

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 7 |
| 1.1 | Skepticism about ethics | 7 |
| 1.2 | Starting points for ethics | 8 |
| 1.3 | Defining morality | 8 |
| 1.4 | Methods | 9 |
| 2 | Well-being | 11 |
| 2.1 | Good lives vs. Bad lives | 11 |
| 2.2 | Prudential value, instrumental value, intrinsic value | 12 |
| 2.3 | What we are trying to do | 13 |
| 2.4 | Why should we care? | 13 |
| 3 | Hedonism | 15 |
| 3.1 | Introduction | 15 |
| 3.2 | Hedonism | 15 |
| 3.2.1 | Two forms of pleasure | 16 |
| 3.3 | Arguments for hedonism | 17 |
| 3.3.1 | Best explanation for a particular correlation | 17 |
| 3.3.2 | Best explanation for the overlap between welfare subjects and pleasure/pain beings | 17 |
| 3.3.3 | Best explanation for motivation | 18 |
| 3.3.4 | Compatible with the experience requirement | 18 |

| | | |
|------------|--|-----------|
| 3.3.5 | Sensitive to human differences | 19 |
| 3.4 | Arguments against hedonism | 20 |
| 3.4.1 | The experience machine argument | 20 |
| 3.4.2 | The false beliefs argument | 21 |
| 3.4.3 | The argument from autonomy | 22 |
| 3.4.4 | Base pleasures argument | 22 |
| 3.4.5 | Base pleasures equality argument | 23 |
| 3.4.6 | Argument from harm | 23 |
| 3.5 | Conclusion | 24 |
| 4 | Desire-fulfilment theory | 25 |
| 4.1 | Introduction | 25 |
| 4.2 | Arguments for DFT | 26 |
| 4.2.1 | The correlation argument for DFT | 27 |
| 4.2.2 | The experience machine argument for DFT | 27 |
| 4.2.3 | The individual-dependence argument | 27 |
| 4.2.4 | The motivation argument | 28 |
| 4.3 | Arguments against DFT | 29 |
| 4.3.1 | Scope problem: Something good for you but you do not desire it | 29 |
| 4.3.2 | Scope problem: Satisfying a desire but with no change to your life | 29 |
| 4.3.3 | Scope problem: Desiring something bad for you | 30 |
| 4.4 | Objections to arguments against DFT | 30 |
| 4.5 | Arguments against IDFT | 31 |
| 4.5.1 | Lack of predictive value | 31 |
| 4.5.2 | IDFT is worse than DFT | 31 |
| 4.6 | Conclusion | 32 |
| 5 | Objective list theories | 33 |
| 5.1 | Introduction | 33 |
| 5.2 | Objective List Theory | 33 |
| 5.2.1 | Points of clarification | 33 |
| 5.2.2 | Developments of OLT | 34 |
| 5.3 | Arguments for objective list theory | 34 |
| 5.3.1 | The experience machine argument for OLT | 34 |
| 5.3.2 | The argument from scope | 35 |
| 5.3.3 | The argument from intuitions (desire-independent goods) | 35 |
| 5.3.4 | The argument from particular theories | 36 |
| 5.4 | Arguments against objective list theory | 36 |
| 5.4.1 | The argument from arbitrariness | 37 |
| 5.4.2 | The argument from alienation | 38 |
| 5.5 | Conclusion | 39 |

| | | |
|-----------|--|-----------|
| 6 | Ethics and Religion | 40 |
| 6.1 | Introduction | 40 |
| 6.2 | Argument 1: Religious belief and moral motivation | 40 |
| 6.3 | Argument 2: Divine Command Theory | 41 |
| 6.3.1 | Objections to DCT | 42 |
| 6.4 | Argument 3: Moral wisdom through God | 43 |
| 7 | Natural Law Theory | 45 |
| 7.1 | The theory and its arguments | 45 |
| 7.1.1 | Argument 1: Objectivity | 45 |
| 7.1.2 | Argument 2: Morality as applying to humans | 45 |
| 7.1.3 | Argument 3: Origin of morality | 46 |
| 7.1.4 | Argument 4: Answers Hume's argument against morality | 46 |
| 7.2 | Conceptions of human nature | 47 |
| 7.2.1 | Human nature as animal nature | 47 |
| 7.2.2 | Human nature is what is innate | 47 |
| 7.2.3 | Human nature is what humans have in common | 47 |
| 7.3 | Natural design and human purposes | 48 |
| 7.4 | Objections to Natural Law Theory | 49 |
| 8 | Psychological Egoism | 50 |
| 8.1 | Introduction | 50 |
| 8.2 | Psychological Egoism | 50 |
| 8.2.1 | Ethical consequences of psychological egoism | 51 |
| 8.3 | Arguments for psychological egoism | 51 |
| 8.3.1 | Argument from desire | 51 |
| 8.3.2 | Argument from expected results | 52 |
| 8.4 | Arguments against psychological egoism | 53 |
| 9 | Ethical Egoism | 54 |
| 9.1 | Ethical egoism explained | 54 |
| 9.2 | Arguments for Ethical Egoism | 55 |
| 9.2.1 | Argument from social welfare | 55 |
| 9.2.2 | Argument from reasons | 56 |
| 9.3 | Arguments against ethical egoism | 57 |
| 9.3.1 | Argument against common sense moral beliefs | 57 |
| 9.3.2 | Argument from lack of difference | 57 |
| 10 | Utilitarianism | 59 |
| 10.1 | Introduction | 59 |
| 10.2 | What is Utilitarianism? | 59 |

| | | |
|-------------|---|-----------|
| 10.3 | Three core arguments against Utilitarianism | 62 |
| 10.3.1 | Argument from injustice | 62 |
| 10.3.2 | Argument from supererogation | 63 |
| 10.3.3 | Argument from partiality | 64 |
| 10.4 | Other arguments against Utilitarianism | 65 |
| 10.4.1 | Argument from the past | 65 |
| 10.4.2 | Argument from distribution | 65 |
| 10.4.3 | From intentions | 65 |
| 10.4.4 | Argument from experience machine | 65 |
| 10.5 | Defending Utilitarianism | 66 |
| 10.5.1 | Response to the argument from injustice | 66 |
| 10.5.2 | Response to impartiality | 66 |
| 10.5.3 | Response to the argument from supererogation | 67 |
| 10.5.4 | A Modification to Utilitarianism: Act vs. Rule | 67 |
| 10.5.5 | Defense 3, Utilitarianism and Common Sense | 68 |
| 10.6 | Utilitarianism: Applications | 68 |
| 11 | Deontology | 70 |
| 11.1 | Introduction | 70 |
| 11.2 | The good will | 70 |
| 11.2.1 | Problem 1: whether the good will is the only thing that matters | 71 |
| 11.2.2 | Problem 2: what does it mean to have a good will | 72 |
| 11.2.3 | Point 6: Returned, the moral law as a categorical imperative | 75 |
| 11.2.4 | The good will: a summary and a theory | 76 |
| 11.2.5 | A problem with the categorical imperative | 77 |
| 11.3 | The principle of humanity | 77 |
| 11.3.1 | Problems with the the humanity version | 78 |
| 12 | Contractarianism | 81 |
| 12.1 | The theory | 81 |
| 12.1.1 | Contractarianism | 81 |
| 12.1.2 | The student's dilemma and the prisoner's dilemma | 82 |
| 12.1.3 | Contractarianism again | 84 |
| 12.2 | Arguments for social contract theory | 84 |
| 12.2.1 | Morality as a social phenomenon | 84 |
| 12.2.2 | Proof of immoral acts even to the egoist | 85 |
| 12.2.3 | Explains moral exceptions | 85 |
| 12.3 | Arguments against contractarianism | 86 |
| 12.3.1 | Argument from free riders | 86 |
| 12.3.2 | Argument from non-consent | 86 |
| 12.3.3 | Argument from the scope of the moral community | 87 |
| 13 | Pluralistic Absolutism | 89 |
| 13.1 | Ethical monism | 89 |
| 13.2 | Ways to object to ethical monisms | 89 |

| | | |
|-------------|--|------------|
| 13.3 | Pluralistic absolutism | 90 |
| 13.3.1 | Argument from disaster prevention | 90 |
| 13.3.2 | Response: the doctrine of double effect | 91 |
| 13.3.3 | The argument from moral conflict | 92 |
| 14 | Prima facie duties and ethical particularism | 94 |
| 14.1 | Ross theory of prima facie duties | 94 |
| 14.1.1 | Arguments for Ross's theory | 95 |
| 14.1.2 | Arguments against Ross's theory | 96 |
| 14.2 | Ethical particularism | 96 |
| 14.2.1 | Arguments against ethical particularism | 97 |
| 15 | Virtue Ethics | 98 |
| 15.1 | Defining moral good | 98 |
| 15.2 | Arguments for virtue ethics | 100 |
| 15.2.1 | The argument from the right amount of precision | 100 |
| 15.3 | Arguments against virtue ethics | 101 |
| 15.3.1 | Virtue theory is circular | 101 |
| 15.3.2 | Virtue ethics leads to contradiction | 102 |
| 16 | Care Ethics | 103 |
| 16.1 | Care ethics | 103 |
| 16.2 | Arguments for care ethics | 105 |
| 16.2.1 | Argument from partiality | 105 |
| 16.3 | Arguments against care ethics | 106 |
| 16.3.1 | The scope of the moral community | 106 |
| 16.4 | Further Readings | 106 |
| 17 | Ethical Relativism | 107 |
| 17.1 | Introduction: moral nihilism and moral relativism | 107 |
| 17.2 | Cultural relativism | 108 |
| 17.2.1 | Argument for cultural relativism from descriptive cultural relativism | 110 |
| 17.2.2 | Argument against cultural relativism: cultural nihilism | 110 |
| 17.3 | Individual ethical relativism (ethical subjectivism) | 111 |
| 17.3.1 | Arguments for individual (and cultural) relativism: from moral facts and knowledge | 111 |
| 17.3.2 | Arguments against relativism: moral disagreement is impossible | 112 |
| 18 | Moral Nihilism | 113 |
| 18.1 | Moral nihilism | 113 |
| 18.2 | Error theory | 114 |
| 18.2.1 | Argument for error theory | 114 |
| 18.2.2 | Against error theory: error theory does not meet the burden of proof | 115 |

| | | |
|-------------|---|------------|
| 18.3 | Expressivism | 116 |
| 18.3.1 | Arguments for emotivism | 116 |
| 18.3.2 | Argument against emotivism: cannot explain the denial of past moral beliefs | 117 |

1. Introduction

This course is divided into three main units.

