If I touch this wire, I get an electric shock” and “If I touch

this wire, I do not get an electric shock”. This is however exactly what we can honestly say by applying material implication.

Hence, by using if-then clauses in everyday language, we do not always consider the rules of material implication. Rather sometimes, we connect the antecedent and the consequent in such a way that the “if” is thought as “because”. For in everyday life, we consider happenings in order to understand what someone says. We do not calculate truth-values of sentences. Now we can see why the compatibility of two if-then clauses with the same antecedent but with two contrary consequents is counterintuitive. By replacing the “if” with “because” we actually get incompatible sentences. The touching of a specific wire cannot be the cause of a shock I am getting when, at the same time, it causes the contrary as well. However, this incompatibility is not a logical one but one considering happenings.

In summary, my remarks indicate that in everyday life, while applying if-then clauses, we think in causes (since we consider happenings), not in material implication (since we do not calculate truth-values).

Question 2

The second question is concerned with the discussion between Joe and Jim about the meaning of the sentence “If Carr is out, then, if Allen is out, Brown is in”. As we have seen, truth-functional logic regards these sentences as equivalent to the sentence “If Carr and Allen are both out, Brown is in”. But Joe does not accept this equivalency, which indicates that it is not a matter of course. In the following, I am going to point out that, regarding natural language, it is in fact not that clear whether those sentences mean the same. As well as the first issue of the Barber-Shop Paradox, this second issue cannot be solved entirely by means of truth-functional logic, for we have to differentiate between a logical understanding and one considering our world.

Take the sentence which Joe uses as his premiss (1) “If Carr is out, then, if Allen is out, Brown is in”. What does this sentence mean? For my part, I have to admit that I do not really understand it intuitively. In truth-functional logic, it means the same as “If Carr and Allen are out, Brown is in”. But Joe’s premiss (1) and this sentence could mean something different in everyday life. As we live in a world where time passes, it is possible that the absence of one of the three barbers has to be thought as taking place before the absence of another barber. This is due to the fact that the barbers cannot just be in or out of a shop as they would in a timeless universe. Thus, it might be that in

our world the absence of Carr is temporally before the absence of Allen. For example, it might be that Allen only went out because Carr did the same thing before him. For instance, because the sister of Carr came in crying and told Allen that Carr's daughter had a terrible crash whereupon Allen went out to look for Carr because Carr's sister was so depressed that she had to stay there and lay down for a moment. Thus, considering everyday life, the sentence "If Carr is out, then if Allen is out, Brown is in" does not have to be understood in the same way as the sentence "If Carr and Allen are out, Brown is in". Since it might be that the absence of Carr is thought to be before the absence of Allen. Thus, the understanding of if-then clauses in everyday language is not detached from time as is the understanding of if-then clauses in a truth-functional way. On the one hand, the "then" in the sequence "If Carr is out, then if Allen is out" can be understood with an aspect of time. In this case, we have to imagine that Carr is out at a time when Allen is still in, and that after a while Allen goes out as well. On the other hand, the "then" in the same sequence can be understood more logically, which is considering the possibility that Carr and Allen are both out. In a truth-functional logical way of thinking, this possibility goes without any aspect of time, for in truth-functional logic we consider only information about truth-values in order to calculate other truth-values.

As my remarks show, if-then clauses can be understood in a logical manner or in a way which considers time. Since in our world time passes, it is always possible that something happens between two events and that the first event has to occur in order that the second event occurs as well. By differentiating logical thinking (which considers every imaginable possibility) from everyday thinking (which considers those possibilities that are imaginable within the world we live in, thus within a world where time passes) it should now be intelligible why we are able to understand if-then clauses in diverse ways.

Question 3

The third question pointed out by Carroll is whether if-then clauses with a false antecedent are legitimate. Although, within truth-functional logic, they probably are regarded as legitimate (since in truth-functional logic one works with such if-then clauses), we should take a closer look at it in order to answer this question entirely. Firstly, we need to find a criterion which makes sentences legitimate at all. Secondly, as soon as we have found a suitable criterion for the legitimacy

of sentences, we need to estimate whether all if-then clauses with a false antecedent satisfy this criterion.

In truth-functional logic, it only makes sense to work with sentences that carry a truth-value. Hence, only such sentences may be regarded as legitimate in this field. Since truth-functional logic ascribes to all if-then clauses with a false antecedent the truth-value *true*, they all carry a truth-value, and thus all such clauses may be regarded as legitimate. Therefore, if the criterion for the legitimacy of sentences is that we may ascribe a truth-value (either *true* or *false*) to them, we have to estimate whether this can actually be done.

Take the sentence which Joe utters in order to persuade Jim about the legitimacy: "If you were to run from here to London in five minutes, you would surprise people". At first glance, this seems to be true in Carroll's story even if the addressee does not run from the given place to London in five minutes. However, this is not enough to ascribe the truth-value *true* to all if-then clauses with a false antecedent. For we need to have a look at the context in which such sentences are expressed. If Joe says this to Jim in Oxford (maybe next to the Carfax Tower) and means with "London" the City Centre of London, then we could accept this sentence. If the two men are standing 300 feet next to the city boundaries of London and if with "London" the place behind these boundaries is meant, then the whole sentence does not seem to be true even if the antecedent is false. The person in question actually could run from there to London without great effort, he or she just does not do it. So, nobody would probably be surprised, if he or she did it. Hence, on the one hand, one and the same sentence seems to be true, and on the other hand, it does not. Whether we consider this sentence to be true, depends on the context in which it is expressed. This means that we can neither attribute "true" nor "false" to if-then clauses with a false antecedent without knowing something about the circumstances. But, if the circumstances are relevant in order to determine the truth-value of this kind of sentences, we cannot determine whether we regard this type of sentences as legitimate in the context of our world, in a general way.

All in all, we need to have more information than a formalised if-then clause with a false antecedent in order to determine whether it should be regarded as legitimate. For if we cannot ascribe either *true* or *false* to a sentence (what we cannot honestly do without context), it cannot be regarded as legitimate. We need to have information about the context in which such a sentence is expressed in order to answer the respective question.

2.3 A Different Kind of Solution

Even though truth-functional logic delivers a solution to the Barber-Shop Paradox, there remains an unsolved conflict. This conflict is due to two different kinds of understanding if-then clauses, as my argumentation in section 2.2 shows. In the story of the three barbers, a logical understanding is confused with an everyday understanding. Whereas the former is a formal approach, the latter depends on causes, time, and context and is rather intuitive. (In the following, I will mostly use the term “intuitive understanding” instead of “everyday understanding”. Furthermore, I will use the terms “logical understanding” and “rational understanding” interchangeable.) In this section, I will point out that the differentiation between these two understandings shall help to solve the Barber-Shop Paradox in a satisfying way. For this purpose, I will show on the one hand, the merit of truth-functional logic with regard to Carroll’s Paradox and on the other hand, that this merit is not enough to solve the difficulties of if-then clauses Carroll was concerned with. Thereafter, I will give an account of how my approach manages to solve these difficulties.

The merit of the logical solution is that it brings Joe’s errors of reasoning to light. The first and the second question which are discussed within the Barber-Shop Paradox contain such an error.

Regarding the first question, Joe’s error is the following: He claims that two if-then clauses with the same antecedent but two contrary consequents cannot be possibly true together. However, as shown by means of truth-functional logic, there is a way in which such clauses may be true together. Also, Jim tries to explicate this to Joe. Joe does not accept Jim’s explication, which is seen through his response “How can one and the same protasis prove two contradictory apodoses?”. It is this response that contains an error of reasoning. For the possibility that two if-then clauses with the same antecedent but with two contradictory consequents are jointly true does not mean that the antecedent *proves* two contrary consequents. The antecedent “Allen is out” does neither prove that Brown is out nor that Brown is in even though the sentences “If Allen is out, then Brown is out” and “If Allen is out, then Brown is in” may be true together. In order to prove something, we need context. So, in order to prove, for example, that Brown is out, the following context is needed:

1. It has to be the case that: Allen is out.
2. It has to be the case that: If Allen is out, then Brown is out.

Whereas 2. is granted by Jim and Joe, we have no information in regard of 1.; we do not know whether Allen is in or out. Thus, the central issue which Joe neglects, and which is brought to light by means of truth-functional logic (and Jim’s explication) is that, since we do not know whether Allen is in or out, no proof can be given of the whereabouts of Brown. Thus, Joe’s error of reasoning lies in his (unconscious) assumption, that an antecedent proves a consequent without any further information. But, in order that an antecedent may prove its consequent, the antecedent has to be the case. Regarding the second question, Joe’s error of reasoning can be shown as follows: Joe does not accept premiss (1) of Jim’s argument (“If Carr and Allen are both out, Brown is in”) as to be logically equivalent to his own premiss (1) (“If Carr is out, then if Allen is out, Brown is in”). But if we reason correctly, we have to accept this equivalency. As shown in section 1.2, Joe and Jim both build their arguments out of the two facts that (a) always, someone (we assume that it is one of the three barbers) has to be in the shop and that (b) Allen does not leave the shop without Brown. Considering these two facts, there are the following possibilities regarding the occupation of the barbershop:

1. Carr is out; Allen and Brown are in.
2. Carr and Brown are out; Allen is in.
3. Carr and Allen are in; Brown is out.
4. Carr is in; Allen and Brown are out.
5. Carr, Allen, and Brown are all in together.

Because Joe wants to prove by the method *Reductio ad Absurdum* that Carr is in, he assumes the contrary: Carr is out. At this point, there are still two possibilities remaining (1. and 2.). Thus, even if Carr is out, Joe cannot generate a contradiction out of the two facts ((a) and (b)), and therefore his assumption may be the case. So, Joe cannot conclude that Carr is in the shop. Considering all the possibilities listed above, it can be seen that Jim’s interpretation of Joe’s premiss (1) is correct in a rational sense (for it takes all possibilities into consideration), and that Joe does not recognise this, because he neglects some possibilities regarding the occupation of the barbershop.

Although a logical way of thinking brings Joe’s errors of reasoning to light and shows why Joe cannot prove logically that Carr is in the shop, it does not solve the problems with regard to if-then clauses which Carroll was concerned with. For in everyday language, we sometimes understand if-then clauses with an aspect of time and cause. Differentiating between a rational way

and an intuitive way of understanding if-then clauses could solve the problems in a satisfying way, as I intend to show with the following explanation:

Since the crux of the Barber-Shop Paradox is a clash between a rational way and an intuitive way of understanding if-then clauses, it cannot be tackled by means of logic. For logic cannot capture the intuitive way of understanding entirely. The nature of logic lies in a rational kind of understanding. However, understanding intuitively is not always rational. Intuitively, we would probably consider two sentences with the same antecedent but two contrary consequents as incompatible. For example, expressing the sentence “If I touch this wire, I get an electric shock and if I touch this wire, I do not get an electric shock” seems to be rather silly. However, only in the context of our world may this sentence be regarded as silly. While understanding this sentence intuitively, we consider (unconsciously) the consequents as to be dependent from the antecedent. This means that the former is thought to be an effect of the latter. Yet in a logical way of understanding if-then clauses, the consequent is not an effect of the antecedent, for something can only be an effect of something else if things are happening one after the other and if happenings may (in some way or another) be causes of other happenings, thus if we consider time and causality. However, a rational, logical way of handling if-then clauses considers *all* possibilities, even such that are assuming a timeless world, in which causality does not play a role, for example a pure mathematical world which does not have any content but is completely formal. Hence, the rational way of understanding if-then clauses cannot correspond entirely with the intuitive one. The reason for this is that whereas the latter considers (unconsciously) the world we live in and sometimes neglects very improbable possibilities, the former does not.

The differentiating of these two kinds of understanding if-then clauses does solve the Barber-Shop Paradox in a more satisfying way than a purely logical solution, for it respects the intuitive understanding just as the rational one. Respecting both of these understandings is important in order to solve the Barber-Shop Paradox because it seems—as my approach and the following diary entry of Carroll indicates—that the difficulties concerning if-then clauses arise from these two different kinds of understandings:

“My night’s thinking over the very puzzling subject of ‘Hypotheticals’ seems to have evolved a new idea—that there are two kinds,

(1) where the Protasis is *independent* of the Hypothetical, (2) where it is dependent on it.” (Carroll 1977, 448)

It seems to me that the two understandings which Carroll had in mind while writing this entry, accord with the two understandings which my approach reveals. Within the rational understanding the consequent does not depend on the antecedent, whereas within the intuitive understanding it does. Thus, the identification of the two understandings, which I suggest to be the base of the Barber-Shop Paradox, was the most important step in order to find satisfying answers. The next step is simply to present the answers—which arise from the two understandings—to the three questions discussed above:

1. **Question:** Are two if-then clauses, of the forms “if p , then q ” and “if p , then not q ” compatible?

Answer: In a rational way of thinking, they are, since it is possible that both of these clauses are true together (as it is shown in section 2.1). They are, however, not compatible within an intuitive understanding because the consequent seems to depend on the antecedent, and in this case, it is not possible that both of these clauses are true together (as it is shown in section 2.2).

2. **Question:** How do the following propositions differ (if they do)?

- (a) p , q and r cannot be all true at once
- (b) If p and r are true, q is not true
- (c) If r is true, then if p is true, q is not true
- (d) If p is true, then if r is true, q is not true

Answer: Rationally, they do not differ, since a rational understanding considers all possibilities, without any aspect of time. In this case, all of these propositions are equivalent (as it is shown in section 2.1). Intuitively, however, they differ in a way in which the propositions containing the word “then” may be understood with an aspect of time (as it is shown in section 2.2).

3. **Question:** Can an if-then clause, whose antecedent is false, be regarded as legitimate?

Answer: Yes, it can. Nonetheless, we cannot answer this question in the same way for all if-then clauses, since the answer depends on the context in which such sentences are uttered. Furthermore, it depends on what we would like to achieve with such if-then clauses. If we would like to solve logical or mathematical puzzles, if-then clauses with a false antecedent seem to be legitimate. Yet, if our

aim is to find out something about the world we live in, we should not consider if-then clauses with a false antecedent to be legitimate, provided that we know (or strongly assume) that the antecedent is false. Since, if the antecedent of an if-then clause is known (or strongly assumed) to be false, this clause is not useful for gaining knowledge about our world.

Conclusion

Lewis Carroll pointed out the Barber-Shop Paradox at the end of the 19th century—during a time when the logical understanding of if-then clauses was more heavily disputed than it is today. In virtue of the more mathematical approaches of the early 20th century, truth-functional logic offers a solution to this paradox. Since the truth-functional logical approach is clear, simple and consistent, it might bring the solution which Carroll had in mind while looking for an agreement among logicians. However, even if we might have found the best logical approach with regard to the Barber-Shop Paradox, the actual problem cannot be solved by means of logic, for a mathematical approach is not able to grasp our everyday, intuitive understanding of if-then clauses entirely. We cannot calculate all things that are inherent to our world. There are some things—like time and cause—which a formal system is not able to capture.

The aim of this essay was to show that the Barber-Shop Paradox is not a logical problem but a confusion between a logical (or rational) understanding and an everyday (or intuitive) understanding of if-then clauses. Thus, in order to solve the Barber-Shop Paradox entirely, we need to differentiate between these two understandings, and give answers to the relevant questions with regard to them. Therefore, the solution I am proposing to the Barber-Shop Paradox is not a logical solution, since this paradox is not a logical problem. Rather, my approach solves the Barber-Shop Paradox by disentangling a conflict between intuition and rational thinking.³

References

- Bobzien, Susanne. 2019. "Dialectic School." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2019. <https://plato.stanford.edu/archives/spr2019/entries/dialectical-school/>
- Carroll, Lewis. 1894. "A Logical Paradox." *Mind* 3 (11): 436–440.
- Carroll, Lewis. 1977. *Lewis Carroll's Symbolic Logic, Part I, Elementary, 1896, Fifth Edition, Part II, Advanced, never previously published, Together with Letters from Lewis Carroll to eminent nineteenth-century Logicians and to his "logical sister," and eight versions of the Barber-Shop Paradox*. Edited by William Warren Bartley. New York: Clarkson N. Potter, Inc.
- Edgington, Dorothy. 2014. "Indicative Conditionals." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2014. <https://plato.stanford.edu/archives/win2014/entries/conditionals/>
- Klement, Kevin C. 2019. "Propositional Logic." In *The Internet Encyclopedia of Philosophy*, edited by James Fieser and Bradley Dowden. Accessed September 5, 2019. <https://www.iep.utm.edu/prop-log/>
- Sedley, David. 2018. "Diodorus Cronus." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2018. <https://plato.stanford.edu/archives/win2018/entries/diodorus-cronus/>
- Smith, Robin. 2019. "Aristotle's Logic." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2019. <https://plato.stanford.edu/archives/sum2019/entries/aristotle-logic/>
- Zalta, Edward N. 2019. "Gottlob Frege." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2019. <https://plato.stanford.edu/archives/win2019/entries/frege/>

Vera Moser schreibt eine strafrechtsphilosophische Dissertation am Institut für Strafrecht und Kriminologie der Universität Bern und studiert im 7. Semester Philosophie. Sie interessiert sich vorwiegend für die Schnittstelle zwischen Philosophie und Wissenschaft. In ihrem Studium beschäftigt sie sich mit Wissenschaftsphilosophie, Methoden der Philosophie, Logik und Strafrechtsphilosophie.

³ I would like to express my very great appreciation to Michael Müller for critical remarks and valuable advice.

Die Rolle der Selbstbestimmung in der Rechtfertigung passiver und aktiver Sterbehilfe*

1. Zur Debattenlage

Wann und auf welche Weise dürfen Ärzte am Tod eines Patienten mitwirken? In der Auseinandersetzung über diese Frage wird oft zwischen verschiedenen Arten von Fällen unterschieden, für welche die beiden folgenden paradigmatische Beispiele sind:

Fall 1: Ein sterbenskranker Patient leidet in aus seiner Sicht unerträglichem Maße an starken Schmerzen und den Einschränkungen der Kommunikations- und Bewegungsfähigkeit, die seine Erkrankung unweigerlich mit sich bringt. Er erklärt darum, dass er nicht wiederbelebt werden möchte, sollte sein Herz-Kreislauf-System versagen. Es kommt zu einem solchen Versagen und sein Arzt unternimmt nichts, um das Herz des Patienten wieder zum Schlagen zu bringen, woraufhin dieser verstirbt.

Fall 2: Ein sterbenskranker Patient leidet in aus seiner Sicht unerträglichem Maße an starken Schmerzen und den Einschränkungen der Kommunikations- und Bewegungsfähigkeit, die seine Erkrankung unweigerlich mit sich bringt. Er bittet darum seinen Arzt, sein Leben auf möglichst rasche und schmerzfreie Weise zu beenden. Der Arzt kommt dieser Bitte schließlich nach und verabreicht dem Patienten eine tödliche Injektion, woraufhin dieser verstirbt.

In beiden Fällen handelt es sich um ärztliche Sterbehilfe. Darunter sollen hier alle Vorgänge verstanden werden, in denen ein Arzt am Herbeiführen des Todes eines Patienten beteiligt ist,¹ dies dem Willen des Pati-

enten entspricht und der Arzt in Kenntnis dieses Patientenwillens sowie der (voraussichtlich) tödlichen Folgen für den Patienten handelt. Ein Herbeiführen des Todes liegt dann vor, wenn der Zeitpunkt des Todes des Patienten so beeinflusst wird, dass dieser signifikant früher eintritt, als andernfalls zu erwarten gewesen wäre. Natürlich können nicht nur Ärzte, sondern etwa auch Angehörige Sterbehilfe leisten. Dennoch beschränke ich mich hier auf die Diskussion ärztlicher Sterbehilfe, da in den meisten Fällen bei der Beurteilung der Situation, der Aufklärung des Patienten sowie der Mitwirkung am Herbeiführen des Todes medizinische Fachkenntnisse erforderlich sind und Ärzte folglich mit der Entscheidung für oder gegen ein solches Mitwirken besonders häufig konfrontiert sind. Zudem wirft der Umstand, dass eine bestimmte Handlung durch einen Arzt ausgeführt wird, zusätzliche Fragen auf, die, wie wir noch sehen werden, im Kontext der Sterbehilfe besonders relevant sind.

Der Unterschied zwischen den beiden genannten Fällen wird insbesondere in der öffentlichen Diskussion häufig mit dem Begriffspaar aktiv/passiv in Verbindung gebracht. In Fall 1 leistet der Arzt demzufolge *passive* Sterbehilfe, während es sich bei Fall 2 um *aktive* Sterbehilfe handle. Über die prinzipielle Zulässigkeit der ersten Form von Sterbehilfe besteht dabei breites Einvernehmen. Kontroverser ist dagegen die moralische Bewertung von aktiver Sterbehilfe, wie sie Fall 2 illustriert. Vertreter einer restriktiven Position halten lediglich die passive Sterbehilfe für zulässig, lehnen aktive Sterbehilfe aber als grundsätzlich moralisch un-

* Dieser Text ist erstmalig erschienen im *Jahrbuch für Wissenschaft und Ethik*, Band 20, Heft 1 (2016), S. 5–27 (<https://doi.org/10.1515/jwiet-2016-0103>). Wiederabdruck mit freundlicher Genehmigung des Verlags.

1 „Beteiligung“ soll hier zunächst so weit verstanden werden, dass auch der Arzt an der Herbeiführung des Todes eines Patienten beteiligt ist, der ein Mittel (beispielsweise eine tödliche Dosis ei-

nes Betäubungsmittels) bereitstellt, mit dessen Hilfe sich der Patient selbst tötet. Ärztliche Suizidbeihilfe gilt demnach als eine Form von Sterbehilfe. Davon ist unbenommen, dass die Art oder der Grad der Beteiligung des Arztes – d.h. das, worin sich das Bereitstellen eines tödlichen Medikaments vom Injizieren eines solchen Medikamentes unterscheidet – einen Unterschied für die moralische Bewertung seines Handelns machen kann. Ob dies so ist, wird im Folgenden jedoch nicht eigens untersucht.

zulässig ab. Andere beziehen eine liberale Position, der zufolge auch aktive Sterbehilfe in bestimmten Situationen moralisch zulässig sein kann.²

Grundsätzlich gilt, dass wenn zwei Handlungen sich in ihrem moralischen Status unterscheiden, auch ein moralisch relevanter Unterschied zwischen diesen Handlungen bestehen muss. Handlungen, die sich in keiner moralisch relevanten Hinsicht unterscheiden, müssen auch hinsichtlich ihrer moralischen Zulässigkeit gleich beurteilt werden. Die restriktive Position, die ja die Zulässigkeit passiver Sterbehilfe nicht bestreitet, setzt somit voraus, dass die Unterscheidung zwischen aktiver und passiver Sterbehilfe einen moralisch relevanten Unterschied markiert. Daraus ergibt sich, dass eine Möglichkeit zur Begründung der liberalen Position darin besteht zu zeigen, dass zumindest in bestimmten Situationen kein moralisch relevanter Unterschied zwischen aktiver und passiver Sterbehilfe besteht. Lässt sich diese Äquivalenzthese verteidigen, so kann aus ihr, zusammen mit der Zulässigkeit passiver Sterbehilfe, angesichts des soeben erläuterten Zusammenhanges gefolgert werden, dass auch aktive Sterbehilfe in bestimmten Situationen zulässig ist. Dieses Äquivalenzargument für die Zulässigkeit aktiver Sterbehilfe hat einen festen Stand in der gegenwärtigen Debatte.³

Das Äquivalenzargument ist aber natürlich nicht die einzige Option für die Begründung einer liberalen Position. Alternativ kann man auch direkt für die Zulässigkeit aktiver Sterbehilfe zu argumentieren versuchen. Eine solche Argumentationsweise rekurriert in ihren Prämissen nicht auf die Zulässigkeit passiver Sterbehilfe; sie benennt vielmehr die für und gegen die aktive Sterbehilfe sprechenden Gesichtspunkte und beansprucht zu zeigen, dass erstere letztere in man-

chen Situationen überwiegen. Dabei wird meist in der einen oder anderen Weise auf das Selbstbestimmungsrecht bzw. die Autonomie des Patienten Bezug genommen.⁴ Aktive Sterbehilfe ist demzufolge zulässig, wenn und insoweit sie dem (autonomen) Patientenwillen entspricht, den es auch im Falle von Entscheidungen über den eigenen Tod zu respektieren gelte. Eine solche Argumentation kann zugestehen, dass es einen moralisch relevanten Unterschied zwischen der aktiven und der passiven Sterbehilfe gibt, bestreitet jedoch, dass dieser stets den Ausschlag gegen die Zulässigkeit aktiver Sterbehilfe gibt.

