Rational Animal Ethics

Effective in Means, Consistent in Ends

By

Stijn Bruers

Rational Animal Ethics: Effective in Means, Consistent in Ends

By Stijn Bruers

This book first published 2024

Ethics International Press Ltd, UK

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

Print Book ISBN: 978-1-80441-941-0

eBook ISBN: 978-1-80441-942-7

Table of Contents

List of tables

Short description

Our treatment of non-human animals is plagued by moral illusions. These illusions are persistent moral judgments that violate our deepest moral values. Rational ethics attempts to avoid such moral illusions, such that our ends and values become consistent and the means to achieve those ends become most effective. This book constructs a rational animal ethic. The first part presents a consistent moral theory of non-discrimination and equality that includes animals. It starts with the most fundamental principle in ethics, to avoid unwanted arbitrariness. This principle delivers a strong foundation for animal rights and equal treatment of all sentient beings. These rights can be derived from a new, mild welfarist ethic that is less demanding than classical welfarist or utilitarian theories, and more concerned about welfare than classical rights-based or deontological theories. After the theory, the second part deals with the practice of animal ethics. In particular it discusses the most effective means to promote animal welfare and rights. Topics in high-impact practical animal ethics include effective animal advocacy, veganism, veganmodernism, the welfare footprint, alternative protein, wild animal suffering, insect welfare, the harm cascade and herbivorising predators.

About the Author

Stijn Bruers has a PhD in theoretical physics (thermodynamics of ecosystems), a PhD in moral philosophy (animal ethics) and a PhD in economics (behavioral economics cost-effectiveness analyses). He is president of Effective Altruism Belgium and founded the Center for Rationality and Ethics. He is involved in animal rights activism, won the SKEPP-price for critical thinking and wrote more than 50 opinion pieces and 10 books.

Introduction

What is rational animal ethics?

Most people care about animal welfare. Most people believe that factory farmed chickens have a negative welfare. Yet, those people (in high income countries) still eat factory farmed chicken meat. But they are not able to personally grab and kill a chicken without feeling emotional stress. Yet, they are able to grab and kill a cauliflower without feeling discomfort. When they do kill an animal, they have an urge to say "Sorry", but not when they kill a lettuce. They do not have a problem seeing a documentary about fruit cultivation, but feel very reluctant to watch a documentary about slaughterhouses. When they eat chicken meat, they would get angry against a vegan at the table who says to them nothing more than what they already believe: "That piece of meat on your plate comes from a chicken that suffered." Yet, when eating a tomato, they would not get angry and not start giving justifications but instead frown their eyebrows when someone says to them with exaggerated accusations and demands: "Did you know tomatoes are round and juicy and when you cut them like that they lose their shape and juice and that is highly disrespectful and immoral. You should stop it right now! It is forbidden to cut tomatoes!" Those people also think that eating dogs or cats is disgusting, inappropriate and impermissible, even when those animals would not be factory farmed. Yet, none of them is able to properly explain (without contradictions or circular reasoning) why it is allowed to eat chickens and cows but not dogs and cats. As children, before having seen Disney movies, they would pet dogs, cats, chickens, pigs, calves and other animals, but not pet carrots and other plants. They can quickly develop an emotional connection with a chicken, but not with a broccoli. They would feel upset when they see a chicken or rabbit being killed, but not when grains are harvested. Their parents don't want them to see undercover videos from slaughterhouses, yet those parents feed them chicken meat.

This does not make any sense. It appears as if people eat things that they do not really want to eat. Eating meat seems to be irrational, as there are healthy, tasty, cheap and sustainable animal-free alternatives to meat available. The case of meat consumption is a prime example of irrationality in

our relations with and attitudes towards animals. But also many vegan animal welfare advocates and animal rights activists hold opinions and ideas about animals that do not make much sense, for example when it comes to the welfare of wild animals and the problem of predation. It seems like we are all vulnerable to moral illusions: persistent intuitive moral judgments that violate our deepest moral values.

This book aims for a rational animal ethic: a relationship with and behavior towards animals that best fits our strongest moral values. Rational animal ethics is rational ethics applied to the question of how to treat animals. 'Rational' is the opposite of 'irrational', not of 'emotional' or 'subjective'. There is certainly room for subjective emotions in rational ethics. Our moral emotions and intuitions can be valid and can play an important role, but we cannot always trust them because they can lead to cognitive biases and fallacies. A rational ethicist wants to avoid those cognitive biases or moral illusions, by applying critical thinking in ethics.

Rational ethics can be summarized by the slogan: "consistent in ends, effective in means". Hence, this book deals with the choices of ends (the intrinsic values we want to pursue) and means to achieve those ends. That explains why this book has two parts.

The first part, about consistency in ends, involves more abstract or theoretical animal ethics and presents some fundamental ideas in moral philosophy, such as unwanted arbitrariness, mild welfarism, dual moral theory and moral illusions.

The second part, about effectiveness in means, involves more applied or practical animal ethics and presents the most effective things we can do to improve the state of animals. Borrowing ideas from effective altruism, it discusses effective animal advocacy, veganism, veganmodernism, the moral footprint, the harm cascade, wild animal suffering and the predation problem.

What is rationality?

Consistent in ends, effective in means, accurate in beliefs. These are the three tenets of rationality. In technical terms, they refer to axiological, instrumental and epistemic rationality respectively.