1. Value theory: what is the nature of well-being? what is worth pursuing for its own sake? what is the nature of prudential value?
2. Normative ethics: what are our fundamental moral obligations? what habits, dispositions, traits are virtuous as opposed to vicious?
3. Metaethics: what are the status of moral propositions? Can propositions concerning morals be true or false? What is the relation between moral truths and our motivation to be moral?

1.1 Skepticism about ethics

Many people express skepticism concerning normative ethics. Shafer-Landau (p.4) offers five arguments against the possibility of there being objective truth in ethics.

1. Argument from disagreement: Notice that there is rampant disagreement about what is morally right and wrong. If there were objective truth, then we would expect, at the very least, that those who inquire into ethics to agree about the ethical truths. But, there is no agreement. Therefore, there is no objective truth.
2. Argument from God's existence: There are universal moral truths if and only if God exists. But God does not exist. Therefore, there are no universal moral truths.
3. Argument from science: The only method that can give objective truths is the scientific method. But, the scientific method does not tell us that moral propositions are true or false. The only explanation for this is that moral propositions are neither true nor false. Therefore, there are no objective moral truths.
4. Argument from non-imposition: If there were objective moral truths, then it would be morally permissible to impose these truths on others. But, it would not be morally permissible to impose these truths on others. Therefore, there are no objective moral truths.
5. Argument from absolutes: If there were objective moral truths / rules, then it would be morally

impermissible to break these rules. But there are always exceptions to every rule. Therefore, there are no objective moral truths.

Exercise 1.1 Form a group of at least four people. In the group, pick one of the arguments against there being objective moral truths. Next, formulate that argument in your own words and create a diagram or example to help illustrate that argument. Finally, put your argument on the board.

1.2 Starting points for ethics

Now that we have considered several arguments that aim at undermining the possibility of objective truths in ethics, it is worthwhile to consider various points that may incline us to believe that there are moral truths.

1. Law and Tradition: We frequently criticize the law and tradition on moral grounds, e.g. unjust laws
2. Fallibility: We acknowledge that sometimes we make moral errors, e.g. If I lie to a friend, I may feel bad because I have wronged her.
3. Exceptions: We think moral rules must be sensitive to who we are as individuals. If a “moral rule” asks us to do the impossible, then it is not a real moral rule.
4. Non-universality: We don’t think moral standards apply to everyone equally, e.g. they apply to adults more than to children, they apply more to those who are well-o
5. Vast moral agreement: We think that heinous acts of violence are morally wrong (or require moral justification)
6. Admiration of moral saints: we admire individuals who engage in heroic acts that involve saving others or charity
7. Rights: we believe that various rights are worth protecting for their own sake, e.g. right to privacy, right to free speech, etc.

None of the above are aimed at proving that there are objective moral truths. Instead, they are aimed at showing that the view that there are no objective moral truths is highly controversial. That is, an individual who believes that there are no objective moral truths would need to show that our criticisms of law and tradition are misplaced, that our purported feelings of moral regret are not based on having done something morally wrong, and so on.

Exercise 1.2 Given the arguments against objective moral truths and the various points that suggest there may be objective moral truths, do you think that there are objective moral truths?

1.3 Defining morality

Much of this course is about morality. In this section, we formulate a working definition of what we will call **critical morality**.

First, it is helpful to distinguish what critical morality **is** from what **it is not**.

Critical morality is not equivalent to

1. the law
2. etiquette (rules of polite behavior)
3. self-interest
4. tradition
5. conventional morality

First, if morality and law were equivalent, then what would be true of the former would also be true of the latter, and vice versa. But this is not the case for consider that there are some things that are

moral but not legal and some things that are legal but not moral.

Exercise 1.3 Can you think of anything that is (i) moral but not legal and (ii) legal but not moral?

Second, if morality and etiquette were equivalent, then what would be true of the former would also be true of the latter, and vice versa. But this is not the case for consider that there are some things that are perfectly moral but against the rules of proper etiquette.

Exercise 1.4 Can you think of anything that is in accordance with the rules of proper etiquette but that is immoral?

Third, if morality and what is in one's self interest were equivalent, then what would be true of the former would also be true of the latter, and vice versa. But this is not the case for consider that there are some things that are perfectly moral but against your own self interest (e.g. sacrificing your life for your country) and things that would benefit your self-interest but are not moral (e.g. assuming you knew you would get away with it, killing a person in order to acquire five dollars).

Fourth, if morality and tradition were equivalent, then what would be true of the former would also be true of the latter, and vice versa. But this is not the case for consider that there are some things that are moral but against traditional practices (e.g. father raising child instead of mother) and some traditional practices that are not moral (e.g. gender stereotyping).

Fifth, if morality and conventional morality (the moral rules adopted by a country, culture, or group) were equivalent, then what would be true of the former would also be true of the latter, and vice versa. But this is not the case for consider that there are some things that are moral but against certain forms of conventional morality (e.g. homosexual relationships) and some things that are conventional moral but are not moral (e.g. slavery)

Definition 1.1 — critical morality. Critical morality is a set of propositions (sentences capable of being true or false) that are not equivalent to (i) law, (ii) etiquette, (iii) standards of self-interest, (iv) tradition, and (v) conventional morality that state (i) whether an action is permissible (allowed) or impermissible (not allowed) and (ii) the scope of one's obligations

In short, critical morality tells us how we ought to live but does not depend upon legal rules or a number of social conventions and thus can be used to criticize both law and social behaviors.

In contrast to critical morality is normative ethics. Here we define normative ethics as the science that aims to what is morally right and wrong.

Definition 1.2 Ethics is a branch of philosophy that inquires into standards of right (morally good) and wrong (morally bad) conduct. Ethics as an academic discipline relies on common experience and critical activity to discern the truth or falsity of certain claims about moral rightness and wrongness. Insofar as it (i) aims to discern the truth of things, (ii) relies on experience, and (iii) employs reason to do so, ethics is a science (although distinct from sciences like physics, chemistry, and biological in that it does not rely on special tools or observations to discern its results).

In short, you can think of critical morality as a set of true propositions about what is morally right and wrong, while normative ethics is a field of inquiry that aims to identify these true propositions.

1.4 Methods

The manner in which we will approach various theories is an **argumentative** one. The presentation will typically involve three steps.

First, a theory about values, morality, or the nature of ethics will be presented. The goal in presenting the theory will be to work toward a definition of a theory and to illustrate this theory with examples.

Second, the theory will be supported by arguments. Topics concerning values, morality, and the nature of ethics are not uncontroversial and so they need to be supported with reasons. For example, suppose Tek thinks that all charity is morally acceptable. He might support this view as follows:

P1: Helping others is always morally acceptable.

P2: Charity is a form of helping others.

C: Therefore, charity is morally acceptable.

In the above, Tek has not merely asserted that charity is morally acceptable but has, instead, supported this view by giving reasons (indicated by P1 and P2) in the form of an argument.

Third, either (i) various arguments put forward by the theory will then be critically evaluated or (ii) various objections will be directed at the theory itself.

Criticism of the arguments will occur in at least one of two ways. First, it may be the case that the premises (reasons) are true and the conclusion is false. This way of criticizing the argument is to point out a problem in the logic of the argument, and involves showing that the conclusion does not follow. Second, it may be the case that the premises (reasons) are false. This way of criticizing the argument does not assert that the conclusion does not follow but instead that the premises are not true (or that there are reasons to believe that the premises are not true).

Objections to the theory itself usually take the following form. First, the theory is presented. Second, it is argued that some implausible consequence follows from the theory. Third, the theory is rejected because it implies something this implausible consequence.

P1: If theory T is true, then X is true.

P2: X is not true.

C: Therefore, theory T is false.

At times, we may add a fourth, fifth, or sixth step. A fourth step would involve responding to various criticisms of the theory (found in the third step) either by showing problems in the critical arguments or by revising the theory to make it resistant to those arguments. Critics of the theory can then turn around and raise new objections to either the responses to the theory or the new, revised theory. Generally, however, due to space, we will ignore these additional steps.

2. Well-being

In this first unit, we will investigate what it means for one's life to go well versus go badly. We will try to craft theories that give a general account of what it means to say that a life is going well versus going badly.

2.1 Good lives vs. Bad lives

We may know (or can at least imagine) individuals whose life has gone (or is going) well versus those whose life has gone (or is going) poorly.

Example 2.1 — Tek's miserable life. Tek grew up in crippling poverty. This poverty caused him not to be able to fulfill his dreams. In addition, it caused him not to be able to acquire medication to deal with a very painful medical condition. He suffered most of his life and was angry about his suffering. He often took his anger out on others and this ultimately resulted in Tek dying, alone and in pain.

Example 2.2 — Liz's wonderful life. Liz grew up with modest means but never had any desires /wants that she couldn't meet either through hard work or by chance. She dreamt of discovering a new dinosaur in South Dakota, a dream she ultimately fulfilled. Liz has a cheery disposition and that caused her to meet many people, some of whom become very good friends. She found love in her late 20s, after her career was started, had two children, and is happy with her partner, children, and career. She has had a few setbacks in life but nothing she hasn't been able to endure.