Hier zeigt sich, dass das Bestreiten der Äquivalenzthese zwar notwendig, aber allein nicht hinreichend für die Begründung einer restriktiven Position ist. Zunächst muss der moralisch relevante Unterschied gewissermaßen auch in die richtige Richtung weisen, d.h. es muss gelten, dass aktive Sterbehilfe aufgrund dieses Unterschiedes moralisch problematischer als passive Sterbehilfe ist, oder dass zur Rechtfertigung der letzteren mehr argumentative Ressourcen zur Verfügung stehen als für erstere. Vor allem aber müssen zwei Handlungen, die sich in solcher Weise unterscheiden, sich deshalb nicht auch hinsichtlich ihrer moralischen Zulässigkeit unterscheiden. In beiden Fällen kann schließlich so viel für die Handlung sprechen, dass es die problematischen Gesichtspunkte überwiegt und die Handlung folglich zulässig ist, obwohl die gegen oder für die Handlung sprechenden Gesichtspunkte im einen Fall gewichtiger bzw. weniger gewichtig sind als im anderen. Der Vertreter einer restriktiven Position ist also nicht nur darauf festgelegt, dass es einen moralisch relevanten Unterschied zwischen aktiver und passiver Sterbehilfe gibt, sondern auch darauf, dass dieser moralisch entscheidend ist, dass also anders als im Falle der passiven Sterbehilfe bei der aktiven Sterbehilfe die moralisch problematischen Gesichtspunkte nicht durch die zur Verfügung stehenden Rechtfertigungsressourcen aufgewogen werden können.

Diese Darstellung der dialektischen Situation in der Diskussion über die Zulässigkeit aktiver Sterbehilfe macht deutlich, dass die Frage, ob die Unterscheidung zwischen aktiver und passiver Sterbehilfe einen moralisch relevanten Unterschied markiert, sowohl für Vertreter einer liberalen als auch für Vertreter einer restriktiven Position von zentraler Bedeutung ist. Für erstere würde eine negative Antwort die entscheidende Prämisse des Äquivalenzargumentes für die Zulässigkeit der aktiven Sterbehilfe bereitstellen, während

2 Beide hier genannten Positionen betreffen die moralische Zulässigkeit von Sterbehilfe, nicht die Frage, ob diese rechtlich erlaubt oder verboten werden sollte. Die moralische Bewertung von Handlungsweisen ist zwar gewiss nicht irrelevant für deren rechtliche Regelung. Es ist aber weder inkonsistent, sich gegen ein rechtliches Verbot von moralisch unzulässigen Handlungen (z.B. Vertrauensbrüchen) auszusprechen, noch für ein rechtliches Verbot von Handlungsweisen zu plädieren, die man in bestimmten Fällen für moralisch zulässig hält (etwa weil andernfalls damit zu rechnen wäre, dass solche Handlungen vielfach auch in Situationen ausgeführt würden, in denen sie nicht zulässig sind). Im Folgenden geht es allein um die moralische Bewertung verschiedener Formen von Sterbehilfe.

3 Vgl. etwa Wolf (1993, 414), Rachels (1989, 261ff.), Birnbacher (1995, 348ff.), Schöne-Seifert (1997, 217ff.) und Sumner (2011, 23). Grundsätzlich erlaubt die Äquivalenzthese natürlich auch eine Argumentation in entgegengesetzter Richtung, d.h. von der Unzulässigkeit aktiver für die Unzulässigkeit passiver Sterbehilfe. Angesichts des Umstandes, dass der philosophische wie auch der gesellschaftliche Dissens die Zulässigkeit der aktiven und gerade nicht die der passiven Sterbehilfe betrifft, wird diese Option, der es zudem in der gegenwärtigen Diskussion an ernst zu nehmenden Vertretern mangelt, hier jedoch nicht weiter berücksichtigt.

4 Vgl. Dworkin (1993, Kap. 7), Quante (1998 sowie 2002, Kap. 6) und Gutmann (2002).

für letztere eine positive Antwort als erster Schritt für die Begründung der Unzulässigkeit aktiver Sterbehilfe erforderlich ist. Die Beantwortung der Frage wird jedoch dadurch erschwert, dass die aktiv/passiv-Unterscheidung in der öffentlichen Diskussion sowie in den verschiedenen Fachdebatten ganz unterschiedlich verstanden wurde und wird und somit die Gefahr besteht, dass die Unterscheidung, die der eine für moralisch irrelevant hält, gar nicht die Unterscheidung ist, die der andere zur Begründung seiner Position heranzuziehen gedenkt.

Vor diesem Hintergrund gehe ich im vorliegenden Beitrag der Frage nach, ob und wie sich die These der Unzulässigkeit aktiver Sterbehilfe angesichts des Konsenses über die Zulässigkeit passiver Sterbehilfe begründen lässt. Dafür stelle ich zunächst die gegenwärtig verbreitetste und überzeugendste Explikation der aktiv/passiv-Unterscheidung vor und grenze sie von anderen, gelegentlich noch immer für Verwirrung sorgenden Verständnisweisen ab. Anschließend wird untersucht, inwiefern die so verstandene Unterscheidung einen moralisch relevanten Unterschied markiert. Hierfür ist eine detaillierte Diskussion der Rolle, die die Einwilligung des Patienten in der Rechtfertigung aktiver und passiver Sterbehilfe spielt, sowie des Verhältnisses solcher Einwilligungen zu anderen Elementen der Selbstbestimmung des Patienten erforderlich. Die Ergebnisse dieser Diskussion zeigen zum einen, dass durchaus ein Unterschied in den zur Verfügung stehenden Rechtfertigungsressourcen für aktive und passive Sterbehilfe besteht, sodass die Äquivalenzthese als zentrale Prämisse des Äquivalenzargumentes zurückgewiesen werden muss. Zum anderen aber machen diese Ergebnisse deutlich, dass Vertreter der restriktiven Position ein direktes Argument für die Zulässigkeit aktiver Sterbehilfe, das an der Einwilligung als Ausprägung der Selbstbestimmung des Patienten ansetzt, nur durch Rekurs auf nicht plausibel begründbare Thesen zurückweisen können. Schließlich zeige ich, dass ein angemessenes Verständnis der Rolle von Selbstbestimmung es außerdem ermöglicht, zwei verbreitete Einwände gegen ein solches direktes Argument zurückzuweisen.

2. Sterbehilfe und Behandlungsverzicht

In der medizinethischen und rechtlichen Fachdiskussion ist es üblich geworden, die verschiedenen Arten von Sterbehilfe, für die Fall 1 und Fall 2 jeweils paradigmatisch sind, danach zu unterscheiden, ob sie mit dem Verzicht auf eine medizinische Behandlung ver-

bunden sind, die für den Patienten lebensnotwendig bzw. lebensverlängernd wäre.⁵ Von einigen ist diese Unterscheidung anhand des Behandlungsbezugs als ein Ersatz für die als irreführend oder zu unklar wahrgenommene aktiv/passiv-Unterscheidung vorgeschlagen worden.⁶ Angesichts der fest verankerten Position, die letztere Unterscheidung insbesondere in der öffentlichen Debatte hat, erscheint es jedoch sinnvoller, nicht zu versuchen, die etablierten Begrifflichkeiten abzuschaffen und durch andere zu ersetzen, sondern das Kriterium des Behandlungsbezuges als eine Explikation der aktiv/passiv-Unterscheidung zu betrachten und an der Rede von aktiver und passiver Sterbehilfe schon aus pragmatischen Gründen festzuhalten. Passive Sterbehilfe besteht danach im Herbeiführen des Todes durch den Verzicht auf oder den Abbruch einer lebenserhaltenden bzw. -verlängernden medizinischen Maßnahme.⁷ Aktive Sterbehilfe dagegen liegt vor, wenn der Tod des Patienten in einer Weise herbeigeführt wird, die keinen solchen Bezug zu einer lebenserhaltenden Maßnahme hat.

Diese Explikation der Unterscheidung zwischen aktiver und passiver Sterbehilfe ist von anderen Differenzierungsversuchen abzugrenzen, die ebenfalls mit der aktiv/passiv-Unterscheidung in Verbindung gebracht werden. Dies ist zum einen die Unterscheidung zwischen Tun und Unterlassen als verschiedenen performativen Modi menschlichen Handelns im weiteren Sinne.⁸ Fälle wie der folgende machen klar, dass diese sich nicht mit der Unterscheidung zwischen aktiver und passiver Sterbehilfe deckt:

Fall 3: Ein Arzt schaltet in Übereinstimmung mit dem Willen eines schwerstkranken Patienten dessen Beatmungsmaschine aus, woraufhin dieser verstirbt.⁹

5 Vgl. exemplarisch Schöne-Seifert (2007, 114) und Ach et al. (2012, 234) sowie Nationaler Ethikrat (2006, 54), Bundesgerichtshof (2010, Rn. 31) und Bundesärztekammer (2011a, A346).

6 Vgl. etwa Nationaler Ethikrat (2006, 54) und Bundesgerichtshof (2010, Rn. 31).

7 Die Rede von „Maßnahmen“ anstelle von „Behandlungen“ erlaubt es, Komplikationen zu vermeiden, die sich zum einen daraus ergeben, dass eine Behandlung einen behandlungsbedürftigen Zustand voraussetzen scheint und darum in Frage gestellt werden kann, ob es sich beispielsweise beim Einstellen einer künstlichen Ernährung um einen Behandlungsabbruch handelt, vgl. Höfling/Rixen (2003, 888f.). Zum anderen kann auch der Verzicht auf eine für die erfolgreiche Durchführung einer lebensverlängernden Behandlung erforderliche diagnostische Maßnahme den Tod des Patienten herbeiführen.

8 Zu dieser Unterscheidung und ihrer Relevanz im Kontext der Sterbehilfe-Problematik vgl. ausführlich Birnbacher (1995).

9 Wie die Fälle 1 und 2 dient auch dieser Fall dazu, spezifische, für den Argumentationsgang zentrale Aspekte verschiedener Arten der Sterbehilfe herauszustellen. Dies macht es erforderlich, von zahlreichen anderen Aspekten zu abstrahieren, wodurch die Be-

Hier wird der Tod des Patienten durch ärztliches Tun herbeigeführt. Dennoch werden solche Fälle gemeinhin – und in Übereinstimmung mit der oben vorgestellten Explikation – als passive Sterbehilfe klassifiziert.

Auch der Kontrast zwischen Formen der Sterbehilfe, bei denen es sich um eine *Tötung* des Patienten handelt, und solchen, für die dies nicht gilt, ist bei der Explikation der Unterscheidung zwischen aktiver und passiver Sterbehilfe wenig hilfreich.¹⁰ Versteht man diesen Kontrast so, dass eine Person nur dann eine andere tötet, wenn sie dadurch eine moralische Pflicht verletzt und also zumindest *prima facie* moralisch unzulässig handelt,¹¹ dann eignet er sich nicht, um eine unterschiedliche moralische Bewertung der entsprechenden Handlungen zirkelfrei zu begründen. Meint man dagegen, dass nur derjenige einen anderen tötet, der den Tod dieser Person (mit-)verursacht,¹² läge auch in Fall 3 eine Tötung vor.¹³

Eine weitere Unterscheidung, die in der Diskussion über die Zulässigkeit verschiedener Formen von Sterbehilfe eine Rolle spielt, betrifft die mentalen Zustände des Handelnden: Beabsichtigt dieser, den Tod des Patienten herbeizuführen, oder nimmt er dessen Tod lediglich als unbeabsichtigte, aber nicht vermeidbare Nebenfolge seines Handelns in Kauf? Hier ist oft von direkter bzw. indirekter Sterbehilfe die Rede. Diese Unterscheidung steht quer zu der zwischen aktiver und passiver Sterbehilfe, denn der Verzicht auf eine intensivmedizinische Behandlung kann ebenso wie das Verabreichen eines medizinisch indizierten Schmerzmittels sowohl mit der Absicht erfolgen, Leid zu mindern, als auch mit der, den Tod zu beschleunigen. Als Explikation des durch Fall 1 und Fall 2 illustrierten Unterschiedes zwischen aktiver und passiver Sterbehilfe eignet sich die Unterscheidung zwischen Beabsichtigen und Inkaufnehmen folglich nicht.

Manche halten diese Unterscheidung allerdings für die eigentlich moralisch relevante und beurteilen da-

her nur direkte aktive Sterbehilfe als unzulässig. Zur Begründung der unterschiedlichen Bewertung von Beabsichtigen und Inkaufnehmen des Todes berufen sie sich dabei auf die Doppeleffekt-Doktrin. Dieser zufolge hängt die Zulässigkeit einer Handlung mit sowohl guten als auch schlechten Folgen davon ab, ob der Handelnde die schlechten Folgen (als Zweck oder Mittel) zu realisieren beabsichtigt oder sie nur als unvermeidbaren Nebeneffekt der Realisierung der guten Folgen in Kauf nimmt. Auch abgesehen von grundsätzlichen Zweifeln an der Plausibilität dieser Doktrin¹⁴ ist jedoch fraglich, inwiefern sie in den zur Diskussion stehenden Fällen überhaupt angewendet werden kann. Schließlich würde sie damit, dass es sich beim Tod des Patienten auch dann um eine *schlechte* Folge handelt, wenn dieser der Vermeidung von großem Leid dient und dem Patientenwillen entspricht, etwas voraussetzen, was viele Befürworter einer liberalen Position bestreiten.¹⁵ Angesichts dieses und weiterer Probleme eines solchen Differenzierungsversuchs¹⁶ steht hier deshalb die Unterscheidung zwischen aktiver und passiver Sterbehilfe im eingangs explizierten Verständnis sowie die Frage nach ihrer moralischen Relevanz im Mittelpunkt.

3. Ist die Unterscheidung moralisch relevant?

Ist die so verstandene Unterscheidung zwischen passiver und aktiver Sterbehilfe moralisch signifikant? Bringt die Unterscheidung einen Unterschied in den Rechtfertigungsressourcen mit sich, die für passive und aktive Sterbehilfe zur Verfügung stehen? Die Antwort auf diese Frage ist einigermaßen kontrovers. Das liegt unter anderem daran, dass man zwei verschiedene Aspekte der medizinischen Maßnahme, auf die im Rahmen passiver Sterbehilfe verzichtet werden soll, in den Vordergrund stellen und auf ihre moralische Relevanz hin untersuchen kann, nämlich einerseits, *was* die Maßnahme bewirken würde, und andererseits, *wie* sie es bewirken würde.

Angesichts der Explikation von passiver Sterbehilfe als Verzicht auf lebensverlängernde Maßnahmen (oder als deren Abbruch) liegt es nahe, den Blick zunächst

schreibungen dieser Fälle an Nähe zu konkreten Situationen einbüßen. Gleichwohl sind viele reale Situationen ihnen in den relevanten Hinsichten ähnlich.

10 „Töten“ wird in diesem Zusammenhang oft mit „Sterbenlassen“ kontrastiert. Der Aspekt des bloßen Geschehenlassens des Todes zeichnet in der Tat die passive Sterbehilfe aus; seine Relevanz wird unten in Abschnitt 3 diskutiert.

11 So etwa Gert et al. (2006, 322f.), die allerdings auch nicht vorschlagen, die aktiv/passiv-Unterscheidung mit der zwischen Töten und Sterbenlassen zu identifizieren.

12 So etwa Brody (1989, 302), vgl. auch Rachels (1989, 261f.).

13 Zudem erlaubt es dieser Vorschlag u.U. nicht einmal, zwischen Fall 1 und Fall 2 zu differenzieren, denn einer plausiblen Kausalitätskonzeption zufolge können auch Unterlassungen ein Ereignis (mit-)verursachen, vgl. Birnbacher (1995, Kap. 3).

14 Einen aktuellen Überblick über die Diskussion bietet McIntyre (2014).

15 Dies gilt beispielsweise für Anhänger einer Konzeption des Unwertes des Todes, wonach dieser nur insoweit schlecht ist, als das durch ihn verhinderte weitere Leben selbst ein Gut darstellt, letzteres jedoch nicht immer der Fall ist. Vgl. dazu exemplarisch Williams (1973), Wittwer (2003, 241ff.) und Sumner (2011, 11ff.).

16 Vgl. die kritischen Diskussionen bei Birnbacher (1995, 153ff. und 345f.), Schöne-Seifert (1997, 220ff.), Quante (2002, 248f.) sowie Sumner (2011, Kap. 3).

auf den erstgenannten Aspekt und insbesondere auf den folgenden Umstand und dessen Rechtfertigungspotential zu richten: Bei der passiven Sterbehilfe wird der Tod des Patienten wesentlich (wenn auch nicht allein) durch einen unabhängigen Prozess verursacht, der bereits vor Beginn der Maßnahme seinen Anfang genommen hat. Dieser könnte durch die Maßnahme allerdings aufgehalten oder zumindest verzögert werden. Wer passive Sterbehilfe leistet, lässt somit lediglich einem unabhängigen Prozess seinen Lauf, d.h. er lässt etwas geschehen, wobei ein solches Geschehenlassen sowohl durch Unterlassen (z.B. einer Wiederbelebung) als auch durch Tun (z.B. Ausschalten einer Beatmungsmaschine) erfolgen kann.¹⁷ Im Gegensatz dazu wird bei der aktiven Sterbehilfe der zum Tod des Patienten führende Prozess erst in Gang gesetzt – hier stirbt der Patient somit nicht in Folge eines unabhängigen oder, wie es dann häufig heißt, nicht in Folge eines *natürlichen* Sterbeprozesses. Macht dies einen moralischen Unterschied? Hier ist Skepsis angebracht. Erstens ist fraglich, ob der Verweis auf die vermeintliche Natürlichkeit des Todes bei der passiven im Gegensatz zur aktiven Sterbehilfe etwas auszutragen vermag. Zunächst ist alles andere als klar, unter welchen Bedingungen der Prozess, der zum Tod eines Patienten führt, als ein „natürlicher“ gelten kann, und ob dies bei passiver Sterbehilfe stets der Fall ist: Inwiefern ist beispielsweise der durch ein ärztlich verabreichtes oder bereitgestelltes Betäubungsmittel herbeigeführte Tod weniger natürlich als der eines durch einen Sportunfall halsabwärts Gelähmten, bei dem die Beatmungsmaschine abgeschaltet wird? Zwar wird – in einem aus Krimis und ähnlichen Kontexten bekannten Sinne – etwa beim Ausfüllen eines Totenscheins zwischen einem natürlichen und einem nicht-natürlichen Tod unterschieden. Dabei handelt es sich jedoch um eine forensische, d.h. mit einem bestimmten Erkenntnisinteresse getroffene Unterscheidung, die darauf abzielt, einen Suizid oder ein Fremdverschulden des Todes auszuschließen. Sie klassifiziert darum auch den Tod infolge des Abbruchs einer künstlichen Beatmung als „nicht natürlich“, wenn die Beatmung beispielsweise durch eine bei einer Schlägerei erlittene Hirnverletzung erforderlich geworden ist. Zudem und vor allem aber ist nicht ersichtlich, warum und inwiefern es besser oder weniger problematisch sein sollte, wenn jemand infolge eines natürlichen Prozesses stirbt. „Die“ Natur ist weder mitfühlend noch wohlwollend oder gnädig, und natürliche Prozesse werden oft als ein

Übel empfunden, das wir gerade auch mit den Mitteln der Medizin aufzuhalten versuchen.¹⁸

Zweitens macht der Umstand, dass ein Tod Folge eines unabhängigen, nicht-aufgehaltenen Prozesses und nicht eines durch den Arzt angestoßenen Prozesses ist, zumindest in manchen Kontexten keinen Unterschied für die moralische Bewertung. Dies illustrieren die folgenden zwei Fälle. Im ersten versäumt es ein Arzt, die Infektion eines Patienten rechtzeitig zu diagnostizieren, weil er die Ergebnisse eines Bluttests in der Krankenakte übersieht, so dass der Patient an den Folgen der Infektion stirbt. Im zweiten Fall infiziert der Arzt einen Patienten versehentlich mit einem Erreger, weil er sich vor der Untersuchung die Hände nicht desinfiziert, und auch hier verläuft die Infektion tödlich. Es erscheint mir nicht plausibel, hier moralisch differenzieren zu wollen. In beiden Fällen verletzt der Arzt seine Sorgfaltspflicht, und in beiden Fällen kommt es so zu einem Todesfall, den der Arzt leicht hätte vermeiden können. Aber nur im ersten Fall kann der Tod auf einen unabhängigen Prozess zurückgeführt werden, den der Arzt lediglich nicht aufhält. Dieser Umstand ist hier also anscheinend moralisch irrelevant. Daraus folgt zwar nicht zwingend, dass dieser Unterschied auch in anderen Kontexten moralisch nicht signifikant ist.¹⁹ Es zeigt aber, dass sich die These, im Falle der Sterbehilfe sei an dieser Stelle moralisch zu differenzieren, nur aufrechterhalten lässt, wenn zugleich begründet werden kann, warum der Unterschied im Falle von Sterbehilfe eine unterschiedliche moralische Bewertung rechtfertigen sollte, in den genannten Fällen versehentlichen tödlichen Handelns aber nicht. Es ist meines Erachtens nicht zu sehen, wie eine solche Begründung aussehen könnte.²⁰

Die verbreitete Skepsis bezüglich der moralischen Relevanz des Umstandes, dass der Tod eines Patienten durch das Geschehenlassen eines unabhängigen Prozesses herbeigeführt wird, erscheint mir daher berechtigt. Folgt daraus, dass die Unterscheidung zwischen

18 Vgl. auch Birnbacher (1995, 123ff.), Schöne-Seifert (1997, 219) sowie Gutmann (2002, 179).

19 Dies zeigt Kagan (1988).

20 Ein wesentlicher Unterschied zwischen den in diesem Absatz geschilderten Fällen und solchen der Sterbehilfe besteht natürlich darin, dass das Herbeiführen des Todes im Falle der Sterbehilfe mit dem Willen des Patienten übereinstimmt. Dies deutet darauf hin, dass eine moralische Differenzierung zwischen Fall 1 (Unterlassung von Wiederbelebensmaßnahmen) und Fall 2 (Verabreichen einer tödlichen Injektion) in den unterschiedlichen Auswirkungen des übereinstimmenden Patientenwillens in diesen beiden Fällen begründet liegen könnte. Wie ich im Folgenden darlegen werde, halte ich diese Überlegung in der Tat für überzeugend; sie ist jedoch davon unabhängig, dass in Fall 1 der Tod des Patienten die Folge eines nicht aufgehaltenen, unabhängigen Prozesses ist.

