



Kant, Korsgaard und die Tiere

Sven Ellmers

Angenommen: 18. Januar 2023 / Online publiziert: 13. Februar 2023
© Der/die Autor(en) 2023

Zusammenfassung Christine M. Korsgaard ist eine der einflussreichsten Stimmen der Gegenwartsphilosophie. In ihren Arbeiten zur Moralphilosophie Kants verteidigt sie diese gegen zwei Einwände: erstens, dass der kategorische Imperativ auf externe Werte angewiesen ist, mithin als formales Testverfahren scheitert, und, zweitens, dass Kant empfindungsfähige Wesen, die kein Vernunftvermögen besitzen, aus seiner Ethik ausgeschlossen habe und ausschließen musste. Den ersten Einwand glaubt Korsgaard durch eine spezifische Interpretation des kantischen Widerspruchsverständnisses zumindest partiell zurückweisen zu können. Wie ich zeigen werde, unterschätzt Korsgaard allerdings die Probleme, die die Allgemeine Formel des kategorischen Imperativs mit sich bringt. Den zweiten Einwand wiederum glaubt Korsgaard in Form einer immanenten Kritik entkräften zu können: Kant selbst hätte auf Grundlage seines transzendentalen Arguments für die Selbstzweckformel zu dem Schluss kommen können, dass Tiere Zwecke an sich sind. Meine Thesen lauten, (a) dass Korsgaards Kant-Auslegung nicht vom Text gedeckt ist, (b) ihr *Regress-upon-the-conditions*-Argument nicht zur Begründung einer interpersonellen Moral taugt und es (c) auch tierethisch nicht zum gewünschten Ziel führt.

Schlüsselwörter Kant · Kategorischer Imperativ · Universalisierungsformel · Selbstzweckformel · Korsgaard · Tierethik

✉ Sven Ellmers
Institut für Philosophie, Fakultät IV, Carl von Ossietzky Universität Oldenburg, 26111 Oldenburg,
Deutschland
E-Mail: Sven.Ellmers@uol.de



Kant, Korsgaard and the animals

Abstract Christine M. Korsgaard is one of the most influential voices in contemporary philosophy. In her works on Kant's moral philosophy, she defends it against two objections: first, that the categorical imperative is dependent on external values, i.e. that it fails as a formal test procedure, and, second, that Kant excluded and had to exclude sentient beings, who do not possess the capacity to reason, from his ethics. Korsgaard believes she can at least partially reject the first objection through a specific interpretation of Kant's understanding of contradiction. As I will show, however, Korsgaard underestimates the problems caused by the Formula of Universal Law. The second objection, on the other hand, Korsgaard believes she can refute in the form of an immanent critique: Kant himself could have concluded that animals are ends in themselves on the basis of his transcendental argument for the Formula of Humanity. I argue that (a) Korsgaard's interpretation of Kant is not covered by the text, (b) her regress-upon-the-conditions argument is not suitable for justifying interpersonal morality, and (c) it does not lead to the desired goal in terms of animal ethics either.

Keywords Kant · Categorical Imperative · Formula of Universal Law · Formula of Humanity · Korsgaard · Animal Ethics

1 Der kategorische Imperativ als formales Testverfahren

Der kategorische Imperativ (im Folgenden: *KI*) ist ein formales Testverfahren für subjektive praktische Grundsätze, die Kant *Maximen*¹ nennt. Jeder Begriff dieser Definition ist erläuterungsbedürftig.

Dass *Maximen subjektive praktische Grundsätze* sind, bedeutet, dass ihre Quelle und Geltungsreichweite das (einzelne) Subjekt ist. Eine *Maxime* kann, muss aber nicht mit dem übereinstimmen, was die praktischen *Gesetze* verlangen, die für *jedes* sinnlich-vernünftige Wesen verbindlich sind.

Dass *Maximen subjektive praktische Grundsätze* sind, bedeutet wiederum, dass das Subjekt, das sie sich setzt, nicht instinktgesteuert ist. Vielmehr kann es eine reflexive Distanz zu seinen unmittelbaren Neigungen aufbauen, sie im Lichte neigungsabhängiger oder neigungsunabhängiger Maßstäbe bewerten und auf dieser Grundlage eine Entscheidung treffen – oder kurz: *handeln*.

Dass *Maximen subjektive praktische Grundsätze* sind, bedeutet schließlich zweierlei. Erstens bedeutet es, dass mit ihnen das Handeln einer *ausnahmslosen Regel* unterworfen werden soll. Eine *Maxime* bringt die feste Absicht zum Ausdruck, nicht nur heute, sondern auch in Zukunft auf eine bestimmte Weise zu agieren. Zweitens zielt der Begriff des Grundsatzes darauf, dass *Maximen* einen hohen Allgemeingrad aufweisen: Sie haben „mehrere praktische Regeln unter sich“ (AA V 19). So können unter die von Kant erwähnte *Maxime*, „keine Beleidigung ungerächt zu erdulden“ (ebd.), verschiedene (wenn auch nicht beliebige) konkrete Regeln fallen.

¹ Siehe hierzu Höffe (1977, 361–363).



Etwa: „Wenn mich mein Chef beleidigt, arbeite ich weniger“, „Wenn mich mein Frau beleidigt, rede ich drei Tage nicht mit ihr“ oder „Wenn mich mein Sohn beleidigt, bekommt er eine Woche kein Taschengeld.“

Soweit zum Gegenstand des Prüfverfahrens (ausführlicher: Albrecht 1994). Aber inwiefern ist der KI eine *formale* Prüfung? Formal ist sie insofern, als sie kein Kriterium verwendet, welches der moralischen Gesetzmäßigkeit *äußerlich* wäre. Nehmen wir das oberste Sittengesetz des Utilitaristen Bolzano: „Wähle unter allen dir möglichen Handlungen immer diejenige, die, alle Folgen erwogen, die Tugend und Glückseligkeit des Ganzen am meisten befördert.“ (RW 19) Mit dem Optimum an *Tugend und Glückseligkeit des Ganzen* werden Handlungen an einem *externen* Kriterium gemessen. Aber wie sollte es auch anders sein? Die Rede von einem moralischen Maßstab impliziert doch, Handlungen oder Handlungsregeln *an etwas* zu messen. Das oberste Sittengesetz muss also – wie *jegliches* Wollen – einen bestimmten Inhalt haben. Und das wusste selbstverständlich auch Kant (vgl. AA V 34). Da ihm zufolge jedoch alle denkbaren externen Maßstäbe nur *bedingt* gut sein können (vgl. hierzu AA IV 393–396) und folglich als *oberster* – das heißt *unbedingter* – moralischer Maßstab ausscheiden, kann dieser Inhalt nichts anderes als „die bloße Gesetzmäßigkeit überhaupt“ sein, „die allgemeine Gesetzmäßigkeit der Handlungen überhaupt“, „d. i. ich soll niemals anders verfahren als so, daß ich auch wollen könne, meine Maxime solle ein allgemeines Gesetz werde.“ (AA IV 402) Der KI fordert uns zu einem Gedankenexperiment auf. Wir sollen uns vorstellen, was passiert, wenn die fragliche Maxime nicht nur ein subjektiver Handlungsgrundsatz ist, sondern zugleich *zwingend* das Handeln *jedes* Vernunftwesens bestimmt. Anders gesagt nötigt uns der KI zu der Fiktion, dass die Gesetzmäßigkeit, die bisher lediglich subjektive Gültigkeit beanspruchte, nun auch notwendig und allgemein wie ein Naturgesetz wirkt: Niemand ist in der Lage, anders zu handeln, als es von der Maxime vorgesehen ist. Kurz: Der KI prüft subjektive Maximen auf ihre Tauglichkeit, subjektiv und (simultan) transsubjektiv zu gelten.

Bei Maximen, die gegen die moralische Pflicht verstoßen, führe dieser Universalisierungstest zu einem Widerspruch. Kant drückt dies auch so aus, dass die pflichtwidrige Maxime „sich selbst zerstören“ (AA IV 403) oder „sich selbst aufreiben“ (AA V 28) müsse. Zur Selbstzerstörung der Maxime kann es nach Kant auf zwei verschiedene Weisen kommen – und diese Differenzierung wiederum verwendet er für die Einteilung in vollkommene bzw. unvollkommene Pflichten: Eine Maxime kann entweder so beschaffen sein, dass sie widerspruchsfrei „nicht einmal als allgemeines Naturgesetz *gedacht* werden kann“ (AA IV 424). In diesem Fall verstößt sie gegen eine vollkommene Pflicht. Oder die Maxime ist von solcher Art, dass man sie sich zwar widerspruchsfrei als allgemeines Gesetz vorstellen, sie jedoch nicht *wollen* kann, „weil ein solcher Wille sich selbst widersprechen würde“ (ebd.). In diesem Fall verstößt sie gegen eine unvollkommene Pflicht.²

² Eine vollkommene Pflicht ist gemäß der GMS – zur Typologie der Pflichten in der MdS siehe Alves (2010) – eine solche, „die keine Ausnahme *zum Vortheil der Neigung*“ (ebd., Herv. von mir) gestattet. Eine unvollkommene Pflicht, so der naheliegende Schluss, ist eine Pflicht, die man in Abhängigkeit von seinen subjektiven Neigungen erfüllen oder nicht erfüllen kann. Kant – und darin folgt er den Naturrechtslehren von Grotius, Pufendorf und Thomasius – will einen Kernbereich der „strengen oder engeren (unnachlässlichen) Pflicht“ abgrenzen von der „weiteren (verdienstlichen) Pflicht“ (AA IV 424) oder „zufälligen



2 Korsgaards Replik auf Hegels Formalismuskritik

Die Zahl der Interpretationen, worin genau der Widerspruch beim Widerspruch-im-Denken (WD) und beim Widerspruch-im-Wollen (WW) besteht, ist Legion. Von Hegel stammt die Auffassung, der KI generiere nur einen Widerspruch, solange Kant, ohne dies transparent zu machen, materiale Zwecke voraussetze: „ein Widerspruch kann sich nur mit etwas ergeben, das ist, mit einem Inhalt, der als festes Prinzip zum Voraus zu Grunde liegt. In Beziehung auf ein solches ist erst eine Handlung entweder damit übereinstimmend, oder im Widerspruch.“ (Hegel 2009, 133, § 135 A) So führen die unmoralischen Maximen, zu morden oder zu stehlen, nur dann zu einem Widerspruch, solange unterstellt werde, dass es Eigentum und Menschen geben solle (vgl. ebd.). Kants Anspruch, mit dem KI ein rein formales Verfahren, das heißt ein Verfahren, das keine externen Werte voraussetzt, vorgelegt zu haben, lasse sich nicht einlösen: Aus der normativen Unbestimmtheit des KI lasse sich keine normative Bestimmtheit hervorzaubern.

Korsgaard hat in ihrem Aufsatz *Kant's Formula of Universal Law* (Korsgaard 1996a) zwei Interpretationen des kantischen Widerspruchskonzepts vorgelegt, die den Einwänden Hegels mehr oder weniger standhalten sollen. Eine Maxime, die den Diebstahl zulässt, gehört zur Gruppe unmoralischer Maximen, die, so Korsgaard, eine soziale Konvention voraussetzen (in diesem Fall: Eigentum) – und bei solchen Maximen könne bereits mit dem *logischen* Widerspruchsverständnis eine Kontradiktion nachgewiesen werden, ohne auf externe Werte zurückgreifen zu müssen. Sie verdeutlicht dies an Kants Depositum-Beispiel: Stelle man sich die Maxime, ein anvertrautes Gut (Depositum) nicht zurückzugeben, als allgemeines Gesetz vor, habe dies in der fiktiven Welt zur Folge, dass es kein Depositum gäbe, womit es wiederum unmöglich wäre, gemäß der Maxime zu handeln; die soziale Konvention, Wertgegenstände bei anderen Personen zu verwahren (ohne dass dadurch der Eigentumstitel auf sie überginge), würde schlicht aussterben, wenn permanent gegen die Pflicht verstoßen würde, die Wertgegenstände auf Verlangen zurückzugeben (vgl. ebd., 86). Der logische Widerspruch bestehe also darin, dass die Maxime als privater Handlungsgrundsatz eine soziale Konvention als existent voraussetzen muss, die gar nicht existieren kann, wenn jeder nach der Maxime handelte; denkt man die pflichtwidrige Maxime als allgemeines Gesetz, untergrabe sie die Bedingung ihrer Möglichkeit. Der Akteur, der sie sich zu eigen macht, will also einerseits auf eine bestimmte Weise handeln – z.B. unaufrichtig ein Versprechen abgeben, wenn es dem eigenen Vorteil dient – und muss deshalb auch die *Voraussetzung* dieser Handlung wollen – in diesem Fall: dass einem Versprechen Glauben geschenkt wird –, auf der anderen Seite führt seine (universalisiert gedachte) Handlung zu einer Situation, in der diese Voraussetzung seiner Handlung nicht mehr gegeben ist. Er will A und, „at the same time“ (ebd., 95), non-A.³ Anders gesagt kann der Akteur nicht logisch wi-

(verdienstlichen) Pflicht“ (AA IV 430). Verdienstlich – gebräuchlicher ist heute das Adjektiv *verdienstvoll* – ist eine lobenswerte / anerkennungswerte Handlung. Man rechnet sie dem Handelnden auch deshalb hoch an, weil man sie nicht zwingend von ihm erwarten konnte. Sie untersteht, wie Kant es schon in den *Träumen eines Geistesehers* ausdrückte, dem „schwächere[n] Gesetz der Gütigkeit“, nicht dem „starke[n] Gesetz der Schuldigkeit“ (AA II 335).

³ So auch schon Kersting (1983, 410ff.).



derspruchsfrei wollen, dass sein privater Grundsatz auch von allen anderen Akteuren befolgt wird.⁴ Denn wollte er, dass alle anderen Akteure die *Maxime* ebenfalls übernehmen, würde er *zugleich* ihre Ermöglichungsbedingung wollen und nicht wollen. Mithin scheitert eine pflichtwidrige *Maxime* an der *Simultanitätsbedingung*: Es ist unmöglich, sie im selben Augenblick als eigene Handlungsregel und als allgemeines Gesetz zu wollen.⁵ Mit dem Nachweis, dass man sie logisch konsistent nur unter der Voraussetzung als eigene Handlungsregel wollen kann, dass man sie zugleich nicht als allgemeines Gesetz will, liegt ihr *Parasitismus* offen zu Tage: Pflichtwidrige *Maximen* beruhen auf *einseitigen* oder *unfairen* Voraussetzungen (vgl. Korsgaard 1996a, 95, 100).