Example 2.3 Think of two examples of people, one whose life you know is going well and another you know whose life is going poorly. Try to describe each as much as possible.

Comparing Tek to Liz, intuitively we think that Liz's life has gone better than Tek. In saying this, we are not saying:

1. that Liz is a better person from (a productive point of view) than Tek (she might be but that isn't what we are trying to say)

2. that Liz is a better person (from a moral point of view) than Tek (she might be but that isn't what we are trying to say)

Let's consider each of these in turn.

First, we are not considering how much an individual contributes to society. Yes, Liz may be more productive than Tek, but this does not entail her life is going better than his. If Liz were hyper-productive but stressed and suicidal, we might say she is extremely productive but her life is going poorly. In contrast, if Tek is lazy and mooches off of society, but is not sick, is always in a good mood, and fulfills his limited desires, we would say he is unproductive but his life is going well. Our focus is on how well things are going for a person, not how well they live up to economic-societal standards.

Second, in investigating what it means for a life to go well, we are not considering (at least directly) what it means for an individual to do ethical good or obey ethical rules. In investigating what it means for a life to go well, we are not concerned with which individuals are moral saints or sinners for it is possible (at least at the outset) that an individual who is evil lives a good life, and possible (at least at the outset) that an individual who is morally good has a bad life. Our focus is on how well things are going for a person, not how well they live up to moral standards.

Exercise 2.1 Try to describe an individual who is unproductive and unethical but has a good life.

2.2 Prudential value, instrumental value, intrinsic value

Our investigation will be into what it means for a life to go well. In contrasting Liz and Tek, we are saying that Liz's life has something (or more of something) that Tek's life does not. Her life has more prudential value. She has more of whatever it is that makes her life go well.

Definition 2.1 — prudential value. Prudential value refers to a property of something that is good for a person, and thus it is a property that makes one life go better than another. It is what makes a life go well or better than another.

Prudential value often goes by a number of other names:

- well-being
- welfare
- utility

Thus, all of the following are equivalent:

“being healthy increases an individual's well-being” is equivalent to “being healthy is prudentially valuable to an individual's well-being”.

“being healthy increases public welfare” is equivalent to “being healthy is prudentially valuable to the public”

“being healthy has more utility than being unhealthy” is equivalent to “being healthy has more prudential value than being unhealthy”

In discussing the nature of prudential value, it is important to make a distinction between instrumental value and intrinsic value.

The first is between intrinsic and instrumental value. First, something is instrumentally valuable if and only if it has value for acquiring something else of value. For example, money has no real worth on its own; its value is purely instrumental insofar as it can be used to obtain other things of value.

Definition 2.2 — instrumental value. Something is instrumentally valuable if and only if it has value for acquiring something else of value.

Second, something is intrinsically valuable if and only if it is valuable in itself. That is, a thing has intrinsic value if and only if all other things are done for the sake of that thing. For some people, friendship might be considered intrinsically valuable in that they value being friends with people not because of how it makes them feel or for status or for some further reason.

Definition 2.3 — intrinsic value. Something is intrinsically valuable if and only if it is valuable in itself.

One way of getting at the types of things that might be regarded as intrinsically valuable is to ask a series of why or what for questions.

Example 2.4 Tek: Why do you want A?
 Liz: Because having A would allow me to acquire B.
 Tek: Well why do you want B?
 Liz: Because having B would allow me to acquire C?
 Tek: Well why do you want C?
 Liz: For no other reason than I think C is valuable in itself.
 Tek: Ah! Then C has intrinsic value to you.

2.3 What we are trying to do

In investigating what it means for a life to go well, our aim is to develop a theory of prudential value. We want to be able to look at a life, and use our theory to say if a life is going well or if it is going poorly. We want to know what it is that makes a life go better than another. We want to know what we have to acquire in order to make our lives better.

In putting forward this theory, we are concerned not with what is instrumentally valuable but what is intrinsically valuable. Thus, our goal is to articulate a theory of intrinsic prudential value.

2.4 Why should we care?

There is a relatively simple and straightforward reason why we would want a theory of prudential value. The reason is this: if we think having a good life is better than having a bad life (prudentially speaking), it is reasonable (all other things being equal) to try to increase that which makes lives go better, but we don't have a clear sense of what exactly makes one life go better than another.

Thus, a theory of prudential value would be beneficial for:

1. yourself
2. parents of children
3. political officials
4. anyone with a dependent who cannot make decisions for themselves
5. economists

The primary obstacle to constructing a theory, however, is that we do not know what makes one life go better than another.

Exercise 2.2 For each of the five individuals, try to imagine how their lives might be going very poorly:

1. Tek, super rich
2. Liz, tons of friends
3. Jon, super productive
4. Ana, super famous
5. Ted, super great doctor

3. Hedonism

3.1 Introduction

In this chapter, the hedonist theory of well-being is articulated, argued for, and then critically evaluated.

3.2 Hedonism

People often say “do whatever makes you happy” or “you should do what makes you happy”. This sort of advice seems to imply that all that matters is happiness. One way of understanding happiness (although not the only way) is to equate it with pleasure. Thus, we can view people's suggestions about how we ought to live as we ought to do whatever increases our pleasure in life and decreases our pain.

Let's call one's relative balance of pleasure and pain, one's hedonic level .

Definition 3.1 — hedonic level. The relative balance of one's pleasure and pain

And so, people are saying that a good life is one that has a high hedonic level and a bad life is one that has a low hedonic level.

Definition 3.2 — the hedonist theory of well-being. The hedonist theory of well-being asserts that (i) the only thing having any prudential value is pleasure while the only thing having prudential disvalue is pain and (ii) a person's well-being is entirely determined by their amount of pleasure relative to their amount of pain.

The hedonist theory thus makes at least three claims. First, it asserts only pleasure and pain are intrinsically valuable. Second, it asserts that only pleasure and pain are the only things that are

relevant for determining an individual's well-being (pleasure and pain are the only things with prudential value). Third, it asserts that one's well-being is determined by the relative balance of pleasure and pain.

3.2.1 Two forms of pleasure

In making pleasure and pain the only things that have intrinsic value, one question that immediately arises is what is meant by pleasure and pain.

First, let's distinguish between two different forms of pleasures: pleasures of the intellect and pleasures of the flesh. Let's say that pleasures of the intellect are pleasures that somehow depend upon our cognitive capacities or the development of those cognitive capacities, whereas pleasures of the flesh are pleasures derived from our mostly non-cognitive, physical capacities.

Second, let's assert that certain pleasures of the intellect are greater in terms of quality than pleasures of the flesh. That is, the former pleasures make more of a contribution to our well-being than the latter.

Finally, let's consider two arguments put forward by John Stuart Mill in support of this qualitative form of hedonism (namely the view that pleasure is the only thing having intrinsic value and pleasures of the intellect are of greater quality than pleasures of the flesh).

The first argument is that individuals who are acquainted with both of the pleasures would (more often than not) choose the pleasures of the intellect over the pleasures of the flesh.

Second, Mill contends that except in cases of extreme depression human beings would never consent to be turned into a creature that was fully satisfied by the pleasures of the flesh.

“Few human creatures would consent to be changed into any of the lower animals, for a promise of the fullest allowance of a beast's pleasures; no intelligent human being would consent to be a fool, no instructed person would be an ignoramus, no person of feeling and conscience would be selfish and base, even though they should be persuaded that the fool, the dunce, or the rascal is better satisfied with his lot than they are with theirs.”¹

Thus, the hedonist can deny resist the misinterpretation that our well-being only has to do with physical pleasures and say that the quality of our pleasures also matter when evaluating our well-being.

Exercise 3.1 Are pleasures of the intellect (pleasures that depend upon your higher, developed capacities) of a greater quality than pleasures of the flesh? Let's test this claim. In a group, list a few different things that make you (or others) feel good. Next, consider Mill's claim that assuming we are experienced with both kinds of pleasures, if we had to choose, we would always choose pleasures that depend upon our higher capacities. Turning to your list of pleasures, how do you rank them? Would you be content to live a life wholly occupied by the pleasures your ranked highest on the list in complete absence of those lower on the list rather than the reverse scenario (live a life wholly consumed by the pleasures lowest on the list in absence of those you ranked highest)?

¹Elsewhere Mill writes “It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied.” Mill saw this feature of the theory to be a part of the Epicurean theory for he writes “But there is no known Epicurean theory of life which does not assign to the pleasures of the intellect, of the feelings and imagination, and of the moral sentiments, a much higher value as pleasures than to those of mere sensation.”

3.3 Arguments for hedonism

3.3.1 Best explanation for a particular correlation

The first argument for hedonism is that it offers the best explanation for the correlation of individuals with hedonist levels and individuals with high levels of well-being.

P1: If we take the individuals who have high hedonic levels, it is also the case that these individuals have a high level of well-being.

P2: If we take the individuals with high levels of well-being, it is also the case that these individuals have a high hedonic levels.

P3: The best explanation of this correlation is that one's hedonic level determines one's level of well-being.

C: Thus, hedonism is true.