17 Vgl. Birnbacher (1995, 110ff.) und Gesang (2001).

aktiver und passiver Sterbehilfe aus moralischer Sicht nicht relevant ist?²¹ Dieser Schluss wäre voreilig, denn er setzt voraus, dass der Status des Prozesses, auf den der Tod des Patienten zurückgeführt werden kann, der einzige Aspekt ist, der an der so verstandenen aktiv/passiv-Unterscheidung von Belang ist. Dass dem nicht so ist, wird deutlich, wenn man der Frage nachgeht, warum Fälle passiver Sterbehilfe überhaupt gerechtfertigt sind. Dies führt uns zu dem zweiten der oben genannten Aspekte.

4. Nicht aufhalten vs. nicht eingreifen

Bei den Maßnahmen, die für die Erhaltung oder Verlängerung des Lebens des Patienten notwendig sind und auf die im Falle einer passiven Sterbehilfe verzichtet wird, handelt es sich in aller Regel um Eingriffe in den Körper oder die Freiheit des Patienten, die seiner Einwilligung bedürfen, um zulässig zu sein. Die zentrale Relevanz der informierten Einwilligung als wesentliche Ausprägung des Selbstbestimmungsrechtes des Patienten ist in der Medizinethik und auch in der Rechtswissenschaft unbestritten.²² Demnach ist ein eigenmächtiges Behandeln ohne (oder gar gegen) den Willen des Patienten bis auf wenige Ausnahmen unzulässig.²³ Dieses Einwilligungserfordernis gilt unabhängig davon, ob der in Frage stehende Eingriff medizinisch indiziert oder gar überlebensnotwendig ist. Voraussetzung für eine wirksame Einwilligung ist stets, dass diese bestimmten Gültigkeitskriterien genügt. Dazu gehört, dass der Patient über alle wesentlichen Merkmale des Eingriffs, zu denen insbesondere auch die mit dem Eingriff sowie mit seiner Unterlassung verbundenen Chancen und Risiken gehören, informiert ist, dass die Einwilligung freiwillig und ohne Zwang erfolgt sowie dass der Patient zum Zeitpunkt der Entscheidung für oder gegen eine Einwilligung kompetent ist, eine solche Entscheidung zu

treffen. Zudem kann eine einmal erteilte Einwilligung jederzeit wieder zurückgezogen werden.²⁴ Anlass zu Zweifeln an der Kompetenz zur Einwilligung können beispielsweise kognitive Beeinträchtigungen oder akuter emotionaler Stress geben. Der bloße Umstand, dass ein Patient die Einwilligung in einen lebenserhaltenden bzw. -verlängernden Eingriff verweigert (oder zurückzieht) und sich so für eine Verkürzung seines Lebens entscheidet, ist allerdings kein hinreichendes Kriterium, um diesem Patienten die erforderliche Kompetenz abzusprechen. Folglich ist auch eine medizinische Maßnahme, die der Erhaltung oder Verlängerung des Lebens eines Patienten dient, in der Regel unzulässig, wenn keine gültige Einwilligung des Patienten vorliegt.

Der Verzicht auf oder Abbruch einer lebenserhaltenden Maßnahme im Rahmen einer passiven Sterbehilfe ist also schon allein deshalb gerechtfertigt, weil eine solche Maßnahme unzulässig ist, wenn der Patient seine Einwilligung verweigert bzw. zurückzieht.²⁵ Es kommt bei der Rechtfertigung passiver Sterbehilfe somit gar nicht darauf an, dass ein unabhängiger oder in irgendeiner Hinsicht natürlicher, zum Tod führender Prozess nicht länger aufgehalten wird. Dies erklärt lediglich, warum der Patient infolge des Verzichts bzw. Abbruchs stirbt und es sich darum um einen Fall von Sterbehilfe handelt. Für die Rechtfertigung des Verzichts auf die Maßnahme (bzw. ihres Abbruchs) ist nicht erheblich, *was* durch sie aufgehalten würde, oder dass überhaupt etwas aufgehalten würde, sondern *wie* sie es aufhalten würde, nämlich durch einen Eingriff, der ohne die wirksame Einwilligung des Patienten unzulässig ist.

Hierin unterscheiden sich passive und aktive Sterbehilfe. Passive Sterbehilfe ist legitim, weil die zu unterlassende oder abzubrechende Maßnahme einwilligungsbedürftig ist, aber nicht bzw. nicht länger dem Willen des Patienten entspricht. Während bei der passiven Sterbehilfe also dem Willen des Patienten entsprechend darauf verzichtet wird, in dessen körperliche Unversehrtheit einzugreifen, besteht die aktive Sterbehilfe in genau so einem Eingriff. Auf diesen Unterschied haben unter anderem Norbert Hoerster sowie Bernard Gert und Kollegen hingewiesen.²⁶ Befür-

21 Zu diesem Ergebnis kommen u.a. Birnbacher (1995, 348ff.) und Schöne-Seifert (1997, 217ff.).

22 Vgl. nur Beauchamp/Childress (2013, 120ff.), Schöne-Seifert (2007, 39ff.), Schroth (2010, 23 ff.) sowie Bundesärztekammer (2011b, §§ 7 I, 8).

23 Die Ausnahmen betreffen Notfallsituationen, in denen der Wille des Patienten beispielsweise wegen Bewusstlosigkeit aktuell nicht festzustellen, aber eine Behandlung notwendig ist, um großen Schaden vom Patienten abzuhalten. Aber auch in solchen Fällen ist die Behandlung nur dann und insofern gerechtfertigt, als sie dem mutmaßlichen Willen des Patienten entspricht und also nicht gegen seinen Willen erfolgt. Wenn ohne jeden Zweifel bekannt ist, dass der Patient Maßnahmen wie Bluttransfusionen etwa aus religiösen Gründen grundsätzlich ablehnt, dann wäre dies auch im Rahmen einer Notfallbehandlung des bewusstlosen Patienten unzulässig.

24 Auch ein solches Zurückziehen muss jedoch bestimmten Kriterien genügen, um wirksam zu sein.

25 Die zentrale Rolle, die die Einwilligungsbedürftigkeit aller medizinischen Maßnahmen bei der Legitimation von Sterbehilfe durch Behandlungsverzicht bzw. -abbruch spielt, ist weithin anerkannt, vgl. Nationaler Ethikrat (2006, 41), Bundesgerichtshof (2010, Rn. 35), Roxin (2010, 92f.) und Sumner (2011, Kap. 2).

26 Siehe Hoerster (1986, 65f.) und Gert et al. (2006, 315f.)

wörter des Äquivalenzargumentes lassen ihn jedoch zumeist außer Acht. Sie können darauf verweisen, dass der Eingriff bei der aktiven Sterbehilfe wiederum dem Willen des Patienten entspricht, der diese ja freiwillig erbittet. Wie Befürworter der aktiven Sterbehilfe zu Recht betonen, spielt somit auch in der Rechtsfertigung aktiver Sterbehilfe eine zentrale Rolle, dass diese mit dem Patientenwillen übereinstimmt. Es liegt deshalb die Frage nahe, wie tief der hier identifizierte Unterschied zwischen aktiver und passiver Sterbehilfe tatsächlich geht. Geht es nicht in beiden Fällen darum, das Selbstbestimmungsrecht des Patienten zu achten?²⁷ Um diese Frage beantworten zu können, ist es hilfreich, die moralische Relevanz des Patientenwillens und die spezielle Rolle, die Einwilligungen als Ausdruck dieses Willens spielen, etwas allgemeiner in den Blick zu nehmen. Dies wird es mir erlauben zu zeigen, dass hier tatsächlich ein moralisch relevanter Unterschied zwischen passiver und aktiver Sterbehilfe besteht.

5. Die Rolle von Einwilligungen

In der Tat kommt auch aktive Sterbehilfe überhaupt nur in Betracht, wenn sie dem Willen des Patienten entspricht. Die Rolle der Einwilligung ist hier allerdings eine ganz andere: Sie wird gegeben, nicht verweigert. Während sich aus der Verweigerung der Einwilligung in einen Eingriff allerdings dessen Unzulässigkeit und somit ein Anspruch auf sein Unterlassen ergibt, leitet sich aus der Einwilligung in eine Handlung im Allgemeinen noch kein Anspruch auf die Handlung ab. Eine Einwilligung hat – sofern sie den Gültigkeitskriterien genügt (s. oben) – die Erlaubtheit der Handlung zur Folge, sorgt aber im Allgemeinen nicht dafür, dass sie moralisch gefordert ist.²⁸ Diese asymmetrische Relevanz von Einwilligungen zeigt sich auch in nicht-medizinischen Bereichen, in denen die Einwilligung eine zentrale Rolle für die moralische Bewertung von Handlungen spielt. Jemanden ins Gesicht zu schlagen ist ohne dessen Einwilligung moralisch wie rechtlich unzulässig; wer es dennoch tut, verletzt einen legitimen Anspruch des Geschlagenen. Willigt eine Person jedoch etwa im Rahmen eines Boxkampfes in Schläge ein, sind diese zulässig (wiederum: sofern die Ein-

willigung den Gültigkeitskriterien genügt). Gefordert sind sie allerdings nicht, zumindest nicht aufgrund der Einwilligung. Einem Fremden auf der Straße meine Einwilligung zu geben, mich zu schlagen, berechtigt mich schließlich nicht, anschließend die Schläge von ihm einzufordern. Gleiches gilt für die Einwilligung in sexuelle Handlungen: Ohne Einwilligung sind diese unzulässig und es besteht ein Anspruch auf Unterlassung, aber natürlich gibt weder die eigene Einwilligung noch die des anderen jemandem ein Recht, sexuelle Handlungen einzufordern.

Dies hängt damit zusammen, wie die Einwilligung des von einer Handlung Betroffenen die normative Situation des Handelnden ändert. In Fällen wie dem eines Schläges ins Gesicht spricht etwas Gewichtiges gegen die Handlung, wenn der Betroffene nicht eingewilligt hat: Sie verursacht Schmerzen und greift in die körperliche Unversehrtheit ein. Gleiches gilt für medizinische Maßnahmen. Auch diese gehen in der Regel mit Schmerzen oder Eingriffen in die körperliche Integrität einher, weshalb sie in der deutschen Rechtsprechung üblicherweise als Körperverletzung gewertet werden.²⁹ Die Einwilligung des Betroffenen sorgt nun allerdings nicht dadurch für die Zulässigkeit solcher Handlungen, dass sie diesen Gründen gegen die Handlung solche entgegensetzt, die für die Handlung sprechen. Stattdessen hat sie zur Folge, dass einige oder alle der zuvor problematischen Gesichtspunkte außer Kraft gesetzt werden und somit nicht länger gegen die Handlung sprechen. Metaphorisch gesprochen: Die Einwilligung fügt nicht der „Für“-Waagschale etwas hinzu, das dann den Inhalt der „Gegen“-Waagschale überwiegt, sondern nimmt etwas aus der „Gegen“-Waagschale heraus, so dass diese nicht mehr schwerer wiegt. Zwei Überlegungen machen dies deutlich.

Erstens ist es bei anderen leidensmindernden medizinischen Eingriffen wie beispielsweise dem Einsetzen eines künstlichen Kniegelenkes irreführend, diese so zu beschreiben, dass ohne die Einwilligung des Patienten noch nicht genug für den Eingriff spricht, mit der Einwilligung dagegen schon. Und auch wenn ein Patient aus religiösen Gründen seine Einwilligung in eine lebensrettende Operation verweigert, weil damit eine Bluttransfusion verbunden wäre, wäre es unangemessen, die Situation als eine zu beschreiben, in der noch nicht genug für den Eingriff spricht, um dessen Durchführung insgesamt zur moralisch besseren Option zu machen. Stattdessen spricht ohne die Einwilligung et-

27 Diese Überlegung findet sich etwa bei Sumner (2011, 90).

28 In einzelnen Fällen ist es allerdings durchaus möglich, dass eine Handlung, die ohne (tatsächliche oder mutmaßliche) Einwilligung unzulässig wäre, bei vorliegender Einwilligung gefordert ist, etwa das Verabreichen einer bereitstehenden, lebensrettenden Blutkonserve. Dies liegt jedoch nicht in der Einwilligung selbst begründet, sondern in den gewichtigen Gesichtspunkten, die in einem solchen Fall für die Handlung sprechen.

29 Vgl. Fischer (2014, § 223 Rn. 17).

was sehr Gewichtiges gegen diese Eingriffe, das durch die Einwilligung allerdings außer Kraft gesetzt wird. Nach Erteilen der Einwilligung spricht also nicht unbedingt *mehr für*, sondern vor allem *weniger gegen* die entsprechende Handlung.

Zweitens ist der bloße Verweis auf die Einwilligung des Betroffenen einer Handlung nicht geeignet, diese Handlung auch zu rechtfertigen. Angenommen, A willigt ein, dass B ihn ins Gesicht schlägt. Wenn B dies nun tut und anschließend von C gefragt wird, warum er A geschlagen habe, ist „weil er eingewilligt hat“ keine geeignete Antwort, und zwar nicht bloß, weil sie auf einen Grund verweist, der nicht hinreichend ist, um die Handlung zu rechtfertigen (wie es der Fall wäre, wenn A nicht eingewilligt hätte und B mit „weil ich Lust hatte“ antworten würde). Die Antwort benennt überhaupt keinen Grund und ist deswegen grundsätzlich ungeeignet, eine solche Handlung zu rechtfertigen.³⁰ Würde die Einwilligung selbst einen Grund für den Schlag liefern, so würde B, wenn er trotz der Einwilligung nicht zuschlägt, nun einem Grund zuwiderhandeln. Es spricht aber augenscheinlich trotz Einwilligung überhaupt nichts dagegen, auf den Schlag zu verzichten, solange nicht ein anderer Gesichtspunkt benannt werden kann, der für den Schlag spricht. Die Einwilligung des A stellt also selbst keinen positiven Grund bereit, der B's Schlag rechtfertigen könnte.

Man kann dies so zusammenfassen, dass die Einwilligung in eine Handlung keinen positiven, sondern lediglich einen negativen Beitrag zu deren Rechtfertigung leistet, indem sie einige oder alle gegen die Handlung sprechenden Gesichtspunkte, die es durch eine positive Rechtfertigung aufzuwiegen gilt, außer Kraft setzt. Das erklärt, warum sich allein mit einer Einwilligung in eine einwilligungsbedürftige Handlung noch keine Forderung begründen lässt, denn eine solche Forderung setzt voraus, dass eine positive Rechtfertigung für die Handlung verfügbar ist, was die Einwilligung selbst jedoch nicht garantiert. Das Zurückziehen einer solchen Einwilligung hingegen setzt die problematischen Gesichtspunkte der Handlung, auf die sich eine Unterlassungsforderung berufen kann, wieder in Kraft.

³⁰ Sie mag abhängig vom Kontext freilich implizit auf einen nicht genannten Grund hinweisen. Ist es zwischen B und C geteiltes Wissen, dass B den A nicht leiden kann, und beantwortet B die Frage von C in der genannten Weise, begleitet von einem Grinsen, so mag dies darauf verweisen, dass B den A schon immer mal ins Gesicht schlagen wollte. Auch dann rechtfertigt aber nicht die explizit benannte Einwilligung des A, sondern (wenn überhaupt) der implizit kommunizierte Wunsch B's die Handlung.

Diese asymmetrische Relevanz von Einwilligungen wird verschleiert, wenn davon die Rede ist, dass eine medizinische Maßnahme oder deren Abbruch „dem Willen des Patienten entspricht“, oder dass bei der Durchführung solcher Maßnahmen die „Entscheidungen des Patienten“ zu respektieren sind, denn dadurch wird der Unterschied zwischen dem Geben und dem Verweigern einer Einwilligung ausgeblendet.³¹ Solche Formulierungen laden außerdem zu einer Verwechslung von zwei Aspekten der Selbstbestimmung einer Person ein, die es zu unterscheiden gilt. Damit, dass es A's Willen oder seiner Entscheidung entspricht, von B geschlagen zu werden, kann zweierlei gemeint sein, nämlich zum einen A's Einwilligung in den Schlag und zum anderen sein Wunsch, von B geschlagen zu werden (den er zudem als Bitte äußern mag). Letzterer mag B einen Grund für den Schlag liefern, denn anders als Einwilligungen können Wünsche durchaus einen positiven Beitrag zur Rechtfertigung von Hilfe bei ihrer Erfüllung leisten. Wünschen fehlt allerdings der negative Rechtfertigungsbeitrag von Einwilligungen: Sie setzen Gesichtspunkte, die gegen ihre Erfüllung sprechen, nicht einfach außer Kraft. Wichtig ist zudem, dass die Einwilligung in eine Handlung zwar häufig – und im Kontext medizinischer Behandlungen wohl fast immer –, aber nicht notwendigerweise mit dem Wunsch, dass diese Handlung ausgeführt wird, einhergeht. Dies zeigt das Beispiel des Boxers, der in die Schläge seines Gegners eingewilligt hat, aber dennoch nicht von ihm geschlagen zu werden wünscht und seinen Schlägen darum zu entgehen versucht.³² Es ist deshalb zwar angesichts der weitgehenden Übereinstimmung zwischen Wunsch und Einwilligung im Bereich der Medizin durchaus nachvollziehbar, dass gerade im Kontext der Medizinethik die moralische Relevanz von Einwilligungen mit der Relevanz der Wünsche und Bitten des Patienten gleichgesetzt wird. Dadurch wird aber genau jene Asymmetrie unterschlagen, die erklärt, warum es zwar keinen prinzipiellen moralischen Unterschied zwischen der Nichterfüllung eines Wunsches nach Unterlassung einer Maßnahme und der eines Wunsches nach Durchführung einer Maßnahme gibt, es aber dennoch sehr viel problematischer ist, eine Maßnahme ohne Einwilligung des Patienten durchzuführen, als trotz Einwilligung des Patienten von einer Maßnahme abzusehen. Ist dieses Missver-

³¹ Vgl. Gert et al. (1994).

³² Vgl. dazu ausführlicher Müller (2018, § 3). Auf den Unterschied zwischen der Einwilligung in eine Handlung und dem Wunsch nach oder der Bitte um die Ausführung der Handlung weisen auch Owens (2012, 173) bzw. Dempsey (2013, 17) hin.

ständnis einmal offengelegt, ist auch ersichtlich, dass es einen bedeutenden moralischen Unterschied macht, ob die Willensäußerung des Patienten im Erteilen oder Verweigern einer Einwilligung besteht, auch wenn sie in beiden Fällen mit einem entsprechenden Wunsch des Patienten einhergeht. Dieser Unterschied trennt aktive und passive Sterbehilfe, obwohl man sich in der Rechtfertigung beider auf den Willen und das Selbstbestimmungsrecht des Patienten berufen kann.

6. Zwischenfazit

Die vorangegangenen Überlegungen machen deutlich, dass es sehr wohl einen moralisch relevanten Unterschied zwischen aktiver und passiver Sterbehilfe gibt, denn nur bei der passiven Sterbehilfe wird auf die (weitere) Durchführung einer einwilligungsbedürftigen Maßnahme verzichtet, weil der Patient seine Einwilligung verweigert bzw. zurückzieht. Hieraus ergibt sich ein legitimer Anspruch des Patienten auf den Abbruch der Maßnahme und damit auf die Durchführung der Sterbehilfe, die in solchen Fällen somit nicht bloß zulässig, sondern geboten ist – der Patient hat ein Recht darauf, dass die Maßnahme nicht durchgeführt bzw. abgebrochen wird, wenn er seine Einwilligung verweigert oder zurückzieht. Die Zulässigkeit passiver Sterbehilfe lässt sich folglich unter Rückgriff auf eine Rechtfertigungsressource begründen, die für Fälle aktiver Sterbehilfe nicht zur Verfügung steht, denn das Unterlassen der zum Tod führenden Handlung im Falle der aktiven Sterbehilfe (z.B. das Vorenthalten der tödlichen Spritze) ist nichts, das der Einwilligung des Patienten bedarf. Damit hat sich gezeigt, dass die Äquivalenzthese bezüglich aktiver und passiver Sterbehilfe zurückgewiesen werden muss; zwischen ihnen besteht ein moralisch relevanter Unterschied. Folglich scheitert auch das Äquivalenzargument für die Zulässigkeit aktiver Sterbehilfe – daraus, dass passive Sterbehilfe zulässig ist, folgt nicht, dass auch aktive Sterbehilfe in manchen Situationen zulässig sein muss.

Was heißt das für die Zulässigkeit der aktiven Sterbehilfe? Es scheint, als sei mit der Zurückweisung der Äquivalenzthese zugleich der erste Schritt zur Begründung der restriktiven Position getan. Tatsächlich hat sich damit zwar eine der Voraussetzungen der restriktiven Position – dass es einen moralisch relevanten Unterschied zwischen aktiver und passiver Sterbehilfe gibt – als richtig herausgestellt. Mit Blick auf das Ziel, die Unzulässigkeit aktiver Sterbehilfe zu begründen, sind Vertreter der restriktiven Position allerdings keinen Schritt weiter, denn der Unterschied besteht nicht darin, dass im Falle der aktiven Sterbehilfe ein (weiterer) gegen sie sprechender

Gesichtspunkt vorliegt, von dem sich fragen ließe, ob er moralisch entscheidend ist und den Ausschlag gegen die Zulässigkeit aktiver Sterbehilfe gibt. Der Umstand, dass es sich bei der aktiven Sterbehilfe nicht um das Respektieren der Verweigerung der Einwilligung in einen Eingriff handelt, spricht ja nicht gegen sie. Mehr noch: Die Ergebnisse des letzten Abschnittes erlauben es, die Herausforderung, der sich Verteidiger der restriktiven Position ausgesetzt sehen, zu verschärfen und zugleich zu zeigen, wie ein direktes Argument, das die Zulässigkeit der aktiven Sterbehilfe unter Verweis auf das Selbstbestimmungsrecht zu begründen beansprucht, präzisiert und gegen zwei verbreitete Einwände verteidigt werden kann.

7. Selbstbestimmung und die Zulässigkeit aktiver Sterbehilfe

Wir haben gesehen, dass nur die Verweigerung der Einwilligung in eine medizinische Maßnahme den Patienten automatisch mit einem Anspruch auf den Verzicht auf die Maßnahme ausstattet und deshalb ein Unterschied in den zur Verfügung stehenden Rechtfertigungsressourcen zwischen aktiver und passiver Sterbehilfe besteht. Die Kehrseite dieser Feststellung ist, dass Einwilligungen grundsätzlich Handlungen legitimieren können, die andernfalls moralisch unzulässig wären. Sonst wäre nicht ersichtlich, warum künstliche Beatmung, Wiederbelebung und andere Maßnahmen mit Einwilligung des Patienten zulässig sind. Wie in Abschnitt 5 erläutert, bewirken Einwilligungen eine solche Veränderung der normativen Situation des Handelnden, indem sie gegen die Handlung sprechende Gesichtspunkte außer Kraft setzen. Dies alles muss auch der Verfechter einer restriktiven Position akzeptieren, der zur Rechtfertigung seiner unterschiedlichen Beurteilung von aktiver und passiver Sterbehilfe die asymmetrische Relevanz von Einwilligungen heranzuziehen versucht.