Axiological rationality deals with our ends: our intrinsic values and ultimate goals we want to pursue. We value intrinsic values for their own sake. These values should be consistent or coherent, just like the laws of nature. We can have multiple values, just as our universe has multiple forces such as gravity and electricity. Our values may even conflict with each other, just like gravity and magnetism may act in opposite directions. Different forces working in different directions does not yet make our universe inconsistent. What is not allowed in the laws of nature, is for example a gravitational force that is ambiguous or that both pulls and pushes at the same time. Or an object that has several masses at the same time. Or having three objects A, B and C such that A is heavier than B, B is heavier than C and C is heavier than A. The same goes for our values. They should be consistent in the sense that they should not contain contradictions or ambiguities.

Instrumental rationality deals with our means: the options that we should choose to reach our ends. It would be irrational to pick some means that are counterproductive and do more harm than good. Rationality requires that we choose the most effective means, i.e. the means with which we can respect, promote or satisfy most of our strongest values.

Epistemic rationality deals with our beliefs: our knowledge of the world. In order to find the most effective means, we need an accurate and reliable view of the world. We need science such that our beliefs are in line with reality.

In short: if goals are not coherent, means are not effective, or beliefs are not reliable, then there is irrationality.

Rationality has one crucial characteristic: it abhors problematic arbitrariness. To understand rationality, we first have to understand arbitrariness. Arbitrariness means choosing an option from a set of options, without using a rule. For example, randomly picking one or a few elements of a set, as if you are choosing blindly. “Arbitrary” is the opposite of “rule-based” or “regular”. To test whether there is arbitrariness, we can ask the question: “Why choosing X and not for example, Y or Z?” If this question is meaningful, which means that Y and Z belong to the same set or category as X (and are therefore not something completely different), if the question is non-trivial, which means that Y and Z are not simply “non-X”, and if this

question cannot be answered by a rule which does not explicitly refer to X, then the choice for X was arbitrary.

Arbitrariness is not always problematic. Sometimes arbitrariness is unavoidable. Sometimes you have to make a choice even though you have no good rule or decision procedure to help you make that choice. Arbitrariness becomes problematic only when it is avoidable, and it involves some conflict or dispute. Because there are three components of rationality, there are also three ways in which arbitrariness can be conflicting or disputed. Conflicting arbitrariness means that the consequences of the arbitrary choice of ends are experienced as undesirable by at least one individual[1], when the consequences of the arbitrary choice of means violate at least one end, or when the conclusions of the arbitrary choice of evidence for a belief are contradicted by at least one piece of evidence. In short, our ends, means and beliefs become irrational when they contain avoidable problematic arbitrariness.

Examples of irrational ends are dictatorship and discrimination. Dictators arbitrarily select themselves as the ones who are allowed to make all decisions, and discriminators arbitrarily select a group of individuals who get certain rights, at the exclusion of others. The groups can be anything: living beings of a certain species or people with certain physical characteristics, beliefs, preferences, descents or places of birth. The choice of group always contains arbitrariness, because you can always ask the question why this group was chosen and not another (except if the group includes everything and everyone and there is therefore only one group). An example of discrimination that will be discussed at length in this book, is speciesism: discrimination based on perceived[2] species membership.

Second, to achieve your goals, you must choose means. When you choose a means arbitrarily or blindly, when you have at least one goal that is harmed by that means, and when it is possible to choose another means that better

---

[1] In this book, an 'individual' always refers to a sentient being, which can be a human or non-human animal. A 'person' most often refers to a human, but can sometimes also refer to a non-human animal. Similarly, 'individuals' refers to sentient beings, whereas 'people' most often refers to humans, but occasionally also to non-human animals.

[2] There is no unique, objective way to classify individuals into species. Hence, the notion of species will be a 'perceived' species, i.e. as perceived by the speciesist.

respects or satisfies your goals, then your choice of means contains avoidable problematic arbitrariness. Then your choice of means is irrational. Perhaps you choose your means based on your gut feelings. It seems as if you are following a rule: choose the means that feels intuitively right to you. But this still contains arbitrariness, because you chose to listen to your gut feeling instead of someone else's gut feeling. Choosing means based on gut feelings can therefore quickly become irrational, because there is a high probability that a randomly or intuitively chosen means conflicts with at least one of your goals and that with a little more thought you can choose a better means that better respects all of your goals.

Third, to find the effective means, you need reliable knowledge. All too often, we choose ineffective means because of inaccurate beliefs that are based on pseudoscience. Some characteristics of pseudoscience are: using anecdotal evidence (i.e. not considering all relevant data and reliable counter-evidence), cherry-picking (selecting a few studies and giving arbitrarily less weight to other, conflicting studies), rejecting a scientific consensus (placing an arbitrary higher trust in experts who hold a minority opinion) and ad hoc reasoning (choosing a random explanation, without being able to give a good justification for it, or adjusting a theory in a few arbitrary places without being able to explain why those places in particular may be adjusted). These examples of pseudoscience show that it is based on conflicting or disputed arbitrariness. One arbitrarily chooses data or studies that are in conflict with other scientific evidence, or arbitrarily trusts experts who disagree with the majority of experts.