Not only does the truth of hedonism give us an explanation of this correlation, it would give us an explanation of several other phenomena:

1. if an individual's pleasure increases (all other things being equal), they are better-off than they were before. This is because increasing a person's hedonic level makes them better off
2. if an individual's pain increases (all other things being equal), they are worse-off than they were before. This is because decreasing a person's hedonic level makes them worse off

Objection 3.3.1 — Correlation but not causation. While one's hedonic level may be correlated with one's well-being, it does not follow that one's hedonic level causes or determines one's well-being. For consider that:

1. there are completely spurious correlations. There is a correlation between the US spending on science, space, and technology and the number of suicides by hanging, strangulation, and suffocation. This a correlation with the two variables have nothing to do with each other.²
2. the causal variable can be misidentified in a correlation. There is a correlation between driving at 56mph and my speedometer reading 56mph (the two are correlated), but it would be wrong for me to assume that my speedometer reading 56mph causes my car to move at 56mph.
3. there causal variable may not be a part of the correlation. Suppose there is a correlation between being of a particular race and being a criminal. This might lead some to assert that being of a particular race makes it more likely (causes you) to commit crimes. But this would be mistaken since it is another set of variables (perhaps relating to education, location, poverty, etc.) that is responsible for a group committing crimes.

Applying the above to the correlation argument for hedonism, it may be the case that (i) pleasure /pain and well-being have nothing to do with each other, (ii) the two are correlated but being better-off causes one to have high levels of pleasure or (iii) some third factor determines both high-levels of pleasure and being better off.

3.3.2 Best explanation for the overlap between welfare subjects and pleasure/pain beings

Let's consider another argument. Consider that the types of beings that can have well-being (e.g. people, animals, etc.) are also the same types of beings that can experience pleasure/pain. The hedonist might argue that this should count as evidence for the hedonist theory being true.

Objection 3.3.2 — one property among many. The evidence is too weak. Consider that the types of beings that can have well-being (e.g. people, animals, etc.) are also the same types of beings that can form social bonds. Using the same type of reasoning invoked by the hedonist, one could argue that the social bonds determine well-being.

²For a number of amusing spurious correlations, see [Spurious Correlations](#)

3.3.3 Best explanation for motivation

Another argument for hedonism is based on the idea that people are egoistic (self-interested) in nature and that the only thing that can motivate an individual to act is pleasure. In short:

P1: Human beings aim to maximize their personal self-interest

P2: An individual is only motivated to act by that which will bring them pleasure or remove pain.

C: Therefore, hedonism is true.

In support P1, individuals often seem to be preoccupied with activities that will advance their interests. People vote for politicians that reflect their political interests. People aim to advance their status in the workplace. And, very often, people are reluctant to help others unless there is "something in it for them".

With respect to P2, it often seems true that people are motivated to increase their own personal pleasure and decrease their own pain. If I poke you in the arm with a needle, you will be motivated to remove it. And, more generally, people often seek to remove of x things that cause them distress, whether it be conflicts in the workplace, arguments with friends /families, psychological issues.

Objection 3.3.3 — The argument isn't valid. It is possible for the premises to be true and the conclusion false and so the conclusion doesn't logically follow from the premises. We can imagine that individuals do aim to maximize their self-interest and they are solely motivated by that which brings them pleasure and pain, but hedonism is false (something other than pleasure would make the life go better).

For example, suppose that it is in my self-interest to brush my teeth (P1) but I'm not motivated not to do it since I find it boring (painful) (P2), but I'm forced to do it by my parents. Arguably I would be better off even though my hedonic level decreased.

Objection 3.3.4 — P1 is false. Individuals often engage in pure acts of charity, viz., they undertake activities that do not benefit them at all.

Objection 3.3.5 — P2 is false. Individuals are often motivated to act in ways that do not take into account whether it will increase their pleasure or decrease their pain.

Exercise 3.2 Consider the objections to P1 and P2. In a group, write down three original examples that help to illustrate these objections.

3.3.4 Compatible with the experience requirement

Another argument for hedonism is that it is compatible with the experience requirement on any theory of well-being. That is, it might be argued that any theory of well-being / prudential value ought not to conflict with certain requirements or conditions, one of which being the experience requirement.

Definition 3.3 — Experience requirement on well-being. The requirement that any increase or decrease in an individual's well-being must be correlated with a change in their experience.

The implication of this requirement is that you can't be better-off or worse-off without some experience that contributes to this change in well-being. For example, Tek cannot be well-off one day and worse-off the next without at least some change in how he feels.

The argument then from hedonism is the following:

P1: Any theory of well-being must meet the experience requirement

P2: The hedonist theory meets the experience requirement because to have an increase/decrease

in pleasure and/or pain is to have a certain type of experience.

C: Therefore, the hedonist theory is true.

Objection 3.3.6 — Invalid. The conclusion does not logically follow from the premises. Simply meeting one of the requirements is not sufficient for it is possible that other theories may meet the experience requirement. For example, Fletcher (p.14) points out that perhaps having a feeling of self-respect meets the experience requirement and may be a contributing factor to one's well-being. Thus, what is needed is the premise that "no other theory of well-being can meet the experience requirement" but this is not clearly proven.

The hedonist thus needs a stronger version of P2. Namely, rather than P2, they need P2*: the hedonist theory is the only theory that meets the experience requirement.

Objection 3.3.7 — the experience requirement is false. We can imagine situations where an individual may be better or worse off without knowing or experiencing that they are better or worse off.

Tek may not pay attention to his salary and his employer may be skimming portions of Tek's check. Tek may not feel better or worse because of this but we would regard Tek as being worse off before the employer began stealing from Tek.

3.3.5 Sensitive to human differences

We might imagine that what makes an individual better off will vary somewhat. That is, there is variation concerning what makes individuals better off.

Example 3.1 Tek may be better off with 100K, married, and working in a competitive industry while Liz may be better off with 50K, single, and working in a setting where social relationships come before work life.

Let's formulate the following general requirement for a theory of well-being. A requirement that we will call the individual-dependence (variation) requirement on well-being.

Definition 3.4 — individual-dependence requirement on well-being. The individual-dependence requirement on well-being is a requirement for any theory of well-being. This requirement states that every theory of well-being must take into account the fact that what makes one individual better off might not be the same thing that makes another individual better off.

One argument for hedonism then is that hedonism meets the individual-dependence requirement. This requirement states that any theory of well-being must build in a kind of variation depending upon features found in the person. Hedonism does this by noting that the things that bring one person pleasure (and thus increase their well-being) may not bring another person pleasure.

Example 3.2 Tek finds pleasure in making 100K, being married, and working in a competitive industry and not find pleasure in making only 50K, being single, and working in a setting where social relationships come before work life.

We can thus formulate the argument for hedonism as follows:

P1: Any theory of well-being must meet the individual-dependence requirement

P2: The hedonist theory meets the individual-dependence requirement because there is variation concerning what increases an individual's hedonic level.

C: Therefore, the hedonist theory is true.

Objection 3.3.8 — Invalid. Again, much like the experience-requirement, the conclusion does not logically follow from the premises. Simply meeting one of the requirements is not sufficient for it is possible that other theories may meet the experience requirement

The hedonist thus needs a stronger version of P2. Namely, rather than P2, they need P2*: the hedonist theory is the only theory that meets the individual-dependence requirement.

3.4 Arguments against hedonism

In the previous sections, objections to specific arguments in support of hedonism were considered. These objections aimed to show that the arguments put forward in support of hedonism are not convincing. In this section, objections to hedonism are considered. These objections aim to show that hedonism as a theory of well-being is false.

3.4.1 The experience machine argument

One objection to hedonism involves considering a thought experiment. That is, it involves considering a hypothetical example to test the consequences of hedonism. Let's consider this thought experiment and then consider the argument against hedonism:

Suppose there were an experience machine that would give you any experience you desired. It would maximize your pleasure and minimize your pain. Or, if you think that experiencing pain is necessary ingredient to experiencing pleasure, it would give you enough pain to maximize your pleasure. You would not actually be doing the things that cause you to feel this pleasure. Instead, you would be floating in a tank, with electrodes attached to your brain. Of course, if you find the idea of being hooked up to an experience machine distressing, the machine would cause you to forget that you were hooked up. While in the machine, you would believe you were in the real world. Others can also plug in to have the experiences that they want, so there's no need to stay unplugged to serve them. (Ignore problems such as who will service the machines if everybody plugs in.)³

If we put the idea of the experience machine and hedonism together, there are two consequences:

1. according to hedonism, pleasure and pain are the only things that have any intrinsic value; thus, if the experience machine produced more pleasure than pain, it would be rational to hook up to the experience machine and irrational not to hook up.
2. according to hedonism, if the experience machine produced the same amount of pleasure as "real" life, there would be no rational reason why you should prefer "real" life over the experience machine.

Some find these consequences implausible. Namely, they contend that there is a rational reason for preferring real life over the experience machine even if the experience machine The argument can be put straightforwardly as follows:

P1: If hedonism is true, then there would be no rational preference for "real" life over a machine that produces experiences (the experience machine).

P2: There is a rational preference for "real" life over the experience machine.

C: Therefore, hedonism is false.

The experience machine objection thus criticizes the hedonist idea that only pleasure contributes to our well-being. If pleasure was all that mattered, then we would be perfectly content with having our pleasures produced by a machine. However, so the objection goes, other things matter and this is drawn out by our preference for real experiences over those produced by a machine.

³Based on Robert Nozick, "The Experience Machine," pp. 42-5 from *Anarchy, State and Utopia*. New York: Basic Books, 1974.

Exercise 3.3 The experience machine objection contends that there is a rational reason for preferring real life over the experience machine (see P2). While many have the intuition that there is a rational reason, it is notoriously difficult to express this rational reason clearly. Nozick contends “First, we want to do certain things, and not just have the experience of doing them.” In addition, “we want to be a certain way, to be a certain sort of person.” Is there a rational reason to prefer real life over the experiences produced by the experience machine? If so, what is it?

How might they hedonist respond to the above argument?

Objection 3.4.1 — status quo bias. A status quo bias is a bias (or preference) for the current state of affairs (the status quo) and involves drawing the inference that any change from the status quo is negative regardless of the change's merit. Simply, it is a preference for the status quo because it is the status quo.