Auf die Möglichkeit, eine ansonsten unzulässige Handlung zu erlauben, indem man als Betroffener dieser Handlung in sie einwilligt, verweisen jedoch auch die Vertreter eines direkten Argumentes für die Zulässigkeit der aktiven Sterbehilfe. Sie berufen sich oft auf das Selbstbestimmungsrecht des Patienten, das es auch bei Entscheidungen über Leben und Tod zu respektieren gelte. Die Rede von Selbstbestimmung ist jedoch, wie wir im fünften Abschnitt gesehen haben, mehrdeutig. Es ist deshalb sinnvoll, nicht von einem Selbstbestimmungsargument, sondern von einem Einwilligungsargument für die Zulässigkeit aktiver Sterbehilfe zu sprechen. Die Einwilligung sorgt im oben

diskutierten Falle des Boxers dafür, dass die entsprechende Handlung moralisch zulässig ist, und auch sexuelle Handlungen sind im Allgemeinen zulässig, wenn eine gültige Einwilligung aller Beteiligten vorliegt. Warum nun, so die dem Einwilligungsargument zugrundeliegende Überlegung, sollte die Einwilligung des Patienten in Fall 2 nicht ebenfalls ausreichen, um die tödliche Injektion zwar nicht moralisch gefordert, aber doch zumindest zulässig zu machen?

Verfechter einer restriktiven Position können dieser Konsequenz nur entgehen, indem sie die Reichweite dieses Einwilligungsvermögens – und damit die Selbstbestimmung des Patienten – beschränken. Genauer gesagt: Sie müssen annehmen, dass die Einwilligung des Patienten manche, aber nicht alle Gesichtspunkte, die gegen eine Maßnahme sprechen können, außer Kraft zu setzen vermag. Im Gegensatz beispielsweise zu den Verletzungen des Boxers durch die Schläge seines Gegners liegen einige solcher Gesichtspunkte demzufolge jenseits der Reichweite wirksamer Einwilligungen. Zudem müssen sie zeigen, dass ein solcher Gesichtspunkt in allen Fällen aktiver Sterbehilfe die Gründe, die für die Sterbehilfe sprechen, überwiegt. Aktive Sterbehilfe ist also trotz Einwilligung des Patienten nur dann grundsätzlich unzulässig, wenn in allen Fällen, in denen diese zur Entscheidung steht, ein Gesichtspunkt vorliegt, für den gilt, dass er zum einen einwilligungsresistent ist und zum anderen alle Gründe, die in der jeweiligen Situationen für die Sterbehilfe sprechen, zu überwiegen vermag.

Welcher Gesichtspunkt kann diese anspruchsvolle Rolle spielen? Ist es der Umstand, dass aktive Sterbehilfe ärztliches Handeln beinhaltet, welches den Tod des Patienten zeitlich unmittelbar kausal zur Folge hat? Dies gilt auch für das Ausschalten einer Beatmungsmaschine. Oder ist es der Umstand, dass der Patient nicht infolge eines natürlichen Prozesses stirbt? Die moralische Relevanz dieses Gesichtspunktes hat sich bereits in Abschnitt 3 als zweifelhaft herausgestellt. Verfechter der restriktiven Position mögen an dieser Stelle einfach darauf pochen, dass menschlichem Leben ein unbedingter und nicht abzuwiegender Wert zukommt, der es selbst für den, dessen Leben es ist, gänzlich unverfügbar macht. Eine solche These bedarf jedoch einer plausiblen Begründung. Dass eine solche Begründung angesichts der Bedingung, dass sie, um allgemein überzeugen zu können, auf weltanschaulich-religiöse Annahmen verzichten müsste, nicht in Sicht ist, ist an anderer Stelle überzeugend dargelegt worden.³³

Zudem würde dieser Gesichtspunkt auch gegen passive Sterbehilfe wie in Fall 3 sprechen. Auch wer ein lebenserhaltendes Beatmungsgerät abstellt, verfügt über das Leben eines anderen, wenn auch auf dessen Geltendmachen eines berechtigten Anspruchs hin.³⁴ Wer dies für zulässig hält, gesteht also zu, dass ein solcher Wert zumindest den Anspruch auf das Abbrechen einer lebenserhaltenden Maßnahme, der sich aus der Einwilligungsbedürftigkeit dieser Maßnahme ergibt, nicht zu überwiegen vermag, sondern vielmehr selbst durch diesen überwogen wird. Daraus folgt zwar nicht schon, dass jener Wert auch in Fällen aktiver Sterbehilfe überwogen werden kann, denn ein solcher Anspruch kann bei aktiver Sterbehilfe ja gerade nicht zur Rechtfertigung herangezogen werden. Warum kann dieser Wert dann aber nicht auch etwa durch das als unerträglich empfundene und nur durch Herbeiführung des Todes vermeidbare Leid eines Patienten überwogen werden?

Hier drängt sich der Verdacht auf, dass der Vertreter der restriktiven Position sich die spezifische Relevanz dieses Wertes *ad hoc* so zurechtlegt, dass er genau die Rolle spielt, die im Rahmen seiner Position erforderlich ist. Eine überzeugende Begründung dieser Position müsste stattdessen erklären, warum dieser Wert alle anderen Gesichtspunkte *bis auf genau einen* überwiegen können und warum er anders als so viele andere Gesichtspunkte grundsätzlich einwilligungsresistent sein sollte. Der Vertreter der restriktiven Position ist hier in der Bringschuld: Kann er nicht plausibel machen, dass gegen die aktive Sterbehilfe derart schwergewichtige Gesichtspunkte sprechen, die grundsätzlich außerhalb der Reichweite wirksamer Einwilligungen liegen, gibt es vor dem Hintergrund des in Abschnitt 5 dargelegten Verständnisses der Rolle von Einwilligungen keinen überzeugenden Grund, die Gültigkeit des Einwilligungsargumentes für die Zulässigkeit zumindest mancher Fälle aktiver Sterbehilfe zu bezweifeln.

³⁴ Gleiches gilt für den Versuch, die grundsätzliche Unzulässigkeit der aktiven Sterbehilfe mithilfe eines *Rechts auf Leben* des Patienten zu begründen, über das dieser (anders als etwa über sein Recht auf körperliche Unversehrtheit) nicht verfügen kann. Auch dieses würde in Fall 3 verletzt. Zudem ist schwer zu sehen, wie ein solches Recht gewissermaßen gegen den Willen des Rechteinhabers gewendet werden kann, um dessen Verfügungsfreiheit einzuschränken und ihm im Ergebnis eine Pflicht zum Leben aufzuerlegen. Dies scheint mit einem Verständnis dieses Rechts als eines Abwehrrechts inkompatibel zu sein und stattdessen vorauszusetzen, dass es nicht die Freiheit des Inhabers, sondern einen objektiven Wert von der oben diskutierten Art schützt.

³³ Vgl. Kuhse (1987), Dworkin (1993, Kap. 3) und Quante (2002, 239 ff.).

8. Zwei Einwände

Gegen derartige Argumente, die sich auf das Selbstbestimmungsrecht des Patienten berufen, sind jedoch andere Einwände vorgebracht worden, die auch viele derer skeptisch machen, die aktive Sterbehilfe in manchen Situationen für zulässig halten. Denn, so der erste Einwand, wenn allein die Einwilligung des Patienten hinreicht, um eine aktive Sterbehilfe zu rechtfertigen, dann muss dies auch in den Fällen gelten, in denen ein Patient nicht aufgrund von Schmerzen oder anderem Leid, sondern beispielsweise aus bloßer „Lebenssatttheit“ sterben will. Dies aber geht vielen zu weit.

Dieser Einwand verliert jedoch an Schlagkraft, wenn man sich in Erinnerung ruft, dass eine Einwilligung keine positive Rechtfertigung für eine Handlung bereitstellt, sondern lediglich einen negativen Beitrag zu ihrer Rechtfertigung leistet, indem sie gegen die Handlung sprechende Gesichtspunkte außer Kraft setzt.³⁵ Gerade so schwerwiegende Eingriffe wie etwa eine tödliche Injektion aber erfordern eine positive Rechtfertigung durch gute Gründe. Nun ist auch in problematischen Fällen wie bloßer Lebenssatttheit anzunehmen, dass der Patient nicht nur in die Sterbehilfe einwilligt, sondern diese auch wünscht und darum bittet; und anders als Einwilligungen können solche Wünsche und Bitten grundsätzlich zur positiven Rechtfertigung einer Handlung beitragen (vgl. Abschnitt 5).

Sie tun dies jedoch nur, wenn sie bestimmten Bedingungen genügen, die nicht in jedem Fall bei einer grundsätzlich einwilligungsfähigen Person erfüllt sind. Insbesondere sind die Bedingungen dafür, dass die Wünsche eines anderen mir einen Grund geben, aktiv zu ihrer Erfüllung beizutragen, anspruchsvoller als die Bedingungen, die erfüllt sein müssen, damit ich seine Wünsche respektieren muss und ihn nicht bei ihrer Erfüllung behindern darf. Aus Erwägungen des Wohlwollens etwa ergibt sich im ersten Fall nur dann ein Grund zu helfen, wenn anzunehmen ist, dass die Wünsche mit dem Wohlergehen der Person übereinstimmen, was nicht immer der Fall ist. Diese Bedingung muss meines Erachtens jedoch nicht erfüllt sein, um Wünsche im genannten Sinne respektabel zu machen. Angenommen ein Freund wünscht, eine großflächige Tätowierung seiner Unterarme vornehmen zu lassen, weil er glaubt, dass dies ihn insgesamt glücklicher machen wird, liegt damit jedoch in vor-

hersehbarer Weise falsch. Dann kann es geboten sein, seinen Wunsch dennoch zu respektieren und nicht in paternalistischer Weise zu versuchen, die Tätowierung zu verhindern. Der Wunsch gibt mir jedoch keinen Grund, bei seiner Erfüllung etwa durch die Finanzierung der Tätowierung behilflich zu sein. Ebenso wenig gibt der Wunsch zu sterben jemandem einen Grund, bei der Herbeiführung des Todes behilflich zu sein, wenn er offensichtlich einer Fehleinschätzung des zu erwartenden künftigen Wohlergehens entspringt.

Doch selbst wenn der Sterbewunsch derartige Bedingungen nicht verletzt, heißt das noch nicht, dass damit das entsprechende Handeln eines Arztes bereits als unproblematisch ausgewiesen ist. Dies kann zumindest zurückweisen, wer die Gründe, aus denen ein Arzt handeln darf, gegenüber denen anderer Personen für beschränkt hält. In der Tat halten wir in vielen Kontexten das Handeln aus manchen Gründen, die für andere eine Handlung zu rechtfertigen vermögen, im Falle von Ärzten für unzulässig oder zumindest problematisch. Beispiele hierfür sind Profitmaximierung oder die Befriedigung sexueller Bedürfnisse. Aus solchen Gründen, so scheint es, sollten Ärzte nicht (be-) handeln. Viele lehnen zudem ein Bild von Ärzten als bloßen Wunscherfüllern ab, die den Bitten eines Patienten um die Durchführung bestimmter Maßnahmen unabhängig davon nachkommen, ob sie diese für medizinisch sinnvoll halten. Vor dem Hintergrund eines so verstandenen „ärztlichen Handlungsethos“ ließe sich also eine Begrenzung der Zulässigkeit aktiver Sterbehilfe begründen, die Fälle bloßer Lebenssatttheit grundsätzlich ausschließt.

Der Sorge, eine auf dem Einwilligungsargument aufbauende Rechtfertigung aktiver Sterbehilfe laufe unweigerlich darauf hinaus, dass diese zulässig sei, wann immer ein Patient sie wünscht und seine Einwilligung erteilt, kann somit begegnet werden. Die Zulässigkeit aktiver Sterbehilfe in Fällen, in denen Leid und Schmerzen, die aus Sicht des Patienten unerträglich sind, Anlass für seinen Wunsch zu sterben sind, wird durch solche Überlegungen jedoch nicht in Frage gestellt. Schließlich gehört das Beenden oder Verhindern von Schmerzen und Leid seit jeher zu den Kernaufgaben eines Arztes. In solchen Fällen liegen also durchaus paradigmatische Gründe für ärztliches Handeln vor.

An diese Überlegungen lässt sich jedoch ein zweiter Einwand anschließen: Wenn das Beenden von Leid und Schmerz auch dann für ein ärztliches Handeln spricht, wenn dieses im Herbeiführen des Todes eines Patienten besteht, ist es dann nicht auch prinzipiell

³⁵ Natürlich leistet eine Einwilligung dies auch nur dann, wenn sie den Gültigkeitskriterien genügt. Es ist meines Erachtens zu bezweifeln, dass dies in vielen der Fälle zutreffen wird, in denen ein Patient ohne (aktuelle oder antizipierte) Schmerzen oder andere Leiden um aktive Sterbehilfe bittet.

möglich, dass es die gegen solche Tötungen sprechenden Gesichtspunkte überwiegt, wenn diese nicht durch die Einwilligung des Patienten außer Kraft gesetzt werden? Läuft dies also darauf hinaus, dass auch unfreiwillige „Mitleidstötungen“ gerechtfertigt sein können?

Dieser Befürchtung lässt sich zum einen entgegensetzen, dass nichts des bisher Gesagten das Prinzip in Frage stellt, demzufolge es grundsätzlich unzulässig ist, eine Person *gegen ihren Willen* zu töten, selbst dann, wenn sie dadurch von einem Leid „erlöst“ würde. Zum anderen ignoriert die Befürchtung die Rolle, die der Wille des Patienten bei der Bestimmung seines Leidens spielt. Dass ein Patient einem Zustand zu entgehen wünscht, ist zwar nicht hinreichend, um davon zu sprechen, dass er unter diesem Zustand leidet (andernfalls wäre jede Erfüllung eines Wunsches zugleich das Beseitigen eines Leids). Andererseits bestehen hinsichtlich der Überzeugung, dass ein Patient an einem Zustand leidet, erhebliche Rechtfertigungsschwierigkeiten, wenn der Patient in diesem Zustand zu verbleiben wünscht. Dass eine Person etwas Vermeidbares dennoch wünscht, spricht in epistemisch *prima facie* ausschlaggebender Weise dagegen, dass sie darunter leidet. Wer eine Person gegen ihren Willen aus Mitleid tötet, macht also in der Regel den Fehler, die Person hinsichtlich Ihres vermeintlichen Leidens zu bevormunden, und ist schon deshalb in seinem Handeln nicht gerechtfertigt.³⁶ Auch dieser zweite Einwand gegen das Selbstbestimmungs-Argument für die Zulässigkeit aktiver Sterbehilfe geht somit fehl.

9. Schlussbemerkung

Der hier vorgelegte Text untersucht die Rolle, die verschiedene Aspekte dessen, was man gemeinhin als Selbstbestimmung bezeichnet, bei der Beantwortung der Frage spielen, ob und unter welchen Bedingungen es moralisch zulässig ist, dass Ärzte an der Herbeiführung des Todes von Patienten mitwirken. Diese Untersuchung ist zu einem insbesondere für Vertreter einer restriktiven Position bezüglich der Zulässigkeit aktiver Sterbehilfe zweischneidigen Ergebnis gekommen. Einerseits hat sich gezeigt, dass das gegenwärtig verbreitete Verständnis der Unterscheidung zwischen

aktiver und passiver Sterbehilfe, das auf das Kriterium des Behandlungsbezuges abstellt, entgegen mancher Vertreter einer liberalen Position durchaus einen moralisch relevanten Unterschied markiert. Deren Äquivalenzargument für die Zulässigkeit auch aktiver Sterbehilfe scheitert somit.

Andererseits setzt die Begründung, warum hier ein moralisch relevanter Unterschied besteht, ein Verständnis der moralischen Rolle von Einwilligungen voraus, das denjenigen, der im Gegensatz zur passiven die aktive Sterbehilfe für grundsätzlich unzulässig hält, in Bedrängnis bringt. Es zeigt sich nämlich, dass er, wenn er sich diese Begründung zu eigen macht, zugleich den Ausgangspunkt eines direkten Argumentes für die Zulässigkeit aktiver Sterbehilfe akzeptiert. Dessen Konklusion könnte er nur noch entgehen, indem er eine überzeugende und nicht auf weltanschaulich-religiösen Annahmen beruhende Begründung für die These vorlegt, dass in allen Fällen aktiver Sterbehilfe etwas gegen deren Durchführung spricht, das anders als andere solcher Gesichtspunkte grundsätzlich einwilligungsresistent ist und zudem alle Gründe, die für eine solche Maßnahme sprechen können, überwiegt. Eine solche Begründung, die zugleich der Zulässigkeit passiver Sterbehilfe ohne Ad-hoc-Annahmen Rechnung zu tragen vermag, ist jedoch nicht in Sicht.

Außerdem können mithilfe eines differenzierteren Verständnisses der moralischen Relevanz von Einwilligungen und Wünschen als zwei zu unterscheidenden Aspekten von Selbstbestimmung zwei verbreitete Einwände zurückgewiesen werden: Weder legitimiert eine vom Selbstbestimmungsrecht des Patienten ausgehende Argumentation eine aktive ärztliche Sterbehilfe notwendigerweise auch in Fällen, in denen diese aus bloßer „Lebenssattheit“ gewünscht wird, noch zwingt sie uns, in der Konsequenz sogar unfreiwillige „Mitleidstötungen“ für rechtfertigbar zu halten. Insgesamt zeigen diese Überlegungen zur Rolle der Selbstbestimmung also, dass die Zulässigkeit aktiver Sterbehilfe sich zwar nicht schon mithilfe des Äquivalenzargumentes aus der Zulässigkeit passiver Sterbehilfe ergibt, eine liberale Position aber dennoch erheblich besser begründet ist.³⁷

³⁶ Die Situation ist komplizierter, wenn das Herbeiführen des Todes eines leidenden Patienten nicht *gegen*, sondern lediglich *ohne* seinen erklärten Willen erfolgen würde, etwa weil dieser zu einer solchen Willenserklärung nicht (mehr) in der Lage ist und auch keine frühere Erklärung vorliegt. Eine hinreichend differenzierte Auseinandersetzung mit der Frage, ob in solchen Fällen auch eine aktive Sterbehilfe unter bestimmten Bedingungen moralisch zulässig sein kann, würde den Rahmen dieses Textes jedoch überschreiten.

³⁷ Für hilfreiche Diskussionen und Kommentare zu früheren Versionen dieses Textes danke ich Johann S. Ach, Susanne Boshammer, Annette Dufner, Jan Gertken, Benjamin Kiesewetter, Felix Koch, Peter Schaber, Thomas Schmidt, Konstantin Schnieder, Bettina Schöne-Seifert, Moritz Schröder, Héctor Wittwer sowie den Teilnehmerinnen und Teilnehmern der Kolloquien von Susanne Boshammer in Osnabrück und Thomas Schmidt an der HU Berlin.

Literatur

- Ach, J., Wiesing, U. und G. Marckmann. 2012. „Sterbehilfe: Einführung.“ In *Ethik in der Medizin*, 4. Auflage, herausgegeben von U. Wiesing, S. 232–243. Stuttgart : Reclam.
- Beauchamp, T. L. und J. F. Childress. 2013. *Principles of Biomedical Ethics*, 7. Auflage. New York: Oxford University Press.
- Birnbacher, D. 1995: *Tun und Unterlassen*. Stuttgart: Reclam.
- Brody, B. A. 1989. „Der vegetable Patient und die Ethik der Medizin.“ In *Medizin und Ethik*, herausgegeben von H.-M. Sass, S. 296–310. Stuttgart : Reclam.
- Bundesärztekammer. 2011a. „Grundsätze der Bundesärztekammer zur ärztlichen Sterbebegleitung.“ *Deutsches Ärzteblatt* 108(7): A346–348.
- Bundesärztekammer. 2011b. (Muster-)Berufsordnung für die in Deutschland tätigen Ärztinnen und Ärzte in der Fassung der Beschlüsse des 114. Deutschen Ärztetages 2011 in Kiel. [http://www.bundesaerztekammer.de/downloads/MBO_08_20112.pdf].
- Bundesgerichtshof. 2010. *Urteil vom 25.06.2010 [Sterbehilfe durch Behandlungsabbruch]*. BGH 2 StR 454/09.
- Dempsey, M. M. 2013. „Victimless Conduct and the Volenti Maxim: How Consent Works.“ *Criminal Law and Philosophy* 7: 11–27.
- Dworkin, R. 1993. *Life's Dominion*. New York: Random House.
- Fischer, T. 2014. *Strafgesetzbuch mit Nebengesetzen*. 61. Auflage. München: Beck.
- Gert, B., Bernat, J. L., und R. P. Mogielnicki. 1994. „Distinguishing between Patients' Refusals and Requests.“ *Hastings Center Report* 24(4): 13–15.
- Gert, B., Culver, C. M. und K. D. Clouser. 2006. *Bioethics: A Systematic Approach*. 2. Auflage. New York: Oxford University Press.
- Gesang, B. 2001. „Aktive und passive Sterbehilfe – Zur Rehabilitation einer stark kritisierten deskriptiven Unterscheidung.“ *Ethik in der Medizin* 13: 161–175.
- Gutmann, T. 2002. „Der eigene Tod – Die Selbstbestimmung des Patienten und der Schutz des Lebens in ethischer und rechtlicher Dimension.“ *Ethik in der Medizin* 14: 170–185.
- Hoerster, N. 1986. *Sterbehilfe im säkularen Staat*. Suhrkamp, Frankfurt a. M.
- Höfling, W., Rixen, D. 2003. „Vormundschaftsgerichtliche Sterbeherrschung?“ *Juristenzeitung* 58: 884–894.
- Kagan, S. 1988. „The Additive Fallacy.“ *Ethics* 99(1): 5–31.
- Kuhse, H. 1987. *The Sanctity-of-Life Doctrine in Medicine*. Oxford: Clarendon Press.
- McIntyre, A. 2014. „Doctrine of Double Effect.“ in: Zalta, E. N. (Hg.). *Stanford Encyclopedia of Philosophy*. Winter 2014 Edition. [<http://plato.stanford.edu/archives/win2014/entries/double-effect/>].
- Müller, A. 2018. „Die Verbindlichkeit von Patientenverfügungen: Welcher Wille zählt?“ In: *Zeit – eine normative Ressource?*, herausgegeben von F. Dietrich, J. Müller-Salo und R. Schmücker. Frankfurt a.M.: Klostermann.
- Nationaler Ethikrat. 2006. „Selbstbestimmung und Fürsorge am Lebensende.“ *Stellungnahme, Berlin*. [<http://www.ethikrat.org/dateien/pdf/selbstbestimmung-und-fuersorge-am-lebensende.pdf>].
- Owens, D. 2012. *Shaping the Normative Landscape*. Oxford: Oxford University Press.
- Quante, M. 1998. „Passive, indirekt und direkt aktive Sterbehilfe – deskriptiv und ethisch tragfähige Unterscheidungen?“ *Ethik in der Medizin* 10: 206–226.
- Quante, M. 2002. *Personales Leben und menschlicher Tod*. Frankfurt a. M.: Suhrkamp.
- Rachels, J. 1989. „Aktive und Passive Sterbehilfe.“ in: Sass, H.-M. (Hg.). *Medizin und Ethik*. Stuttgart: Reclam. S. 254–264.
- Roxin, C. 2010. „Zur strafrechtlichen Beurteilung der Sterbehilfe.“ In: *Handbuch des Medizinstrafrechts*, 4. Auflage, herausgegeben von C. Roxin und U. Schroth, S. 75–121. Stuttgart: Boorberg.
- Schöne-Seifert, B. 1997. „Die Grenzen zwischen Töten und Sterbenlassen.“ In: *Jahrbuch für Wissenschaft und Ethik, Band 2*, herausgegeben von L. Honnfelder und C. Streffer, S. 205–226. Berlin: De Gruyter.
- Schöne-Seifert, B. 2007. *Grundlagen der Medizinethik*. Stuttgart: Kröner.
- Schroth, U. 2010. „Ärztliches Handeln und strafrechtlicher Maßstab.“ In: *Handbuch des Medizinstrafrechts*, 4. Auflage, herausgegeben von C. Roxin und U. Schroth, S. 21–50. Stuttgart: Boorberg.
- Sumner, L. W. 2011. *Assisted Death*. New York: Oxford University Press.
- Williams, B. 1973. „The Makropoulos Case: Reflections on the Tedium of Immortality.“ In: *Problems of the Self*, Ders. New York: Cambridge University Press.
- Wittwer, H. 2003. *Selbsttötung als philosophisches Problem*. Paderborn: Mentis.
- Wolf, J.-C. 1993. „Aktive und passive Euthanasie.“ *Archiv für Rechts- und Sozialphilosophie* 79(3): 393–415.