Also religion, like pseudoscience, is plagued by arbitrariness. There are thousands of possible gods (God, Yahweh, Allah, Brahma, Zeus, Jupiter, Osiris, Iluvatar, Krishna, Ame-no-Minakanushi, Kukulkan, Huitzilopochtli,…), and the (lack of) evidence for the existence of those gods is equal. So what would make God so special that only he would exist? In fact, all theistic believers are 99,9999…% atheist: every believer does not believe in all the other possible gods. It is arbitrary to believe in just one of the possible gods that have equal evidence. We can avoid this arbitrariness by either believing in all possible gods at once, but that is not feasible, or by not believing in any god. The same goes for the holy scriptures: why would the Bible be the basis of morality, instead of for example the Bhagavat Gita? It is arbitrary to pick one book as holy if there are many possible holy books

and if all those holy books have equal evidence. Scientific books have more evidence than religious books, so we can pick scientific books non-arbitrarily. Hence, religious beliefs are irrational, and rationality leads to atheism.

This book is about rational ethics, which means it focuses on the first two tenets: consistency in ends and effectiveness in means. Elements of epistemic rationality (accurateness in beliefs) will also be addressed in the second part of this book about effective means.

Part 1

Consistency in ends

The first part of this book contains theoretical, philosophical ideas in animal ethics. It presents the most fundamental ethical principles and discusses their implications for our treatment of non-human animals. The aim is to construct a coherent ethical system based on a consistent set of ends and values.

The first, anti-arbitrariness principle is a procedural principle: for all choices we make, from what to eat to which moral theory to adopt, we have to follow the procedure to avoid unwanted arbitrariness as much as possible. This implies for example that we should not discriminate against animals, i.e. no speciesism (discrimination on the basis of species membership).

The second, mild welfarist principle is a substantive principle: welfare is crucially important, except in certain specific cases. We should improve everyone's welfare, but everyone has the right to discount someone else's welfare under specific conditions. As animals are sentient beings and hence experience a welfare, this theory of mild welfarism has direct implications for how we treat animals. It implies for example that we should not use animals as a means against their will, and should not breed animals for food, not even if the animals have a positive welfare.

Both the procedural and substantive principles have a certain duality in common, because they involve two parties: the decision-makers versus the affected individuals and the right-holders versus the duty-bearers. The dual moral theory shows how many ethical principles, rules and values are unified.

The fourth chapter of this first part on theoretical animal ethics discusses moral illusions. In particular, six moral illusions that appear in our relationships with animals and distract us away from a rational ethic, are presented.

Chapter 1

Unwanted arbitrariness

The anti-arbitrariness principle is the most important principle in ethics.[1] Simply put, it says that everyone who makes a choice has to avoid unwanted arbitrariness as much as possible. This is the most fundamental principle in ethics, in the sense that it applies to all choices, including the choice of moral theory. This anti-arbitrariness principle is the reason why dictatorship and discrimination are morally wrong, why rights are universal and why you should not do unto others what you don't want done unto you (the golden rule). No moral principle may contain unwanted arbitrariness, except if it is really unavoidable.

But what is unwanted arbitrariness exactly?

Unwantedness means being incompatible with someone's largest consistent set of that individual's strongest subjective preferences. A subjective preference is a conscious value judgments or evaluation that has a subjective strength (to be distinguished from e.g. a mere unconscious behavioral disposition). For example, being told a lie is incompatible with a preference for knowing the truth. If something is not logically compatible with the largest consistent set of your strongest preferences, it cannot be consistently wanted by you. Everything that is compatible with that set, can be consistently wanted by you. Saying that you cannot consistently want something is the same as saying that you can reasonably object to it. If you want something because you incorrectly believe it is good for you whereas in fact it is inconsistent with the largest consistent set of your strongest preferences, then you cannot consistently want that thing.

Arbitrariness means selecting an element (or subset) of a set without using a selection rule. A selection rule is a rule that logically determines the selection. It is an if-then statement that consists of a set of conditions. For example: "If element X has conditions A and B or not C, then select X." If the question "Why selecting element X instead of element Y?" has no answer that refers to a selection rule (for example if the only answer is "Therefore!"), then selecting X is arbitrary.

Combining the above definitions of unwantedness and arbitrariness, we can define *unwanted arbitrariness as making a choice without following a rule, whereby the consequences of that choice cannot be consistently wanted by at least one individual*. Here, a choice can be defined as a conscious decision. Making a choice means consciously selecting an element from a choice set, the set of eligible options.

To avoid arbitrary exclusion of choices, this anti-arbitrariness principle applies to all possible choices, including very specific actions ("Sit at seat 5 on bus 42 at 1 pm Friday"), to more general choices ("Use public transport"), to justifications ("Take a seat when the seat is empty and you paid for a ticket"), to higher level moral choices ("Choose the action allowed by a contractualist ethic"), to even very basic choices of premises and logical deduction rules used in justifications ("Use deontic logic to determine the validity of an argument"). It also even applies to the choice of selection rule itself ("Pick a color from the color spectrum, then take a ball of that color"). For practical reasons, we do not have to consider impossible choices ("Avoid unavoidable unwanted arbitrariness").

But the above mentioned anti-arbitrariness principle does not yet say what happens if we don't avoid unwanted arbitrariness. Also, the 'as much as possible' hints at the possibility that sometimes unwanted arbitrariness may not be avoidable. Therefore, we can give a more exact formulation of the anti-arbitrariness principle, in a strong and a weak version.

Anti-arbitrariness principle, universal formulation, strong version. If you do not avoid avoidable unwanted arbitrariness when making a choice, you are not allowed to make that choice.