The hedonist might object that individuals prefer real life over the experience machine simply because it is the status quo.

Objection 3.4.2 — pain. While we may intuitively prefer pleasure caused by the real world over pleasure caused by the experience machine, if there were a rational reason to prefer life outside of the experience machine, then we should prefer real pain over experience-machine pain. But this is not the case. We find pain to be equally bad no matter what its source.

Objection 3.4.3 — evolutionary bias. The hedonist might argue that the only reason we would prefer the experiences found in real life over those in a machine is due to an evolutionary bias. The idea here is that there are a number of false beliefs that are the result of evolutionary processes that may be advantageous for survival or the propagation of the species. For example, suppose I am more inclined to believe that someone is watching me even when no one is watching me. This may be an advantageous inclination as it would make me hyper-aware of predators (even though it may routinely produce false beliefs).

The hedonist may argue that our belief that real experiences, real friendships, living in the real world have intrinsic value is nothing more than a false belief produced by a non-truth-tracking evolutionary process.

3.4.2 The false beliefs argument

What the experience machine argument seems to show is that there is something unrelated to pleasure and pain that is important to well-being. One difficulty for the non-hedonist, however, is articulating what has prudential value beyond pleasure and pain. Perhaps one suggestion is that what is important is being right, viz., having true beliefs.

Example 3.3 Suppose that Jon and Tek are identical in every way except for one. That is, they both have the exact same level of happiness but while Jon believes his wife is faithful and she is, Tek believes his wife is faithful and she is not. One objection to hedonism then is that even if they have the same hedonic levels and thus should have the same level of well-being, Jon is better off than Tek.

P1: If hedonism is true, then two individuals with the same hedonic levels necessarily have the same level of well-being

P2: It is not the case that two individuals with the same hedonic levels necessarily have the same level of well-being for sometimes having certain true beliefs increase one's well-being even when those true beliefs do not increase our hedonic levels.

C: Therefore, hedonism is false.

What supports P2 then are cases where individuals have the same hedonic levels but one individual has true beliefs while the other has false beliefs.

3.4.3 The argument from autonomy

Again what the experience machine argument seems to show is that there is something unrelated to pleasure and pain that is important to well-being. And, again, one difficulty for the non-hedonist, however, is articulating what has prudential value beyond pleasure and pain. Perhaps one suggestion is that what is important is being autonomous, viz., having the capacity to self-legislate.

Definition 3.5 — autonomy. Autonomy refers to the capacity of an individual to determine the principles or plan by which they wish to live.

Example 3.4 Suppose that Tek has a greater hedonic level than Liz. However, suppose the reason that Tek's hedonic level is so high is because he regularly takes a designer drug that cause him to make decisions that increase his overall level of pleasure. On the other hand, Liz prefers to make her own decisions, even if they would not increase her overall level of pleasure. The non-hedonist might contend that Liz is actually better off than Tek, even if Tek has a greater hedonic level, simply because Liz is autonomous.

P1: If hedonism is true, then two individuals with the same hedonic levels necessarily have the same level of well-being

P2: It is not the case that two individuals with the same hedonic levels necessarily have the same level of well-being for sometimes being autonomous increases one's well-being even when being autonomous does not increase our hedonic levels.

C: Therefore, hedonism is false.

What supports P2 then are cases where individuals have the same hedonic levels but one individual is autonomous but the other is not autonomous.

3.4.4 Base pleasures argument

Another argument against hedonism is that the theory asserts that all pleasures contribute to our well-being. But, the argument is that this is false since there are a number of pleasures that do not contribute at all to an individual's well-being. That is, there are a number of pleasurable experiences that do not make our lives better off at all. Fletcher (p.19) provides three such pleasures:

1. evil or immoral pleasures (e.g. pleasure from torturing someone)
2. base pleasures (e.g. pleasure from eating a whole tub of ice cream)
3. undignified pleasures (e.g. pleasure from sulking or being sad)

Thus, the argument against hedonism from base pleasures might be formulated as follows:

P1: If hedonism is true, then all pleasures increase well-being

P2: If all pleasures increase well-being, then evil, base, and undignified pleasures increase well-being.

P3: Either evil, or base, or undignified pleasures do not increase well-being.

C: Therefore, hedonism is false.

Objection 3.4.4 — base pleasures do increase well-being. The hedonist can object by denying P3. Namely, the hedonist can argue that the reason we think base pleasure does not increase well-being is because we are preoccupied with any of the following:

1. the possible painful future consequences of the actions, or
2. painful side effects of having a character trait that derives base pleasure from certain types of actions

For example, the individual who derives pleasure from torture may one day be caught, and we think of the pain associated with being caught and so infer that the act of torture does not make the torturer better off. Or, the individual who derives pleasure from torture is likely mentally ill and we

imagine that being mentally ill is painful.

Nevertheless, the hedonist objects that while these may factor into how better or worse off the individual is overall (their hedonic level), they do not undermine the fact that the pleasure derived from torture makes the torturer better off.

Objection 3.4.5 — confusion of moral value with prudential value. The hedonist might respond that one reason people believe P3 is true is because they confuse moral value with prudential value. But the hedonist will argue that this is mistaken because things can be morally evil but have prudential value for the individual. For example, suppose Jon was once living a happy life but now he is suffering. He needs an organ transplant or he will die. He kidnaps another individual and steals that individual's organs. His act is, no doubt, evil, but Jon is certainly better off than he was before. He will go on living another 30 years and living happily.

Thus, while an individual might argue that base pleasures do not make one better off, they may simply be judging those pleasures morally rather than whether they do, in fact, increase the well-being of the individual. That is, they are thinking about what would be best overall to do rather than what is best for the individual to do.

3.4.5 Base pleasures equality argument

Another type of objection to hedonism involves comparing different types of pleasures. The argument is that hedonism is false because of its view on how different sorts of pleasures contribute to an individual's well-being. That is, it seems to imply that two equal-sized pleasures will contribute to an individual's well-being to the same extent. But this is false since other things beyond pleasure play a role in determining how a pleasure impacts a person's well-being.

More precisely:

P1: If hedonism is true, then all pleasures of equal size have the same sized contribution to an individual's well-being.

P2: If same-sized pleasures make a same-sized contribution to well-being, then a base-pleasure and a non-base-pleasure of the same size will contribute equally to well-being.

P3: Base pleasures do not contribute as much as non-base pleasures of the same size.

C: Therefore, hedonism is false.

To illustrate, suppose you could spend ten hours sitting in your room sulking or five hours learning to play music. Furthermore, suppose that the contribution each makes is the same size in terms of the pleasure it produces. According to hedonism, both of these should contribute the same amount to your well-being. They would make an equal contribution to how better-off you would be. But, the argument contends (at P3) that this is false. You would be better off learning to play music than sitting around moping.

Objection 3.4.6 — good for you vs. good overall. The argument confuses what is good for the individual versus what is good overall, e.g. good for society.

Objection 3.4.7 — intrinsic vs. instrumental. The argument confuses instrumental and intrinsic value. That is, we might suppose that learning to play music might be instrumentally valuable for other obtaining other pleasures. And so, it is more instrumentally valuable for your future well-being to learn to play music. But this does not take away from the fact that both, at the time, have the same contribution to the individual's well-being.

3.4.6 Argument from harm

If hedonism is true, then an individual can only be harmed (either physically, emotionally, or frustrated in terms of satisfying their goals) if the supposed harm decreases their hedonic level.

Example 3.5 Suppose Tek plays the piano. He has played for years in hopes of becoming a professional piano player. However, Tek is injured and his hands are broken, never to be fully repaired. According to the hedonist, Tek is worse off if and only if this injury decreases his hedonic level.

However, it might be argued that what makes an individual worse off is not the decrease in one's hedonic level but the harm itself. That is, what makes Tek worse off is not that his hedonic level decreases due to the injury. Rather, he is worse off because the injury no longer allows him to pursue his goals or his purpose in life.

P1: If hedonism is true, then an individual can only be harmed (either physically, emotionally, or frustrated in terms of satisfying their goals) if the supposed harm decreases their hedonic level.

P2: Individuals can be harmed in ways not strictly tied to a decrease in their hedonic levels.

C: Therefore, hedonism is false.

3.5 Conclusion

In this chapter, we looked at the hedonist theory of well-being. The hedonist, to put it simply, equates prudential value with one's hedonic level (one's relative balance of pleasure /pain). We saw that there were a number of arguments in favor of hedonism, e.g. the fact that hedonic level is correlated with well-being, that it applies to non-human beings capable of having well-being, that it follows from truths about the motivation of human beings, and that it meets the experience requirement. In addition, we saw that the hedonist theory faces a number of serious challenges, e.g. that it is susceptible to the experience machine objection and that it makes the implausible claim that all pleasures (regardless of their type) increase well-being.

4. Desire-ful llment theory

4.1 Introduction

Hedonism contends that increasing one's pleasure and minimizing one's pain is what is ultimately valuable. It thus takes no serious account of the relation one has with one's goals, desires, or aims . That is, suppose Tek wants to get an A on his philosophy paper. Assuming he would feel the same way about cheating and getting an A versus writing the paper and getting an A, hedonism cannot say that one option is better than the other. If both produce the same amount of pleasure, both contribute the same to how well Tek's life is going.

According to the desire-ful llment theory, the focus on pleasure as the only good overlooks the complexity of how one's well-being is related to one's goals. The desire-ful llment theory of prudential value asserts that ful lling or satisfying one's desires makes you better o , while not ful lling those desires makes you worse o .

De nition 4.1 — Desire-ful llment theory (DFT). The desire-ful llment theory contends that (i) something is intrinsically good for a subject S because it ful lls an intrinsic desire of S and (ii) something is intrinsically bad for a subject S because it interferes with (or obstructs) the ful llment of the intrinsic desire of S.