Obgleich Korsgaard dem logischen Widerspruchsverständnis attestiert, unmoralische *Maximen* treffsicher zu identifizieren, wenn sie die Existenz sozialer Konventionen ausnutzen, stößt es ihr zufolge an seine Grenzen bei *natürlichen Handlungen*, das heißt bei Handlungen, „die von den Naturgesetzen ermöglicht werden“ (T 169), insbesondere bei Akten der Gewalt: „The Logical Contradiction interpretation works well for immoral conventional actions, but it is not clear how it can handle immoral natural actions“. (Korsgaard 1996a, 85) So führe die offenkundig unmoralische *Maxime* einer Mutter, ihre Neugeborenen immer dann zu töten, wenn sie mehr als üblich in der Nacht schreien und ihr so den Schlaf rauben, zu keiner logischen Kontradiktion, wenn man sie verallgemeinere (ebd., 82).

Um die Anwendung der Universalisierungsformel auf natürliche Handlungen zu ermöglichen, schlägt Korsgaard eine andere Lesart des formalen Prüfverfahrens vor. Eine pflichtwidrige *Maxime* sei nicht logisch, sondern insofern *praktisch* widersprüchlich, als sie universalisiert gedacht genau die Bedingungen untergräbt, die erforderlich sind, *um den mit ihr verfolgten Zweck zu erfüllen*. Der Fokus liegt nicht länger darauf, dass etwas nicht mehr existiert, sondern darauf, dass die *Maxime* untauglich im Hinblick auf die ihr zugrundeliegende Zwecksetzung ist. Im Hinblick auf soziale Konventionen ist der Unterschied zwischen dem logischen und dem praktischen Selbstwiderspruch offenkundig nur marginal, denn die Tatsache, dass der Zweck der unmoralischen *Maxime* verfehlt wird, sobald die *Maxime* als allgemeines Gesetz gedacht wird, ist ja gerade darin begründet, dass die Konvention nicht länger existiert:

When we are dealing with an action that falls under a practice, the two views [the Logical and the Practical Contradiction Interpretation, S.E.] are readily confused, because the reason the action no longer works is *because* it no longer exists. But on the Practical Contradiction Interpretation it is the failure of efficacy, not the non-existence, that really matters. (ebd., 97)

Die bei sozialen Konventionen noch unerhebliche Akzentverschiebung erweist sich Korsgaard zufolge im Hinblick auf *Maximen*, die die Anwendung von Gewalt zulassen, jedoch als entscheidend. Sie demonstriert dies an einem Arbeitssuchenden, der, um eine Stelle zu ergattern, einen aussichtsreicheren Bewerber tötet. „Can this be universalized? Killing is a natural act, not a conventional one. We cannot say that

⁴ Vgl. Hoerster (2003, 109).

⁵ Darauf hat mit Nachdruck Kleingeld (2017) hingewiesen.



if this sort of action is abused the practice will die out, for that makes no sense whatever.“ (ebd., 98) Anders gesagt komme es bei dieser Maxime zu keinem logischen Selbstwiderspruch. Jedoch widerspreche sich die Maxime, sobald sie verallgemeinert wird, in *praktischer* Hinsicht: Der Zweck, eine Anstellung zu finden, werde konterkariert durch die Bedrohung, die nun von den anderen Bewerbern ausgeht. Zwar seien der Zweck, der in der Maxime enthalten ist (den Job zu bekommen), und die Konsequenz der universalisierten Maxime (meine Konkurrenten wollen mich töten) in diesem Fall nicht identisch, aber da das eigene Leben die Voraussetzung dafür ist, einer Arbeit nachzugehen, erweise sich die Maxime letztlich doch als in praktischer Hinsicht selbstwidersprüchlich (vgl. ebd., 99).

Allerdings muss Korsgaard selbst zugestehen, dass die um die Idee des transzendentalen Guts Sicherheit erweiterte Idee praktischer Widersprüchlichkeit nicht jede gewaltbefürwortende Maxime als unmoralisch ausschließt: „A harder kind of case would be somethink like killing for revenge, or out of hatred. In these cases it is not some enduring condition that the agent wants to achieve – he wants the immediate result – so the security consideration will not help us here.“ (ebd., 100) Man könnte ergänzen, dass auch andere Maximen, die Gewalt vorsehen, sich nicht als praktisch widersprüchlich erweisen. Die Mutter, die Korsgaard anführte, um zu zeigen, dass die logische Lesart des Selbstwiderspruchs es nicht vermag, Gewalt als unmoralisch einzustufen, handelt auch im Sinne der praktischen Kontradiktionslesart nicht selbstwidersprüchlich, denn die Tatsache, dass nicht nur sie, sondern auch alle anderen Mütter ihre übermäßig oft schreienden Neugeborenen töten, vereitelt nicht den von ihr verfolgten Zweck: Sie kann weiter morden und in Ruhe schlafen.

3 Warum Korsgaard die Probleme der Universalisierungsformel unterschätzt

Doch darin erschöpfen sich die Schwächen der Allgemeinen Formel des KI nicht. Bereits Bernard Bolzano, Wegbereiter der Analytischen Philosophie, hat anhand des Depositum-Beispiels nachgewiesen, dass der WD-Test selbst im Bereich der sozialen Konventionen nicht dazu taugt, die moralischen Pflichten zu begründen, die Kant mit ihm begründen zu können glaubt. Zur Erinnerung: Kant behauptete in der KpV, dass die verallgemeinerte Maxime, „daß jedermann ein Depositum ableugnen dürfe, dessen Niederlegung ihm niemand beweisen kann“, zur Konsequenz hätte, „daß es gar kein *Depositum* gäbe“ (AA V 27). Bolzano bezweifelt mit guten Gründen, dass die Verallgemeinerung der Maxime die von Kant behauptete Konsequenz hat, denn „so ist es ja nicht völlig wahr, was hier gesagt wird, daß es gar keine anvertrauten Güter mehr gäbe, wenn es nicht Pflicht wäre, sie wieder zurückzustellen“ (RW 36). Kant verwechsle die *Erlaubnis*, anvertraute Güter nicht zurückzugeben, mit der *Pflicht*, sie nicht zurückzugeben. „Denn wenn es nur nicht Pflicht wäre, sie vorzuenthalten: so könnte ja Mancher noch hoffen, das Gut [...] wieder zurück zu erhalten“ (ebd.). Als allgemeines Gesetz gedacht untergräbt die Maxime nicht zwingend die Bedingung ihrer Möglichkeit: Weil sie nicht verlangt, dem rechtmäßigen Eigentümer das anvertraute Gut vorzuenthalten, könnte es weiterhin Deposita geben –

womit keine (logische oder praktische) Kontradiktion vorliegt und das Vorenthalten des Depositums moralisch zulässig wäre.

Ähnlich ließe sich auch im Hinblick auf andere Maximen argumentieren, die Kant zufolge einen Widerspruch im Denken generieren. Nehmen wir die Situation, die in *Über ein vermeintes Recht aus Menschenliebe zu lügen* geschildert wird. Ein Hausbesitzer wird von einem Mordlustigen gefragt, ob sich bei ihm eine Person versteckt hält, die er verfolgt. Soll der Hausbesitzer ehrlich sein oder soll er lügen, um den Verfolgten zu schützen? Laut Kant ist es „ein heiliges, unbedingt gebietendes, durch keine Convenienzen einzuschränkendes Vernunftgebot: in allen Erklärungen wahrhaft (ehrl)ich zu sein“ (AA VIII 427), denn eine Maxime, die Ausnahmen vom Wahrhaftigkeitsgebot vorsieht, sei eine, die sich, als allgemeines Gesetz gedacht, „geradezu selbst widerspricht“ (AA VIII 430). Die vollkommene Pflicht zur Wahrhaftigkeit ergibt sich also durch den Nachweis, dass eine gegen sie verstoßende Maxime – eine Maxime, die unter bestimmten Bedingungen eine Lüge zulässt – zu einer fiktiven Welt führt, die man nicht denken könne. Nur warum soll man sie nicht denken können? Sie gleicht eher der Welt, in der wir tatsächlich leben: Man kann sich nicht sicher sein, ob der andere ehrlich ist oder sich in einer Ausnahmesituation wähnt, die ihm eine Lüge gestattet. Keineswegs ist die fiktive Welt aber eine solche, in der ausnahmslos gelogen wird und folglich „Aussagen (Declarationen) überhaupt keinen Glauben finden“ (AA VIII 426).

Oder nehmen wir ein Beispiel aus der GMS: „wenn ich mich in Geldnoth zu sein glaube, so will ich Geld borgen und versprechen es zu bezahlen, ob ich gleich weiß, es werde niemals geschehen.“ (AA VI 422) Nach Kant würde die Universalisierung dieser Maxime „das Versprechen [...] selbst unmöglich machen“ (ebd.). Dieter Birnbacher (2013, 148) fragt zu Recht, wie Kant zu seiner Schlussfolgerung gelangt. Die Maxime lautet ja nicht „Ich bezahle geliehenes Geld niemals zurück“, sondern sie knüpft den Betrug an die Bedingung (subjektiv empfundener) materieller Not. Dass Kant aus dem Notlügner *de facto* einen notorischen Lügner macht, zeigt Birnbacher zufolge, dass Kant den Universalisierungstest auf eine problematische Weise denkt: In ihm wird nicht nur (zurecht) die Beschränkung auf das einzelne Handlungssubjekt aufgehoben, sondern es wird auch der in der Maxime noch angegebene *Situationstyp* getilgt. Vom KI werde dies jedoch nicht ausdrücklich verlangt. Auch der Verweis auf die Naturgesetzformel helfe hier nicht weiter, denn „Naturgesetze [...] enthalten in der Regel Spezifikationen der Bedingungen, unter denen sich Naturobjekte in bestimmter Weise verhalten und sind nicht in dem Sinne bedingungslos oder ‚unbedingt‘, dass sie auf beliebige Situationen zutreffen.“ (ebd.) Tilge man die in der Maxime angegebene Bedingung jedoch nicht, komme es auch zu keiner Kontradiktion.

Aber nehmen wir an, die zu prüfenden Maximen eröffnen weder Handlungsspielräume (wie im Depositum-Beispiel) noch enthalten sie Einschränkungen auf gewisse Situationstypen (wie in den beiden anderen Beispielen). Ist der WD-Test wenigstens für diese Teilklasse subjektiver Handlungsgrundsätze ein zuverlässiges Prüfverfahren? Testen wir die Maximen, die bereits Hegel (2009, 118, § 135 A) heranzog: „Ich eigne mir Gegenstände an, ohne auf rechtliche Eigentumstitel Rücksicht zu nehmen“ und „Ich töte wahllos“. Auf den ersten Blick generieren beide Handlungsgrundsätze genau den Widerspruch, den Kant im Sinn hatte. Einerseits implizieren die Vorsätze,



zu stehlen oder wahllos zu morden, das Wollen der Ermöglichungsbedingung des Diebstahls (rechtliches Eigentum) und des Mords (lebende Menschen), andererseits führt das Wollen der Maxime als allgemeines Gesetz dazu, dass es kein rechtliches Eigentum und damit keinen Diebstahl bzw. keine lebende Menschen und damit keinen Mord geben kann. Jedoch kommt es zu diesem Widerspruch nur solange, wie der handelnden Person eine bestimmte Intention unterstellt wird – die Intention, von Voraussetzungen profitieren zu wollen, ohne selbst etwas zu deren Existenz beizutragen (Parasitismus-Intention). Aber wäre es nicht auch denkbar, dass die Maxime genau das bewirken soll, was sie als allgemeines Gesetz gedacht bewirken würde? Warum soll die Maxime nicht ein Mittel zum Zweck sein können (zum Zweck der Aufhebung der bestehenden Eigentumsordnung oder der Auslöschung der Menschheit)? In diesem Fall nimmt sich eine Person vor, auf eine gewisse Weise zu handeln, aber sie will gerade *nicht* die Ermöglichungsbedingung dieser Handlung, denn die Existenz des Eigentums bzw. die Existenz der Menschheit gilt ihr als das mittels der Maxime zu lösende Problem: Sie will gegen die Eigentumsordnung verstoßen, *weil* dies die Eigentumsordnung unterminiert – und je mehr Menschen sich die Maxime zu eigen machen, umso sicherer wird das Ziel erreicht. Der KI als Detektor für einseitige Ausnahmen⁶ kann in diesem Fall nicht anschlagen, weil sich die private und die zum Gesetz erhobene Maxime *komplementär* zueinander verhalten.

Fazit: Selbst man den WD-Test auf Maximen begrenzt, die dem Akteur weder Handlungsspielräume eröffnen noch einschränken auf gewisse Situationstypen enthalten, erfüllt der WD-Test seine Aufgabe nicht zuverlässig. Zuverlässig funktioniert er lediglich als Filter für strikten/unbedingten Parasitismus. Unsittliche nicht-parasitäre Maximen passieren den Filter hingegen mühelos.⁷

4 Welche Kritik an der Universalisierungsformel ihren Gegenstand verfehlt und welche nicht

Viele Kant-Interpreten geben die Probleme, die mit dem WD- und WW-Test verbunden sind, unumwunden zu. So schreibt Wood (2006, 345), die Universalisierungs-

⁶ „What Kant wants to rule out with the Categorical Imperative is that one aims to make an exception. In the case of duties towards others, one wants them to adhere to a law, but wants to make an exception for oneself. In the case of duties towards self, one wants to make an exception to a law that one regards as valid.“ (Sensen 2014, 172).