The weak version can be derived from this strong version. Suppose unwanted arbitrariness is unavoidable. You necessarily have to make a choice that involves unwanted arbitrariness. What about other people making other choices? Are you allowed to determine the choices of others, that is to impose your choice on them? Are you allowed to choose who may make the choice? Choosing yourself as the dictator who dictates the choices of others would involve unwanted arbitrariness again. To avoid this new unwanted arbitrariness of dictatorship, you are not allowed to be the dictator. You have to accept the choices made by other people.

Anti-arbitrariness principle, universal formulation, weak version. If you cannot avoid unwanted arbitrariness when making a choice, you are allowed to make that choice, but other people may make other choices from the same choice set (i.e. you have to tolerate that other people make other choices).

The above formulations are universal, in the sense that everyone and everything must abide by this principle. No arbitrary exceptions are allowed. The principle applies to everyone and everything that is able to make choices based on selection rules. It also applies for example to artificial intelligent machines. Of course, when someone cannot make a choice, that is an exception, but not an arbitrary exception because it is justified using an 'ought implies can' rule: "If you cannot do something, you have no obligation to do it."

We can give another, personal formulation of the anti-arbitrariness principle:

For every choice you make, you have to be able to give a justification rule such that you and everyone can consistently want that everyone follows that rule in all possible (including hypothetical) situations (i.e. you and everyone can accept the consequences of a universal compliance by everyone of the justification rule).

This is a personal formulation, because it refers to what you can personally want. Hence, this formulation applies to everyone who is not only able to make choices, but also able to want something, i.e. someone with personal preferences. Whereas the first, universal formulation referred to selection rules, this second, personal formulation refers to justification rules. A justification rule is a selection rule that is used in moral reasoning, to justify one's choices to other people. Therefore, a justification rule for (im)permissibility of a choice should be used in a logical deduction. That means a justification rule is basically an if-then statement that consists of a set of conditions: "If conditions C apply, then it is permissible to choose X."

The above personal formulation does not yet say what to do when you are not able to formulate a justification rule. Therefore, as with the first, universal formulation, we have to make this second, personal formulation of the anti-arbitrariness principle more precise. And as with the universal formulation, this personal formulation also comes in two versions, of which the weak one can be derived from the strong version.

Anti-arbitrariness principle, personal formulation, strong version. If, when making a choice, you cannot give a justification rule of which you would accept universal compliance, then you are not allowed to make that choice nor follow that rule.

Anti-arbitrariness principle, personal formulation, weak version. If, when making a choice, you cannot give a justification rule of which everyone would accept universal compliance, then you must accept or tolerate that other people make other choices from the same choice set and follow other justification rules for making those choices.

Notice the difference between the strong and the weak version: the strong version dictates that you should accept universal compliance, the weak version dictates that everyone should accept universal compliance. As the condition of universal acceptance of universal compliance is stricter than the condition of personal acceptance of universal compliance, the second version is weaker.

There are many similarities between the universal and the personal formulations of the anti-arbitrariness principle, such that they can be said to be roughly equivalent.

First, there is a correspondence between the selection rule and the justification rule. The first formulation works with a selection rule to avoid arbitrariness. In the second formulation, arbitrariness is avoided by the justification rule and by the idea that if you may follow that rule in a specific situation, then everyone may follow that rule in all possible situations. Suppose that the "everyone" and "all possible situations" were no requirements. Replacing them by "some people" and "some situations" would introduce arbitrariness, because arbitrary subsets of the sets of all people and all situations can be chosen.

Second, both formulations look for what can be consistently wanted. The condition "everyone can consistently want that everyone follows that rule in all possible situations" is the opposite of unwanted arbitrariness. Suppose you choose option A arbitrarily and other people are in a position P in which they cannot consistently want that arbitrary choice. If we consider everyone and all possible situations, this includes the situation where those people choose A and you are in the same position P that those people were in, in which case you cannot consistently want A.

A third similarity between the two formulations is that they both come in a weak and a strong version. Unwanted arbitrariness may not always be avoidable, because there may always be someone who cannot consistently want a choice that cannot be based on a selection rule. Similarly, it may not be possible to find a justification rule of which everyone can accept universal compliance. In these cases, people must tolerate that other people make other choices.

A final similarity between the two formulations of the anti-arbitrariness principle, is that they both apply to all possible choices, including the choice of selection and justification rules (in particular the choice of conditions in those rules). That means a selection meta-rule should be given to select the selection rule from the set of all selection rules. Similarly, a justification meta-rule should be given that justifies the chosen conditions in a justification rule. With the application to all possible choices and the resulting necessary inclusion of such meta-rules (and higher order meta-meta-rules), the anti-arbitrariness principle becomes in a sense the most fundamental principle in ethics.

An example might give some clarification. Consider the situation of taking a seat on the bus. If you choose to take a seat, the rule could be: "If you are white, you may take the seat," or "If you have permission by person X, you may take the seat." But the choice of these conditions is arbitrary (they refer to an arbitrary skin color or person). A better rule would be: "If the seat is empty and you have permission by the people who have a special relationship with the seat, you may take the seat." We have to specify what counts as a special relationship. This can again be done by considering relationships of which everyone can consistently want that they are part of the conditions in the justification rule. Examples of such a special relationship could be 'being the owner of the bus' or 'having reserved the seat'. Having permission could mean 'having paid for a ticket'.

Why is unwanted arbitrariness bad?

If your choices involve some avoidable unwanted arbitrariness, you indicate that unwanted arbitrariness is permissible. That means an arbitrary other person is allowed to choose for some unwanted arbitrariness as well.