First, note that DFT refers to intrinsic (or non-instrumental) desires.

De nition 4.2 — intrinsic desire. An intrinsic desire is something you desire for its own sake rather than for the sake of something else.

Suppose Tek wants a girlfriend to give the appearance of being well-liked. In this example, Tek has an instrumental desire for a girlfriend. He desires a girlfriend for the sake of something else. Now suppose that Tek wants to be well-liked and not for some additional purpose. It is one of his ultimate desires. In this example, the desire to be well-liked would be intrinsic for Tek. He simply wants to

be well-liked and not for some further reason.

Second, note that DFT asserts that something is prudentially good for you because it causes you to fulfill your desires. This theory is different from hedonism which says that the only thing that is prudentially good for you is pleasure. For example, suppose that hard work helps Tek buy a home and feel safe (the feeling of safety is one of his intrinsic desires). According to DFT, hard work is a trait that makes Tek better off because it causes him to fulfill his desires. It isn't merely an instrumental good that is only good because it yields to the ultimate good of pleasure.

Third, one of the more interesting aspects of DFT is that what matters is the actual fulfillment of the desire rather than the perceived fulfillment. That is, according to DFT, what matters is that the world corresponds to the way that you desire it, not whether you believe or feel that the world corresponds to the way that you desire it.

Example 4.1 You have a child named Renna. You haven't heard from her in two months. You desire that she is safe. You also desire to know that she is safe because not knowing creates anxiety in you. You receive a phone call from her and she tells you that she is fine. You feel a sense of relief. According to DFT, you are better off now because she is safe and because you know she is safe. Her being safe fulfills your desire and therefore, according to DFT, makes you better off.

But there is also a couple other points worth drawing out of DFT here:

1. you are also better off before you knew she was safe because you (i) desired her to be safe and (ii) she was safe.
2. if she told you she was safe but she was not safe, you would not be better off even if you believed she was safe because (i) you desire her to be safe, and (ii) she is not safe.

This third point provides a real point of contrast between DFT and hedonism. For hedonism, what matters is simply how you feel and that is wholly guided by your perception of things. For DFT, your well-being is partly determined by external factors (the way the world is, how things actually are) since you often desire for the world to be a certain way rather than appear to you to be a certain way.

A fourth point is that DFT – as a general theory of well-being – leaves open various questions about the fulfillment of various desires and their relation to well-being. One such question has to do with weighting or the intensity of the desire-fulfillments. At first glance, we might say that the more intense the desire-fulfillment, the greater the contribution to one's well-being. If my strongest desire is to find love, then fulfilling this desire causes me to have a significantly better life. On the other hand, this might not be the case since perhaps my the fulfillment of my slight /weak desire to have a successful surgery with zero-complications has a greater overall impact on how well my life goes.

Exercise 4.1 One significant feature of DFT is that an individual's well-being is determined by how things actually are rather than merely how the individual perceives them to be. In discussing this distinction, an example is provided. Create your own example that helps to illustrate this same point.

Exercise 4.2— DFT as a theory. We have gone through DFT as a theory, but haven't talked through why you should accept DFT as theory of well-being. Before we do this, take a moment to articulate DFT for yourself and come up with a few examples to help illustrate the theory (that is, use some examples to help make it clear).

4.2 Arguments for DFT

As a theory, DFT has a lot of intuitive force. One of the reasons it has so much force is because how well one's life is going seems to be partially determined by the individual. Who can say how well a

life is going besides the individual themselves? If John is poor, sick, isolated, and in pain but does not desire to be well-off, healthy, or social, or alleviate his pain, it seems presumptuous to say his life is going poorly.

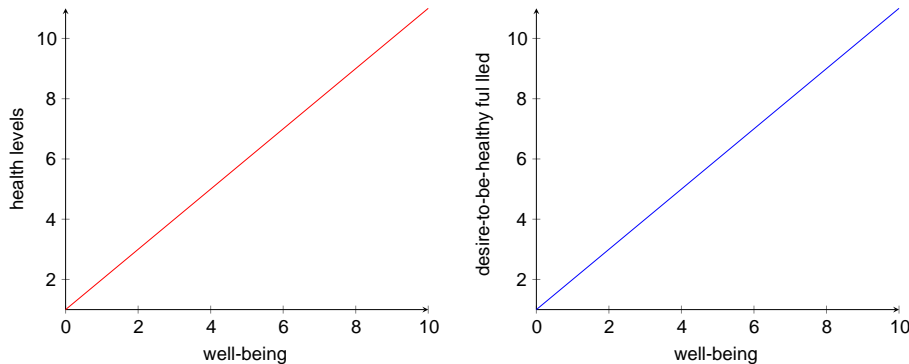
In this section, various arguments are given in support of DFT.

4.2.1 The correlation argument for DFT

The first argument given in support of DFT is that the fulfillment of our desires correlates with things that are straightforwardly good for us. Intuitively, being healthy is better than not being healthy. Individuals who are healthy are better off than those who are not. Now consider that

there are individuals who have their desire to be healthy fulfilled
there are individuals those who have their desire to be healthy frustrated

According to DFT, those that have their desire to be healthy fulfilled are better off than those who have their desire to be healthy frustrated. And so, there is a correlation between the healthy individuals who are intuitively think are better off and those who have their desire to be healthy fulfilled.



4.2.2 The experience machine argument for DFT

A second argument in support of DFT is that it is immune to the experience machine objection.

Recall that the experience machine objection was an objection directed at hedonism. The objection was that if pleasure is the only thing that matters to an individual's well-being then this could be produced in a machine that elicits those experiences. What matters is the feeling one gets from being in a friendship rather than friendship itself. DFT, by contrast, makes the well-being of an individual depend (at least in some cases) on the status of the world. If one desires to have a friend, the fulfillment of that desire depends upon actually being in a friendship rather than feeling that one is in a friendship.

Example 4.2 If Tek desires that Renna is safe, it is not sufficient for his well-being that he feels she is safe or feels happy believing she is safe. Such a feeling could be produced by an experience machine. Instead, what matters to Tek's well-being is that she really is safe.

4.2.3 The individual-dependence argument

We might imagine that what makes an individual better off will vary somewhat. That is, there is variation concerning what makes individuals better off.

Example 4.3 Tek may be better off with 100K, married, and working in a competitive industry while Liz may be better off with 50K, single, and working in a setting where social relationships

come before work life.

Let's call the condition that well-being is partly dependent upon the peculiar features of the individual the individual-dependence requirement on well-being. Recall the individual-dependence requirement.

One argument for DFT then is that DFT meets the individual-dependence requirement. This requirement states that any theory of well-being must build in a kind of variation depending upon features found in the person. DFT does exactly this! What improves an individual's well-being will vary based upon what it is they desire. What is prudentially good for Tek will be what fulfills his desires while what is good for Liz will be what fulfills her desires.

Exercise 4.3 Let's consider a possible criticism of this argument. Consider hedonism and then consider whether it meets the individual-dependence requirement on well-being. If so, what sort of criticism can we make of this argument for DFT?

4.2.4 The motivation argument

Next, let's consider what we will call the motivation argument. Suppose that we add an additional requirement to any theory of well-being. Let's call it the motivation requirement.

Definition 4.3 — motivation requirement on well-being. The motivation requirement on well-being states that any theory of well-being must be able to give a reasonable explanation of the relation between our motivations and what is prudentially valuable.

The motivation argument for DFT is that DFT provides the best explanation for the relation between our motivations and what is prudentially valuable.

P1: If something makes you better off, then (according to DFT) that thing will satisfy your desires.

P2: If something satisfies your desires, then you will be motivated to acquire the thing that satisfies your desires.

C: Therefore, according to DFT, if anything makes you better off, then you are motivated to acquire that thing. And so, DFT can explain the relation between our motivations and what is prudentially valuable.

Let's consider an example.

Example 4.4 Suppose that getting an A in a course would make Tek better off. Any theory of well-being needs to explain the relation between getting an A in a course and one's motivations for action.

DFT can do this for according to DFT, getting an A would satisfy Tek desire to get an A. And, if Tek desires to get an A, then he will be motivated (at least a little) to obtain an A. Thus, DFT can explain the relation between what makes us better off and our motivation.

Objection 4.2.1 — P1 is false. Some might assert that there are things that make us better off that do not satisfy any of our desires.

Objection 4.2.2 — P2 is false. We are not always motivated to acquire the thing that satisfies our desires. Suppose I desire to play QB for Penn State football, but I realize that this will require a tremendous amount of effort and so I instantly lose motivation. Here I may desire to play QB but not have any motivation to satisfy that desire.

4.3 Arguments against DFT

We now turn to several criticisms of DFT. These criticisms draw out counter-intuitive consequences of DFT and reason that because these consequences are false, DFT is a flawed theory.

P1: If DFT is true, then X is the case.

P2: X is not the case.

C: Therefore, DFT is false.

4.3.1 Scope problem: Something good for you but you do not desire it

Let's consider two examples where you do not desire something, but acquiring that thing would be good for you.

Example 4.5 — Jane gets a massage. Suppose Jane could really benefit from a massage or a run or eating better. She does not desire these things, but intuitively they would make her better off. The massage would make her relaxed, the run decrease her blood pressure, and eating better would make her healthy.

If Jane fails to desire these things, then DFT says they don't make her better off. But, intuitively, they would make her better off. DFT thus underdetermines the number of cases that should count as improvements to well-being.

Example 4.6 — Jane receives a fortune. Suppose I hack into Jane's financial accounts and I deposit one million dollars into her account. She never desired to be a millionaire and never really thought about whether she would like to be rich. Nevertheless, we would say that this financial windfall has made her better off even if she did not desire it.