⁷ Schönecker/Wood treiben die Kritik noch etwas weiter: Selbst wenn der WD-Test sämtliche pflichtwidrigen Maximen erfolgreich aussonderte, würde er immer noch nicht das leisten, was man von ihm erwartet: dass er uns sagt, welche Maximen wir uns setzen *sollen*. Tatsächlich gebe er jedoch nur darüber Auskunft, welche Maximen wir uns geben *dürfen* oder *nicht* geben dürfen. Selbst wenn sich beispielsweise nachweisen ließe, dass die Maxime, lügenhaft zu versprechen, eine Kontradiktion nach sich zieht, folge daraus „nicht, uns ihr kontradiktorisches Gegenteil als Maxime zu eigen zu machen (also etwa überhaupt Versprechen zu geben). Denn aus der Tatsache, daß wir einer bestimmten Maxime *nicht* folgen dürfen, folgt ja nicht direkt, daß wir einer anderen Maxime folgen *müssen*.“ (Schönecker/Wood 2011, 140f.) Das ist einerseits richtig. Aus dem Wissen, kein Versprechen lügenhaft abgeben zu dürfen, folgt nicht, dass man sich die gegenteilige Maxime setzen muss – denn als Alternative käme immer noch der Grundsatz in Betracht, gar keine Versprechen abzugeben. Andererseits: Was Kant sehr wohl gezeigt hätte, wäre, dass ein Versprechen, *so man denn eines abgeben will*, nur in der Absicht abgegeben werden darf, es auch zu halten. Praktisch bedeutungslos wäre die kantische Ethik also nicht.



formel wäre „radically defective as a general moral criterion, since it systematically yields both false positives and false negatives“. Sven Nyholm (2015) hat dem widersprochen. Die Universalisierungsformel führe regelmäßig zu Problemen, weil viele Interpreten sich nicht darüber im Klaren seien, was sie überhaupt testen: Viele Maximen, die den Kontradiktionstest nicht bestehen, seien gar keine Maximen im Sinne Kants, sondern situationsspezifische Handlungsregeln/-pläne. Es sei kein Wunder, dass das Testverfahren nicht funktioniere, wenn es ständig mit Input gefüttert werde, auf den es nicht ausgelegt ist. Nyholms Kritik trifft etwas. Tatsächlich ist die von Dietrichson (1969, 188) diskutierte Maxime, ein Neugeborenes zu töten, wenn es weniger als 6 Pfund wiegt, keine Maxime, sondern eine (überaus) konkrete Handlungsregel. Das gleiche gilt für den Vorsatz, den die schlafbegierige Mutter bei Korsgaard verfolgt – oder für den von Parfit (2011, 289) diskutierten Vorsatz, allen weiß gekleideten Frauen die Geldbörse zu stehlen, die Erdbeeren essen, während sie die letzten Seiten von Spinozas *Ethik* lesen. Maximen, so der berechtigte Einwand Nyholms, sind keine konkreten Handlungsregeln, sondern *leitende Grundmuster für solche Regeln*.

Sind damit auch die Einwände obsolet, die ich im vorherigen Abschnitt diskutiert habe? Überprüfen wir dies anhand der von Benjamin Constant konstruierten Situation, in der ein Mörder Auskunft erbittet. Ich hatte argumentiert, dass die Person, die vor der Wahl steht, wahrhaftig zu sein oder zur Notlüge zu greifen, einer Maxime folgen kann, die eine Ausnahme vom Gebot der Wahrhaftigkeit vorsieht, ohne dass dies zu einem Widerspruch im Denken führen würde: Weil in der fiktiven Welt nicht permanent, sondern nur in Ausnahmefällen gelogen wird, käme es nicht zu der von Kant imaginierten Situation, in der niemand mehr dem anderen Glauben schenkt und somit dem privaten Handlungsgrundsatz der Boden entzogen wäre. Folgt man Nyholms Spur, liegt der Verdacht nahe, dass ich keine Maxime, sondern eine situationsspezifische Handlungsregel überprüft habe. In diesem Fall: „Steht ein Mörder vor meiner Tür, der eine Person verfolgt, der ich Unterschlupf gewährt habe, so greife ich zur Lüge.“ Ist möglicherweise jeder Vorsatz, der in diesem Fall etwas anderes vorsieht als bedingungslos zu lügen oder bedingungslos wahrhaftig zu sein, eine situationsspezifische Handlungsregel – und damit keine Maxime? Wäre dem so, hätte dies eine merkwürdige Konsequenz: Abgesehen von den beiden direkt entgegengesetzten Geboten – unbedingte Wahrhaftigkeit vs. unbedingte Lüge – gäbe es keine einzige wahrhaftigkeitsbezogene Maxime, die sich mittels der Universalisierungsformel testen ließe. Der praktische Nutzen des KI wäre gleich null – denn wer käme schon auf die Idee, dass die universelle Lüge ethisch geboten sein könnte?

Nun ist es aber problemlos möglich, sich eine Handlungsregel vorzustellen, die eine Ausnahme vom Wahrhaftigkeitsgebot zulässt, ohne dass sie deshalb aufhörte, ein leitendes Grundmuster für situationsspezifische Handlungsregeln zu sein. Nehmen wir folgendes Beispiel: „Ich sehe es als meine vorrangige moralische Verpflichtung an, Menschen in existentieller Not zu helfen.“ Anders als in den von Nyholm monierten Beispielen gibt diese Handlungsregel keine spezifischen Mittel für „particular ends“ (Nyholm 2015, 290) an. Der angegebene Zweck – nämlich der moralischen Pflicht Genüge zu tun – ist ebenso grundsätzlicher Natur wie der Zweck der Selbstliebe in dem von Kant diskutierten Vorsatz, sich ab einem gewissen Punkt das Leben zu nehmen (vgl. AA IV 421 f.). Es handelt sich also um eine *Maxi-*



me, um einen subjektiven praktischen Grundsatz. Dieser schließt andere moralische Verpflichtungen nicht aus. Freunde und Familie mögen der Person, die sich den Grundsatz gibt, überaus wichtig sein. Und auch dem Gebot der Wahrhaftigkeit mag sie gewissenhaft Folge leisten wollen. Nur sind diese Verpflichtungen im Kollisionsfall nachrangig. Genau dies reflektieren die situationspezifischen Regeln, die sich zu dieser *Maxime* denken lassen: „Ich engagiere mich mehrmals in der Woche bei der örtlichen Tafel, auch wenn ich dann kaum noch Zeit für meine Freunde habe“. Oder: „Ich gewähre Geflüchteten, die abgeschoben werden sollen, obwohl sie in ihrer Heimat in großer Gefahr wären, heimlich bei mir Unterschlupf. Sollte ich in Verdacht geraten, streite ich dies ab.“ Obwohl die *Maxime* implizit (und in Gestalt ihrer situationspezifischen Handlungsregel auch explizit) eine Notlüge zulässt, ist nicht zu sehen, aus welchem Grund sie den *WD-Test* nicht bestehen sollte.

Aber überprüfen wir auch das. Möglicherweise ließe sich argumentieren, dass die *Maxime* den *WD-Test* nicht besteht, weil sie, als allgemeines Gesetz gedacht, insofern die Bedingung ihrer Möglichkeit unterminiert, als es existentielle Not in dieser fiktiven Welt höchstens noch für gewisse Zeiträume geben könnte – denn sollte jemand in existentielle Not geraten, ist sofort eine (andere) helfende Hand zur Stelle. Gegen den Versuch, auf diese Weise eine Selbsterstörung der *Maxime* nachzuweisen, spricht jedoch schon, worauf bereits Bolzano hinwies (vgl. RW 36): Jede *Maxime* enthält implizit oder explizit eine *Anwendungsbedingung*, in diesem Fall die Bedingung, dass sich ein Mensch in existentieller Not befindet. Integriert man sie in die *Maxime*, so lautet diese: „*Erhalte ich davon Kenntnis, dass sich ein Mensch in existentieller Not befindet, sehe ich es als meine vorrangige moralische Pflicht an, ihm zu helfen.*“ Der Umstand, dass die *Anwendungsbedingung* nicht (oder nur vorübergehend) erfüllt ist, erzeugt in meinem Beispiel jedoch keinen Widerspruch. Zu einem Widerspruch zwischen dem privaten und dem allgemein verfolgten Handlungsgrundsatz käme es nur, wenn die Person, die sich die *Maxime* gibt, Motive verfolgt, die Kant dem Prinzip der Selbstliebe zugeordnet hätte: z.B. das Motiv, anderen Menschen helfen zu wollen, weil man Freude dabei empfindet oder weil man für seine Hilfsbereitschaft von anderen gelobt werden will. Die Hilfe wäre dann ein notwendiges Mittel zum Zweck. Bei unserer Person ist dies jedoch nicht der Fall. Sie will helfen, weil sie dies als ihre *moralische Pflicht* erachtet. Gibt es aber keine existentielle Not, gibt es auch keine vorrangige moralische Pflicht, der sie nachkommen müsste. Anders als für denjenigen, der helfen möchte, weil er gelobt werden will, ist dies für sie jedoch kein Problem. Im Gegenteil, das Ausbleiben der moralischen Nötigung, zu helfen, ermöglicht es ihr, *die* moralischen Pflichten erfüllen zu können, die sie ansonsten hätte hintanstellen müssen.

Fazit: Wie Nyholm zeigt, gehen viele Kant-Kritiken fehl, weil sie *Maximen* mit beliebigen Handlungsplänen identifizieren. Die Reichweite dieses Einwands ist jedoch begrenzt, denn selbst wenn man, wie Kant (in §1 KpV), unter *Maximen* handlungsleitende Grundmuster für situationspezifische Handlungsregeln versteht, führt die Universalisierungsformel des KI immer noch nicht zu den von Kant gewünschten Ergebnissen.

5 Die Selbstzweckformel

Wie bereits erwähnt wird die Ansicht, dass der Universalisierungstest *für sich genommen* ein unzuverlässiges Verfahren wäre, von vielen Kant-Interpreten geteilt. Das Urteil falle jedoch deutlich positiver aus, sobald man die Universalisierungsformel im Lichte der Menschheit-als-Selbstzweck-Formel deute (so u.a. Schönecker/Wood 2011, 127). Zwar sei es richtig, dass Kant sowohl in der GMS (AA IV 402 und 420) als auch der KpV (AA V 27) aus dem Begriff eines kategorischen Imperativs dessen Universalisierungsformel ableitet. Ebenso sei es nicht von der Hand zu weisen, dass Kant die Universalisierungsformel als „allgemeine Formel“ (AA IV 436) bezeichnet und den anderen Versionen des KI explizit vorzieht: „Man thut aber besser, wenn man in der sittlichen Beurtheilung immer nach der strengen Methode verfährt und die allgemeine Formel des kategorischen Imperativs zum Grunde legt“ (ebd.). Ungeachtet dieser Einschätzung durch den Autor sehen viele Kant-Interpreten in der Selbstzweckformel die bedeutendere Formel, vor deren Hintergrund alle anderen Formeln gelesen werden sollten. Tatsächlich hatte Kant selbst betont, dass beide Formeln denselben Sachverhalt zum Ausdruck bringen:

Denn daß ich meine Maxime im Gebrauche der Mittel zu jedem Zwecke auf die Bedingung ihrer Allgemeingültigkeit als eines Gesetzes für jedes Subject einschränken soll [Universalisierungsformel, S.E.], sagt eben so viel, als: das Subject der Zwecke, d. i. das vernünftige Wesen selbst, muß niemals bloß als Mittel, sondern als oberste einschränkende Bedingung im Gebrauche aller Mittel, d. i. jederzeit zugleich als Zweck, allen Maximen der Handlungen zum Grunde gelegt werden [Selbstzweckformel, S.E.]. (AA IV 438)

Davon ausgehend nimmt die Kantverteidigung nun eine ganz andere Richtung. Anstatt sich an dem Nachweis zu versuchen, dass der KI als formales Prüfkriterium belastbar ist, wird (auch von Korsgaard) die These vertreten, dass Kants Moralphilosophie gar nicht so inhaltslos und leer sei, wie Hegel glaubte. Vielmehr fuße sie auf der Idee der menschlichen *Würde*. Was versteht Kant unter diesem Begriff? Bevor ich auf Korsgaards Interpretation eingehe, werde ich zunächst mein eigenes Verständnis darlegen.

In der GMS grenzt Kant die Würde vom Preis ab (vgl. AA IV 434f.). Alles, was einen Preis hat, kann durch ein Äquivalent ersetzt werden. Waren haben einen *relativen* Wert, denn sie wechseln den Besitzer, sobald der Käufer die Geldsumme aufbringt, die der Verkäufer veranschlagt hat. Hingegen kann kein Geld der Welt etwas ersetzen, das einen *absoluten* Wert hat: Man brächte ihm nicht die Achtung entgegen, die man ihm entgegenzubringen hat, wenn man unter gewissen Bedingungen (bei einem bestimmten Preis) bereit wäre, es gegen ein anderes Gut einzutauschen. Denselben Sinn hat es, wenn Kant vom *inneren* Wert spricht: Dass etwas einen inneren Wert oder Würde hat, ist ein anderer Ausdruck dafür, dass etwas *unbedingt* zu respektieren ist.

Grund der Würde ist das Vermögen zur sittlichen Autonomie. Die „Würde der Menschheit besteht“ in der „Fähigkeit, allgemein gesetzgebend, obgleich mit dem Beding, eben dieser Gesetzgebung zugleich selbst unterworfen zu sein“ (AA IV 440). Der Mensch ist ein Wesen, das sich die Gesetze seines Handelns *selbst* auferlegen



kann. Wie ist das zu verstehen? Zunächst verfügen Menschen mit der empirisch-praktischen Vernunft über das Vermögen, ihr Handeln gewissen Grundsätzen zu unterwerfen, weil diese ihr subjektives Wohlbefinden fördern. „Der Mensch ist ein bedürftiges Wesen, so fern er zur Sinnenwelt gehört, und so fern hat seine Vernunft allerdings einen nicht abzulehnenden Auftrag von Seiten der Sinnlichkeit, sich um das Interesse derselben zu bekümmern und sich praktische Maximen, auch in Absicht auf die Glückseligkeit [...], zu machen.“ (AA V 61) Nehmen wir eine Person, die sich entschließt, sich ein Aquarium zuzulegen, weil sie den Anblick von Fischen beruhigend findet. Die Person trifft die Entscheidung in dem Wissen, dass es erforderlich sein wird, regelmäßig das Wasser zu wechseln, die Filteranlage zu reinigen und die Fische zu füttern, das heißt, sie unterstellt ihr zukünftiges Handeln einem Grundsatz, der die Befolgung gewisser Regeln impliziert. Einerseits bedeutet das nicht, dass man sich ungeachtet der Umstände sklavisch an die Regeln halten müsste. Ist der Aquarianer schwer erkältet, kann er den Wasserwechsel auch mal eine Woche aussetzen. Andererseits ist damit ebenfalls nicht gemeint, dass das Wasser nur dann gewechselt werden müsste, wenn ihm gerade der Sinn danach steht. Wenn er sich das vorbehalte, hätte er sich letztlich *nicht* dazu entschieden, ein Aquarium zu betreiben; streng genommen hätte er sich zu gar nichts entschieden, sondern folgte jeden Tag aufs Neue dem momentanen Verlangen (vgl. T 187). Um was für eine Form der Selbstverpflichtung handelt es sich in diesem Beispiel? Ist sie bereits hinreichend für das, was Kant Autonomie nennt? Nein, das ist sie nicht, denn die Entscheidung für das Aquarium steht noch im Dienste einer individuellen Neigung, das heißt der Handlungsgrundsatz ist seinem *Inhalt* nach nicht *selbstgegeben*, sondern *vorgegeben*. Allein, Kant zufolge ist der Mensch

doch nicht so ganz Thier, um gegen alles, was Vernunft für sich selbst sagt, gleichgültig zu sein und diese bloß zum Werkzeuge der Befriedigung seines Bedürfnisses als Sinnenwesens zu gebrauchen. Denn im Werthe über die bloße Thierheit erhebt ihn das gar nicht, daß er Vernunft hat, wenn sie ihm nur zum Behuf desjenigen dienen soll, was bei Thieren der Instinct verrichtet (AA V 61).