In particular, that other person is allowed to arbitrarily do something that another arbitrary person, for example you, do not want, without needing justification (because a justification rule was not necessary). As you cannot possibly want to be treated in such a way, you cannot consistently want a world where unwanted arbitrariness is permissible and is not avoided as much as possible. In other words: when you permit avoidable unwanted arbitrariness, you have no arguments to defend yourself against you being arbitrarily treated by others. When you are treated arbitrarily in ways that you do not want, you cannot complain. You could use force to prevent such treatment, but then you step outside the realm of ethics. The realm of ethics consists of moral arguments and not the law of the strongest.

Does the choice for the anti-arbitrariness principle itself contain avoidable unwanted arbitrariness such that the principle defeats itself? No: it is not even arbitrary. Of course we can always ask the trivial question: "Why be against arbitrariness and not against non-arbitrariness?" But any other nontrivial question becomes meaningless. For example: "Why be against arbitrariness and not against apples or bananas?" Apples and bananas do not belong to the same category as arbitrariness. However, the 'unwantedness' condition involves some arbitrariness: why 'unwanted by at least on individual' and not 'unwanted by everyone' or 'unwanted by no-one'? But this arbitrariness cannot be unwanted by anyone. The condition 'unwanted by everyone' is too strict and will make the anti-arbitrariness principle relevant in only a few, trivial cases. The 'unwanted by no-one' entails that unproblematic arbitrariness is never allowed, and that is unsatisfactory as well.

How to avoid unwanted arbitrariness?

We can study unwanted arbitrariness by the most simple but sufficiently general structure: a choice set containing two elements {X,Y}. One could choose both elements, in which case there is no arbitrary selection of elements (there is only one way to select both elements). Or one could choose one element, either X or Y. This allows room for arbitrariness: if X is chosen, one could ask for the selection rule why X instead of Y is chosen. Finally, one could choose none of the elements, in which case there is no arbitrariness possible. All the options can be grouped together in the power

set of all subsets: the universal set {X,Y}, the singleton sets {X} and {Y} and the empty set {}. This power set has a hierarchy, with several levels:

- Top level (no arbitrariness possible): {X,Y} (the full set of all elements)
- Intermediate level (arbitrariness possible): {X} or {Y} (the subsets of individual elements)
- Bottom level (no arbitrariness possible): {} (the empty set)

Only at the intermediate level is arbitrariness possible. This arbitrariness can be called first-order or horizontal arbitrariness, because there is another, meta-level arbitrariness possible, namely the choice of the level. We can consider the set of levels: {Top level, Intermediate level, Bottom level}. If one chooses the top level without following a selection rule, that choice is arbitrary. This second-order arbitrariness can be called vertical arbitrariness. One could use a selection rule, such as 'choose the level that does not allow for horizontal arbitrariness and contains at least one element', that uniquely selects the top level. Now the choice for the top level is no longer arbitrary (i.e. no vertical nor horizontal arbitrariness), but the choice of the selection rule can be arbitrary, because one could equally choose a selection rule such as 'choose the level that does not allow for horizontal arbitrariness and contains no elements' (which selects the bottom level) or 'choose the highest level where horizontal arbitrariness is possible' (which selects the intermediate level). Hence, there is a third-order arbitrariness. Avoiding this arbitrariness requires a fourth level, where a fourth-order arbitrariness occurs. This indicates that there will always be some arbitrariness: there will always be some level n with an n-order arbitrariness. It is impossible to avoid all arbitrariness.

The issue of vertical and horizontal arbitrariness can be illustrated with the example of the moral community: the subset of all entities in the universe that have moral status (in the sense of e.g. having moral rights). Consider only living beings for simplicity. According to the biological classification, we can classify living beings in a vertical taxonomic hierarchy, with the taxonomic rank 'life' at the top (i.e. the set of all living beings), followed by ranks such as 'domains' (e.g. eukaryotes), 'classes' (e.g. mammals), 'orders' (e.g. primates), 'species' (e.g. humans) and finally the taxonomic

rank 'populations' (races, subspecies) at the bottom. A white supremacist first chooses the lowest level in this hierarchy (the populations or ethnic groups), and then picks a subset at this level (the ethnic group of whites). Similarly, a speciesist first selects the level of the species, and then selects a specific species (e.g. *Homo sapiens*) as the moral community. If no selection rules were followed, these two choices involve respectively vertical and horizontal arbitrariness. We can first ask the non-trivial question: "Why choosing a species and not e.g. a biological order or a phylum?" And at the level of the species, we can ask: "Why choosing *Homo sapiens* (humans) and not e.g. *Sus scrofa* (pigs)?" One could answer: "Because most humans have the capacity for moral thought", but it is possible that this answer also applies to some levels up or down in the hierarchy. If for example there are less than 14 billion primates alive, containing more than 7 billion humans with the capacity for moral thought, then the majority of primates have this capacity. Hence, one could equally well first select the level of orders and then the order of primates. By selecting a biological group as a moral community, it is not easy to avoid double arbitrariness.

Horizontal arbitrariness involves choosing an element from a choice set. One way to avoid unwanted horizontal arbitrariness is by choosing the full set of choices (the top level) or choosing the empty set (the bottom level). However, it may not always be possible to choose the full or the empty set, because of some logical inconsistency. It may also be less desirable to choose the top or the bottom level. This undesirability happens in a general sense when at least someone cannot consistently want the full set or the empty set, or it happens in a more strict sense of 'preference dominance': when those who cannot consistently want the intermediate level also cannot want the top or bottom level, and at least one individual who can consistently want the intermediate level cannot consistently want the top or bottom level (in this case the top or bottom level is preference dominated by the intermediate level). We can categorize the situations where choosing the intermediate level is unavoidable or more desirable. There are four cases.