Andreas Müller ist Assistenzprofessor (Tenure Track) für praktische Philosophie mit Schwerpunkt Ethik am Institut für Philosophie der Universität Bern. Er arbeitet zu Fragen der Metaethik und der Theorie praktischer Gründe, sowie zu Grundlagenfragen der normativen Ethik und deren Zusammenhang mit ausgewählten Problemen der angewandten Ethik, etwa im Zusammenhang des Phänomens der Einwilligung.

Pornography's Authority to Subordinate Women

Introduction

pornography, n.:

the explicit description or exhibition of sexual subjects or activity in literature, painting, films, etc., in a manner intended to stimulate erotic rather than aesthetic feelings; printed or visual material containing this. (Oxford English Dictionary 2019)

They way in which the Oxford English Dictionary defines pornography is also how it is mostly perceived by society; as something used to satisfy an urge with, or “masturbation material”, as Catharine MacKinnon once called it (1993, 17). However, it is much more than that. After already being heavily criticised during the 1980's by notable feminists like Catharine MacKinnon and Andrea Dworkin, the main issue emerged: what was depicted in pornography was the subordination of women. Yet, MacKinnon's and Dworkin's criticism did not end there. They furthermore concluded that pornography not only depicts subordination but is a form of subordination itself. It therefore takes an active role in the discrimination against woman and is an advocate for women's standing as sex objects, as having no rights and as wanting to be victims of sexual violence. MacKinnon made the point the following way: “Pornography defines what a woman is through conditioning the male sexual response to that definition, to the unilateral sexuality pornography is part of and provides” (1987, 181). In the last thirty years, pornography's influence has only grown, and the topic is more pressing than ever: the reach pornography has gained through the internet and portable devices, such as smartphones and tablets, has exploded into dimensions MacKinnon and Dworkin could not have imagined in the 1980's. Nowadays, children younger than ten years old are able to access videos depicting sexual interactions that have

blown out of proportion in their depiction of physical and verbal violence. In 2010 already, studies showed that in best-selling pornography, more than 88% of the scenes analysed depicted physical violence, with the aggressors being mostly men and the victims being overwhelmingly women (Bridges et al. 2010). This trend has only increased in the last years, with pornography's willingness to break social taboos and to venture further into extremes growing steadily (Makin and Morczek 2015). For all these unsettling reasons, it is today more important than ever to talk about pornography and to show how exactly it is also responsible for the subordination of women.

This thesis will be based on a theory by Rae Langton, who combined MacKinnon's argument that pornography is subordinating women with John Austin's Speech Act Theory. In her theory, defining pornography as speech, she argues that pornography subordinates, as well as silences, women. She draws on Austin's idea that certain utterances are themselves actions (he called them *performatives*) and establishes a theory, in which it is shown that if pornography has the necessary authority, it is able to deprive women of their rights, rank them as objects, and legitimate sexual violence towards them, as well as make certain speech acts, like the one of refusing sexual intercourse, unspeakable for them. The issue she faces, though, is the difficulty to prove that pornography has this sort of authority, since an answer to this question would have to include empirical data.

In this thesis, I will argue that today pornography has all the authority it needs to subordinate women. In my venture to do this, I will begin by elaborating on Austin's Speech Act Theory and explain the terms which are necessary to understand this thesis. This will be followed by a close analysis of Langton's theory, after which I will try to fill in some missing points in her ar-

gumentation. Most importantly, I will introduce a new category of speech acts to help resolve some of the issues in her theory. The next step in my argumentation will be to show why exactly pornography has the authority it needs to make judgements about women in our society. Drawing on both Austin and Langton, I will argue that pornography makes specific verdicts about women to which society tries to live up to, therefore making them count in all the ways that matter. My argument will conclude with the idea that pornography subordinates because it is a powerful enough authority to be able to construct reality.

1. The theory of “Speech Acts and Unspeakable Acts”

1.1 Terms of “How to do things with words”

The main theory Rae Langton supports her argument with, goes back to John L. Austin’s *How to do things with words* (1962). In his theory of speech acts, he presents the argument that sentences can be more than just the stating of a fact or the description of said fact. Since the terminology of Austin and Langton is very specific and complex, all necessary terms shall be explained in more detail to prevent confusion at later stages of the paper.

Austin divides statements into two different categories: *constative* and *performative* utterances. *Constatives* are mostly descriptive and can either be true or false (Austin 1962, 3). Austin, however, focuses on the *performatives* since he finds that they have not been discussed sufficiently, especially in philosophy of language. He furthermore believes that the discussion has ignored the actions which can be performed with the help of specific utterances. Examples of this sort of utterances might be to say “I do” and by saying so marrying another person or “I name this ship the *Queen Elizabeth*” in the context of a ceremony with the goal of naming a ship (Austin 1962, 5). Performatives, according to Austin, therefore neither describe or constate a fact, nor can they be true or false. In fact, they are part of performing an action—hence the names ‘performatives’ and ‘speech acts’ (Austin 1962, 7). Speech acts are thus a sub-category of actions. Such utterances which are actions have three different aspects, according to Austin, namely the *locutionary act*, the *illocutionary act* and the *perlocutionary act*. The locutionary speech act is “the act of ‘saying something’” (Austin 1962, 94), so the term merely describes what is being said with an utterance. The illocutionary speech act is the “perfor-

mance of an act *in* saying something” (Austin 1962, 99) and therefore has an *illocutionary force* which means that something is being done or intended with this utterance. It also describes the use of the locution. Lastly, there is almost always a certain effect produced by an utterance and these effects may well be intended. This sort of speech act Austin calls the perlocutionary speech act or simply the *perlocution*. Since these terms are now quite abstract, an example by Austin himself shall shed some light: he establishes a situation in which someone says to me “Shoot her!” and I then kill said person. Austin then continues to explain which speech acts are done with this utterance:

Act (A) or **Locution**

He said to me ‘Shoot her!’ meaning by ‘shoot’ shoot and referring by ‘her’ to *her*.

Act (B) or **Illocution**

He urged (or advised, ordered, etc.) me to shoot her.

Act (C) or **Perlocution**

He persuaded me to shoot her.

(Austin 1962, 101)

The perhaps most difficult, but definitely most interesting speech act, is that of the illocution. The question emerges what actually is necessary to be able to successfully perform an illocutionary speech act since they can fail like any other action. Most of all, the illocution has to satisfy certain felicity conditions. There are conventional rules or procedures that must be followed and be carried out correctly by all participants (Austin 1962, 14–15).¹ A speaker must also have the authority to carry out a speech act, for example for an order to be successful, it must be given by a commanding authority (Austin 1962, 28). Another important term here is “uptake”: the hearer of the illocutionary speech act has to recognize the speech act’s intention or its illocutionary force. The speaker therefore has to “secure uptake” (Austin 1962, 138), or otherwise he will fail. Lastly, Austin makes a distinction between what kind of illocutionary forces there can be. He distinguishes between five different groups but only two, namely the *verdictives* and the *exercitives*, shall be explained at this point because they are most important for this paper.² Verdictives are characterized “by the giving of a ver-

¹ For example, if someone wants to marry, they have to say “I do”; if they do not follow that rule, then they do not marry.

² Other illocutionary forces are the *commissives*, the *behabitives* and the *expositives* (Austin 1962, 150). The actions which are carried out with these are not the kind of actions which are important for this paper.

dict [...] by a jury, arbitrator, or umpire” (Austin 1962, 150). Cases of verdictives might be illocutions that rank, value, find (as a matter of fact) or place (Austin 1962, 152). Exercitives, as Austin states, “are the exercising of powers, rights, or influence” (Austin 1962, 150). Cases of exercitives are illocutions that order, enact, grant or proclaim (Austin 1962, 155).³ After having defined the most important terms for this paper, let us turn to Langton’s theory.

1.2 Introducing Langton’s two theses

Rae Langton (in “Speech Acts and Unspeakable Acts”, 1993) begins by giving some general information about the two theses she wants to defend. Her goal is to show that these two theses can in fact be defended from a philosophical point of view, unlike some opponents have argued. The theses are the following:

- **Thesis 1:** Pornography subordinates women.
- **Thesis 2:** Pornography silences women.

Both theses take as a premise that pornography can carry out an illocutionary speech act. This stems from the following realization: “Pornography depicts subordination and causes it. That—in Austin’s terms—is to describe its locutionary and perlocutionary dimensions” (Langton 1993, 296). What is missing in this is the kind of illocutionary force pornography has. The illocution of pornography has been ignored until this point because the overall consensus was that pornography is speech depicting subordination (locution) which has certain undesirable effects on the viewer and on society (perlocution). This opinion was challenged most notably by Catharine MacKinnon and Andrea Dworkin who introduced the idea of pornography being subordination itself as well.⁴ Supporting herself on Austin, Langton now strives to defend MacKinnon’s and Dworkin’s theory.⁵

3 Since Austin died before he could publish his work as a book, *How to do things with words* is a composition of his lectures. For this reason, the theory is not always perfectly clear. An example of illuminating secondary literature might be *Foundations of speech act theory: philosophical and linguistic perspectives*, edited by Savas L. Tsohatzidis, which was published in 1994.

4 The definition of pornography that was introduced by Catharine MacKinnon and Andrea Dworkin shows how the issue with pornography was newly defined as it being a form of subordination. The following definition was established originally by Dworkin in 1981 (*Pornography: Men Possessing Women*) and can be found in *Feminism Unmodified* (1987) by MacKinnon: “Pornography is the graphic sexually explicit subordination of women, whether in pictures or in words,” (262) where “women are presented as sexual objects for domination, conquest, violation, exploitation, possession, or use, or through postures or positions of servility or submission or display” (274). This phrasing is an abbreviated version of their legal definition.

5 Langton’s argument takes as a premise that pornography is

Before one can dive into Langton’s two theses, some of her remarks on them have to be mentioned as well. Concerning the first thesis (I shall call it the *subordination-thesis*), there is a conflict between liberty and equality, or, more precisely: a conflict between “the liberty of men to produce and consume pornography, and the rights of women to equal civil status” (Langton 1993, 298). On the other hand, the second thesis (let us call it the *silencing-thesis*) poses a conflict within liberty itself which results in a conflict between being free to produce and consume pornography and the liberty of women to speak freely. Langton also notes that the silence mentioned in the silencing-thesis has a literal instead of a metaphorical meaning (Langton 1993, 293–300).⁶

Thesis 1: Pornography subordinates women

Langton explains her two theses by always discussing the general ideas of subordinating and silencing speech acts first before applying the concepts to pornography. Following her structure, it shall at this point be explained what *subordinating speech acts* are.

In order to find examples for subordinating speech acts, Langton begins with the utterance “Blacks are not permitted to vote” (Langton 1993, 302). The utterance, if carried out by an authority like the Apartheid government, has certain repercussions. It is a locutionary speech act which has as its content that Blacks are not permitted to vote. It is also a perlocutionary speech act since it has the effect that Blacks are not permitted to vote. Nonetheless, the act of uttering this sentence in this context is an act that changes something about the world and therefore, it has an illocutionary force. According to Langton, illocutionary acts of subordination are acts that unfairly *rank* the worth of someone, *legitimate* discriminatory behaviour towards someone and unjustly *deprive* someone of rights and power (Langton 1993, 304). In the Apartheid example, Black people are unjustly deprived of their right to vote. Langton also reaches back to the idea of exercitives and verdictives already introduced by Austin and which aptly describe illocutionary speech acts which have the power of carrying out acts of subordination (Langton 1993, 302–305).

speech. If one does not want to support this line of thought, an alternative argument is done by Melinda Vadas (1987) who argues that pornography subordinates without including the idea that pornography is speech.

6 Some have criticised the idea that pornography silences and argued that it was a confusion, meaning that one could only mean silencing in a metaphorical way (e.g. Ronald Dworkin 1991). However, Langton speaks of literal silencing because certain speech acts have been made unspeakable.

The question remains how pornography can perform such speech acts of subordination. Langton begins with the explanation why neither the locutionary content of subordination nor the perlocutionary effects of pornography resulting in subordination are the issue at hand. Firstly, although pornography has the locutionary content portraying the subordination of women, there are other forms of media which have the same locutionary content, for example a documentary about the subordination of women. Such a documentary, however, would not include an illocutionary act of subordination—its message or intention would be something else. Secondly, the perlocutionary effects pornography has on its viewers lead to an increase in sexual violence towards women because of pornography's behaviour-altering properties.⁷ Langton classifies the harm caused by this as part of women's subordination status and therefore states the following: "If pornography has sexual violence as its effect and sexual violence is an aspect of women's subordination, then pornography is a perlocutionary act of subordination" (Langton 1993, 307). Nevertheless, this still does not capture the active role that pornography plays in the subordination of women. The illocutionary aspect goes beyond these issues and Langton claims that pornography carries out exercitive, as well as verdictive speech acts about women. She states:

"Pornography is, first, verdictive speech that *ranks* women as sex objects, and, second, exercitive speech that *legitimizes* sexual violence. Since sexual violence is not simply harm, not simply crime, but discriminatory behaviour, pornography subordinates because it legitimizes this behaviour." (Langton 1993, 307–08, emphasis added)

With this, Langton admits, the issue arises whether pornography is close enough to a paradigm case of illocutionary speech acts so one can firmly say that pornography is in fact carrying out this illocutionary act of subordination. In a paradigm case, all felicity conditions would have to be fulfilled and uptake be secured for the illocution to succeed. One of the most important felicity conditions in this context would be whether the authority of the speaker is given or not. In

this context, Langton introduced the idea of a "game of sex" (Langton 1993, 311) in which pornography has a similar authority to carry out illocutions as the umpire has in the game of tennis. However, Langton lacks the empirical data to support her thesis but she nonetheless postulates the theory that "[i]n this domain, and for these hearers, it may be that pornography has all the authority of a monopoly" (Langton 1993, 312).⁸ Furthermore, it remains unclear whether uptake has been secured since the concept does support itself on the hearer's perception which is difficult to attain from an outside perspective.

Although Langton struggles to demonstrate that pornography has the authority to subordinate, she has nonetheless shown that if pornography has said authority, then it is capable of subordinating women. She has also succeeded in establishing the subordination-thesis as a thesis that is philosophically defensible, which was her main goal (Langton 1993, 305–14).

Thesis 2: Pornography silences women

Turning to her second thesis, Langton introduces the concepts of both silenced and silencing speech acts. They both follow Austin's idea of the three different aspects of a speech act.

Concerning *silenced speech acts*, Langton firstly presents what she calls "simple silence" (Langton 1993, 318) in which the locutionary speech act is silenced (because people are too intimidated to speak or think that their speech would be in vain). In these cases, there is no utterance at all. Secondly, speakers can experience *perlocutionary frustration* where they are able to speak but fail to achieve the effects they intended.⁹ Lastly, silence can happen on an illocutionary level, which Langton calls *illocutionary disablement*. In these cases, the speaker experiences a failure to perform the intended action. This can happen for multiple reasons: either the speaker does not have the required authority (a slave commands his master), he has not secured uptake (an actor screams "Fire!" and his viewers think it part of the play¹⁰) or the felicity conditions are not satisfied (a black South African's vote is not counted during the Apartheid regime). The important thing about these cases is that the speech acts not only fail to have the intended effects, but they fail altogether,

7 Studies have shown that the consumption of pornography can lead to an increase in violent behaviour, as it was shown for example by Donnerstein, Linz and Penrod in *The Question of Pornography: research findings and policy implications* (1987). Their studies also concluded that the material does not have to be sexually explicit to have this effect.

8 We shall come back to this idea later.

9 An example of this might be that someone invites some people to a party, but no one shows up (Langton 1993, 315).

10 This example goes most famously back to Donald Davidson in "Communication and Convention" in *Inquiries into Truth and Interpretation* (1984).

meaning that the illocutionary act is not possible for these speakers. They can utter the words and carry out a locutionary speech act, but they do not perform an action (Langton 1993, 314–17).

In accordance to the silenced speech acts, the *silencing speech acts* put forward by Langton follow a similar pattern. The general idea is that one way in which speech acts can be silenced is through other speech acts: one can order someone to silence, therefore causing locutionary silence. And one can conduct speech acts that prevent perlocutionary effects from happening, resulting in perlocutionary frustration. Most importantly, though, one can carry out speech acts which cause illocutionary disablement by making certain other speech acts unspeakable. At this point, Langton argues that since the felicity conditions for illocutionary speech acts are mainly set by conventions that have been introduced by speech acts (e.g. laws like the one preventing black South Africans from voting), it seems not that far-fetched to say that speech acts can define the range, scope, and space of other speech acts and therefore determine which illocutionary acts are possible (Langton 1993, 318–20).

To apply these categories of speech acts to pornography, Langton introduces two important examples, namely “Refusal” and “Protest”¹¹. In “Refusal”, she describes a situation in which someone wants to force sex onto a woman and the woman refuses with the locution “No” but her attacker proceeds to rape her. According to Langton, the woman’s speech could have failed in two different ways: either she experiences perlocutionary frustration (she has secured uptake and her hearer recognizes her illocution—but ignores it) or illocutionary disablement (she has not secured uptake and her illocution is not recognized) (Langton 1993, 320–22).

In accordance with the categories explained above, there are three ways in which pornography can silence women. To achieve simple or locutionary silence, pornography can be used to threaten or blackmail someone. And to prevent women from achieving certain intended effects with their speech acts, they can be perlocutionary frustrated or illocutionary disabled. To explain how pornography can perlocutionary frustrate women, we return to the example “Refusal”, where in one case a women’s illocution of refusing was recognized but her hearer continued to rape her in spite

of or perhaps even because of her refusal. Langton argues that pornography legitimates this behaviour “by sexualizing the use of force in response to refusal that is recognized as refusal. Such pornography eroticizes refusal itself, presenting the overpowering of a woman’s will as exciting” (Langton 1993, 323). Let us turn to the second way in which the woman’s refusal could have failed in this example, namely illocutionary disablement: this happens when a woman refuses sex, but her illocution is not even recognized. Remembering the possible reasons for a failed illocution, it seems that some felicity conditions are not satisfied: “Whatever the conventions governing sexual interactions may be, they can mean that intending to refuse [...] is not enough. The rules fixing possible moves in the language game of sex are such that saying ‘no’ can fail to count as making a refusal move” (Langton 1993, 323–24). The question remains how pornography can achieve such silence in women. Langton answers this by explaining that—since felicity conditions can be set by other speech acts—pornography can also conduct such speech acts which define what other speech acts there can be. In the case of the example “Refusal”, it seems that pornography does not only legitimate rape through eroticising refusal. Even more strikingly, pornography gives the notion of refusal no space at all in its depictions of sex, as it portrays women as only being able to consent. This means that a hearer of the locution “no”, who has only learned from pornography about the conventions concerning sexual interactions may not even recognize refusal or just plainly interpret it as consent. In this situation, the illocution of refusal has been made unspeakable for women and is destined to fail. As a result, in the example “Refusal”, the woman has not only been unable to secure uptake, she is also unable to attain the authority over her life and body (Langton 1993, 322–26).

An important premise in the argument of saying that pornography does indeed silence women demands it having the necessary authority. Langton states: “if pornography is speech that determines the kind of speech there can be—then it is exercitive in Austin’s sense, for it is in the class of speech that confers and removes rights and powers” (Langton 1993, 326). As already stated before, exercitive speech needs to be spoken by an authority to work—if pornography has this sort of authority is again a question of empirical data. Nevertheless, Langton was able to show that her silencing-thesis is philosophically defensible, same as her subordination-thesis.

11 “Protest” mainly describes a situation in which a woman’s protest against pornography is only recognized as more pornography, therefore illocutionary disabling her. It shall not be explained furthermore but had to be mentioned for the sake of completeness.

2. Problems with Langton

In the following, I will elaborate on some points in Langton's theory which appear to be absent. There is especially a case that has not been discussed by Langton at all but demands attention, nonetheless. Furthermore, two cases seem to have been categorised incorrectly by Langton. At this point, these issues will only be put forward, a solution to them will be presented later in this thesis. In this chapter, I shall also explain a critical argument against Langton by Leslie Green.

2.1 A case missing

First of all, Langton only includes locutionary silencing as in the case of pornography being used as a means to a certain goal. In that case, pornography is either used to threaten or blackmail someone into silence.¹² However, the case in which pornography actively silences locutionary speech acts is excluded. There is in fact another case of locutionary silencing when the locutionary speech act (e.g. during a rape scenario) is not conducted out of fear or self-protection (e.g. a woman who decides not to refuse sex because she fears that she will be beaten or killed otherwise). In this case, pornography was not used by someone as a means to threaten someone else. Moreover, if one follows Langton's line of argument, the felicity conditions have been altered by pornography so that the woman herself thinks it may be more dangerous to refuse sex than to just endure it. Therefore, there seems to be more to silencing a locutionary act than Langton admits. Since she only includes cases in which the locutionary speech act is impossible to conduct because pornography was used as a tool to threaten someone, she misses an important point: The inefficiency of the refusal can go as far as being something that is not only useless but outright dangerous. It therefore seems that the locution "no" to decline sex is in this case not only silenced but is deemed to have an even more negative impact, if carried out.

2.2 Cases in the wrong category

Both of the following cases belong to the example "Refusal" which was introduced by Langton. They are not necessarily abbreviations of the original example but are used to elaborate on some aspects.

Let us first take a look at the perlocutionary frustration which is one way in which a woman's refusal can

malfunction in example "Refusal", the idea being that the problem here is in fact another one. In this example, the illocutionary speech act of refusing is recognised but is chosen to be ignored which, according to Langton, is perlocutionary frustration; the illocutionary act, therefore, does not have the desired effect. However, the problem, as I argue, lies much deeper: although uptake is secured it is deemed irrelevant. This tells us something about the standing of the woman in this situation, namely that her attacker recognizes her authority to refuse him but deems it irrelevant. "Recognize" in this sense only means that he is able to perceive her authority but does not value and respect it. The illocutionary act is therefore considered worthless. This shows us that—although the woman has the authority to state the refusal of sex—her authority is still overthrown by the desire of her attacker. Thus, the problem is not only that the woman's illocutionary speech act does not have the desired effect (which results in perlocutionary frustration) but that her illocution of refusing is being reduced to a point where it can be ignored.

With illocutionary disablement there is another issue at hand: the woman's uptake can be completely missing which, according to Langton, results in illocutionary disablement (no illocution is carried out) (Langton 1993, 318–19). The problem is that although there seems to be a disablement of the illocution of refusal, there is, at the same time, another illocution which is supposed to take its place, namely the illocution of consent. The question remains what kind of standing this illocutionary speech act has in this situation. First and foremost, it is misattributed to the woman which might not have been done intentionally by her attacker. Instead, it is a result of him being conditioned to misinterpret the illocution "refusal" as "consent". Although the attacker might not do this on purpose, it is nonetheless what is happening and the very fact that the felicity conditions are such that he is in a position to change her illocution, speaks of a certain entitlement. The attacker's entitlement to interpret the illocutionary act as seen fit is an example of a higher standing of his person in this context. It also shows that in this situation, the woman's illocution has become subordinated because it can be interpreted as seen fit.