Die Vernunft hat nicht nur eine empirisch-praktische Seite, denn sonst kompensierte sie lediglich die Instinktarmut des Menschen. Sie hat auch eine *neigungsunabhängige*, von Kant als *rein* bezeichnete Seite: Menschen sind dazu veranlagt, ihren Willen „unabhängig von dem Naturgesetz der Erscheinungen, nämlich dem Gesetz der Causalität“ (AA V 29), zu bestimmen, das heißt *frei* zu sein – frei zu sein nicht nur im Hinblick auf die äußeren Handlungen, sondern auch im Hinblick auf den Willen, der die äußeren Handlungen veranlasst. Die Freiheit des Willens wiederum hat Konsequenzen für seinen Inhalt. Der Gedanke, der für die Klassische Deutsche Philosophie zentral werden wird, lautet, dass Freiheit als indeterminierte Selbstgesetzgebungskompetenz zwangsläufig eine selbstreflexive Struktur und damit einen notwendigen (nicht-kontingenten) Inhalt hat. Der freie Wille will sich selbst – und das bedeutet: die Freiheit aller Vernunftwesen.⁸ Damit steht hinter den verschiede-

⁸ Dass die Wertschätzung der Freiheit der Grund unserer moralischen Verpflichtung ist, steht auch im Zentrum der Kant-Auslegung von Paul Guyer (1998, 2000).

nen Formeln des KI stets derselbe Grundgedanke: Praktische Regeln, die um der Lust willen befolgt werden, sind aufgrund ihres subjektiv-kontingenten Charakters nicht dazu geeignet, ein praktisches Gesetz abzugeben, das heißt sie können nicht *jedes* vernünftige Wesen auf „ein und dasselbe Object“ (AA V 28) verpflichten. Positiv formuliert besteht die Aufgabe praktischer Gesetze darin, den für jedes vernünftige Wesen zustimmungsfähigen Kernbereich der Verbindlichkeit zu etablieren, der differente Lebensentwürfe allererst ermöglicht; im Bereich des Handelns sollen sie genau das leisten, was im Bereich der Erscheinungen die Naturgesetze vollbringen: Sie sollen „alles einstimmig“ (ebd.) machen (vgl. Willaschek 1992, 195 ff.). Indem sie die Gefahren ausschalten, die aus dem Pluralismus der materialen Bestimmungsgründe des Willens noch für den Pluralismus selbst erwachsen, bilden sie das tragende Fundament des *Reichs der Zwecke*:

Ich verstehe aber unter einem Reiche die systematische Verbindung verschiedener vernünftiger Wesen durch gemeinschaftliche Gesetze. Weil nun Gesetze die Zwecke ihrer allgemeinen Gültigkeit nach bestimmen, so wird, wenn man von dem persönlichen Unterschiede vernünftiger Wesen, imgleichen allem Inhalte ihrer Privat Zwecke abstrahirt, als Ganzes aller Zwecke (sowohl der vernünftigen Wesen als Zwecke an sich, als auch der eigenen Zwecke, die ein jedes sich selbst setzen mag), in systematischer Verknüpfung, d. i. ein Reich der Zwecke, gedacht werden können, welches nach obigen Principien möglich ist. (AA IV 433)

Es hat den Anschein, dass die Einwände, die auf den Formalismus der kantischen Ethik zielen, damit entkräftet sind. Bolzano et al., so der sich aufdrängende Eindruck, fokussieren Probleme, die sich nur stellen, weil sie die Universalisierungsformel dekontextualisieren. Sie erweisen sich damit als gute Logiker aber schlechte Interpreten. Ein solcher Schluss wäre jedoch voreilig. Denn erstens ist die Frage, *wie* sich praktische Gesetze identifizieren lassen, immer noch nicht beantwortet. Wenn sich der WD-Test und der WW-Test als untauglich erwiesen haben, um den Würde-Gedanken zu operationalisieren, welcher Test tritt dann an ihre Stelle? Die Selbstzweckformel fordert uns zwar auf, die Würde, das heißt den absoluten Wert vernünftiger Wesen zu respektieren, aber es bleibt unklar, *wie* wir Maximen erkennen können, die diesen Respekt vermissen lassen (vgl. Schönecker/Wood 2011, 153). Zweitens sind nicht alle Kant-Kritiker schlechte Interpreten, denn viele von ihnen sind sich durchaus bewusst, dass die Freiheitsidee das organisierende Zentrum der kantischen Ethik ist. Was aber haben sie dann noch an ihr auszusetzen?

6 Bolzanos Kritik der Selbstzweckformel

Man könnte vermuten, dass sie die Selbstzweckformel und Reich-der-Zwecke-Formel ablehnen, weil sie die Freiheit des Willens für eine Illusion halten. Und tatsächlich ist dies nicht selten der Fall. Der bereits erwähnte Logiker Bolzano ist jedoch ein gutes Beispiel dafür, dass dies nicht für alle Kant-Kritiker gilt. Denn wie schon Eduard Winter (1932) feststellte, ist die deterministische Position, zu der der späte Bolzano neigte (wie in der *Athanasia*), das Resultat einer Entwicklung, die unter



dem Einfluss von Steinbart und Kant mit einem entschiedenen Indeterminismus begann:

Ich theile die Summe meiner Vorstellungen oder der Objekte meines Bewußtseins in zwei Klassen: bei einigen fühle ich mich *leidend* und auf diese wend' ich den Begriff von Ursache und Wirkung an [...]. Andre fühle ich als *mein Werk*, nenne sie meine Handlungen, und denke mich als den alleinigen Grund derselben, so daß kein Grund vorhanden, der erst *mich bestimmt*, daß ich so handle [...]. Und dieß ist eben der Begriff von Freyheit, welcher also den Satz vom Grunde geradezu aufhebt. Wo der Eine gilt (in der *Erscheinungswelt*) gilt nicht der andre (in den *Handlungen*). (Bolzano 2009, 106 f.)

Freiheit bestehe in dem Vermögen, zwischen den *Wünschen* des Glückseligkeitstriebs und dem *Sollen* als einem Urteil der Vernunft zu wählen (vgl. RW 34). Der dabei entstehende Wille könne nicht das (determinierte) Resultat einander entgegengewirkender kausaler Kräfte sein, weil das Urteil der Vernunft und die Wünsche des Glückseligkeitstribs *inkommensurabel* seien. Es handele sich nicht um gleichartige Kräfte, von denen sich sagen ließe, die eine sei größer als die andere gewesen und habe sich *deshalb* durchgesetzt (vgl. Bolzano 1981, 55 f., 120 f.). Die Parallelen zu Kants Freiheitsbegriff der *Religionsschrift* sind unverkennbar. Ob wir dem moralischen Gesetz folgen oder nach dem Prinzip der Selbstliebe handeln, hatte Kant erklärt, sei Ausdruck eines ursprünglichen Wahlakts, der sich kausal nicht erklären lasse: „absolute Spontaneität der Willkür“ (AA VIII 24).

Dass Bolzano im *Lehrbuch der Religionswissenschaft* die Selbstzweckformel des KI nicht als oberstes Sittengesetz akzeptiert, ist also nicht in unterschiedlichen Ansichten zum Thema Willensfreiheit begründet. Beide argumentieren hier vielmehr deckungsgleich. Warum übernimmt er dann aber nicht die Selbstzweckformel als obersten Moralgrundsatz? Das Hauptproblem ist für ihn der Ausschluss aller Lebewesen, die nicht vernunftbegabt sind. Der „gesunde Menschenverstand“ verlange,

unter zwei Handlungen, die einen gleich vortheilhaften Einfluß auf vernünftige Wesen haben, deren die Eine aber einem Thiere Schmerzen macht, während dieß bei der andern nicht der Fall ist, die letztere vorziehen sollen. Es ist also gewiß unrichtig zu behaupten, daß man lebendige Wesen, wenn sie vernunftlos sind, als bloße Mittel gebrauchen dürfe, und nicht verpflichtet sey, auch in ihnen einen Zweck an sich anzuerkennen. (RW 37)

Tatsächlich zieht Kant – wie vor ihm bereits die Stoiker und Theologen wie Thomas von Aquin – eine scharfe Trennlinie zwischen Tier und Mensch. Alleinige Quelle der Würde sei das Vernunftvermögen, weshalb die instinktgesteuerten Tiere auch keinen absoluten, sondern einen nur relativen, das heißt auf menschliche Zwecke bezogenen Wert hätten:

Die Wesen, deren Dasein zwar nicht auf unserm Willen, sondern der Natur beruht, haben dennoch, wenn sie vernunftlose Wesen sind, nur einen relativen Werth, als Mittel, und heißen daher Sachen, dagegen vernünftige Wesen Personen genannt werden, weil ihre Natur sie schon als Zwecke an sich selbst, d. i.

als etwas, das nicht bloß als Mittel gebraucht werden darf, auszeichnet, mithin so fern alle Willkür einschränkt. (AA IV 428)

Den Gedanken, dass Tiere der menschlichen Willkür unterworfenen Sachen sind, wiederholt Kant in der *Anthropologie in pragmatischer Hinsicht*:

Daß der Mensch in seiner Vorstellung das Ich haben kann, erhebt ihn unendlich über alle andere auf Erden lebende Wesen. Dadurch ist er eine Person und vermöge der Einheit des Bewußtseins bei allen Veränderungen, die ihm zustoßen mögen, eine und dieselbe Person, d. i. ein von Sachen, dergleichen die vernunftlosen Thiere sind, mit denen man nach Belieben schalten und walten kann, durch Rang und Würde ganz unterschiedenes Wesen (AA VII 127).

Wie insbesondere in *Mutmaßlicher Anfang der Menschengeschichte* deutlich wird, hängt für Kant die Idee, Menschen seien Zwecke an sich, eng mit der Idee zusammen, Tiere seien bloße Mittel. Den letzten Schritt in der Entwicklung des Vernunftvermögens habe der Mensch erst getan, als er „begriff, er sei eigentlich der Zweck der Natur“, als er seines „Vorrechtes“ gegenüber den Tieren inne wurde, als er sie „nicht mehr als seine Mitgenossen an der Schöpfung, sondern als seinem Willen überlassene Mittel und Werkzeuge zur Erreichung seiner beliebigen Absichten ansah. Diese Vorstellung schließt [...] den Gedanken seines Gegensatzes ein: daß er so etwas zu keinem Menschen sagen dürfe“ (AA VIII 114). Wie der antike Demos auf der Abgrenzung von den sprachbegabten Werkzeugen (Sklassen) beruhte, beruht die moralische Gemeinschaft auf der Abgrenzung von den empfindungsfähigen Werkzeugen.

In anderen Schriften schlägt Kant einen leicht anderen Ton an. Der Mensch, so argumentiert er in der *MdS*, habe Tieren gegenüber keine moralischen Pflichten, weil moralische Pflichten das Resultat einer reziproken Nötigung seien: Der Mensch hat „sonst keine Pflicht, als bloß gegen den Menschen (sich selbst oder einen anderen); denn seine Pflicht gegen irgend ein Subject ist die moralische Nötigung durch dieses seinen Willen.“ (AA VI 442) Trotzdem solle er keine Tiere quälen oder durch Arbeit zu Tode schinden (vgl. AA VI 433). Diese Einschränkung der Willkür erfolgt jedoch nicht um der Tiere, sondern um der Menschheit willen: „Allein weil alle Thiere nur als Mittel da sind [...], der Mensch aber der Zweck ist [...], so haben wir gegen die Thiere unmittelbar keine Pflichten, sondern die Pflichten gegen die Thiere sind indirecte Pflichten gegen die Menschheit.“ (Kant 2004, 345) Die Pflicht, gutherzig zu den Tieren zu sein, ergibt sich nicht etwa aus ihrer Würde – die haben sie nach Kant nicht –, sondern aus zwei Gründen, die logisch voneinander unabhängig sind. Erstens – ich nenne es das Selbstachtungsgesetz – ist es *unter unserer* Würde, Tiere schlecht zu behandeln. „Es erniedrigt uns selbst eine Handlung, womit wir Thiere martern oder Noth leiden lassen“ (AA XXVII 710).⁹ Und zweitens – das Verrohungsgesetz – schlägt sich unser Verhältnis zu den Tieren in Charakterdispositionen nieder und beeinflusst somit das zwischenmenschliche Verhältnis positiv oder negativ: Um zu vermeiden, dass Menschen grausam zu Menschen sind, sollen Menschen nicht grausam zu Tieren sein (vgl. Kant 2004, 345 f.; AA VI 443).

⁹ Siehe hierzu Baranzke (2002, 199 ff.).



Insofern man nach Kant überhaupt von Pflichten gegenüber Tieren sprechen kann, handelt es sich lediglich um abgeleitete Pflichten. „Selbst die Dankbarkeit für lang geleistete Dienste eines alten Pferdes oder Hundes (gleich als ob sie Hausgenossen wären) gehört indirect zur Pflicht des Menschen, nämlich in Ansehung dieser Thiere, direct aber betrachtet ist sie immer nur Pflicht des Menschen gegen sich selbst.“ (AA VI 443) Der Mensch hat Pflichten *in Ansehung* oder *hinsichtlich* der Tiere, aber *gegenüber* den Tieren hat er keinerlei Pflichten; es gibt Schuldigkeiten *in Bezug auf* Tiere, man schuldet ihnen jedoch nichts.