Example 4.7 — Liz – brainwashed. Suppose Liz is conditioned by her parents to think she is worthless. She not only thinks she is worthless but does not desire many things that would make her better off. She reasons, I don't even desire to be have self-esteem because I'm such a miserable wretch.

According to DFT, having self-esteem would not make Liz better off since she does not desire it.

P1: If DFT is true, then Jane is not better off after getting a massage (or the million dollars)

P2: Jane is better off after getting a massage (or the million dollars).

C: Therefore, DFT is false.

4.3.2 Scope problem: Satisfying a desire but with no change to your life

Suppose a decision of mine is satisfied but there is absolutely no change to my life. According to DFT, in virtue of the fact that my desire is satisfied, I am better off. But this is counterintuitive for how can I be better off if no part of my life has changed at all.

Example 4.8 — desire-fulfillment with no effect on well-being. It is the end of the semester and a student wishes me well. I reciprocate and wish the student well. We both desire that the other has a happy life. After the semester we go our separate ways and never meet again, see each other again, or even think about each other again. The student lives a happy life and so my desire is fulfilled. And, I go on and live a happy life and so their desire is fulfilled.

If DFT, then the student and I are better off. But, we do not seem any better off since there is no change to our lives at all. If the student went on to have a horrible life, I would go on living in the exact same way as if they had a happy life. Therefore, DFT must be false.

Putting the above argument in the form above, we can write:

P1: If DFT is true, then the student and I are better off because our desires are fulfilled.

P2: Neither the student nor I are better off.
 C: Therefore, DFT is false.

So, the moral of the argument against DFT is that it makes the scope of desire-fulfillment larger than the scope of well-being improvement. That is, some things may fulfill our desires but not improve our well-being. In other words, DFT overdetermines the number of cases that should count as improvements to well-being.

4.3.3 Scope problem: Desiring something bad for you

DFT says that if something fulfills your desires, then it is good for you. One set of counterexamples are cases where the fulfillment of your desires actually makes you worse off. That is, one set of counterexamples involve, not a case where a desire-fulfillment fails to cause an change in one's well-being, but one where it is detrimental to one's well-being.

Example 4.9 — Tek – the drug addict. Suppose Tek is a first-time drug user. He doesn't desire to be addicted to drugs. He is simply interested to know what using recreational drugs feels like. After a few uses, Tek is hooked. His sole desire in life now is to acquire drugs to become high. According to DFT, feeling high is good for Tek because it fulfills his non-instrumental desires.

But this is counter-intuitive. Tek is not better off in continually being high even if that is his entire desire.

P1: If DFT is true, then Tek is better off because he has fulfilled his desire to use drugs and be high.
 P2: Tek is not better off.
 C: Therefore, DFT is not-true.

This argument points to the same basic conclusion. DFT overdetermines the number of cases that should count as improvements to well-being.

Exercise 4.4 Consider another example where DFT entails a case where someone's desires are fulfilled but they are not better off because of this fulfillment.

4.4 Objections to arguments against DFT

In this section, we consider responses to arguments against DFT. That is, we consider various attempts by a proponent of DFT to save the theory.

Objection 4.4.1 — P2 is false. Overlooking small increases in well-being. P2 is false. Both the student and I are better off. We just don't really appreciate being better off because the increase in our well-being is very small. For example, suppose that unknown to me my bank fixed an error in my checking account. In fixing that error, they added one cent to my account. I am better off with the one cent than without it, but not by an amount I would notice.

One response to this objection is that while it would address cases where the desire is not very strong, it seems as though it would fail to address cases where the desire is significant. Consider the following example.

Example 4.10 Tek hopes Renna is having a happy life. It is his strongest and deepest desire. He never hears from her again. Even though she is safe, there is no real change to Tek's life.

In the above example, it seems implausible to say that the reason we don't recognize that Tek is better off is because the fulfillment of Tek's desire only has a small, easily-overlooked effect on his well-being.

Objection 4.4.2 — P1 is false. Reining the view.. P1 in the argument against DFT states “if DFT is true, then X is the case” (where X is some problematic cases where well-being is not improved by desire-ful llment). One response is that P1 is false when DFT is understood in a more sophisticated (idealized) way. That is, P1 is false when DFT is replaced by an idealized desire-ful llment theory.

De nition 4.4 — idealized desire-ful llment theory (IDFT). Something makes an agent (i) better o if and only if and because it ful ls the desires they would have if they were (a) have full information about the objects of their desires and (b) instrumentally rational in that they take the means that would improve their lives overall and (ii) worse o if and only if and because it frustrates their desires they would have if they were both (a) and (b).

In order to better understand this theory, let's test it against the problematic cases considered earlier.

First, concerning the student and the instructor who both wish each other well at the end of the semester, if we assume that both agents have full information about the objects of their desires (as IDFT says), then they will know that their desire has been fulfilled and this knowledge does make a difference to their well-being.

Second, concerning Tek who is addicted to drugs, if we assume that Tek is instrumentally rational and fully knowledgeable about the objects of his desires (as IDFT says), then we can say either of the following:

1. Tek is lacking in instrumental rationality because he is not taking the means to satisfy his ultimate desires. For example, perhaps he has other desires but lacks the will-power or strength to fulfill these desires because the desire to be high is all-consuming. If he were more powerful, he could also satisfy those other desires.
2. Tek is lacking in knowledge. For example, perhaps he is unaware of other things he might desire more than using drugs.

In short, the DFT theory can be modified to an idealized version of the theory where we assume that what makes something good for someone is what they would desire given that they are fully rational and fully knowledgeable.

4.5 Arguments against IDFT

Let's assume that DFT is revised to IDFT. That is, we assume that what makes people better off is the fulfillment of the desires they would have in an idealized (rational and knowledgeable) state. In this section, we consider arguments against IDFT.

4.5.1 Lack of predictive value

The first argument against IDFT involves two related criticisms. First is that the theory lacks any predictive value since it is difficult to determine how we would know what some idealized person would desire. For example, how do we know that Tek would desire differently if he were a completely rational being. Second, how do we know that Tek's desires in a highly idealized state are better for him than those in a less idealized state? For example, perhaps being fully rational and knowledgeable leads one to desire things contrary to one's well-being.

4.5.2 IDFT is worse than DFT

A second argument against IDFT criticizes the theory for making a mess of cases that DFT handled in a straightforward manner. Consider that there were a number of cases that DFT seemed to give

the right responses. Why is eating a sandwich when Tek is hungry good for Tek right now? Well, as DFT says, because it fulfills Tek's desires. However, if IDFT were to no longer be able to explain such cases, we would be highly skeptical of the theory as an account of well-being. Let's consider an example.

Example 4.11 Let's consider Tek and idealized-Tek. Tek is Tek while idealized-Tek is the version of Tek that is fully rational and knowledgeable. It is very well possible that the desires of Tek are not the same as idealized-Tek. Tek desires spending nights with his wife watching on Net ix, he enjoys a greasy cheeseburger every week, and walking in the woods. Idealized-Tek does not desire these things. Instead, idealized-Tek only watches documentaries, enjoys a good salad, and spends most of his time listening to ne classical music.

The argument against IDFT then is that IDFT says that the desire-ful llments of idealized-Tek determine what makes Tek better o . And this is supposed to be true even if Tek does not desire these things and even desires not for these events to occur. That is, watching documentaries, enjoying salad, and listening to ne classical music is good for Tek even if he desires not to do these things.

4.6 Conclusion

5. Objective list theories

5.1 Introduction

5.2 Objective List Theory

Objective list theory contends that there are some prudential values that are independent from individual desire.

Definition 5.1 — objective list theory (OLT) of well-being. Objective list theory contends that there is at least one thing that is good for agents that is good for them (prudentially valuable) even if they do not desire that thing.

5.2.1 Points of clarification

Before we get into variations of this general theory, it is worthwhile to make a few points of clarification.

First, OLT does not say that there is something that is good for you (makes you better off) and you do not desire it. Instead, it asserts that there is at least one thing that is good for you whether you desire it or not. That is, it asserts that there is something that is good for you (prudentially valuable) regardless of whether you desire it.

Example 5.1 Suppose friendship is prudentially valuable in the sense that is meant by OLT. According to OLT, being in a friendship is good for you regardless of whether or not you desire it. You may desire it, and it is good for you. Or, you may not desire it, and it is still good for you.

Second, OLT involves the denial of desire-theory. According to desire theory, something is good for you if and only if and because you desire it. This means that in order for something to be good for you, it is necessary that you desire it. In contrast, OLT says that it is not necessary for you to desire something in order for it to be good for you.

Third, one version of OLT is equivalent to hedonism. For consider that hedonism asserts that pleasure is good for you and it asserts this independently of whether or not you desire pleasure. Thus, hedonism is a species of OLT.

5.2.2 Developments of OLT

With these points of clarification in place, it worthwhile to note that OLT can be developed in a number of ways.

First, an OLT might specify the number of desire-independent things that are good for you. Along these lines, OLT can be divided into two main types: monistic-OLTs and pluralistic-OLTs.

A monistic-OLT is a type of OLT that says that is one and only one desire-independent thing that is good for you.

Definition 5.2 — monistic objective list theories (m-OLT). A monistic objective list theory contends that there is exactly one thing that is good for agents that is good for them even if they do not desire that thing.

In contrast to m-OLTs, there are pluralistic-OLTs.

Definition 5.3 — pluralistic objective list theories (p-OLT). A pluralistic objective list theory contends that there is more than one thing that is good for agents that is good for them even if they do not desire that thing.

To illustrate the difference, consider the following things that might be considered desire-independent goods (let's assume that the list is exhaustive):

1. friendship
2. pleasure
3. achievement
4. sense of self-worth
5. health
6. knowledge
7. belief in the afterlife
8. love

A p-OLT would consist of at least two goods from the list above while a m-OLT would consist of one and only one good from the list above.