1) The full set and empty set are impossible: these situations often involve a choice set {do X, don't do X}. Of course, choosing both or choosing neither, is impossible.

2) The full set is impossible, the empty set undesirable (i.e. not wanted by at least someone): consider a choice between moral theories {moral theory X, moral theory Y}. Moral theories, such as a utilitarian welfare ethic and a deontological rights ethic, are based on universal principles. We may have a choice between {maximize total welfare, minimize rights violations}. Respecting both principles of both utilitarian and deontological theories is logically impossible: there are cases when maximizing welfare involves violating someone's rights. Choosing none of the principles and moral theories is not impossible, but it is undesirable, because it is likely that at least someone cannot consistently want an 'anything goes' situation without guiding ethical principles.

3) The full set is undesirable, the empty set impossible: suppose that helping both persons X and Y is impossible, and one faces a choice between {don't help X, don't help Y}. It is possible to choose both, but if both individuals want to be helped, this is less desirable than choosing either one of the options.

4) The full set and the empty set are undesirable. An instructive example is the choice of road traffic laws, such as the choice set: {make driving left permissible, make driving right permissible}. Choosing none of the options implies a prohibition of driving, and there are people who want to drive. Choosing both options results in more unwanted traffic accidents. The rule to drive on the right lane is arbitrary (we can ask the question "Why on the right and not the left?"), it can be avoided (e.g. by allowing to drive everywhere), but this arbitrariness is harmless because no-one cares if everyone collectively decides to drive on one lane instead of the other. The only possibilities to avoid this (horizontal) arbitrariness is to say that we can drive nowhere (neither left nor right) or to say that we can drive everywhere (both left and right), because 'nowhere' and 'everywhere' are not directions such as 'left' or 'right'. We strongly prefer to avoid accidents and we strongly prefer to use a vehicle, so a rule to drive on the right is compatible with our strongest preferences and wants. No-one has a value system that is incompatible with a rule to drive on the right lane. Everyone can consistently

> prefer arbitrariness (to drive on the right lane) above a universal prohibition (to drive nowhere) and a universal permission (to drive everywhere) resulting in chaos and accidents.
>
> Another example is: {eliminate starvation by feeding hungry people, eliminate starvation by killing hungry people}. Hungry people cannot consistently want the empty set, because that means not eliminating starvation. And they do not want the full set either, as that involves killing hungry people.

If choosing the intermediate level is unavoidable or more desirable, we might face horizontal arbitrariness, unless we are able to use a selection rule that selects one of the elements at the intermediate level. We can look for a rule 'If a set of conditions C are satisfied, then choose X instead of Y.' Now the challenge becomes choosing a proper set C of selection rule conditions that everyone can consistently want (otherwise, the choice of the selection rule itself generates unwanted arbitrariness). If such conditions cannot be found, then we have truly unavoidable unwanted arbitrariness.

One starting point for the selection rule could be: 'If choosing X can be consistently wanted by most individuals, then choose X.' It is already possible that everyone can consistently want this condition C that represents the majority criterion. If there remain some individuals who can reasonably object against this majority criterion, then they can propose another criterion (i.e. another set of conditions for the selection rule). Now we face the choice of selecting an element from the set {majority criterion, another criterion}. Choosing both elements (the full set) is impossible, choosing the empty set undesirable. To avoid horizontal arbitrariness, we need another, higher level selection rule that selects either the majority criterion or the other criterion. This process can continue to even higher levels. We can go on as far as is feasible, to minimize unwanted arbitrariness. But the further we go, the more important the choice of a higher level selection rule becomes, the more depends on it, and the harder it becomes to have reasonable objections against the choice. The preferred higher level selection rule becomes so fundamental, that one is likely to have a strong preference for it. It is, for example, difficult to have a stronger preference for another criterion than the majority criterion. That means the majority criterion selec-

tion rule is likely consistent with someone's largest consistent set of that person's strongest subjective preferences.

With the above line of reasoning, we can apply the anti-arbitrariness principle to itself. The choice set involves the two options {avoid unwanted arbitrariness as much as possible, don't avoid unwanted arbitrariness as much as possible}. Choosing both or none of the options is impossible. So we are stuck at the intermediate level, where we can arbitrarily pick one of the two options. But picking the second option (not avoiding unwanted arbitrariness) immediately becomes extremely unwanted. Allowing avoidable unwanted arbitrariness has so many ramifications that it is likely in contradiction with anyone's largest consistent set of strongest subjective preferences. So you cannot consistently want the arbitrary choice for the second option.

To see this in more detail, suppose that you disagree with the anti-arbitrariness principle. You say that avoidable unwanted arbitrariness is permissible. But then you cannot give reasonable counterarguments when I allow unwanted arbitrariness in my moral choices. I may follow arbitrary principles that you cannot consistently want. When I impose my choices on you, you are not able to complain. You are not able to give justified arguments against the imposition of my choices, because you acknowledged that unwanted arbitrariness is allowed, and hence that it is permissible to arbitrarily ignore or violate someone else's largest consistent set of strongest preferences.