7 Korsgaards kantisches Plädoyer für Tierrechte. Der Grundgedanke

Kants Gedanke, Tiere seien weniger wichtig als Menschen, wird nicht nur von Bolzano (und zeitgenössischen Utilitaristen wie Peter Singer), sondern auch von Korsgaard abgelehnt. Tiere seien ebenfalls Selbstzwecke. Zu behaupten, sie seien *an sich* nicht so wichtig wie Menschen, ergebe „jenseits eines antiquierten teleologischen Weltbegriffs keinerlei Sinn“ (T 28). Warum? Dass etwas wichtig ist, setzt voraus, dass es wichtig für *jemanden* ist. Wenn die Bedeutung von etwas immer an ein Wesen gebunden ist, dem es etwas bedeutet, stellt sich die Frage: „Für wen sollen menschliche Wesen wichtiger sein? Für das Universum? Für Gott?“ (T 25)

Eine Antwort auf diese Frage, die kein teleologisches Weltbild voraussetzt, könnte lauten, dass Menschen vielleicht nicht *an sich* wichtiger sind (man also keine objektive Rangordnung der irdischen Geschöpfe vertritt), sondern dass der Mensch dem Menschen schlicht mehr bedeutet als das Tier – so wie ein Familienangehöriger den meisten mehr bedeutet als ein Fremder. Für Korsgaard ist diese emotionale Verbundenheit solange kein Problem, wie daraus nicht der Schluss gezogen wird, Tieren einen moralischen Status absprechen zu können. Die moralische Verpflichtung hänge nämlich nicht davon ab, ob man jemandem zugetan ist oder nicht (sie hänge vielmehr davon ab, ob das Geschöpf eine gewisse Eigenschaft aufweist). Allenfalls, so Korsgaard, könnte man argumentieren, der Mensch sei bei elementaren Konflikten – wie der Übertragung einer gefährlichen Krankheit durch Ratten – berechtigt, die eigene Spezies aus Verbundenheit zu bevorzugen (T 24). Aber selbst wenn man die Rechtmäßigkeit einer solchen Form der Parteinahme unter extremen Umständen zugesteht, müsse man noch längst nicht der Ansicht sein, dass Tiere *an sich* weniger wichtig als Menschen sind (und man deshalb mit ihnen tun und lassen könne, was man wolle). „Dinge haben eine Bedeutung für Geschöpfe, aber die Geschöpfe selber stehen in keiner absoluten Rangordnung ihrer Bedeutung.“ (T 25)

Warum sind für Korsgaard Tiere ein Selbstzweck, andere organische Lebewesen (wie Pflanzen) jedoch nicht? Ein lebender Organismus unterscheidet sich von unbelebten Gegenständen durch seine *rekursive* Organisation: Die durch die spezifische Relation ihrer Bestandteile bestimmte organische Einheit strebt fortwährend danach, ihre sich verbrauchenden Bestandteile durch Transformation interner oder externer Elemente zu erneuern, um so den individuellen Fortbestand und den Fortbestand der Gattung zu sichern. „Lebende Organismen kümmern sich um sich selbst“ (T 37), das heißt anders als unbelebte Naturdinge oder Werkzeuge betreiben sie

Autopoiesis.¹⁰ Tiere sind jedoch besondere Organismen. Für Tiere lassen sich nicht nur Bedingungen angeben, die für sie gut sind, um gut zu funktionieren – und das bedeutet: sich und die Gattung zu reproduzieren –, sondern Tiere streben nach dem, was im funktionalen Sinne gut für sie ist, auf eine bewusste, mit angenehmen und unangenehmen Empfindungen verbundene Weise. Sprich, das funktional Gute ist auch gut *für sie*: „Für die meisten Tiere sind Essen und Trinken, sexuelle Aktivität, körperliche Tätigkeiten, Spielen mit dem Nachwuchs oder miteinander, Wärme, Behaglichkeit und Geselligkeit positive Güter.“ (T 39) Gefahren, Schmerzen und Tod hingegen meiden sie. Ihr Leben ist für sie (wie für den Menschen) ein Selbstzweck, ein Zweck an sich.¹¹

Daraus folgt wiederum nicht, dass es keine gewichtigen Unterschiede zwischen Mensch und Tier gebe. Korsgaard zufolge zeichnet sich der Mensch durch Rationalität aus, wobei sie Rationalität von Intelligenz abgrenzt. Intelligent zu sein bedeutet, „aus Erfahrungen zu lernen und Probleme durch Überlegung zu lösen“ (T 59). Obwohl viele Tiere intelligent sind, sind sie keine rationalen Wesen, denn:

- I. Tiere haben zwar Gründe für das, was sie tun – in dem Sinne, dass sie auf Basis ihrer aktuellen Wahrnehmung, ihrer Erfahrungen und ihrer Instinkte agieren –, aber sie sind sich der Gründe nicht *bewusst*, weshalb sie außerstande sind, die Gründe zu *bewerten* und anders zu agieren, als sie es tun. So löst der Anblick eines Löwen bei einer Antilope Angst aus und verleitet sie zur Flucht, aber weder denkt sie darüber nach, dass die Angst der Beweggrund ihrer Flucht ist (vgl. T 61), noch überlegt sie, ob es richtig ist, sich in dieser Situation von der Angst leiten zu lassen. Menschen hingegen verfügen über die Möglichkeit, *reflektiert* zu handeln. Wenn sie es tun, ist das, was sie tun, nicht instinktgeleitet, sondern setzt Selbstkontrolle voraus: Bevor sie agieren, hinterfragen sie, ob die unmittelbaren Wünsche wünschenswert sind, das heißt sie prüfen, ob die Beweggründe der beabsichtigten Handlung wirklich handlungswirksam werden sollen. Um Missverständnisse zu vermeiden: Damit ist nicht gemeint, dass die Erwägung, die der Handlung vorausgeht, eine bestimmte Qualität aufweisen muss, damit

¹⁰ Müssten Roboter, die sich um sich selbst kümmern, indem sie defekte Bestandteile ersetzen, dann nicht auch lebende Organismen sein? Wie verhält es sich mit Software, die in ihrem Programmcode nach Fehlern sucht und korrigiert? Korsgaard spricht das Problem an, geht aber schnell darüber hinweg: „Wir mögen Fortschritte bei der Herstellung von Maschinen machen, die sich auf die eine oder andere Weise selbst instand halten, aber ein Messer schärft sich normalerweise nicht selbst und ein Auto sucht sich nicht das beste Benzin aus.“ (T 36f.).

¹¹ Dass Korsgaard Tieren nur deswegen einen moralischen Status zuspricht, weil sie über zumindest rudimentäre Formen von Selbstbewusstsein verfügen (vgl. T 48ff.), hat zur Folge, dass man dem nicht-empfindungsfähigen Teil der Natur moralisch nichts schuldet. Nussbaum (2010, 490) ist ebenfalls der Ansicht, dass Empfindungsfähigkeit die „Minimalbedingung der Zugehörigkeit zur Gemeinschaft jener Wesen“ ist, „die überhaupt gerechtigkeitsbasierte Ansprüche haben können.“ Dies hindert Schlosberg (2007) und Kortemäki (2017) nicht, an Nussbaums *capabilities approach* anzuknüpfen, um Pflichten gegenüber anderen Teilen der Natur – zum Beispiel gegenüber Ökosystemen – zu begründen. Dabei gehören sämtliche Fähigkeiten, die sie anführen, zum Bedeutungsspektrum der Autopoiesis: Integrität, Selbstregulierung, Entwicklungsfähigkeit und Resilienz sind die zentralen Begriffe. Problematisch scheint mir zu sein, dass die erwähnten Eigenschaften nicht nur auf Ökosysteme, sondern auch auf soziale Systeme zutreffen. Nehmen wir die kapitalistische Ökonomie mit ihrer ausgeprägten Fähigkeit zur Selbstreproduktion. Hat sie einen moralischen Status? Ist es gar unmoralisch, sie durch eine solidarische Wirtschaftsform ersetzen zu wollen?



der Akteur als rational gelten kann. Wichtig ist nicht, *was* er reflektiert hat, sondern *dass* er reflektiert hat. Seine Gründe mögen töricht oder gar verbrecherisch sein, aber im deskriptiven Sinn handelt er immer noch rational. Und Korsgaard zufolge ist Kant der Ansicht, dass bereits dieses gattungsspezifische Vermögen die Selbstzweckhaftigkeit des Menschen begründet: „Kants These über den Wert vernünftiger Wesen bezieht sich auf vernünftige Wesen im deskriptiven Sinne.“ (T 111)

- II. Da Tiere die Gründe ihres Tuns nicht bewerten, verfügen sie erst recht nicht über Rationalität im *normativen* Sinne. Menschen hingegen sind in der Lage, nur solche Gründe als Gründe gelten zu lassen, die für jedes andere (bei Kant: vernünftige) Wesen annehmbar sind. Gründe sind moralisch vertretbare Gründe. Rationale Wesen machen deshalb Ansprüche gegeneinander geltend. Person A kann von Person B begründet fordern, eine bestimmte Handlung einzustellen, weil sie nicht als allgemeines Gesetz gewollt werden kann. Indem Person A Person B daran erinnert, so zu handeln, wie es das praktische Gesetz fordert, gesteht sie Person B ihrerseits zu, begründete Ansprüche an sie zu stellen (im Namen dieses oder eines anderen praktischen Gesetzes). Die moralischen Beziehungen rationaler Wesen zeichnen sich durch *Reziprozität* aus (vgl. T 164f., 192)

Korsgaard leugnet also weder, dass zwischen rationalen und nicht-rationalen Wesen unterschieden werden kann, noch stellt sie in Abrede, dass sich die moralischen Beziehungen, die rationale Wesen untereinander eingehen (in Form reziprok erhobener Ansprüche), von den moralischen Beziehungen zu nicht-rationalen Wesen unterscheiden. Was Korsgaard von Kant trennt, ist die Annahme, dass es moralische Verpflichtungen gegenüber nicht-rationalen Wesen gibt. Und sie ist der Ansicht, dass Kant selbst zu dem Schluss hätte kommen müssen. Warum?

Korsgaard bezieht sich auf die Passage in der GMS, mit der Kant zur Selbstzweckformel des KI überleitet. Die Passage beginnt mit der Feststellung, dass die *Objekte* der Neigungen keinen absoluten Wert haben können. Wir begehren sie nicht, weil sie gut sind, sondern sie sind gut, weil wir sie begehren (ihr Wert ist *bedingt* durch unsere Neigungen). Die Neigungen *selbst* haben ebenfalls keinen unbedingten Wert: „daß vielmehr, gänzlich davon frei zu sein, der allgemeine Wunsch eines jeden vernünftigen Wesens sein muß“ (AA IV 428). Von unbedingtem Wert seien nur die vernünftigen Wesen selbst. Und dies, so Kant, entspricht auch dem Selbstverständnis eines jeden vernünftigen Wesens:

[D]ie vernünftige Natur existiert als Zweck an sich selbst. So stellt sich notwendig der Mensch sein eignes Dasein vor; so fern ist es [das oberste Moralprinzip, S.E.] also ein subjectives Princip menschlicher Handlungen. So stellt sich aber auch jedes andere vernünftige Wesen sein Dasein zufolge eben desselben Vernunftgrundes, der auch für mich gilt, vor; also ist es zugleich ein objectives Princip, woraus als einem obersten praktischen Grunde alle Gesetze des Willens müssen abgeleitet werden können. (AA IV 429)

Kant konstatiert an dieser Stelle, *dass* der Mensch sich als Zweck an sich begreift, aber er legt noch nicht dar, *warum* dies der Fall ist. Für eine Begründung

verweist er auf den dritten Abschnitt der GMS.¹² Obgleich Korsgaard sich bewusst zu sein scheint, dass die Passage, die sie für zentral hält, knapp und argumentativ unvollständig ist – Sensen (2011, 57) bezeichnet sie als „very cryptic“ –, glaubt sie aus ihr herauslesen zu können, dass die Aufforderung, alle vernünftigen Wesen als Zweck an sich anzuerkennen, Kant zufolge eine *Implikation rationalen Handelns* ist: Laut Korsgaard möchte er in dieser Passage zeigen, dass das, was die Selbstzweckformel verlangt, die *Bedingung der Möglichkeit erstopersonalen Wollens* sei. Kant erteile dem metaethischen Realismus damit eine Absage. Er beginne nicht mit der Unterstellung einer metaphysischen Eigenschaft, die einen unbedingten Wert darstelle, jedem Einzelnen als Substanz innewohne und darum eine Verpflichtung bei allen anderen Menschen hervorrufe (vgl. T 175 ff., T 194). Er beginne umgekehrt mit einem basalen Willensverhältnis, das jedoch dazu nötige, sich selbst und schließlich allen vernünftigen Wesen einen unbedingten Wert zuzusprechen (ähnlich argumentiert Wood 2008, 88 ff.). Sehen wir uns an, wie Korsgaard die von Kant erwähnte Begründungslücke zu schließen können glaubt:

[B]ecause we are rational, we cannot decide to pursue an end unless we take it to be good. Most of our ends, however, are simply the objects of our inclinations, and the objects of our inclinations are not, just as such, intrinsically valuable. So we need some further story about why we take them to be good. That further story is that we attribute to ourselves the power to confer value on our ends by rationally choosing them. In so doing, we attribute a fundamental kind of value to ourselves. We attribute value to our own humanity, a property which Kant identifies with our capacity to determine our ends through rational choice. I summed this all up by saying that humanity is the unconditioned condition of all value, and as such, it must be valued. (Korsgaard 2021, 171)

Kants Hinführung zur Selbstzweckformel versteht Korsgaard (1996b, 122) als ein „regress upon the conditions“ des unbedingten Gutseins begehrteter Gegenstände. Sie bedient sich also einer *Präsuppositionsanalyse*. Die Präsuppositionsanalyse ist eine Form der Rechtfertigung, zu der gerne gegriffen wird, wenn andere Rechtfertigungsformen zu keinem Konsens führten: Das, was besonders umstritten oder nur schwer zu begründen ist – in diesem Fall das oberste Moralprinzip in Gestalt der Selbstzweckformel –, wird im Rahmen einer Präsuppositionsanalyse dadurch begründet, dass es sich als stillschweigende Voraussetzung von etwas erweist, das unumstritten oder leicht zu begründen ist.

Betrachten wir Schritt für Schritt, wie Korsgaard Kants Herleitung der Selbstzweckformel versteht (ähnlich Birch 2020, 8):

P1: Wer sich rational entscheidet, einen Zweck zu verfolgen, der hält den Zweck zwangsläufig für unbedingt gut.¹³

¹² Vgl. AA IV 429, Fußnote. Folgt man Kants Hinweis, so findet man in GMS III das Argument, jedes vernünftige Wesen sehe in sich einen Zweck an sich, weil es sich als *frei* begreift (vgl. AA IV 447f.). Zum dritten Abschnitt der GMS siehe die Studie von Schönecker (1999).