2.3 Pornography cannot subordinate

In the following, I will elaborate on one of the biggest issues that Langton is confronted with: in her first thesis, it seems quite difficult to bring about evidence

¹² For example, MacKinnon mentions cases in which children are forced into prostitution by threatening them to leak the pornographic evidence of it to their relatives. The children therefore stay silent (MacKinnon 1987, 180).

that pornography subordinates. This is also due to the fact that the question whether pornography has the authority to subordinate cannot only be answered through philosophical thinking but must also take into consideration the relevant empirical facts. In this section, an example of a counterargument against Langton's subordination thesis shall be presented as well as Langton's reaction to it.¹³

Said counterargument was introduced by Leslie Green in his paper "Pornographizing, Subordinating, and Silencing" (1998) and it shows that Langton's subordination-thesis not only struggles with the lack of empirical data but can be attacked on another basis: according to Green, it is not only the question whether pornography has the authority stipulated by Langton but, furthermore, whether it would even matter to women if it had the required authority to subordinate. Green argues that pornography does not subordinate women because women are not under its sphere of influence. He supports his thesis with the Mick-and-Max-Analogy in which he explains that if Mick and Max are both gay but Mick is Catholic and Max Jewish and the Catholic Church sees homosexuality as an 'intrinsic moral evil' then Mick is subordinated by the Church, but Max is not. He argues that women are similarly not under the jurisdiction of pornography same as Max is not affected by the Catholic Church. Concerning pornography, Green states: "the first chain of command is not found in the authority of pornographers over women, but of pornographers over *men and boys*, or, put another way, the authority of some pornographers (producers) over other pornographers (consumers)" (Green 1998, 295). Green means by this that women are not affected by pornography because they are not addressed by it. With Mick and Max, Green argues, the Church only has a "*purported*" authority that is different from what Green calls "*effective* authority". In addition to that, its homophobic beliefs stand in contrast to what society's general convictions are. Similarly, Green argues that pornography does not have a didactic function and is mainly used as a means for self-pleasure, therefore not pursuing the goal of teaching about how sex should be like, which means that it does not have effective

authority. Furthermore, Green (1998, 296) states that, similarly to the Mick-and-Max-Analogy, the subordination depicted in pornography is confronted with society's beliefs that sexual violence is wrong (he names the "putative social authorities, including the state, the family, and the church"). As long as these authorities stand against the convictions presented in pornography, Green argues, then "pornography has the character not of an authoritative voice of the ruling power but of a private view" (Green 1998, 296). Therefore, pornography is not capable of changing societal norms and conventions about speech acts pertaining sexual interactions (Green 1998, 293–98).

In Langton's response to this ("Pornography's Authority? Response to Leslie Green" in *Sexual Solipsism* from 2009¹⁴) she first of all points out that one has to differentiate between verdictive and exercitive speech in this case. The speech act the Catholic Church carries out in the Mick-and-Max-Analogy is verdictive but has an exercitive dimension because it is often the case that "[w]hen the powerful say 'this is how it is', *the world arranges itself to affirm* what the powerful say" (Langton 2009, 96)¹⁵. Verdictives are therefore just as self-fulfilling as exercitives but they take more time and work in a more subtle way. Langton makes an example of a girl who repeatedly hears the verdictive "Anna is not clever" from her classmates and teachers.¹⁶ This illocution is not immediately fulfilled after it is spoken (as an exercitive would) but nonetheless "with time and the classroom neglect of Anna which [it justifies, it] may become true" (Langton 2009, 95). Similarly, she argues, the Catholic Church in Green's analogy does not have the authority to conduct exercitives but speaks verdictives about homosexuals. The analogy therefore does not work, since the Church does in fact subordinate Max. Furthermore, she doubts that society really condones sexual violence as much as Green thinks it to be the case (she refers to the fact that rape victims are often shamed, and rape is one of the least reported crimes). Langton also points out the fact that women are often under the jurisdiction

13 There are a number of other critics who attack Langton's theory but do so with different intentions. Alexander Bird doubts the philosophical defensibility of Langton's theses and argues that uptake is not required for illocutions ("Illocutionary Silencing" in 2002). Langton was defended by Mari Mikkola in "Illocution, Silencing and the Act of Refusal" (2011). Another line of critique was introduced by Daniel Jacobson who argued that for free speech, only locutions are important and therefore, pornography does not silence ("Freedom of Speech Acts? A Response to Langton" in 1995). Langton has defended her theory with support by Jennifer Hornsby in "Free Speech and Illocution" (1998).

14 A more elaborate version of this would be Rae Langton's paper "Subordination, Silence, and Pornography's Authority" published in 1998.

15 Langton draws on MacKinnon here: "Having power means, among other things, that when someone says, 'this is how it is', it is taken as being that way. [...] When that person is believed over another speaker, what was said becomes proof. Speaking socially, the beliefs of the powerful become proof, in part because the world actually arranges itself to affirm what the powerful want to see" (1987, 164).

16 This example goes back to the Pygmalion studies by Rosenthal and Jacobson (*Pygmalion in the Classroom* in 1968).

of pornography through the men that consume it and are therefore influenced by it¹⁷ (Langton 1993, 93–100). Nonetheless, Langton still cannot give proof that pornography has indeed the authority to carry out these verdictive speech acts that have an exercitive dimension. The question remains if empirical data would support it.

3. Introducing a new category: Subordinated Speech Acts

3.1 Silencing as a form of subordination

After elaborating on some issues Langton experiences in her theory, I will now try to resolve these problems in order to establish the argument that pornography does indeed subordinate. To do so, I will first expand Langton's theory to be able to cope with the missing cases previously described.

It is striking that Langton only includes three types of speech acts (subordinating speech acts, silencing speech acts, silenced speech acts) and uses them to explain her two theses. However, there seems to be a pattern with these categories which has remained incomplete: as there are silenced speech acts which can be a result of silencing speech acts, it appears that there would have to be *subordinated speech acts* that can be the result of subordinating speech acts. With the new category in italics, the following schema aims at visualising this:

Subordinating Speech Acts:

Acts that rank people, deprive them of rights and legitimate discrimination.

Silencing Speech Acts:

Acts that prevent locutionary, illocutionary or perlocutionary speech acts from happening

Subordinated Speech Acts:

Speech acts that matter less than others or are faulty due to unfairly changed felicity conditions

Silenced Speech Acts:

Failure to perform locutionary acts, illocutionary disablement or perlocutionary frustration

Moreover, the schema also shows in what kind of hierarchical relations the different categories stand to each other. Looking at the definitions of subordinating and silencing speech acts, it is not far-fetched to say

that silencing might just be a form of subordination:¹⁸ the definition of subordinating speech acts includes the fact that these speech acts are able to deprive people of their rights. I argue that one of these rights can be the right of (free) speech. This means that speech acts that silence are speech acts that subordinate. The same correlation exists between subordinated and silenced speech acts: silenced speech acts are mainly speech acts in which someone lacks the opportunity to conduct his right of speech and is therefore subordinated. In an attempt to capture the problematic aspects of subordinated speech acts, I have included the following definition in the schema: subordinated speech acts matter less than other speech acts or are somehow faulty due to unfairly changed felicity conditions. Note that the right of speech or the authority to carry out a specific speech act can be included in these felicity conditions. Said conditions also have to be changed *unfairly* to contribute to subordination since in our society, it is necessary for some people's speech acts to matter more than others, without them being cases of subordination. Let me give an example of a soccer game in which a referee is supported by his assistant referees. The assistant referees are able to conduct speech acts that make a verdictive statement about the match. However, although their speech acts count more than that of a spectator shouting "Foul!" (who has no say in this), they can still be overruled by the referee who has more authority in this matter. This shows that the speech acts of the assistant referees are less important than the speech acts of the referee. Yet, that does not make them subordinate in the sense in which Langton talks about subordination (which she also defines as unfair).

For thoroughness' sake, an explanation about silenced speech acts has to be given. One could say that since they are silenced speech, they are not speech anymore. But that would be a hasty conclusion. If we look at them in the context of subordinated speech acts, then it seems that silenced speech acts are the most extreme form of subordinated speech because they cannot exist anymore due to unfair felicity conditions. Although silenced, they still have to be defined as speech, otherwise one would be justified in ignoring the issue by saying that there is no problem remaining. However, a fundamental right has been taken away, namely freedom of speech. Silenced speech acts are still subordinated, and this has to be discussed. I therefore conclude with the statement that

17 Langton's response is more elaborate but had to be broken down due to space issues.

18 Langton does mention this correlation but does not elaborate on it (Langton 1993, 329).

silenced speech acts are just as much a subcategory of subordinated speech acts as silencing ones are of subordinating ones.

3.2 How to cope with problematic cases

In the following, I will look at the cases previously discussed in chapters 2.1 and 2.2 to show that they are in fact cases of subordination instead of silencing—or at least have an aspect of subordination that was left out by Langton.

In the missing case of locutionary silencing, we have looked at a situation in which a woman does not even dare to refuse sex out of fear of something even worse happening to her. In this case, the speech act of saying “no” has been degraded as unhelpful although being one of the most powerful and clear ways to refuse something. Moreover, it is even deemed dangerous. This shows how this speech act is more than locutionary silenced: it is not only faulty but even counterproductive. To grasp the magnitude of this, one has to look at this speech act as also being subordinated, not merely silenced. The right of speech was technically not even taken away (the woman could carry out the locution “no”) but acting within this right would cause her even more harm.

The cases of subordination described in chapter 2.2 are further examples of the subordination aspect in silenced speech acts. The two examples of subordinated speech acts are the cases in which a woman either has secured uptake while refusing sex but experiences perlocutionary frustration or lacks uptake and is therefore illocutionary disabled. Both examples face the problem that either uptake is granted but deemed irrelevant or an illocution is misattributed to the woman. Where uptake becomes irrelevant, the whole illocution of refusal intended by the woman seems to matter less and is ignored. This suggests that although the situation ends up in perlocutionary frustration, the illocutionary speech act of refusing was subordinated because it was unfairly ignored due to felicity conditions that make it seem as if her illocution is of less importance. This becomes even more obvious in the second case in which illocutionary misattribution happens. The illocution of refusal is similarly subordinated but furthermore, it is overridden with another illocution (of consent) which was not the victim’s intention. Her speech act is therefore faulty due to felicity conditions which make it possible for her attacker to (although unintentionally) change the illocution he is hearing. This makes her speech act again a subordinated one first, and a silenced one second.

It has to be added that we can still look at these cases as instances in which silencing happens. However, they end up being cases of subordination anyway as it was established in Chapter 3.1 that silencing is a form of subordination and therefore a subcategory.

4. Pornography’s authority

4.1 Pornography’s status in the domain of sex

Let us now turn to one of the main issues which Langton experienced, namely that it is hard to give proof of pornography’s authority. In this chapter, I will show that especially in our society today, pornography has all the authority it needs to silence and subordinate.¹⁹ In an attempt to do this, I will include some empirical facts to help illustrate how exactly pornography might very well have the “authority of a monopoly” (Langton 1993, 312).

It has to be said that this thesis shall by no means try to give all the empirical data to support Langton’s argument—but a closer look at society and especially how it is talked, taught and learned about sex seems to be illuminating. If someone argues that pornography does not have the required authority to conduct silencing or subordinating speech acts one has to ask oneself exactly what other options there are in the domain of sex to learn about the rules of sexual interactions. Since it seems arbitrary that there are specific rules to sex then exactly what kind of institutions or authorities are there that establish these rules in the domain of sex in our society? Looking back to what Green suggested in his article, where he named the “putative social authorities, including the state, the family, and the church” (Green 1998, 296), I will try to show what kind of authority these institutions have in the domain of sex. While doing this, I will look at some data that may give some insight into their efficiency. To give a frame to the vast expanse of data, I will concentrate on the United States since they have been, for years, the biggest consumer of pornography (taken from Pornhub’s “Year in Review” statistics, 2016–2018).²⁰

Concerning the state, it seems advisable to look at programs funded by the US government that might help

19 Nellie Wieland also defends Langton’s theory by saying that pornography has the authority it needs but grounds her ideas on a different line of argument than mine (2007).

20 Pornhub’s Year in Review’s have been used here because Pornhub.com is one of the largest and most-visited websites for pornography. They provide yearly reviews about the demographics of their visitors. The conclusion that the US are the biggest consumer of pornography stems from the data about the daily traffic on Pornhub.

define the rules in the domain of sex. One option to do this is sexual education. Let us look at what kind of education the US has provided for the last twenty years: what has overwhelmingly been taught is that sexual abstinence is the only way to cope with one's sexuality. In fact, since the early 2000's, the US government has only provided funding for abstinence-only programs with the main goal of preventing teenage pregnancies (Stanger-Hall and Hall 2011). The reason for this is that the US have ranked "first among developed nations in rates of both teenage pregnancy and sexually transmitted diseases" (Stanger-Hall and Hall 2011). Although studies have shown the inefficiency of abstinence-only programs in preventing teenage pregnancies, the US have continued to fund this sort of sexual education with only minor improvements (Stanger-Hall and Hall 2011). With President Trump now again wanting to introduce it as the only sort of sexual education funded by the government, a setback seems imminent (SIECUS 2019). But what kind of implications does this have on the question whether sexual education is a valuable option for authority in the domain of sex? Most importantly, there seems to be a problem with the topics of the available sexual education.²¹ Instead of speaking about the rules of sexual interactions, it is taught that sex is something one should simply not do. The goal of sexual education, namely preventing teenage pregnancies, is further very restrictive and does not promote topics such as consent, sexual violence and pornography itself. Sexual education appears to ignore the rules concerning sexual interactions and therefore cannot be the predominant authority that teaches said rules. This also gives us an image of what exactly the state is doing about pornography: namely ignoring the discussion. The second institution Green talks about is the family. While I will not try to argue against a family's influence in the upbringing of someone, the question remains if a family is always capable of teaching the rules of sexual intercourse, especially if family members had this kind of previously discussed sexual education themselves. In fact, if we look at what sexual education is teaching, there seems to be a predominance of shame and taboos which one can still not talk about.²²

21 Probably also with availability since only 72% of fifth- and sixth-grade teachers reported that sexual education was taught at one or both grade levels (Stanger-Hall and Hall 2011).

22 I am intentionally ignoring the third institution Green proposed, namely the Church. Although it seems that religious beliefs are already present in the kind of sexual education the US government funds (in which teens are urged to wait until marriage instead of learning about the rules of sexual interactions), to evaluate the Church's influence on this topic is far too complex for this thesis and is also not its goal.

I am in line with Langton's argument here: to say that pornography does not have the authority to spread its convictions by saying that there are other authorities that condemn sexual violence is a hasty conclusion. Just because the law prohibits sexual violence, it does not mean that society talks about or condemns it as much as it should.

If we compare what little sexual education and other institutions are doing to teach about the rules of sex with the amounts of pornography consumed daily (92 Million average daily hits worldwide and 207,405 videos watched per minute on Pornhub alone in 2018), the facts are quite sobering. With its easy accessibility and portability through the internet and mobile devices like smartphones and tablets, pornography's reach has massively increased since MacKinnon first called it 'subordination' instead of the 'depiction of subordination'. It therefore is not far-fetched to say that since there are no valuable alternatives where one could learn about the rules in the domain of sex, pornography has established a predominance in the domain. Let me end this chapter with a quote from Langton that fittingly summarizes the issue:

"What is important is whether it is authoritative in the domain that counts—the domain of speech about sex—and whether it is authoritative for the hearers that count: people, men, boys, who in addition to wanting "entertainment," want to discover the right way to do things, want to know which moves in the sexual game are legitimate. What is important is whether it is authoritative for those hearers who—one way or another—do seem to learn that violence is sexy and coercion legitimate." (Langton 1993, 312)

And this seems to be the case.

4.2 Pornographic speech has the authority to silence

At this point it is time to return to Langton's two theses and I will try to show that maybe with a weaker premise, Langton seems to have been right. From there, it will be possible to develop the argument towards the subordination-thesis.

The subordination-thesis brings with it issues and questions about how exactly women are subordinated by pornography, how they are ranked, deprived of rights, and how it legitimates discrimination towards them. The topic is vast and complex and further ar-

gumentation for Langton's subordination thesis itself, therefore, seems difficult at this point. The more obvious way to take is to first defend Langton's silencing-thesis. As I established in the last chapter that pornography has at least some kind of authority, we now face the issue of showing how pornography can use it to silence women. The authority discussed in the last chapter is exactly the kind of authority that can be used to carry out speech acts with illocutionary force. This has to be the case, otherwise those speech acts would not be able to change the rules in the domain of sex. The question remains with what kind of illocutionary force pornography can speak.

Langton was quite certain that pornographic speech had to be exercitive if it were to silence. As a reminder: exercitive speech acts are illocutions that instantly will something into being like the phrase "You are fired" which results in someone losing their job. Langton states the following:

"If pornography sets up the rules in the language games of sex—if pornography is speech that determines the kind of speech there can be—then it is exercitive speech in Austin's sense, for it is in the class of speech that confers and removes rights and powers." (Langton 1993, 326)

To show that it has to be exercitive is in fact the more difficult task. If pornography's speech were exercitive, it would have a similar power as the government when enacting laws. That seems to be quite a leap. To find a solution, a closer look at a later paper by Langton (2009) is helpful. In this, Langton argues that there is verdictive speech that has an exercitive dimension (Langton 2009, 93–96). This concept might be the answer to resolving the issues in her silencing thesis.²³ At this point I would like to elaborate on verdictives a bit more and go back to an example which Austin introduced himself while explaining verdictives: he gave as a standard example the case of the umpire who calls "Out!" during a tennis match. The utterance is a verdict and states something about the game—he is giving a finding. The same happens with a judge who convicts a felon; after the utterance, it has been defined that the felon is now convicted of a crime. What is important about verdictives, though, is that they have a closer

connection to truth and falsity than other illocutionary forces; an umpire can also be wrong when he calls "Out!" (Austin 1962, 150–52).²⁴

Concentrating on one of Langton's examples, namely "Refusal", I argue that pornography makes a similar verdictive judgement by depicting rape scenarios and by defining a woman's illocutionary act of refusal as disabled. Similar to the umpire, pornography makes a judgement about women, ranks them as objects and grades their speech acts as having less value. Most of all, pornography states as a fact that even though a woman refuses, it can be interpreted as consent. This is not merely a portrayal of a specific sexual interaction; it is also a judgement about *how it is*, when talking about her speech acts. Furthermore, this verdictive judgment has an exercitive angle. Langton states the following about this:

"When someone authoritatively says 'this is how it is', that can be a verdictive judgement that something is so, and it is then 'taken as being that way': verdictives construct part of reality, in making something count as thus and so." (Langton 2009, 94)

A similar concept was also introduced by MacKinnon who argued that verdictives have an exercitive angle because the world changes to accommodate these verdictives if stated by an authority (MacKinnon 1987, 164). The question now still remains whether pornography possesses the authority to carry out verdictive speech acts. First of all, it is less strong of a statement than saying that pornography has the authority to carry out exercitive speech acts. By saying that authoritative verdictives are possible it means that pornography is able to make a judgement about society (in this case that women are or should be illocutionary silenced). There is no doubt that pornography does conduct speech acts that state the silencing of women—the only thing needed for them to be a verdictive judgement is enough authority in the relevant domain, namely that of sex. The answer to the question whether pornography has this authority can be found in chapter 4.1. Since there are no real alternatives to what *could* have the authority to carry out verdictive speech acts and since porn seems to be one of the shaping factors of how society talks about and perceives sex, it is not that far-fetched to say that pornography can make judgements and does construct our reality.

²³ Langton does not make this connection herself, she merely proposes it as an answer to Green's counterargument in which he stated that women are not under the jurisdiction of pornography, similarly as the gay Jew is not under the jurisdiction of the Church that sees homosexuality as a "objective disorder" (Green 1998, 293).

²⁴ I will come back to this connection to truth and falsity again.

But what about the close connection to truth and falsity that verdictives seems to have? One could argue that in the case of the umpire who calls “Out!” and is wrong in his judgement, we would then be justified not to believe the umpire, making his verdict worthless. Similarly, pornography’s verdictive illocutionary force could be deemed irrelevant. However, this does not seem to be the way in which the world works. If an umpire calls “Out!” there may very well be a discussion about his judgement. Maybe the player who now loses the game will try to show that the ball was in fact inside the line and will make an attempt at disproving the umpire. The issue he will face is that in the end, the umpire will still have the authority to make the verdict about the ball, unless some other authority overrides him. Maybe the umpire even was at fault. But unless someone with more authority than him gives a different verdict, his judgement will be what counts. In the case of pornography, there is *no other authority* to override its verdicts. Pornographic speech makes the verdictive judgement that women are silenced in sexual interactions and there is no greater authority to speak a different verdict. And due to the exercitive dimension of verdictives, the world tends to adjust to the standards of the powerful force in a domain.

4.3 Pornography has the authority to subordinate

As it was now established that pornography has the authority to carry out verdictive speech acts that silence women (which confirms Langton’s silencing-thesis) let us turn to the implications this has for Langton’s first thesis, namely the one arguing that pornography subordinates.

As I stated before, the reach of Langton’s first thesis is far vaster than that of the second one and far more difficult to support since it includes ideas like pornography depriving women of their rights and legitimating violence. To carry out speech acts which do these things, pornography would have to speak with an exercitive illocutionary force, or so thought Langton. However, if we look at what is happening with the verdictives I discussed in chapter 4.2, it seems that there might be a way in which to build up towards a subordination-thesis that way. First of all, as I established in 3.1, silencing is a subcategory of subordination. Since the verdictive judgements made by pornography end up in women being silenced, it seems that they have been deprived of a right, namely the right of speech. This is the exercitive dimension of verdictives. To illustrate this, let me return to an example

by Austin (1963, 152–53) and elaborate on it to show my point: let us assume a judge who convicts a felon for the crime of murder. Before the judge’s verdict, the man in question may very well have had the right to walk around freely and someone trying to imprison him would have done something illegitimate. The verdict now changes everything. It may deprive the man from the right to go about his daily life and make it legitimate that he now goes to prison. The verdict is therefore, although indirectly, able to deprive someone of a right and legitimate a certain behaviour towards them. The same thing may well happen in the case of pornography; through the verdictive speech acts pornography makes about women, some of their speech acts are disabled, some rights have been taken from them and sexual violence towards them becomes—although not by law but according to the rules in the game of sex—legitimate.

Conclusion

The goal of this thesis was to show that pornographic speech has the necessary authority to subordinate and silence women in the way in which Rae Langton has proposed in her paper “Speech Acts and Unspeakable Acts” (1993). After an analysis of both, Austin’s Speech Act Theory to elaborate on some important terms, and Langton’s theory to lay out the basics with which to work, I discussed some points which remained unclear in Langton’s line of argument. In an attempt to resolve these issues, I introduced a new category of speech acts into Langton’s theory, building on the idea that silencing is a subcategory of subordinating. After a discussion of some cases left out by Langton, the thesis ventured into the task of showing that pornography nowadays has all the authority it needs to establish dominance over the domain of sex. By including empirical data, it was illustrated that there is no alternative institution that is able to act with the necessary authority to overthrow pornography’s dominance. Drawing on Austin and Langton again, I elaborated that Langton may not have been right about pornographic speech having to be exercitive to be able to deprive women and legitimate sexual violence towards them. Instead, pornographic speech is *verdictive with exercitive dimensions*, therefore making judgements about *how it is*. It was established that verdictive judgements have the power to shape reality because society tends to adjust to the statements of the most powerful authority. In a last step, it was illustrated that if pornography has the illocutionary force to silence women, it also subordinates them, since silencing is a form of subordination.