If you permit unwanted arbitrariness, I can say to you that your moral values and judgments are not valid. And if you complain and say that your ethical theory is valid, then I can reply that if you are allowed to arbitrarily exclude other moral views and make an ad hoc exception for your own moral rules, then so am I. So I may even make the exception that everyone's moral views should be respected, except yours. All your objections can easily be bounced back by saying: "If you are allowed to arbitrarily do that, then so am I, and so is everyone. What would make you so special that you are allowed to arbitrarily exclude others, but I am not? You should not arbitrarily pick yourself from the set of all individuals and say that you are the only one who may do that thing."

In summary: rejecting the anti-arbitrariness principle while avoiding irrationality is extremely difficult, if not impossible.

The above discussion applies to the cases where the top and bottom levels are impossible or undesirable. There are two other interesting categories to consider.

1) The full set is possible and not clearly undesirable, the empty set is undesirable or impossible. A prime example is the choice set {I decide, you decide}, or {I have a right to vote, you have a right to vote}. Someone has to decide, and at least someone wants to vote, so the bottom level is impossible or undesirable. But choosing the intermediate level and arbitrarily choosing one of the options results in a kind of dictatorship where one person can decide or vote.

2) The full set is impossible or undesirable, the empty set is possible and not clearly undesirable. Here we deal with choice sets such as {harm person A, harm person B} or {privilege A over B, privilege B over A}. It is undesirable to harm both A and B and it is not possible to privilege A over B and B over A at the same time, so the top level is undesirable or impossible. But choosing the intermediate level and arbitrarily choosing one of the options results in a kind of discrimination where one individual is harmed or disadvantaged.

As the anti-arbitrariness principle deals with choices and rules, we are confronted with two important 'who' questions. Who decides or chooses the choices and rules? And who is affected by the choices and rules? These two questions relate to the dual problems of dictatorship and discrimination. The next two sections discuss how the anti-arbitrariness principle implies the non-dictatorship and non-discrimination principles.

Non-dictatorship

The non-dictatorship principle says that no-one should have the unconditional power to always unilaterally make decisions that negatively affect some other individuals. A vote is a power (or right) to influence a decision

(the outcome of a decision process) made by a group, such that the outcome is more in accordance with one's personal preferences. In a dictatorship, there is at least one individual whose vote is excluded from the decision process and who does not want this exclusion. A dictatorship clearly violates the anti-arbitrariness principle, because the choice for the dictator is arbitrary (as the dictator's power is unconditional, no rule was followed to grant that power), and unwanted (there are affected individuals who do not want the decisions made by the dictator). The dictatorial idea of 'one person, all votes' involves an arbitrary choice for that one person (the dictator) and at least one other person prefers the democratic principle of 'one person, one vote' where all persons are treated equally.

Suppose person X wants to make choice A, but person Y cannot consistently want the consequences of that choice, and hence prefers choice B. Instead of the principle 'might makes right', which is a dictatorship of the most powerful, those people can look for other methods to decide who gets to decide. One such alternative method is generating justifications by giving arguments. Instead of the strongest person winning, now the strongest reason, justification or argument wins. The principle that the best argument wins is also arbitrary, just like the principle that might makes right, but it is less likely to be unwanted, because anyone can give the best argument whereas not everyone can be the strongest.

Person X can simply claim: "I, person X, decides." This is the moral rule: "If the person is X, then that person may choose." Person Y does not want that, and counters: "No, person Y decides." The justification rule proposed by person X refers to X, and that choice should be justified as well. So person X can claim the meta-rule: "Person X decides who decides." But here again, person Y can complain, and the meta-rule arbitrarily refers to person X again. This discussion can go on to infinity. For practical relevance, the anti-arbitrariness principle should state that an infinite regression of justification rules is not allowed.

Constructing coherent moral theories

The non-dictatorship principle can also be applied to moral theories or ethical systems. These theories are logical systems of ethical principles that represent moral intuitions or values. There are different moral theo-

ries, such as a deontological rights ethic, a consequentialist utilitarian welfare ethic, a libertarian ethic, a contractarian ethic or pluralist ethics that combine several ethical principles. But which theory should we choose? We cannot choose the moral theory that simultaneously contains all possible principles, because that results in contradictions. The selection of possible, mutually consistent principles from the set of all available principles is always unavoidably arbitrary. But the anti-arbitrariness principle sets strong constraints on the choice of principles for a moral theory. The theory should be coherent in the sense that it should be constructed following some rules. Here are five rules of thumb to construct a coherent system.

1) One should not arbitrarily limit the ethical principles to an arbitrary group of objects, beings or individuals.

2) One should not arbitrarily give weaker (less strongly felt) moral intuitions a stronger priority. One should not arbitrarily change or exclude basic moral judgments.

3) One should not arbitrarily allow inconsistencies and gaps in the ethical system.

4) One should not arbitrarily introduce ambiguous or vague principles that one can interpret and apply arbitrarily in concrete situations.

5) One should not arbitrarily add artificial, complex, ad hoc constructions and exceptions to save the moral theory from counterintuitive implications.

These rules are comparable to the rules of a crossword puzzle, whereby the words correspond to ethical principles. One should fill in the whole puzzle, one should not write words that do not correspond with the given descriptions, one should not leave a white box empty, one should not write more than one letter in a white box, one should not write letters in the black boxes, and one should not invent new words or make ad hoc changes in the spelling of words. A solves crossword puzzle is a coherent system.

These construction rules for a coherent theory can be consistently wanted. If for example you allow inconsistencies, gaps, ambiguities or arbitrary

exceptions in your theory, then you have to accept that someone else's moral theory also contains such things. With such an incoherent theory, everyone else can easily justify choices that you cannot consistently want. An incoherent theory always contains avoidable unwanted arbitrariness that should be rejected.