¹³ „Kant denkt, dass wir als vernünftige Wesen nicht beschließen können, einen Zweck oder ein Ziel zu verfolgen, solange wir sie nicht für unbedingt gut halten.“ (T 180) Die erstaunliche Aussage, dass wir nach Kant unsere Ziele für *unbedingt* gut halten, ist kein Übersetzungsfehler, denn auch im englischen Original



- P2: Die meisten Zwecke, die wir verfolgen, hängen von unseren individuellen Neigungen ab, das heißt sie *sind* nicht unbedingt gut, nicht intrinsisch wertvoll. Die Zwecke *werden* unbedingt gut, weil wir uns die Autorität zusprechen, ihnen einen von anderen Personen zu respektierenden Wert zu verleihen, *indem wir sie rational wählen*.
- P3: Wer sich die Autorität zuspricht, Zwecken qua rationaler Wahl einen Wert zu verleihen, der anderen Verpflichtungen auferlegt, der sieht in sich selbst einen fundamentalen Wert, das heißt versteht sich als Zweck an sich.
- P4: Wer sich als einen Zweck an sich versteht, weil er Zwecken qua rationaler Wahl einen unbedingten Wert verleiht, der sieht sich genötigt, auch jeden anderen Akteur, der qua rationaler Wahl Zwecken einen unbedingten Wert verleiht, als Zweck an sich ansehen.
- K1: Wer sich rational entscheidet, einen Zweck zu verfolgen, sieht sich genötigt, auch alle anderen rationalen Akteure als Selbstzweck anzuerkennen.

P1-P3 bilden eine Einheit. Sie formulieren ein transzendentes Argument dafür, dass jeder rationale Akteur praktisch genötigt ist, sich fundamentalen Wert zuzusprechen. Ein rationaler Akteur bezieht sich demnach auf sich selbst als Zweck an sich, indem er Zwecke auf Basis seiner Zwecksetzungskompetenz als unbedingt gut deklariert:

Es ist, als ob Sie, wenn Sie eine Entscheidung treffen, sich jedes Mal sagen würden: „Was mir wichtig ist, ist mir wichtig, *Punkt*, unbedingt wichtig, weil ich selbst mir wichtig bin.“ Indem Sie verfolgen, was gut für Sie ist, zeigen Sie also, dass Sie sich als Zweck an sich selbst begreifen (T 183)

When Kant says: „rational nature exists as an end in itself. Man necessarily thinks of his own existence in this way; thus far it is a subjective principle of human actions“ (429/47), I read him as claiming that in our private rational choices and in general in our actions we view ourselves as having a value-conferring status in virtue of our rational nature. We act as if our own choice were the sufficient condition of the goodness of its object: this attitude is build into [...] rational action. (Korsgaard 1996b, 122 f.)

Mit P4 wiederum wird der Übergang von der Erste-Person-Perspektive zur interpersonalen Moral vollzogen: Wer sich als rationaler Akteur Selbstzweck ist, der ist genötigt, auch alle anderen rationalen Akteure als Selbstzweck anzuerkennen.

Nehmen wir an, dieses Argument für die Selbstzweckformel finde sich tatsächlich bei Kant, und nehmen wir (wie Korsgaard) weiterhin an, es handele sich grundsätzlich um einen überzeugenden Ansatz der Moralbegründung: Wie lässt es sich dann einsetzen, um moralische Pflichten gegenüber Tieren zu begründen? Wenn Korsgaards These richtig ist, dass Kant auf Grundlage seines „regress upon the conditions“-Arguments selbst zu der Ansicht hätte gelangen können, dass Tiere Selbstzwecke sind, muss sie eine immanent unplausible Prämisse in dem Argument präsentieren, die die Selbstzweckhaftigkeit unzulässig auf Menschen beschränkt.

heißt es: „Kant thinks that because we are rational, we cannot decide to pursue an end unless we take it to be absolutely good.“ (Korsgaard 2018, 137).



Dies ist aus ihrer Sicht P2. Korsgaard zufolge müssen sich rationale Wesen nicht nur als *rationale* Wesen wertschätzen, sondern sie müssen sich auch als *Naturwesen* wertschätzen, um rational handeln zu können. Kant habe Tiere nicht als Zweck an sich begreifen können, weil er übersah, dass Rationalität die Wertschätzung des oberen *und* des unteren Begehrungsvermögens voraussetze:

Betrachten wir aber meine *ursprüngliche* Entscheidung, irgendeinen Gegenstand meiner Neigung einen Wert beizumessen, also etwas, das gut für mich ist, so zu behandeln, als wäre es unbedingt gut. Diese Entscheidung ist kein Akt der Achtung vor meiner eigenen Autonomie. Denn ich kann meine eigenen Entscheidungen nicht respektieren und tun, was nötig ist, um sie in die Tat umzusetzen, *bevor* ich sie getroffen habe. Dem Gesetz, das ich mir selbst gegeben habe, kann ich erst gehorchen, nachdem ich es mir gegeben habe. Im Augenblick der ursprünglichen Entscheidung betrachte ich mich daher als Zweck an sich selbst in einem Sinne, der nicht von der Vorstellung gedeckt ist, dass ich meinen eigenen Status als Gesetzgeber im Reich der Zwecke achte. Wenn ich die ursprüngliche Entscheidung treffe, [...] habe ich keinen anderen Grund, mein Ziel als unbedingt gut zu begreifen, als die Tatsache, dass es gut *für mich* ist. (T 188)

Wenn wir einen Gegenstand unserer Neigungen so behandeln, als sei er unbedingt gut, dann tun wir dies nicht nur, weil wir unser Vermögen, Zwecke zu setzen, in besonderem Maße wertschätzen. Dies geht schon daraus hervor, dass wir den Gegenstand bereits *vor* der Zwecksetzung für unbedingt gut halten: Ich betrachte den begehrten Gegenstand als unbedingt gut, weil er gut *für mich* ist, das heißt gut für ein *fühlendes* Wesen ist. Diese zweite Argumentation lautet also:

- P1': Wer sich rational entscheidet, einen Zweck zu verfolgen, der hält den Zweck zwangsläufig für unbedingt gut.
- P2': Wer sich rational entscheidet, einen Zweck zu verfolgen, der hält den Zweck für unbedingt gut, *bevor* er sich entscheidet, ihn zu verfolgen.
- P3': Wer einen Zweck verfolgt, den er bereits für unbedingt gut hält, bevor er sich rational entschieden hat, ihn zu verfolgen, sieht in sich *als fühlendes Wesen* einen fundamentalen Wert, das heißt versteht sich als Zweck an sich.
- P4': Wer sich als ein Wesen Selbstzweck ist, das seine Neigungen befriedigt, sieht sich genötigt, auch alle anderen Wesen, die ihre Neigungen befriedigen wollen, als Selbstzweck anzusehen.
- K2: Wer sich rational entscheidet, einen Zweck zu verfolgen, sieht sich genötigt, alle empfindungsfähigen Wesen als Selbstzweck anzuerkennen.

Die erste Prämisse ist in beiden Argumenten identisch. Kants Kardinalfehler in seinem Argument für die auf rationale Wesen beschränkte Selbstzweckformel (Argument 1) steckt Korsgaard zufolge also in P2, das heißt in der Annahme, sämtliche Zwecke seien gleichermaßen bedeutungslos, solange sich nicht ein rationaler Akteur für einen von ihnen entschieden hat. Für empfindungsfähige Wesen ist die Befriedigung ihrer natürlichen Neigungen jedoch bereits ein Gut, bevor der Akteur zwischen ihnen eine Entscheidung trifft. Wer müde und zugleich hungrig ist, wird zwischen



beiden Optionen wählen müssen, aber die Befriedigung beider Neigungen wird nicht erst durch die Wahl ein Gut (deshalb ersetzt Korsgaard P2 durch P2').

8 Korsgaards Kant-Lesart

Bevor ich näher auf Korsgaards (tierethisch motivierte) Modifikation des transzendentalen Arguments für die Selbstzweckformel eingehe, möchte ich mich kurz zur Frage äußern, ob sich dieses Argument bei Kant überhaupt findet. Meines Erachtens ist dies nicht der Fall. Abgesehen davon, dass die kurze Passage, auf die sich Korsgaard hauptsächlich bezieht (AA IV 429), eine derart weitreichende Interpretation nicht stützt, setzt ihr Projekt ein spezielles Verständnis des kantischen Vernunft- und Autonomiebegriffs voraus – und wie aus meiner eigenen Darstellung der Selbstzweckformel schon hervorgegangen sein dürfte, stimme ich mit diesem Verständnis nicht überein.

Unstrittig ist, dass nicht nur Korsgaard, sondern auch Kant zwischen Vernunft im deskriptiven Sinne (empirisch-praktische Vernunft) und Vernunft im normativen Sinne (reine praktische Vernunft) unterscheidet. Ersteres bezeichnet die Fähigkeit, *irgendeinen* Zweck zu setzen oder sich *irgendeinen* Handlungsgrundsatz zu geben, letztere hingegen das deutlich weiter reichende Distanzierungs- und Steuerungsvermögen, subjektive Maximen nur unter dem Vorbehalt ihrer Verallgemeinerbarkeit zu verfolgen – und das bedeutet: den Willen neigungsunabhängig zu bestimmen. Unstrittig ist ebenfalls, dass nach Kant bereits die Vernunft im deskriptiven Sinne die Menschheit vom Tier grundlegend unterscheidet. Die von Korsgaard angeführten Zitate aus der *GMS*, der *Religionsschrift* und aus *Mutmaßlicher Anfang der Menschengeschichte* belegen dies (vgl. Korsgaard 1996b, 110–114). Sie hätte auch auf den Abschnitt „Der Kanon der reinen Vernunft“ in der *Kritik der reinen Vernunft* verweisen können, in dem Kant die empirisch belegbare und mit kausaler Determination kompatible *praktische* Freiheit – im Unterschied zur theoretisch nicht beweisbaren *transzendentalen* Freiheit – heranzieht, um die menschliche Willkür von der tierischen abzugrenzen (vgl. AA III 521). Eine noch prägnantere Stelle findet sich in der *Anthropologie in pragmatischer Hinsicht*: Durch seine technischen, pragmatischen und moralischen Anlagen sei der Mensch „von allen übrigen Naturwesen kenntlich unterschieden, und eine jede dieser drei Stufen kann für sich allein schon den Menschen zum Unterschiede von anderen Erdbewohnern charakteristisch unterscheiden.“ (AA VII 322)

Was diese Zitate jedoch nicht belegen, ist folgende These Korsgaards (1996b, 111): „it is the capacity for the rational determination of ends in general, not just the capacity for adopting morally obligatory ends, that the Formula of Humanity orders us to cherish unconditionally.“ Dass der Mensch ein Zweck an sich ist, sei bereits in der empirisch-praktischen Seite seiner Vernunft begründet, das heißt in der Tatsache, dass Menschen sich nicht instinktiv verhalten müssen, sondern auf Grundlage von Überlegungen wie auch immer geartete Zwecke setzen können:

Angenommen, ich folge einer Neigung, die sich auf irgendetwas richtet, das ich will. Weil ich es will, scheint es mir gut für mich und wird Zweck oder



Ziel meines Handelns. Wenn ich dieses Ziel wähle, erlege ich jedem anderen vernünftigen Wesen eine Pflicht auf: Niemand darf mich, wenn er nicht einen wirklich guten Grund dazu hat, an der Verfolgung dieses Ziels hindern. Und jeder hat ohne einen solchen Grund eine gewisse Verpflichtung, mir dabei zu helfen, falls ich Hilfe brauche. (T 188)

Korsgaard fährt fort, dass aus den neigungsbedingten Entscheidungen nicht nur Verpflichtungen für andere Personen, sondern auch Selbstverpflichtungen entspringen:

Nehmen wir an, ich beschließe, im Garten Gemüse anzubauen, wohl wissend, dass mich das zwingt, regelmäßig Unkraut zu jäten. Damit verpflichte ich mich, künftig in gewissen Abständen Unkraut zu jäten, und zwar auch dann, wenn mir nicht danach ist. [...] In diesem Sinne habe ich mein künftiges Selbst einer Art kategorischen Imperativ unterworfen. Und mein künftiges Selbst erlegt mir seinerseits Pflichten auf. Wenn es das unerlässliche Jäten übernehmen soll, ist es entscheidend, dass ich ihm ein paar Knieschoner besorge und Werkzeuge, die es zum Jäten brauchen wird. Und auch *das* muss ich tun, ob mir danach ist oder nicht. In diesem ganz schlichten Sinne erlege ich mir Pflichten auf, wann immer ich eine Entscheidung treffe. Ich gebe mir selbst zwingende Gründe, etwas zu tun. Wenn ich nach diesen Gründen handle, kann ich davon sprechen, dass ich meine Autonomie achte, indem ich dem Gesetz gehorche, das ich mir selbst gegeben habe. (T 188)

Von Autonomie sei also bereits dann zu sprechen, wenn jemand sein zukünftiges Handeln einem *neigungsbedingten* Grundsatz unterstellt. Diese Kant-Auslegung liegt P2 zugrunde. Und ohne sie ergäbe der gegen Kant gerichtete Einwand, dass wir neigungsabhängige Zwecke für unbedingt gut halten, noch bevor wir uns rational entschieden haben, sie zu verfolgen (weshalb wir uns nicht nur als rationale Akteure, sondern auch als empfindungsfähige Wesen fundamental wertschätzen), überhaupt keinen Sinn. Damit die Selbstzweckhaftigkeit von Wesen, die ihren Neigungen nachgehen, von Korsgaard als eine Idee präsentiert werden kann, zu der Kant selbst hätte kommen können, bedarf es einer Kant-Interpretation, nach der Menschen bereits dann autonom-selbstzweckhafte Wesen sind, wenn sie unter dem Einfluss ihrer Neigungen Zwecke setzen. Kant hat jedoch einen deutlich engeren Begriff von Autonomie:

Autonomie des Willens ist die Beschaffenheit des Willens, dadurch derselbe ihm selbst (unabhängig von aller Beschaffenheit der Gegenstände des Willens) ein Gesetz ist. Das Princip der Autonomie ist also: nicht anders zu wählen, als so, daß die Maximen seiner Wahl in demselben Willen zugleich als allgemeines Gesetz mit Begriffen seien. (AA IV 440; deckungsgleiche Definition: AA IV 444)

Nach Kant verstehen wir uns nicht als Zweck an sich, weil wir über Vernunft im deskriptiven, sondern weil wir über Vernunft im normativen Sinne verfügen, das heißt, weil wir nicht nur kluge Grundsatzentscheidungen im Dienste unserer vorhandenen Neigungen treffen, sondern weil wir die Ansprüche der Neigungen bei



der Bestimmung des Willens in toto suspendieren können, sollten sie die Prüfung des KI nicht bestehen. Vernunft im deskriptiven Sinne ist für Kant lediglich die notwendige, nicht aber die hinreichende Bedingung dafür, dass der Mensch Zweck an sich ist. Noch einmal das obige Zitat, dieses Mal jedoch in voller Länge:

Denn im Werthe über die bloße Thierheit erhebt ihn das gar nicht, daß er Vernunft hat, wenn sie ihm nur zum Behuf desjenigen dienen soll, was bei Thieren der Instinct verrichtet; sie wäre alsdann nur eine besondere Manier, deren sich die Natur bedient hätte, um den Menschen zu demselben Zwecke, dazu sie Thiere bestimmt hat, auszurüsten, ohne ihn zu einem höheren Zwecke zu bestimmen. Er bedarf also freilich nach dieser einmal mit ihm getroffenen Naturanstellung Vernunft, um sein Wohl und Weh jederzeit in Betrachtung zu ziehen, aber er hat sie überdem noch zu einem höheren Behuf, nämlich auch das, was an sich gut oder böse ist, und worüber reine, sinnlich gar nicht interessirte Vernunft nur allein urtheilen kann, nicht allein mit in Überlegung zu nehmen, sondern diese Beurtheilung von jener gänzlich zu unterscheiden und sie zur obersten Bedingung der letzteren zu machen. (AA V 61 f.)