What has to be said at this point is that, of course, this thesis is based on the assumption that pornography is in fact speech. This is something that could be criticized, if need be. Nonetheless, the question remains if pornography were not still responsible for the rules in the domain of sex and capable of defining what kind of speech acts there can be. What could also be analysed further is whether in countries different from the United States there are other institutions with potentially more authority than pornography (maybe in a country where sexual education is better developed and more including of topics like sexual violence and consent). However, with pornography's massive global influence, the chances do not seem high.

So let us, to give a more positive ending to this, turn towards what could be done now to resolve these issues. Langton proposes that her silencing-thesis may create a base for a censure of pornography. She also argues that pornography has to be confronted with more effective female speech (Langton 1993, 328). Concentrating more on pornography's authority, I would like to argue that there are two ways in which one could go about changing how pornography subordinates. One way would be to use pornography's authority and turn it into a force of good. This would have to include that pornography itself changes in order to avoid making verdictive judgements about women and not rank them as sex objects, deprive them of rights, silence their speech acts, and legitimate violence towards them. However, since the trend in pornography only seems to go further towards violence, a change like this seems quite optimistic. Another possible alternative could be the following: since pornography is at the moment the highest authority carrying out speech acts with verdictive illocutionary force, one solution would be to give something else said authority. This would mean that another institution would have to step up and establish dominance in the domain of sex. This could be done through sexual education; it would mean that the education in this field would have to grow immensely. Not only would its reach have to increase, but also the topics discussed would have to include issues like sexual violence, the notion of consent and pornography and its verdictive judgements itself. Whatever way our society chooses to go, the decision has to be made soon. With pornography's authority growing and therefore its influence gaining more momentum every year, there is really no time to be wasted.

References

- Austin, John L. 1962. *How To Do Things With Words. The William James Lectures delivered at Harvard University in 1955*. London: Oxford University Press.
- Bird, Alexander. 2002. "Illocutionary Silencing." *Pacific Philosophical Quarterly* 83 (1): 1–15.
- Bridges, Ana J., Robert Wosnitzer, Erica Scharrer, Chyng Sun and Rachael Liberman. 2010. "Aggression and Sexual Behaviour in Best-Selling Pornography Videos: A Content Analysis Update." *Violence Against Women* 16 (10): 1065–1085.
- Davidson, Donald. 1984. "Communication and Convention." In *Inquiries into Truth and Interpretation*, 265–280. Oxford: Clarendon Press.
- Donnerstein, Eduard, Daniel Linz and Steven Penrod. 1987. *The Question of Pornography: research findings and policy implications*. New York: The Free Press.
- Dworkin, Andrea. 1981. *Pornography: Men Possessing Women*. New York: Putnam.
- Dworkin, Ronald. 1991. "Two Concepts of Liberty." In *Isaiah Berlin: A Celebration*, edited by Edna and Avishai Margalit, 100–110. London: Hogarth Press.
- Green, Leslie. 1998. "Pornographizing, Subordinating, Silencing." In *Censorship and Silencing: Practices of Cultural Regulation*, edited by Robert Post, 285–311. Los Angeles: Getty Research Institute.
- Hornsby, Jennifer and Rae Langton. 1998. "Free Speech and Illocution." *Legal Theory* 4 (1): 21–37.
- Jacobson, Daniel. 1995. "Freedom of Speech Acts? A Response to Langton." *Philosophy & Public Affairs* 24 (1): 64–79.
- Langton, Rae. 1993. "Speech Acts and Unspeakable Acts." *Philosophy & Public Affairs* 22 (4): 293–330.
- Langton, Rae. 1998. "Subordination, Silence, and Pornography's Authority." In *Censorship and Silencing: Practices of Cultural Regulation*, edited by Robert Post, 261–284. J. Paul Getty Trust and Oxford University Press.
- Langton, Rae. 2009. *Sexual Solipsism. Philosophical Essays on Pornography and Objectification*. Oxford: Oxford University Press.
- Longworth, Guy. 2017. "John Langshaw Austin." *The Stanford Encyclopedia of Philosophy*. Accessed June 10, 2019. <https://plato.stanford.edu/entries/austin-jl/>
- MacKinnon, Catharine A. 1987. *Feminism Unmodified. Discourses on Life and Law*. Cambridge (MA): Harvard University Press.
- MacKinnon, Catharine A. 1993. *Only Words*. Cambridge (MA): Harvard University Press.
- Makin, David A. and Amber L. Morczek. 2015. "X Views and Counting: Interest in Rape-Oriented Pornography as Gendered Microaggression." *Journal of Interpersonal Violence* 31 (12): 2131–2155.
- Mikkola, Mari. 2011. "Illocution, Silencing and the Act of Refusal." *Pacific Philosophical Quarterly* 92 (3): 415–437.
- OED Online. 2019. "pornography, n." Accessed July 11, 2019. <https://www.oed.com/view/Entry/148012?redirectedFrom=pornography>
- Pornhub. 2017. „Pornhub's 2016 Year in Review." Accessed June 9, 2019. <https://www.pornhub.com/insights/2016-year-in-review>
- Pornhub. 2018. "Pornhub's 2017 Year in Review." Accessed June 9, 2019. <https://www.pornhub.com/insights/2017-year-in-review>
- Pornhub. 2018. "Pornhub's 2018 Year in Review." Accessed June 9, 2019. <https://www.pornhub.com/insights/2018-year-in-review>
- Rosenthal, Robert and Lenore Jacobson. 1968. *Pygmalion in the Classroom: Teacher Expectation and Pupil's Intellectual Development*. New York: Holt, Rinehart, and Winston.
- Sexuality Information and Education Council of the United States, SIECUS. 2019. "Trump's Teen Pregnancy Prevention Program Shift: A Timeline." Accessed June 9, 2019. <https://siecus.org/resources/trump-shifts-teen-pregnancy-prevention-program/>
- Stanger-Hall, Kathrin F. and David W. Hall. 2011. "Abstinence-Only Education and Teenage Pregnancy Rates: Why We Need Comprehensive Sex Education in the U.S." *PLoS ONE* 6 (10): e24658.
- Tsohatzidis, Savas L., ed. 1994. *Foundations of speech act theory: philosophical and linguistic perspectives*. London: Routledge.
- Vadas, Melinda. 1987. "A First Look at the Pornography/Civil Rights Ordinance: Could Pornography be the Subordination of Women?" *The Journal of Philosophy* 84 (9): 487–511.
- Wieland, Nellie. 2007. "Linguistic authority and convention in a speech act analysis of pornography." *Australasian Journal of Philosophy* 85 (3): 435–456.

Xenia Schmidli (23) hat soeben ihr Masterstudium in Philosophie und Englische Literaturen begonnen. Sie interessiert sich besonders für die Schnittstellen zwischen Sprachphilosophie, Ethik und der feministischen und politischen Philosophie.

Reassessing Credibility Excesses in Epistemic Injustice

Introduction

Miranda Fricker's seminal book *Epistemic Injustice: The Power and Ethics of Knowing* has been and continues to be highly influential in social epistemology. It introduced the concept of epistemic injustice to philosophy by way of a clear analysis and concise arguments. The book sparked interesting conversations, debates and even activism within and outside of the academy. Fricker presents two main ideas: Testimonial Injustice—which is where a speaker's testimony is ignored or not believed because of their belonging to a particular group—and hermeneutical injustice—which is where a group or individual is not provided with adequate conceptual tools to understand some aspect of their personal experience.

This paper will be concerned with the first kind of injustice. In a nutshell, systemic testimonial injustice, as elaborated by Fricker (2007), happens when a hearer assigns less credibility than deserved to a speaker due to unfair prejudices sustained because the speaker belongs to a particular group. The group is subject to prejudices that enduringly follow individuals throughout their lives and are at the root of a variety of injustices committed against that group. Members of the group are hence systematically not trusted when giving testimony. The injustice is that the individual is “wronged specifically in her capacity as a knower” (Fricker 2007, 20).

A crucial role in Fricker's account is given to credibility deficits. The injustice is committed because the speaker is afforded less credibility than deserved. The opposite—credibility excesses—is not seen by Fricker as systematically unjust. For many readers, this conclusion is not intuitive: One would expect that affording someone much more credibility than deserved is a problem we ought to act against. This is important because Fricker's book has been used by many feminists for analyzing social situations on the one hand, and motivating concrete action on the other hand. If her account of

epistemic injustice is used as such a tool, clearly it would be relevant if it overlooked certain phenomena. Let me note that this is a fairly technical issue.

In this paper, I ask what role credibility excesses play in our account of epistemic injustice. My answer is that there are clear examples of epistemic injustice arising from credibility excesses—and hence that Fricker is wrong in claiming there is none. I draw some examples from the current philosophical debate, analyze arguments from Jennifer Lackey and José Medina and come to the conclusion that yes, there is a place for credibility excesses in an account of epistemic justice.

In the following chapters, I will at first briefly summarize the few arguments Fricker presents against the role of credibility excesses in testimonial injustice. I will then outline the central arguments by Medina and Lackey and discuss the strengths of the arguments as they are presented in their papers. Both are arguing not only for a role for credibility excesses but also for the role of the social imaginary (Medina) and a distributive model of credibility (Lackey), respectively. I will present the arguments within those contexts and then try to find what unifies both of them in their assertion that there is a role in epistemic injustice for credibility excesses, and synthesize an argument for the role of epistemic injustice. Unfortunately, Fricker has not written any response to either paper. I will offer some possible objections from a Frickerian perspective. I will hence try to offer some strategies for Fricker. In conclusion, I will assess the status of credibility excesses and mention some open research questions.

1. Fricker's argument against a role for credibility excesses

As mentioned above, Fricker only briefly mentions credibility excesses in her groundbreaking book and

does not discuss it in later work.¹ Her remarks are embedded in the chapter introducing her central case for epistemic injustice². In her account, prejudicial credibility judgments can either be excesses or deficits (Fricker 2007, 17). In face to face interactions, hearers must make credibility judgments; excesses are usually advantageous for the speaker and deficits disadvantageous. She very briefly considers that intuitively, credibility excesses could be seen as unjust:

“Certainly there is a sense of ‘injustice’ that might naturally and quite properly be applied to cases of credibility excess, as when one might complain at the injustice of someone receiving unduly high credibility in what he said just because he spoke with a certain accent.” (Fricker 2007, 19)

Fricker unfortunately does not write more about this, but rather takes this to consider whether credibility is a distributive good. She mentions that the idea in the quote above might be expanded to apply a distributive model of epistemic justice—that someone has gotten “more than his fair share” (Fricker 2007, 19). She mentions that this might be a stretch, but her considerations against treating credibility justice as distributive is twofold:

First, there is no question of what a fair distribution would be: hearers ought to afford as much credibility to the speaker as the evidence for the truth of his testimony warrants. Second, distributive justice usually requires a finite and short supply, which leads to competition for them. For Fricker, credibility is not such a good. However, Fricker does not present an argument for this. The idea seems to be that there is no automatic reduction in credibility afforded to others when someone enjoys a credibility excess.³

This leads to Fricker’s central idea that epistemic injustice is “[to] wrong [someone] specifically in her capacity as a knower” (Fricker 2007, 20). Credibility excesses typically don’t wrong a speaker in this sense because they don’t “undermine, insult, or otherwise withhold a proper respect for the speaker *qua* subject

of knowledge” (ibid.). What she means here is that we should only count something as epistemic testimonial injustice if it wrongs the individual speaker.

Her reasoning here becomes clearer as she considers a possible counterexample. She invites us to imagine a member of an elite family who was systematically afforded a credibility excess to the effect of him possessing great epistemic arrogance, which in turn makes him unable to reach epistemic virtues (Fricker 2007, 20). Would this count as wronging him in his capacity as a knower?

For Fricker, there is a subtle distinction here because our member of the elite was a potential victim of *cumulative* epistemic injustice: There are no individual, token cases of epistemic injustice; rather, many instances taken together form the injustice cumulatively over a long timeframe. Fricker makes it clear that this does not fit in well with her account since she is mostly interested in analyzing individual, token instances (21).

Fricker’s argument against counting credibility excesses as epistemic injustice can be reconstructed thusly:

- (1) For something to count as epistemic injustice, it is necessary but not sufficient that a speaker is wronged in her capacity as a knower in token cases.
- (2) Credibility excesses aren’t wronging the speaker in token cases.
- (3) Therefore, credibility excesses are not instances of epistemic injustice.

Additionally, on the topic of whether credibility is a distributive good, Fricker argues that if it was distributive, there would be a case for unjust excesses. However, Fricker does not believe that credibility is a distributive good.

Unfortunately, Fricker does not have more to say about credibility excesses. It is worth noting that she is explicitly only interested in individual, token moments rather than a bigger context. She is also only looking at the dyadic relationship between speaker and hearer rather than the bigger social groups they belong to—something both Medina and Lackey pick up on.

2. Medina’s contextualist approach to epistemic injustice

José Medina’s paper defends a contextualist view (i.e. the context matters) of epistemic injustice coupled with what he calls a proportional view of epistemic injustice (cf. Medina 2011, 32). Unlike Lackey (see below), he does not think that credibility is a distributive good (19). Rather, he argues that credibility is a comparative,

1 She mentions this in an earlier paper and wrote as if credibility excesses were a form of epistemic injustice (Fricker 2007, 19, Footnote 14) but that she has since changed her view.

2 In what follows I will use “epistemic injustice” and “testimonial injustice” equally to mean the same. This is not quite precise; as mentioned in the introduction, epistemic injustice also encompasses hermeneutic injustices. However, in the literature considered, the terms are often used interchangeably, and in this paper, hermeneutic injustice is not discussed at all.

3 Medina (2011, 16) reads the passage in the same way as I do.

contrastive quality. He argues that credibility excesses can be cases of epistemic injustice if they contribute to a chain of unfair epistemic judgments (cf. p. 17). In his text, the contextualist and proportional parts of his argument are intertwined, but it is worth to look at them separately in order to better understand them.

Let us look at the contextualist arguments first. He criticizes Fricker for focusing only on individual moments (Medina 2011, 16) and for assuming that harm in testimonial exchanges is only done by the hearer to the speaker (18). Part of his analysis of Fricker is that in most cases, there is indeed no immediate harm from credibility excesses. However, he argues that

“The proper analysis of a testimonial exchange requires looking into what happens before and after the exchange, looking into what happens in other exchanges and in society as a whole. We have to follow the trajectories that intersect at (converging in and diverging from) the moment of exchange.” (Medina 2011, 17)

This context is constituted by “thick historicity and sociality” (Medina 2011, 17) which shows up in chains of testimonial exchanges. This is to be understood in opposition to Fricker: Where her analysis only looks at token instances where prejudices sustained over time converge, Medina wants us to look at the entire context and see how token instances are influenced by exchanges that happened before and what happens in the broader context of the testimonial exchange. He argues that there is epistemic harm happening that Fricker ignores.

One way this happens is by directly influencing the privileged speaker during a longer testimonial exchange. Fricker mentions that credibility excesses can lead to a form of cumulative epistemic injustice over the long term whereby the speaker cannot let go of epistemic vices. Medina (2011, 17) argues that in a sufficiently long dialogue, giving excessive credibility to one interlocutor can have immediate consequences by making her become arrogant and dogmatic and ignorant to criticism or even become a bully. It is not quite clear whether Medina sees this as an epistemic injustice or rather as handicapping the speaker. It appears that the harm in question is done towards the hearer and not towards the speaker, whereas Fricker would argue that the cumulative epistemic injustice she looks at is harming the speaker.

Another way are the effects on other interlocutors that may “feel intimidated by the speaker’s authoritative

voice, inhibited to express dissent or raise objections” (Medina 2011, 17–8) because the speaker has been assigned more credibility than deserved; or through hindering dissenting voices by undermining them through backing up a speaker’s authority. Fricker only looks at the effects of assigning less credibility than appropriate to the speaker; her account cannot capture such effects. Those two examples are meant to show that harm is done not only to recipients of a credibility deficit, but also to others around a person enjoying a credibility excess. Medina’s conclusion is that credibility is not assessed in the abstract and isolated from other social interactions and norms, but rather within a rich context (20).

Let us now look at Medina’s argument for the comparative quality of credibility. One example he mentions are inferiority complexes: When the hearer systematically gives excessive credibility to members of another group, she gives an “implicit credibility deficit” to people of her own group (Medina 2011, 18), effectively inhibiting her own voice and impacting her self-trust. The hearer misjudges the credibility of herself and her own group as compared to members of other groups *qua* belonging to that group.

Medina argues that not only in this example but in all credibility judgments, the recipient is not an isolated individual, but “clusters of subjects in particular networks and environments” (2011, 18), that is social groups. Giving too much credibility changes the dynamic of the testimonial exchange. This boils down to his assertion that speakers enjoying a credibility excess are “judged comparatively more worthy of epistemic trust than other subjects” (20): “implicitly, being judged credible to some degree is being regarded as more credible than others, less credible than others, and equally credible as others” (18). As mentioned before, this should not be read as a case of distributive justice. Rather, Medina suggests seeing it as comparative and proportional. It should also be mentioned that he does not think of credibility as a finite good under most circumstances (19).

He introduces the term *proportionality* to analyze this and it seems that he would think this term also captures Fricker’s account. Proportionality is to be understood as such that in a fair judgment, everyone gets the amount of credibility proportional to the “epistemic deserts and credentials” (Medina 2011, 20).

The argument relies on the following premise that is never made quite explicit: the harm is not (only) being done to the speaker but to everyone around her. Judging someone as more worthy of epistemic truth often

means assigning less credibility to others. Like Fricker, Medina does not believe that credibility is a limited, finite good (2011, 19), but unlike her, he argues for this relation between credibility excesses and deficits. What remains unclear is whether Medina thinks this applies to all sorts of situations, or to a narrower group of situations, namely those under oppression. Those considerations lead Medina to his definition of how excesses contribute to injustice:

“A credibility excess constitutes an epistemic injustice when and because it involves the undeserved treatment of an epistemic subject who receives comparatively more trust than other subjects would under the same conditions.” (Medina 2011, 20)

The excess is correlated to the credibility deficits others experience because

“an excessive attribution of it involves the privileged epistemic treatment of some (the members of the comparison class, i.e. those like the recipient) and the underprivileged epistemic treatment of others (the members of the contrast class, i.e. those unlike the recipient).” (ibid.)

Due to this relationship between excesses and deficits, Medina believes a credibility excess automatically leads to others being assigned a credibility deficit (ibid.). Fricker would disagree here, as Medina admits, because for her, it is not clear that all excesses automatically devalue someone else’s credibility, such as giving one’s teacher a credibility excess. Medina’s objection is that this is an empirical matter, but we have good empirical reasons from psychoanalysis and social psychology to assume that the correlation he posits does also exist in such situations—which implies that at least in some cases, Fricker is wrong and credibility excesses are harmful (Medina 2011, 21).

There is another interesting thought in Medina’s paper: his analysis of the Tom Robinson case from *To Kill a Mockingbird* (Lee 1960). Fricker analyses it only in terms of the credibility deficit against Tom Robinson (2007, 23–8). Medina believes this to be mistaken and identifies a credibility excess in favor of Mayella Ewell, the accuser (2011, 23–4)⁴. Not only is Tom Robinson *qua* his blackness assigned lower credibility, he

is also *qua* Mayella’s identity of a white woman seen as less credible than her. On his analysis, it is constitutive of the epistemic injustice at hand not only that he is not seen as credible, but also that she is seen as more credible.

To summarize, Medina’s three main points are that a) assigning excessive credibility has direct and indirect effects not only on the speaker, but also on the assigner and other listeners, b) credibility excesses are correlated with credibility deficits in other groups, c) those considerations lead him to claim that credibility excesses can be constitutive of epistemic injustice.

3. Lackey’s distributive justice

Contrary to both Faulkner and Medina, Jennifer Lackey conceives of epistemic justice as inherently distributive and finite, at least in some sense.⁵ She works out an evidentialist norm of credibility (meaning that the credibility assessment should be reflective of the evidence) in Fricker’s account which could be interpreted in two ways: As *categorically* (that speakers ought to match credibility to the evidence that the speaker is offering the truth) and *conditionally* (that speakers ought to do so if they make a credibility judgment). According to Lackey, the former is untenable because there are plenty of instances when we hear a testimony but do not assess the speaker’s credibility (2018, 147) and hence demands too much: We are not required to e.g. assess testimony about the weather we overhear in passing.

The conditional interpretation, which she then argues against, is however too permissive in her view. Lackey’s central example is one of a group of scientists. She invites us to imagine that the men don’t make any judgments about the testimony of the women due to unconscious sexism. By not even considering what the women are saying, they wrong them *qua* capacity as knowers. A second example is when an oppressed racial group is seen as so sub-human that their testimony is not even seen as candidates for true testimony, thereby not giving them a credibility deficit but nonetheless wronging them. Both examples are supposed to show that an entirely conditional reading can also not be right (Lackey 2018, 148–9).

How does her argumentation work then? First, let us look at her arguments for a distributive model of justice. She invites us to consider again the example of scientists, only this time imagine that the men are gi-

4 There is another dimension in his analysis, the role of the social imaginary. For the present purposes, those considerations are less interesting and therefore omitted.

5 Lackey also presents an interesting argument for epistemic justice having a normative part that is outside the scope of this paper.

ving the women the exact right amount of credibility and are also giving an excessive amount to the men in the group (Lackey 2018, 149). This example of a credibility excess has, according to Lackey, at least two implications: First, the male scientists with their inflated credibility might come to not believe anything the women say since in their eyes, it is impossible that someone with less ascribed credibility could know something they do not. Second, when the women's testimony conflicts with some prior beliefs, the men do not appropriately consider the evidence and change their belief, thereby wronging the women in their capacity as knowers. In this example, there is no credibility deficit, only excess. What the example shows is that even if someone is offered the fair amount of credibility, there can be cases of epistemic injustice stemming from credibility excesses—Lackey (2018, 150) makes clear that it is relevant that the excess happens relative to the other group. To be clear, this is in opposition to Fricker, who thinks that epistemic injustices only stem from deficits.

The wrong that the male scientists commit, in Lackey's (2018, 50–51) analysis, is that they do not form appropriate beliefs based on the female scientists' testimony. This happens not because of a credibility deficit but because of a credibility excess which "driv[es] a wedge between such a [credibility] judgment and the corresponding acceptance of testimony" (ibid.). Failure to be believed, on her account, is clearly harming the speaker epistemically and morally. She goes further and argues that failure to believe might be more harmful than credibility deficits, for example when Alex does not believe Bethany when she says she is suicidal. Alex's non-acceptance of the uttered sentiment as true may lead to very bad outcomes, no matter whether the credibility judgment was essentially correct or not.

The issue here is that credibility should usually be correlated to belief acceptance. Lackey's example of the scientists shows that this is not always the case; there are plenty of cases where belief acceptance is hindered by a credibility excess the hearer assigns to themselves or others—and readers are inclined to think such examples happen in real life too. Lackey calls this a "peer-excess testimonial injustice" (2018, 154). The point is that in such cases, when group A is given excessive credibility, for group B it is not important whether group B has been given a credibility deficit or the appropriate amount of credibility. The consequences for group B are the same, no matter the amount of credibility afforded to group B, precisely because group A was assigned an excess.