Unwanted arbitrariness leads to the search for a coherent ethical system. That is the reason why systemization in ethics is so important. We should construct a coherent ethical system based on fundamental, universal principles. This is very similar to the systemization in science and mathematics. Physics theories that describe our universe and axiomatic systems that describe mathematical structures are examples of coherent systems that are consistent and do not contain avoidable arbitrariness. Scientific laws and mathematical axioms have a clarity and are not ambiguous. They are regular and do not contain arbitrary exceptions. For example, in the system of natural numbers, every number, without exception, has a successor that is also a number. In geometry, you can draw a line between every two points, without exception. In physics, the law of gravity applies everywhere in the universe, and all electrons have an electric charge. Note that all those axioms and laws contain words like 'every' and 'all'. That is why they are called universal principles.

Why is it bad to construct incoherent ethical systems that contain inconsistencies or ad hoc arbitrariness? If you choose to follow an incoherent theory, everyone else is allowed to reject that theory and impose their own theories on you, and you are not able to complain. You are not able to give reasonable or justified counterarguments against the imposition of other ethical principles, because by following your incoherent theory, you are acknowledging that unwanted arbitrariness and hence arbitrary exclusion is allowed. That means it is also permissible to arbitrarily exclude your moral theory and ignore your moral views and ethical principles. You can only give a valid complaint or argument if you accept the anti-arbitrariness principle. Without that principle, any critique becomes invalid, and complaints become impossible.

Let's analyze this in more detail. Suppose you believe in an incoherent ethical system, i.e. a system that contains either inconsistencies (internal contradictions) or avoidable arbitrariness.

Suppose that your ethical system is inconsistent. In that case, you acknowledges that it is permissible to have inconsistencies. If you are allowed to have two opposing views at once, then so are others. Other people are allowed to both have the belief that they cannot simply reject your moral beliefs and ethical system, and at the same time the belief that they can reject the ethical system held by you. Even if these two beliefs are mutually contradictory, you cannot argue against it, because those people can always say that they do not reject your system, even if they do, and they are allowed to have contradictions if you may have contradictions. Or if you don't want others to impose their moral views on you but at the same time say that you are allowed to have inconsistencies, i.e. to have two opposing views at once, then others may simply hold the inconsistent opinion that they may and may not impose their moral views on you. If they then impose their views on you, that is permissible.

Suppose that you hold a consistent but arbitrary ethical system, where the ethical principles include avoidable arbitrary exceptions. Now we can follow a similar strategy: if you prefer an ethical system that contains avoidable arbitrariness, then you acknowledge that such arbitrariness is permissible. That means that it is also permissible to arbitrarily ignore someone else's moral views and ethical systems. Other people can say to you that your moral values and judgments are not valid. And if you complain and say that your moral theory is valid, then they can reply that if you are allowed to arbitrarily exclude other moral views and make an ad hoc exception for your own moral theory, then so are they. So now they may even make the arbitrary exception that everyone's moral views should be respected, except your moral views. Making an exception for you is arbitrary, because they do not have to give a reason why they exclude you. And you are not able to give arguments against that arbitrary exclusion, because all arguments can easily be bounced back by saying: "If you are allowed to arbitrarily do that, then so are we. What would make you so special that you are allowed to arbitrarily exclude others, but we are not?"

This shows that incoherent ethical systems are not allowed. We can reject such incoherent ethical systems even arbitrarily, without having to give a reason, because people who adopt an incoherent system acknowledge that arbitrary exclusions or rejections are permissible. People can only give a valid complaint or argument if they accept the anti-arbitrariness principle.

Without that principle, any critique becomes invalid. The impossibility to complain if one has an incoherent ethical system implies that coherent ethical systems gain a more objective or absolute status.

To avoid dictatorship, everyone is allowed to construct a coherent moral theory that best fits one's moral intuitions and values. Incoherent theories are impermissible. But there are many possible coherent moral theories, just like there are many different coherent number systems and geometrical systems in mathematics. In some number systems, every number has two instead of one successor (e.g. the Gaussian numbers). In some geometrical systems, one can draw several straight lines between two points (like the Earth has many meridians running from the north pole to the south pole).The same goes for ethical systems that have different fundamental principles or ethical axioms.

We do not have a rule that determines which of the many coherent moral theories is the best. If we are against unwanted arbitrariness, we have to recognize that every equally coherent moral theory is equally valid. You cannot say that your coherent theory, based on your moral intuitions and valuations, is better than someone else's if both theories are equally coherent. You prefer your theory, but you cannot impose your theory upon others who have a coherent theory, because what would make you so special that you would be allowed to do that? And the same goes for everyone else. No-one is special. It would be an unwanted kind of arbitrariness if you claim that your moral theory is special without good reason. A rational ethicist is tolerant towards all other coherent ethical systems, no matter how much they go against one's own moral intuitions.

Picking one of the coherent moral theories always involves unavoidable arbitrariness. The non-dictatorship principle says that we should democratically choose which moral theories to apply. And if you follow a coherent moral theory without being able to give a justification rule that selects that theory, you should tolerate that other people follow other coherent moral theories. If you choose a coherent rights-based deontological ethic, you should tolerate someone else choosing an equally coherent welfare-based utilitarian ethic. We should be tolerant towards all other coherent ethical systems, no matter how much they go against our own moral intuitions.