Im Wert über die bloße Tierheit erhebt die Menschheit erst die *reine* praktische Vernunft. In GMS argumentiert Kant deckungsgleich. Die „Würde (Prärogativ) vor allen bloßen Naturwesen“ bestehe darin, „seine Maximen jederzeit aus dem Gesichtspunkte seiner selbst, zugleich aber auch jedes andern vernünftigen als gesetzgebenden Wesens [...] nehmen zu müssen“ (AA IV 438). Für Kant ist „Sittlichkeit und die Menschheit, so fern sie derselben fähig ist, dasjenige, was allein Würde hat“ (AA IV 435). In GMS III – Kant hatte angekündigt, hier zu erläutern, warum sich der Mensch als Zweck an sich begreift – findet sich ebenfalls keine Argumentation, die der von Korsgaard auch nur von weitem ähnelt (vgl. Sensen 2011, 62). Und in der *Metaphysik der Sitten* heißt es unmissverständlich:

Der Mensch im System der Natur (*homo phaenomenon, animal rationale*) ist ein Wesen von geringer Bedeutung und hat mit den übrigen Thieren [...] einen gemeinen Werth. Selbst, daß er vor diesen den Verstand voraus hat und sich selbst Zwecke setzen kann, das gibt ihm doch nur einen äußeren Werth seiner Brauchbarkeit [...]. Allein der Mensch, als Person betrachtet, d.i. als Subject einer moralisch-praktischen Vernunft, ist über allen Preis erhaben; denn als solcher (*homo noumenon*) ist er [...] als Zweck an sich selbst zu schätzen, d.i. er besitzt eine Würde (einen absoluten innern Werth), wodurch er allen andern vernünftigen Weltwesen Achtung für ihn abnöthigt. (AA VI 434 f.)

Würde ist ein Attribut, das nur einem Wesen zukomme, welches das moralische Gesetz zu achten und zu befolgen in der Lage ist. Damit ist Kants Ethik einerseits egalitär und inklusiv: *Jedes* Wesen, das moralisch handeln kann – das heißt über die Disposition der transzendentalen Freiheit verfügt –, ist Teil einer Gemeinschaft, in der sich alle und im gleichen Maße Respekt schulden. Dass Kants Ethik auf die Reziprozität autonomie-befähigter Wesen getrimmt ist, birgt jedoch andererseits ein erhebliches Exklusionspotential: Abgesehen von der unter Kant-Exegeten umstrittenen Frage, ob Angehörige der Gattung *homo sapiens*, die ihren Willen *noch nicht* (Babys, Kleinkinder), *nicht mehr* (wie dauerhaft komatöse Menschen) oder *nie-*



mals (geistig schwer Behinderte) autonom bestimmen können, nach Kants eigenen Maßstäben Würde zukommt oder nicht,¹⁴ sind Tiere *auf jeden Fall* ausgeschlossen. Kant-immanent lässt sich kein Weg aufzeigen, wie sich Tieren Würde zusprechen ließe. Dass Korsgaard hier dennoch eine Möglichkeit sah, ist darauf zurückzuführen, dass sie Kants emphatisches Verständnis von Autonomie mit dem bloßen Vermögen identifiziert, *neigungsbedingten* Grundsätzen zu folgen.

9 Die Prämissen von Korsgaards Argumentation

Allerdings entwertet dies nicht Korsgaards Argumentation als solche. Denn selbst wenn Korsgaards Auslegung des kantischen Autonomie- und Würdebegriff nicht haltbar sein sollte, könnte sich ihre Exegese immer noch als instruktives Missverständnis erweisen: genau dann, wenn sie uns zu einer stichhaltigen anthropozentrischen Moralphilosophie (Argument 1) oder zu einer stichhaltigen pathozentrischen Moralphilosophie (Argument 2) geführt hat. Korsgaards Theorie wäre dann zwar nicht mehr durch Kant gedeckt, aber das ist auch nachrangig, solange sie überzeugt. Um zu prüfen, ob dies der Fall ist, müssen wir uns die Prämissen im Detail ansehen. Ich beginne mit den Prämissen, die mir unproblematisch erscheinen:

P4|P4': Mit der jeweils letzten Prämisse wird der Übergang von der Wertschätzung der eigenen Person zur Wertschätzung aller anderen Personen bzw. empfindungsfähigen Wesen vollzogen. Setzen wir für einen kurzen Moment voraus, dass die Prämissen P1-P3/ P1'-P3' überzeugend sind, dann – aber auch nur dann – lässt sich dieser Übergang mit Alan Gewirths ‚criterion of relevant similarities‘ erläutern:

[I]f some predicate P belongs to some subject S because S has the property Q (where the ‘because’ is that of sufficient reason or condition), then P must also belong to all other subjects S1, S2,...,Sn that have Q. If one denies this implication in the case of some subject, such as S1, that has Q, then one contradicts oneself. For in saying that P belongs to S because S has Q, one is saying that having Q is a sufficient condition of having P; but in denying this in the case of S1, one is saying that having Q is not a sufficient condition of having P (Gewirth 1978, 105).

Wenn S behauptet, es sei ein Zweck an sich (P), weil es seinen Zwecken qua Wahl einen unbedingten – das heißt von Anderen zu respektierenden – Wert verleiht bzw. weil es noch vor jeder rationalen Wahl seine Empfindungsfähigkeit unbedingt wertschätzt (Q), dann kann es anderen Subjekten, die ebenfalls über Q verfügen, nicht absprechen, ebenfalls P zu sein, denn dies widerspräche der eigenen Behauptung, dass Q hinreichende Bedingung für P ist. Die Anerkennung anderer Personen/

¹⁴ Höffe (2002, 132) und Hruschka (2002, 478) sind der Ansicht, dass nach Kant bereits die Zugehörigkeit zur biologischen Gattung dem einzelnen Menschen Würde verleiht. Gutman (2010) und Wood (1998, 194) sehen dafür keinen Ansatzpunkt. Gemäß ihrer Auslegung sind für Kant nur solche Wesen ein Zweck an sich, die *aktuell* über das Vermögen der moralischen Selbstbestimmung verfügen. Horn (2014, 68–111) argumentiert vor dem Hintergrund der rechtsphilosophischen Schriften Kants dafür, dass selbst Menschen, die aktuell zur moralischen Willensbestimmung befähigt sind, ihre Würde einbüßen können, weshalb Kant nicht zu den Ahnherren der modernen Menschenrechtstheorie gezählt werden könne.



Wesen als Selbstzweck ist demnach geboten, weil ansonsten *das Selbstverhältnis inkonsistent begründet wäre*. Dieses Argument scheint mir unanfechtbar zu sein.¹⁵

P3|P3': Auch diese Prämisse ist – wiederum *für sich genommen* – überzeugend. Wer sich attestiert, Zwecken, die lediglich bedingt gut sind, einen unbedingten, das heißt andere Personen verpflichtenden Wert zu verleihen, muss sich selbst als ein Wesen von unbedingtem Wert verstehen.

P2|P2' und P1|P1': Hier beginnen die Probleme. Betrachten wir zunächst P2. Diese Prämisse stellt die Gründe bereit, die Korsgaards Kant-Lektüre zufolge jede Person veranlasst, sich als Selbstzweck aufzufassen (P3). Demnach halten rationale Akteure die von ihnen gesetzten Zwecke für unbedingt gut, weil sie diese Zwecke rational gewählt haben. Wenn Korsgaard betont, ein rationaler Akteur verhalte sich zu dem Gegenstand seines Begehrens so, als sei er *unbedingt* gut, so denkt sie immer auch an die Ansprüche, die mit dieser Zuschreibungspraxis einhergehen: Dass ein Gegenstand vom Akteur als unbedingt gut deklariert wird, meint nicht nur, dass dem Akteur der Gegenstand besonders wichtig ist, sondern weil ihm der Gegenstand als *gewählter* Gegenstand besonders wichtig ist, ergeht an andere Akteure die Aufforderung, ihn nicht an der Verfolgung des Zwecks zu hindern bzw. ihn sogar bei der Verfolgung des Zwecks zu unterstützen (vgl. T 188). Anders gesagt, rationale Akteure machen im Rahmen ihrer individuellen Zwecksetzung bereits einseitig moralische Ansprüche gegen andere rationale Akteure geltend. Ich sehe bei Korsgaard allerdings kein Argument, warum solche Ansprüche eine Implikation rationalen Handelns sein sollen. Nehmen wir den rationalen Egoisten (vgl. Sensen 2011, 74). Dieser schätzt sich selbst. Er mag auch seine Fähigkeit schätzen, Zwecke zu setzen (denn ohne sie könnte er seinen Neigungen weniger effektiv nachgehen). Selbstverständlich ist ihm bewusst, dass andere Akteure sich (und ihre Fähigkeit, Zwecke zu setzen) ebenfalls schätzen. Nur wird er bestreiten, dass deshalb irgendjemand berechtigt wäre, anderen gegenüber moralische Ansprüche geltend zu machen. Das wäre doch anmaßend. Jeder, so der rationale Egoist, solle einfach versuchen, die Ziele, die ihm wichtig sind, zu erreichen. Dies mag zu einer Extremsituation führen, wie sie Hobbes beschrieben hat, sodass es geboten erscheint, einen Modus Vivendi zu finden. Die Ansprüche, die dann geltend gemacht werden können, sind zwangsbewehrte Ansprüche auf Basis einer vertraglichen Übereinkunft, der jedoch keine einseitigen moralischen Forderungen zugrunde liegen, sondern eine bestimmte Interessenkonfiguration (siehe hierzu Stemmer 2002, 13 f.). Um möglichen Missver-

¹⁵ Ob Korsgaard es selbst verwendet, steht auf einem anderen Blatt. In einer neueren Publikation (Korsgaard 2021, 183) erweckt sie zumindest den Anschein. Sie vergleicht hier den Grund, nicht nur sich selbst, sondern auch alle anderen Personen als Selbstzweck anzuerkennen, mit dem Grund, nicht nur sich selbst, sondern auch allen anderen Bürgern das Wahlrecht zuzugestehen. “But now suppose you ask me, ‘In virtue of what do you have the right to vote here?’ and I reply, ‘I am a citizen of this nation.’ Citizenship, as I understand it, is a form of normative standing: it gives its possessor certain normative or moral powers. You might reply, ‘Well, I am a citizen too, so I have the right to vote here as well.’ Notice that it would not make sense for me to respond, ‘no, my own citizenship has that normative implication – but so far as I am concerned, yours does not.’” Korsgaard sagt zwar nicht ausdrücklich, *warum* es keinen Sinn ergibt, anderen Staatsbürgern das Wahlrecht abzusprechen (respektive andere Personen die Anerkennung als Zweck an sich zu verweigern). Jedoch sind die Parallelen zu Gewirths Argument offensichtlich: In der Analogie geht es um eine *Eigenschaft* (Staatsbürgerschaft), die beiden Personen *gleichermaßen* zukommt und deren Besitz die *hinreichende* Bedingung für etwas ist (Wahlrecht).

ständnissen vorzubeugen: Mir geht es nicht darum, der Vertragstheorie das Wort zu reden. Meine Kritik beruht nicht auf externen Voraussetzungen. Ich hinterfrage lediglich, ob Korsgaards Vorschlag, Kant so zu lesen, dass sich interpersonale moralische Verpflichtungen über *stillschweigende Voraussetzungen der deskriptiven Rationalität* begründen lassen, theorieimmanent plausibel ist.

Korsgaard ist sich dieses Problems bewusst. In *The Sources of Normativity* (1996c) und *Self-Constitution* (2009) wendet sie sich gegen zeitgenössische Autoren, die sich entweder auf Hobbes (Gauthier) oder auf Kant beziehen (Gewirth, Nagel). Gemeinsam sei diesen Ansätzen, dass sie versuchen, ausgehend von der erstpersionalen Perspektive eine interpersonal verbindliche Moral zu begründen. Korsgaard zufolge scheitern sie, weil sie den Nachweis schuldig bleiben, dass eine Person, die sich selbst Zweck ist, ebenfalls genötigt ist, in anderen Personen einen Selbstzweck zu sehen; dem Egoismus lasse sich ausgehend von der erstpersionalen Perspektive nicht nachweisen, dass er inkonsistent wäre:

Consistency can force me to grant that your humanity is normative for you just as mine is normative for me. It can force me to acknowledge that your desires have the status of reasons for you, in exactly the same way that mine do for me. But it does not force me to share in your reasons, or make *your* humanity normative for me. It could still be true that I have my reasons and you have yours, and indeed that they leave us eternally at odds. (Korsgaard 1996c, 134)

Wer sich auf die Suche nach Gründen für moralische Pflichten macht und dabei einen *privaten* Begriff von Gründen zugrunde legt, der scheitere zwangsläufig. Oder anders gesagt, akteursneutrale Gründe ergeben sich nicht aus akteursrelativen Gründen. Damit sich die Kluft zwischen dem Erst- und dem Interpersonalen gar nicht erst auftut, schlägt Korsgaard ein anderes Verständnis von Gründen vor: Gründe – hier zieht sie eine Analogie zu Wittgensteins Privatsprachenargument – seien „public in their very essence“ (ebd., 135). Wer auf Basis von Gründen handelt, habe den Privatismus hinter sich gelassen. „To act on a reason is already, essentially, to act on a consideration whose normative force may be shared with others.“ (ebd., 136) Ohne in die Details gehen oder eine Bewertung vornehmen zu können (wie bei Skidmore 2002, 132 ff.; Gert 2002; De Maagt 2019), lässt sich festhalten, dass Korsgaards Theorie offenbar eine Kombination aus erstpersionaler und zweitpersonaler Perspektive darstellt. Korsgaard hat dies selbst bestätigt: „Darwall characterizes me both as someone who thinks all reasons are second personal and also as someone who thinks that ‘moral obligations can be grounded in the constraints of first-personal deliberation alone.’ That may sound paradoxical but it is basically right.“ (Korsgaard 2007, 10) Diese Selbsteinschätzung deckt sich mit P1/ P1': Einen rational gewählten Zweck als *unbedingt* gut anzusehen, beinhaltet die Unterstellung, „dass jedes vernünftige Wesen Grund hat, ihn zu verfolgen“ (T 181). Gemeint ist damit nicht die Forderung, jede Person solle begehren, was ich begehre, sondern gemeint ist eine moralische Erwartungshaltung, die auf der (impliziten) Unterstellung beruht, dass die von mir gesetzten Zwecke verallgemeinerbar, das heißt selbst moralisch vertretbar sind. Das bedeutet aber, dass die Präsuppositionsanalyse rationalen Handelns letztlich auf einem Zirkel beruht: Anders als Korsgaard behauptet, ist die Anerkennung aller anderen Personen als Zweck an sich *nicht* das Resultat



der Bedingungen der Möglichkeit *deskriptiver* Rationalität. Unter der Hand hat sie die deskriptive Rationalität mit der moralischen amalgamiert. Lässt man hingegen die deskriptive Rationalität deskriptiv sein, kehrt das Problem, Selbst- und Fremdwertschätzung überbrücken zu müssen, zurück. Es stellt sich die Frage, warum ein rationaler Akteur berechtigt sein soll, andere Akteure moralisch zu verpflichten. Kurz: Das transzendente Argument ist entweder petitiös oder es führt nicht zum gewünschten Ziel.