Another example of Lackey is credibility excesses assigned to experts. While she does not think that experts are generally given too much credibility, she highlights cases where expert testimony is given so much weight it trumps any other testimony unduly (Lackey 2018, 155–6). Her example of "expert excess" testimonial injustice are cases of Shaken Baby Syndrome where expert testimony in court was factually wrong but regarded so highly that no functional defense could be mounted. In her analysis, experts were given a credibility excess *qua* being experts, not due to the evidence.

What she highlights with the cases of peer-excess and expert-excess testimonial injustice is that "it matters both epistemically and morally not only how I judge you, but also how I judge you in relation to myself and other members of your community" (Lackey 2018, 157). The injustice occurs when due to prejudice, credibility is unequally and unfairly distributed within a context or group. This implies a relational understanding of credibility assessments: whether a credibility assessment is just depends on the assessments of the credibility of others in the same context or group. Following Fricker, this could be argued to occur in conjunction with identity prejudice.

Lackey's second thesis serves the purpose of understanding credibility as finite. Broadly speaking, she argues that all kinds of justice (including testimonial) require not only looking at what one person receives, but also what everyone around them gets (Lackey 2018, 158): If Christa gets 10 units and Daniel gets 20, and Christa receives less simply because she is a woman, that is a clear case of injustice, no matter the context. Regarding testimony, Lackey's prime example involves 'he said—she said' situations such as rape accusations: courts and the police can only believe one party (Lackey 2018, 159). In such cases, there is "only a limited amount of credibility to go around" (ibid.) which implies that regarding one party as credible is to regard the other as not credible. This is not the case in all disagreements but rather in situations where "someone's credibility is on the line" (ibid.). Her second example is expertise: we cannot regard everyone as an expert, and hence assessing someone as an expert implies that others are not.

Those two arguments lead Lackey to present what she calls the wide norm of credibility. It is formally stated as

"Wide Norm (of Credibility) [WN]: For every speaker, S_p , and hearer, H , if H makes credibility assessments of the relevant members of a con-

versational context or community, $S_1, \dots, S_n$, then H should match them to the evidence that H not only has but should have $S_1, \dots, S_n$ are offering the truth, and believe, disbelieve, or withhold accordingly.” (Lackey 2018, 163)⁶

Lackey believes that this should be the norm guiding our credibility assessments—rather than the ‘evidentialist’ norm asking us to match credibility assessment to the available evidence—that should be the basis for testimonial injustice. The implication is that a) we need to see credibility assessments in the wider context to rule out distributive problems such as in the case of the scientists and b) accept or reject beliefs accordingly.

To see this more clearly, consider the case of Lara, Hera, and Joaquin that Lackey analyses. In a nutshell, Lara was raped by her sister Hera’s fiancé, Joaquin. The police did not believe the sisters and indicted both for fraudulent reporting and obstructing justice (charges were eventually dropped). In Lackey’s analysis (2018, 164–5), not only was a credibility deficit at work when the head of police made sure his officers should not unconditionally trust their testimony. There is also a credibility excess when the authorities decided to believe Joaquin which led to the indictment of the sisters—explicitly so when the police admitted that they were trusting Joaquin and should have investigated further before charging the sisters. Once he was assessed as credible, in the minds of the police, the sisters became not credible. In other words, believing him implied not believing them. Lackey takes this as an example where credibility is both in limited supply and finite, since the police automatically assigned less credibility than deserved to one party while the other was believed.

There is another interesting argument in Lackey’s paper. She argues that even in token instances, a credibility excess can be harmful—she calls them context-specific credibility excesses. Her example is of a black man who, because of racial stereotypes, is assessed to be knowledgeable about drugs or guns. In assigning him to be an expert in stigmatized topics, the man can be insulted *qua* his capacity as a knower while also undermining respect towards him (Lackey 2018, 152–3). The man would be rightfully horrified to be seen as a drug or gun expert simply because of his identity. She identifies the possible harms of not getting certain jobs and writes that he is disrespected in his epistemic iden-

tity. This is an objection against Fricker’s assertion that credibility excesses assigned to a speaker are almost never harmful.

To summarize this chapter, Lackey’s main theses are a) that credibility excesses assigned to a speaker can be harmful, b) unjust situations can occur when too much credibility is assigned to one group/interlocutor even if the right amount is assigned to another group/interlocutor, c) there are situations where only a limited amount of credibility is available, and d) that those situations show that credibility is a distributive, finite good and credibility excesses can lead to injustices.

4. Evaluation of the arguments

4.1 Credibility as relational

Let us now assess the arguments by Medina and Lackey with regard to the relation between credibility excesses and deficits. First, it is noteworthy that their argumentation is fairly similar⁷. Both invite us to not only look at the one-on-one token cases of testimonial exchange, but also at the broader context. Both note that credibility excesses assigned to one group can have negative epistemic effects on another group. They differ in how exactly the injustice is constituted in such cases.

For Medina, credibility is assessed comparatively. To grant someone too much credibility correlates with granting others too little; excesses and deficits are intimately linked. Additionally, credibility excesses have negative impacts on speakers and hearers. In his view, which is similar to Fricker, credibility is proportional rather than distributive.

For Lackey, it works differently: her examples arguably show that in certain cases, we have situations where one group is assigned the fair amount of credibility and the other is given a credibility excess. In her example, this leads to the rejection of testimony of the first group (and hence wronging them in their capacity as knowers) even though they were afforded the deserved credibility. Furthermore, she uses examples to show how in some cases of contradicting testimony (notably formal inquiries such as police investigations), only one side can be assessed as credible, leading to the other side being seen as not credible. This implies that, in a sense, credibility is limited or finite.

What both authors show is that if we look further than the one-on-one token case of testimony, both credibility excesses and deficits become ‘linked’ in a sense. This

6 The ‘evidence that H should have’ part refers to Lackey’s argument for normative defeaters mentioned in footnote 3. It does not concern us for the present purposes.

7 This similarity is also noted by Lackey (2018, 157, 158)

is highly plausible and other examples can be found: Populist politicians telling their voters that other politicians or journalists are not credible are cases where (at least in some cases) the speaker is afforded a credibility excess which in turn make hearers assign not enough credibility to others. When the excessive credibility is wielded like a weapon against specific group—especially against groups identified by race, ethnicity or gender—we can clearly see how a testimonial injustice in the Frickerian sense is constituted.⁸

Another illuminating example where we can see, intuitively, a clear testimonial injustice constituted in yet another way is Rachel McKinnon's analysis of gaslighting (2017). The details need not concern us, but her example is of Victoria, a trans woman, complaining to Susan that James mispronounced⁹ her repeatedly. To be clear: the following is not McKinnon's own analysis but she mentions (in footnote 13) that it could be read this way. Susan is convinced that James would never do such a thing because, in her view, he is an ally to LGBTQI people and she has never seen him behave in such ways. This makes her assign a credibility excess to her own experience and a credibility deficit to Victoria. She voices her distrust of Victoria's experiences and discounts them on the basis that she, Susan, has never seen James misbehave in the alleged way.

In this case, credibility excess and deficit are clearly linked. Without trusting James not to do so on the basis of her own experience, Susan would not have discounted Victoria's testimony. This is consistent with both Medina who would analyze this as a case of disproportional credibility assessments, and with Lackey who would see this as a case of peer-excess testimonial injustice.

We need not commit to either's detailed view—i.e. we need not decide here whether credibility is to be understood proportionally or distributively—to see that we have a clear case where the speaker is not trusted because of the hearer-excess. It suffices to highlight the difference to Fricker's (2007, 15) analysis of Marge in *The Talented Mr. Ripley*: In the latter, Marge is not trusted precisely because she is a woman. The men

around her see her testimony through a prejudiced lens. In the case of Victoria introduced by McKinnon (2017), the situation is different in one respect: Susan need not bring any clear prejudices against Victoria to her assessment, she only needs to assign a credibility excess to herself. And even if she had such prejudices, it is the credibility excess she assigns herself that leads her to not believe Victoria. The case becomes a clear, paradigmatic case of epistemic injustice if Susan comes to the conclusion that Victoria is not trustworthy because of her identity as a trans woman. McKinnon (2017, 169) points out that trans women are often stereotyped as more emotional than rational and hence not credible.

Perhaps such stereotypes are at work in Susan's assessment. However, Susan could stereotype Victoria in that way and still believe her (remember how Lackey argues believing is more important than a fair credibility assessment) or at least take appropriate steps to remind James not to mispronoun, even if she is not fully convinced he is acting in such away. I need not be *certain* that Brad was rude towards Erica to remind him not to be rude, it suffices that I take it as credible even if I may believe Erica is overreacting. However, that is not so in Victoria's case. Susan does not form the appropriate belief and take appropriate actions because Susan assigns too much credibility to herself which leads her to misbelieve Victoria. Even if there are no stereotypes at work, it is still intimately linked to identity because the experience of being mispronounced is a decidedly trans experience; the pain and harassment that follows from it is not something cisgender people experience¹⁰. Victoria is clearly wronged in her capacity as a knower because Susan either does not afford her the right amount of credibility because of stereotyping, or because Susan affords her own experiences with James so much credibility that she cannot believe Victoria's testimony.

This is an example that both Lackey and Medina should be able to agree is a case of testimonial injustice constituted in part by a credibility excess. The example is informed by both Lackey's and Medina's accounts of credibility excesses. Note again that we need not commit to either of Medina's or Lackey's view on the proportional or distributive nature of credibility to see how this example highlights a role for credibility excesses.

⁸ Fricker (2017, 54) actually argues this should be distinguished from the epistemic injustice she has in mind: The populist is not misjudging the credibility of their opponents, they are intentionally manipulating the credibility judgments of their followers.

⁹ Mispronouncing is when someone uses the wrong pronouns on a trans person, e.g. referring to Victoria as "he". In McKinnon's words: "Mispronouncing is a serious offence for trans people: it's one of the most common forms of harassment that they face. In many jurisdictions where gender identity is included as a protected class with respect to harassment, mispronouncing counts as gender harassment." (2017, 168)

¹⁰ This is of course not to say cisgender people are never mispronounced. But arguably, McKinnon highlights the different experience of being mispronounced that trans people have.

4.2 Distributive vs. comparative models

I am emphasizing the point that we need not commit to a specific view of credibility because Fricker would have a much easier time arguing against such examples if they were purely seen through the lens of distributive credibility. Conversely, most of Lackey's excellent examples do not rely on a distributive model. The big difference in Lackey's view would be that at least in some situations of divergent testimony, credibility is finite because only one person can be trusted. I do not take it to be a central point of the argumentation; following Lackey in that respect certainly makes a distributive model more plausible. However, we need not commit to that to see the role of excesses for epistemic injustices. To see this, consider what Medina would say about Lackey's example of Lara, Hera and Joaquin. On his account, we would say that in the complex social network of the police, the two sisters, and Joaquin, Joaquin is given excess credibility which devalues Lara and Hera's testimony. Hence, I take it that we need not commit to either model.

4.3 Lackey's argument for harmful credibility excesses

There is another, direct counterargument from Lackey: The case of the black man seen as a drug or gun expert. Fricker (2007, 18) does discuss the possibility of disadvantageous credibility excesses; however, the example she considers is not linked to identity prejudice, but rather some elites being afforded too much credibility. It seems obvious that Fricker's examples do not constitute epistemic injustice, certainly not in token instances. However, Lackey's example of the black man being seen as a drug expert can easily be made stronger if we consider a black man not being given a job at a pharmacy because he is seen, due to the stereotype, as an expert on drugs. That very clearly constitutes an immediate harm because of a credibility excess that Fricker does not account for.

4.4 The summarized argument for a role for credibility excesses

Let me now try to formulate an argument informed by both Medina and Lackey that should be acceptable to both:

- (1) If credibility excesses are directly harming someone in their epistemic capacity due to them belonging to a specific disadvantaged group, they can be constitutive of an epistemic injustice,

- (2) Assigning someone/a group a credibility excess with regard to stigmatized issues is harming them in their capacity as knowers,
- (3) Assigning a speaker a credibility excess can harm hearers in their capacity as knowers,
- (4) In many situations, a credibility excess afforded to one person/group implies credibility deficits to others,
- (5) Affording one group/person a credibility excess can lead to injustice even if the other group/person is assigned the right amount of credibility if testimony from the latter group is not accepted,
- (6) Premises (2)–(5) are examples of possible situations where premise (1) is instantiated,
- (7) Therefore, credibility excesses can be constitutive of epistemic injustices.

The second premise is taken from Lackey's case of the stereotyped black man. The third premise is taken from Medina's example of hearers being intimidated by an authoritative speaker. The fourth is to be found in both Lackey and Medina, albeit in different variations (see above) and it is also found in the case of Victoria sketched above. Premise five is taken from Lackey's example of the scientists.

Note that the first premise and the conclusion are intentionally formulated to say that excesses *can be constitutive* to allow for both cases where the excess alone is the injustice and some where there is an intimate relationship between excess and deficit.

Also note that the first premise mentions that the harm is happening due to someone belonging to a disadvantaged group. This is also fairly vague. In both Medina and Lackey's writings, it is taken as a given that the harmed person or group is, for lack of a better word, disadvantaged. This is equal to Fricker's notion of tracker prejudices (Fricker 2007, 27). This qualification is necessary because neither Lackey nor Medina (nor me) discuss cases where a member of a privileged group—a male scientist, for example—can be harmed in their capacity as a knower such that it constitutes a systematic testimonial injustice.

5. Frickerian objections

What would Fricker say about those arguments? Unfortunately, as mentioned in the introduction and the first chapter, Fricker does not have a lot to say about credibility excesses. Hence, I will use the word "Frickerian" to denote a position Fricker might occupy herself.

Let us first look at how Fricker would argue against the idea that Lackey's case of the black man seen as

a drug expert constitutes an epistemic injustice. This goes against her assertion that “on the whole, excess will tend to be advantageous” (Fricker 2007, 18). A Frickerian should find it much harder to argue that the situation is not an injustice. It does not fall within the category of incidental testimonial injustice (cf. Fricker 2007, 29) since prejudice against the man tracks him through his life but for once results in an excess rather than a deficit. It could be argued that he is not wronged in his capacity as a knower which on Fricker’s account is necessary for an instance to be epistemically unjust. This might be justified by saying that the man does not suffer the same consequences Fricker identifies as following epistemic injustice—the man is not seen as epistemically not trustworthy (cf. Fricker 2007, 45) but rather seen as too trustworthy; there may be practical harm but no epistemic harm such as losing confidence in his belief (47), the man may be degraded, but not degraded *qua* knower (44). This is not a completely fleshed out objection—for reasons of space limits this could not be fully developed here—but it raises some doubts as to whether Fricker should accept Lackey’s case as an example of epistemic injustice. No doubt it is unjust towards the black man to think he knows about drugs or guns, and Lackey (2018, 153) correctly points out that it is insulting to the man both practically and epistemically, but it may come short of being an epistemic injustice.

Let us turn to the thesis that credibility excess and deficit are linked, namely premises 3–5 in the summarized argument sketched above, which is Medina’s idea that assigning credibility excesses to the speaker can harm hearers directly in being intimidated and dissuading objections or fostering inferiority complexes in the hearers. Again, a Frickerian might point out that this is unjust but not the kind of injustice Fricker has in mind—at least in the sense that it does not lead to the same wrongs Frickerian epistemic injustice does. Medina might object that in inhibiting dissent, the primary form of harm (wronging in the capacity as a knower) is committed in the sense that the hearers are implicitly told that their opinions are not valued, with all the harms Fricker identifies associated. It seems that the Frickerian will have a much harder time discounting this form of speaker-excess injustice: the effect that it has on other members of the epistemic community.

What about the assertion (made by both Medina and Lackey) that giving one person a credibility excess implies devaluing someone else’s credibility? This is at the heart of the dispute. Fricker would need to defend

her concept that epistemic injustice only happens to the speaker and the offense is done by the hearer in a token case. Epistemic exchanges around this token instance do not matter much; the context is only important insofar as the credibility was wrongly assessed due to an identity tracker prejudice. There may very well be good reasons for keeping her conception narrow. In a recent paper, she states her goal in the 2007 book as providing “something much more easily recognisable as making sense of the lived experience of injustice in how a person’s beliefs, reasons, and social interpretations were received by others, even conscientious well-meaning others” (Fricker 2017, 56) than the state of feminist theory at the time. This quote should serve to contextualize the underlying notion of her earlier work: a narrow conception useful in the real world, clearly separated from other forms of injustice. Her conception is narrow in the sense that it is the absence of bad intentions that lead to the credibility deficit and the associated wrongs (cf. p. 53–4).

In this lane, a Frickerian could object that Medina and Lackey unduly broaden the concept of testimonial epistemic injustice to include too much, since her examples of Tom Robinson and Marge Sherwood function well without considering credibility excess. I do not think this would be a strong objection. While it is true for those cases, Medina has a good counterargument: in a sufficiently long exchange, self- and peer-assessed credibility excesses may very well lead to immediately harmful epistemic vices in the same exchange as well as inhibiting dissenting voices (see above).

Imagine a situation where Sarah and Thomas are discussing politics. Both Sarah and Thomas give Thomas too much credibility. It is easy to imagine that Sarah will, over the course of a long discussion, start to hold back her own views because she thinks Thomas is more credible than herself. While that may or may not be a case of Frickerian injustice—it could be seen as such, see above—imagine now that Ursula joins the conversation and argues against Thomas. Since Sarah has assigned a credibility excess to Thomas, she will likely hold Ursula’s arguments in a lower regard *qua* holding Thomas in high regard. This is the interactive, comparative nature of credibility which Medina highlights. There are many other arguments Lackey and Medina put forth, but I take them to be too focused on their specific model of credibility to reiterate here. Additionally, the case of Victoria presented above should give the Frickerian pause—it seems impossible to analyze the case without looking at the credibility excess Susan assigns to herself. As mentioned above, it *may*

be that Susan does not fully believe Victoria due to stereotyping, but it is implausible to analyze the situation without taking credibility excesses into account.

A Frickerian may additionally object that the focus *should* only be on the token instances of testimony. The problem with that objection is that Fricker does not do that either. Her treatment of the Tom Robinson court case from *To Kill a Mockingbird* already has all the ingredients for a contextualized approach to epistemic injustice, something which Medina highlights (see above). Other than that, Fricker does not present a fully formed theory of credibility above and beyond what was said in the second chapter of this paper. Hence it is hard to form a Frickerian objection to premise 4.

Lastly, let us consider Lackey's case of the scientists. What can Frickerians say? If we treat this as a special case of the relational credibility, the same considerations as mentioned in the paragraphs above apply. However, if we take it just at face value, Lackey presents a great example where someone (in this case, the group of female scientists) is wronged in their capacity as knowers not because of a credibility deficit assigned to them, but because of a credibility excess assigned to others, a comparative group. I take this to be a blind spot in Fricker's argumentation. The problem for the Frickerian is that if the female scientists were assessed with a credibility deficit, we would have a clear, paradigmatic case of testimonial epistemic injustice. If we construct the example like Lackey does—where the female scientists are assessed fairly but the male scientists are given a credibility excess—it still seems like a clear case of epistemic injustice: the epistemic wrongs as identified by Fricker are present and they arise due to identity tracker prejudice. This analysis holds if we only look at token instances of testimony and we need not—as emphasized above—subscribe to a particular theory of credibility. The example works whether we are committed distributivists, see it as contrastive or consider us evidentialists like Fricker. On the evidentialist reading, the issue is that the credibility given to the female scientists matches the evidence, but the credibility assessment for the male scientists *does not*. In this sense, I cannot see a strong objection from a Frickerian view.

Conclusion

In the past chapters, I have described the arguments by Medina and Lackey with regard to a role for credibility excesses. I attempted to present them within their context and show that some of their examples and ar-

guments work outside of their own, specific views of what kind of justice model applies to credibility. I have sketched an argument informed by both of them and considered possible objections from a perspective inspired by Fricker—the Frickerian—since Fricker herself has not written anything about credibility excesses that goes further than the few pages dedicated to them in her seminal book on epistemic injustice.

Where this leaves us is that we have good examples of a role of credibility excesses in testimonial injustice if we look at more than only token cases. Both Medina and Lackey give us some good analyses of how such excesses may become very important once we look at the broader social context (and not all of that was discussed in this paper); and Lackey's case of the scientists gives us a good example where even in token cases, credibility excesses may lead to (at least on the surface) paradigmatic cases of epistemic injustice. This is a blind spot of Fricker and highlights one way in which credibility excesses can play a role—and be constitutive—of epistemic testimonial injustices.

Another strong example is Lackey's case of the black man assessed to be knowledgeable about drugs or guns because he is black. This is a clear example where one is given too much credibility because of identity tracker stereotypes leading to some form of injustice. As mentioned above, it is less clear that this amounts to *testimonial* injustice.

In terms of the summarized argument sketched above, Frickerians do not have good objections to the example of intimidated hearers (premise 3), nor the example of the scientists (premise 5). The more general case of credibility deficits almost necessarily arising for interlocutors when a credibility excess is afforded (premise 4) is harder to assess; we lack a fully fleshed-out theory of credibility by Fricker to argue explicitly against it. Additionally, Rachel McKinnon's example of Victoria highlights a situation where the analysis is incomplete without considering credibility excesses. All this is a good indication that there is indeed a role for credibility excesses.

What the research is currently lacking is a good account of *when* (and *why*) credibility excesses are constitutive of testimonial epistemic injustice which makes it hard to precisely assess the role of them. Fricker's book is well-regarded because it spells out all the different ways in which speaker-deficits are wronging the speaker; we are currently lacking an analysis of when (and how) excesses, whether assigned to the speaker, the hearer or other parties, are wronging someone above and beyond what Medina and Lackey (and a

small number of other authors) argue for. Neither of them presents the same kind of substantial analysis; rather, they use credibility excesses to argue for another point—Lackey for a distributive model of credibility, Medina for the comparative nature of credibility. To be clear: both are great contributions to the research, but they only highlight the research gap even further. It should be hoped that this research gap will be closed in the future. Fricker gives us the analysis of hearer-deficit and it stands to reason that our overarching theory of epistemic injustice will profit from a detailed account of credibility excesses—much more detailed than this paper could provide.

References

- Fricker, Miranda. 2007. *Epistemic injustice: Power and the ethics of knowing*. Oxford: Oxford University Press.
- Fricker, Miranda. 2017. "Evolving concepts of epistemic injustice". In *The Routledge handbook of epistemic injustice*, edited by Ian James Kidd, José Medina, Gaile Pohlhaus Jr., 53–60. Abingdon: Routledge.
- Lackey, Jennifer. 2018. "Credibility and the Distribution of Epistemic Goods". In *Believing in Accordance with the Evidence*, edited by Kevin McKain, 145–168. Cham: Springer.
- Lee, Harper. 1960. *To Kill a Mockingbird*. Philadelphia: J. P. Lippincott & Co.
- McKinnon, Rachel. 2017. "Allies behaving badly: Gaslighting as epistemic injustice". In *The Routledge handbook of epistemic injustice*, edited by Ian James Kidd, José Medina, Gaile Pohlhaus Jr., 167–174. Abingdon: Routledge.
- Medina, José. 2011. "The relevance of credibility excess in a proportional view of epistemic injustice: Differential epistemic authority and the social imaginary". *Social Epistemology* 25 (1): 15–35.

Dominik Fitze (30) schliesst dieses Jahr sein Masterstudium ab. Er interessiert sich insbesondere für Wissenschaftsphilosophie, Epistemologie sowie die Schnittmenge beider Felder mit Politik und Moral.