Aber lassen wir die einseitig erhobenen moralischen Ansprüche beiseite. Es stellt sich dann die Frage, ob ein rationaler Akteur tatsächlich, wie Korsgaard behauptet, den auf vorgängigen Neigungen basierenden Zwecken eine eminent wichtige Bedeutung beimessen muss. Bleiben wir zunächst bei ihrer Kant-Rekonstruktion (Argument 1). Korsgaard versteht Kant so, dass die Wertschätzung, die der rationale Akteur für seine Fähigkeit empfindet, neigungs basiert Zwecke zu setzen, sich zwangsläufig auf die gewählten Zwecke überträgt: Weil sie das Resultat der von ihm besonders geachteten Zwecksetzungskompetenz sind, sind die neigungs basierten Zwecke für ihn nun ebenfalls *unbedingt* wichtig. Zwingend ist diese Annahme jedoch nicht. Wie schon bei den einseitig erhobenen moralischen Ansprüchen handelt es sich auch bei der Wertschätzung der unter dem Einfluss der Neigungen gesetzten Zwecke *nicht um eine notwendige praktische Implikation*. Im Rahmen einer Präsuppositionsanalyse rationalen Handelns *müssen* die neigungs basierten Zwecke vom rationalen Akteur nämlich keineswegs *als* neigungs basierte Zwecke wertgeschätzt werden; wertgeschätzt werden müssen sie nur als *Mittel der Rationalität*. So ließe sich argumentieren, dass es gewisser Gegenstände bedarf, die begehrt werden, damit überhaupt Zwecke gesetzt werden können. Obgleich rationales Handeln keine bestimmte Neigung voraussetzt, so muss es doch zumindest *irgendwelche* Neigungen geben, damit dem Zwecksetzungsvermögen nicht das Material ausgeht, aus dem es Zwecke formen kann. So betrachtet erfüllen die neigungs basierten Zwecke jedoch eine *instrumentelle* Funktion. Sie müssen dem rationalen Akteur also keineswegs als *unbedingt* gut gelten. Birch (2020, 8) hat den Einwand so formuliert:

The old argument for the formula of humanity rests on considerations about what we *must* presuppose about our own value in order to act rationally at all. I grant that we *can* presuppose ourselves to be of fundamental value qua sentient beings, and not just qua rational beings. I even grant that we sometimes *do* presuppose this. But I doubt that we *must* presuppose this. For your argument to work, you have to show not just that we *can* and *do*, but that we *must*. To do that, you'd have to find at least one clear example of an end that we *must* value even though it is not good for us qua rational beings, and I don't think you can do that.

Mit anderen Worten: Mittels einer Präsuppositionsanalyse *rationalen* Handelns lässt sich kein zwingendes Argument für die unbedingte Wertschätzung des *Empfindungsvermögens* gewinnen. Dies hat unmittelbare Konsequenzen für Korsgaards tierethisches Ziel (Argument 2): Wenn rationale Akteure sich lediglich genötigt sehen, ihre eigenen Neigungen und Empfindungen als Voraussetzungen ihrer Rationalität – das heißt *bedingt* – wertzuschätzen, gibt es keinen nachvollziehbaren Grund, die Neigungen und Empfindungen der *Tiere* wertzuschätzen zu müssen, weil es sich

bei ihnen ja gar nicht um rationale Wesen handelt. Folglich wäre es nicht inkonsistent, die eigenen menschlichen Empfindungen für bedingt wertvoll zu erachten (als notwendiges Material der Zwecksetzung) und doch Tieren Qualen zuzufügen. Korsgaards Versuch, die pathozentrische Kritik der Utilitaristen dadurch zu entkräften, dass sie mit Kant gegen Kant argumentiert, vermag daher nicht zu überzeugen.

Aber worauf läuft dieser Einwand hinaus? Folgt aus ihm, dass der Utilitarismus gegenüber den Moralsystemen der Klassischen Deutschen Philosophie den Sieg davongetragen hat? Sicher nicht. Zum einen, weil der Utilitarismus selbst Angriffsflächen bietet – insbesondere im Hinblick auf die Sicherung von Grundrechten –, zum anderen, weil Korsgaards *Regress-upon-the-conditions*-Argument nur eine Möglichkeit darstellt, im Anschluss an die Freiheitstheorien der Klassischen Deutschen Philosophie Tierethik zu denken.¹⁶

Danksagung Für wertvolle Anmerkungen und Hinweise zu diesem Artikel danke ich Mark Siebel.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Siglen

- AA Kant, Immanuel. 1968. *Kants Werke*, Akademie-Textausgabe. Berlin, New York: De Gruyter.
 RW Bolzano, Bernard. 1994. *Lehrbuch der Religionswissenschaft*. Erster Teil. §§ 1-85. In *Gesamtausgabe*, Reihe 1, Bd. 6. Stuttgart-Bad Cannstatt: frommann-holzboog.
 T Korsgaard, Christine M. 2021. *Tiere wie wir: Warum wir moralische Pflichten gegenüber Tieren haben. Eine Ethik*. München: Beck.

Literatur

- Albrecht, Michael. 1994. Kants Maximenethik und ihre Begründung. In *Kant-Studien*, 85 (2): 129–146.
 Alves, Julius. 2010. Vollkommene Tugendpflichten. Zur Systematik der Pflichten in Kants *Metaphysik der Sitten*. In *Zeitschrift für Philosophische Forschung* 64 (4): 520-545.
 Baranzke, Heike. 2002. *Würde der Kreatur? Die Idee der Würde im Horizont der Bioethik*. Würzburg: Königshausen & Neumann.
 Birch, Jonathan. 2020. The place of animals in Kantian ethics. In *Biology & Philosophy* 35 (8): 1–11. <https://doi.org/10.1007/s10539-019-9712-0>.
 Birnbacher, Dieter. 2013. *Analytische Einführung in die Ethik*. Berlin, New York: De Gruyter.

¹⁶ Einen aussichtsreichen systematischen Ansatz hat jüngst Müller (2022) vorgelegt.



- Bolzano, Bernard. 1981. *Philosophische Tagebücher 1811–1817*. In *Gesamtausgabe*, Reihe 2, B, Bd. 16. Stuttgart-Bad Cannstatt: frommann-holzboog.
- Bolzano, Bernard. 2009. *Philosophische Tagebücher 1803–1810*. In *Gesamtausgabe*, Reihe 2, B, Bd. 14. Stuttgart-Bad Cannstatt: frommann-holzboog.
- De Maagt, Sem. 2019. It only takes two to tango: against grounding morality in interaction. In *Philosophical Studies* 176: 2767–2783. <https://doi.org/10.1007/s11098-018-1150-3>
- Dietrichson, Paul. 1969. Kant's Criteria of Universalizability. In *Kant: Foundations of the Metaphysics of Morals: Text and Critical Essays*. Hrsg. von Robert Paul Wolff, 163–207. Indianapolis: BobbsMerrill.
- Gert, Joshua. 2002. Korsgaard's Private-Reasons Argument. In *Philosophy and Phenomenological Research* 64 (2): 303–324.
- Gewirth, Alan (1978). *Reason and morality*. Chicago: University of Chicago Press.
- Gutmann, Thomas. 2010. Würde und Autonomie. Überlegungen zur Kantischen Tradition. In *Jahrbuch für Wissenschaft und Ethik* 15 (1): 3–34. <https://doi.org/10.1515/9783110222906.3>
- Guyer, Paul. 1998. The Value of Reason and the Value of Freedom. In *Ethics* 109 (1), 22–35.
- Guyer, Paul. 2000. *Kant on Freedom, Law and Happiness*. Cambridge: Cambridge University Press.
- Hegel, Georg Wilhelm Friedrich. 2009. *Grundlinien der Philosophie des Rechts*. In *Gesammelte Werke*, Bd. 14.1. Hamburg: Meiner.
- Hoerster, Norbert. 2003. *Ethik und Interesse*. Stuttgart: Reclam.
- Höffe, Otfried. 1977. Kants kategorischer Imperativ als Kriterium des Sittlichen. In *Zeitschrift Für Philosophische Forschung*, 31(3): 354–384.
- Höffe, Otfried. 2002. Menschenwürde als ethisches Prinzip. In: *Gentechnik und Menschenwürde. An den Grenzen von Ethik und Recht*. Hrsg. Otfried Höffe u.a., 111–141. Köln: DUMONT.
- Horn, Christoph. 2014. *Nichtideale Normativität. Ein neuer Blick auf Kants politische Philosophie*. Berlin: Suhrkamp.
- Hruschka, Joachim. 2002. Die Würde des Menschen bei Kant. In *Archiv für Rechts- und Sozialphilosophie* 88 (3): 463–480.
- Kant, Immanuel. 2004. *Vorlesung zur Moralphilosophie*. Berlin, New York: De Gruyter.
- Kersting, Wolfgang. 1983. Der kategorische Imperativ, die vollkommenen und die unvollkommenen Pflichten. In *Zeitschrift für philosophische Forschung* 37 (3): 404–421.
- Kleingeld, Pauline. 2017. Contradiction and Kant's Formula of Universal Law. In *Kant-Studien* 108 (1): 89–115.
- Korsgaard, Christine M. 1996a. Kant's Formula of Universal Law. In: *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- Korsgaard, Christine M. 1996b. Kant's Formula of Humanity. In: *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- Korsgaard, Christine M. 1996c. *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Korsgaard, Christine M. 2007. Autonomy and the Second Person Within: A Commentary on Stephen Darwall's The Second-Person Standpoint. In *Ethics* 118 (1): 8–23.
- Korsgaard, Christine M. 2009. *Self-Constitution. Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- Korsgaard, Christine M. 2018. *Fellow Creatures. Our Obligations to the Other Animals*. Oxford: Oxford University Press.
- Korsgaard, Christine M. 2021. Valuing Our Humanity. In: *Respect. Philosophical Essays*. Hrsg. Richard Dean and Oliver Sensen, 171–191. Oxford: Oxford University Press.
- Kortetmäki, Teea. 2017. Applying the Capabilities Approach to Ecosystems. In *Environmental Ethics* 39 (1): 39–56.
- Müller, Nico Dario. 2022. *Kantianism for Animals. A Radical Kantian Animal Ethic*. New York: Palgrave Macmillan.
- Nussbaum, Martha C. 2010. *Die Grenzen der Gerechtigkeit. Behinderung, Nationalität und Spezieszugehörigkeit*. Berlin: Suhrkamp.
- Nyholm, Sven. 2015. Kant's Universal Law Formula Revisited. In *Metaphilosophy* 46 (2): 280–299.
- Parfit, Derek. 2011. *On What Matters*. Volume 1. Oxford: Oxford University Press.
- Schlosberg, David. 2007. *Defining Environmental Justice. Theories, Movements, and Nature*. Oxford: Oxford University Press.
- Schönecker, Dieter. 1999. *Kant: Grundlegung III. Die Deduktion des kategorischen Imperativs*. Freiburg: Alber.
- Schönecker, Dieter / Wood, Allen W. 2011. *Kants „Grundlegung zur Metaphysik der Sitten“. Ein einführender Kommentar*. Paderborn: Schöningh.
- Sensen, Oliver. 2011. *Kant on Human Dignity*. Berlin/Boston: De Gruyter.

- Sensen, Oliver. 2014. Universalizing as a Moral Demand. In *Estudos Kantianos* 2 (1): 169–184.
- Skidmore, James. 2002. Skepticism about Practical Reason: Transcendental Arguments and Their Limits. *Philosophical Studies* 109: 121–141. <https://doi.org/10.1023/A:1016217224647>
- Stemmer, Peter. 2002. Moralischer Kontraktualismus. In *Zeitschrift für philosophische Forschung*, Bd. 56 (1): 1–21.
- Willaschek, Marcus. 1992. *Praktische Vernunft. Handlungstheorie und Moralbegründung bei Kant*. Stuttgart: Metzler.
- Winter, Eduard. 1932. Die Entwicklung der Auffassung B. Bolzanos von der Willensfreiheit. Dargestellt auf Grund des handschriftlichen Nachlasses Bolzanos. In *Philosophisches Jahrbuch* 45: 483–499.
- Wood, Allen W. 1998. Humanity as an End in Itself. In: *Kant's Groundwork of the Metaphysics of Morals. Critical Essays*. Hrsg. Paul Guyer, 165–187. Lanham, Boulder, New York: Rowman & Littlefield.
- Wood, Allen W. 2006. The supreme principle of morality. In: *The Cambridge Companion to Kant and Modern Philosophy*. Hrsg. Paul Guyer, 342–381. Cambridge: Cambridge University Press.
- Wood, Allen W. 2008. *Kantian Ethics*. Cambridge: Cambridge University Press.