Exercise 5.1 Construct your own personal OLT. First, state the theory. Second, state whether it is a pluralistic or monistic OLT. Third, indicate which thing(s) have prudential value.

5.3 Arguments for objective list theory

In this section, a number of arguments for OLT are considered.

5.3.1 The experience machine argument for OLT

The first argument for OLT is that versions of OLT are preferable over hedonism in that versions of OLT are not susceptible to the experience machine objection. In short, the argument can be summarized as follows:

- P1: Hedonism is susceptible to the experience machine objections.
 P2: Versions of OLT are not susceptible to the experience machine objection.
 P3: All other things being equal, we should accept the less objectionable theory.
 C: Therefore, we should accept versions of OLT not susceptible to the experience machine objection.

The premise in need of clarification is P2. Suppose that there is some good besides pleasure that is a desire-independent good. Let's say it is friendship. Now let's consider two individuals: Tek and Tek-machine. Tek is outside of the experience machine while Tek-machine is inside the experience machine. Assuming their levels of pleasure are the same, hedonism would assert that they are both equally well off. However, if friendship is a good (not merely the pleasure associated with being in a friendship is a good), then according to OLT, Tek is better off than Tek-machine for Tek has a good that Tek-machine lacks.

So, in short, some versions of OLT depend upon having certain goods outside of the machine (actual events) and so entering the machine would be to lose out on those goods.

It is worth noting, however, that one version of OLT is not susceptible to the experience machine objection. Namely, a monistic-OLT that says pleasure is the only desire-independent would still be susceptible to the experience-machine objection.

Objection 5.3.1 While this argument shows that some versions of OLT are preferable over hedonism, it doesn't show that they are preferable over desire-fulfillment theory.

5.3.2 The argument from scope

If the first argument in support of OLT aimed to show why the theory should be accepted over hedonism, the second argument in support of OLT aims to show why the theory should be accepted over DFT.

- P1: DFT is susceptible to scope objections (an individual desires X but X does not make them better off).
 P2: Versions of OLT are not susceptible to scope objections
 P3: All other things being equal, we should accept the less objectionable theory.
 C: Therefore, we should accept OLT.

Earlier we saw how DFT is susceptible to scope objections by considering various scenarios where an individual desires something but that thing doesn't make them any better off (and in worst cases scenarios, it may make them worse off). The response to this by DFT was IDFT, which came with its own host of problems. Since OLT simply states what is good for the agent and what would make them better off is independent of their desires, there is no scope problem for OLT.

Putting the experience-machine and scope arguments together, we now have a reason for preferring OLT over hedonism and desire-fulfillment theory, respectively.

5.3.3 The argument from intuitions (desire-independent goods)

While we have an argument for OLT over hedonism and desire-fulfillment theory, we don't have any independent reason to accept the theory. In this section, we consider such an argument.

If we ask an individual what they ultimately want for their loved ones, we are often given a list of fundamental goods: happiness, health, pleasure, knowledge, love, etc. If we are asked whether we would want these things for our loved ones even if our loved ones did not desire them, we affirm that these things would be good for them even if our loved ones didn't desire them.

Example 5.2 Tek is the son of Liz. Tek desires to party, to do drugs, to not study, to do whatever he wants when he wants. Liz, however, says she wants best for Tek. She makes him stay home, she gets

him clean, she forces him to study, and she helps to shape her son into a particular type of person.

Exercise 5.2 Think of another concrete example where we would want certain goods or things for someone regardless of whether that person desires them.

The argument then from intuitions is this:

P1: Intuitively, we think things that are good for people independent of whether they desire them.

P2: All other things being equal, we should accept the theory that best accords (ts) with our intuitions.

C: Therefore, we should accept OLT.

Objection 5.3.2 — P2 is false: inconsistency. It is worth noting that there is a problem with accepting P2. Even if there is some intuition that there are some desire-independent goods, there is no agreement on what they are. And, our intuitions often lead us to inconsistency. Thus, our intuitions are not a good guide on this matter and thus should not be used to support OLT.

To illustrate, first consider that if we were to ask two individuals Tek and Liz to list all of the fundamental desire-independent prudential goods. It is likely the case that there would be disagreement between them.

Second, we need not even consider two different individuals. Note that what an individual takes to be a fundamental good may change over time.

Thus, P2 seems true if there is no disagreement but false if individual intuitions are inconsistent with each other.

5.3.4 The argument from particular theories

Another set of arguments for OLT involves citing particular desire-independent goods and arguing for OLT on the basis of these goods.

The basic form of this argument involves considering a particular good, then imagine that the good is taken away. Intuitively, we might say that if the thing is taken away, then individual is harmed (worse off) and so this provides evidence that the thing is a good regardless of whether the individual desires it.

For example, Fletcher (p.54) considers the example of privacy.

Example 5.3 Imagine that your phone is hacked. You don't know it but someone steals your personal information, is spying on your texts, is looking at your private pictures. You might be worse off even if the hacker doesn't do anything with your information.

Exercise 5.3 How might the desire theorist respond to the argument that privacy is a desire-independent good and, in particular, the case of an individual's private information being hacked without them knowing it?

5.4 Arguments against objective list theory

In this section, various objections to OLT are considered.

5.4.1 The argument from arbitrariness

The first argument against OLT is that the desire-independent goods cited by any OLT appears completely arbitrary. That is,

1. there is no explanation why the good appears on the list rather than not on the list
2. when there are multiple goods, there is no non-trivial explanation for what is common about both of the goods
3. when there are multiple goods, there is no explanation for the weighting of the goods

Let's consider each of these in turn.

5.4.1.1 Arbitrary why specific goods are on the list

Let's take (1). This says that there is no explanation for why a good like friendship appears on the list rather than being alone. The goods seem to magically appear on the list. Its appearance on the list is not explained, has no justification, and is arbitrary.

The argument against OLT then is this:

- P1: We should not accept a theory that says things are prudentially valuable for no reason (no arbitrariness)
- P2: OLT posits goods arbitrarily.
- C: Therefore, we should not accept OLT.

To this an OLT theory might give the following response:

Objection 5.4.1 — brute vs. explanatory OLT. The OLT can distinguish between two types of OLT: brute objective-list theories (b-OLT) and explanatory objective-list theories (e-OLT). b-OLT would agree that there is no explanation for why the particular good appears on the list. It is simply a fundamental truth (a truth that is used to explain all other truths) or simply a brute fact. e-OLT would contend that there is at least some explanation for why desire-independent prudential goods appear on the list.

Thus, while b-OLT has a response to the objection (no explanation is needed), e-OLT can provide a response by providing a non-trivial response as to why the prudential good is on the list.

The idea then is that at least potentially e-OLT can respond to the charge that OLT is arbitrary. What we haven't provided here is an explanation.

5.4.1.2 Arbitrary what the goods have in common

Next, (2) demands for an explanation of what all of the goods on the list have in common (besides the property of making people better off). The objection is that the list is arbitrary because it fails to explain what all of the items in the list have in common.

Objection 5.4.2 — b-OLT and e-OLT objections. In response to this objection, b-OLT might contend that there is no explanation for why any two properties together are on the list. It is simply a brute fact.

In contrast, e-OLT might contend that there is some common property (e.g. all of them do X or all of them have X, or all of them are capable of being experienced) held among all of the goods on the list.

5.4.1.3 Complete lack of weighting

Finally, (3) contends that when there are multiple goods, there is no explanation for the weighting of the goods. This is a demand for a hierarchy of goods (which is the most important and which is the least important).

Objection 5.4.3 — unreasonable demand. An OLT might simply respond by stating that this demand is too unreasonable. That is, no other theory can really provide this sort of response. For example, DFT cannot give a better account of the hierarchy of desire-fulfillments than OLT can give a hierarchy of desire-independent goods. Thus, it would be an objection that applies to all theories equally.

Exercise 5.4 The argument from arbitrariness contends that the items on the OLT list are arbitrary. This is unpacked by saying that OLT is lacking in its capacity to explain the items in the list in some way. In response, OLT can divide itself into b-OLT and e-OLT. Which of these two versions of OLT do you find preferable and why?

5.4.2 The argument from alienation

The argument from alienation is that an individual can be well-off according to various versions of OLT but completely alienated. And, no theory should say an individual is well-off but alienated. Therefore, OLT is flawed.

In order to understand this argument more fully, it helps to get a clearer sense of what it means to be alienated. Since it is hard to give an exact definition of “alienation”, it is sometimes more helpful to give some characteristic properties and then further clarify it with examples.

An individual is alienated when:

1. they are not actively engaged in what they do
2. they are unable to take any pleasure or joy in what they do
3. their actions are unrelated from their desires, concerns, or aims

Next, let's consider an example of an individual that is alienated.

Example 5.4 Suppose Tek works in a cubicle. He looks at numbers all day. He hates his job but has developed various strategies to make the work he does automatic. He arrives to work at 9AM, mindlessly works until 5PM, and then leaves. He is paid a salary and the work he does is not appreciated. Nor does he take part in the riches he has made for the company.

In the above example, Tek is alienated in that he doesn't take any joy in what he does, it lacks his active engagement, and his actions seem unrelated to his goals. Namely, his work alienates him since it isn't sensitive to his concerns as an individual.

The argument then against OLT from alienation is that according to OLT an individual can be well-off yet alienated. That is, the desire-independent prudential goods on an OLT can say an individual has a high-level of well-being yet they are completely depressed, miserable, bored, and unhappy.

Example 5.5 Let's consider an example of alienation that applies to a version of OLT. Suppose an m-OLT that says friendship is the only good. On this theory, Tek may have friends but might still be unhappy. He may have desires unrelated to having friends and so while he is extremely well-off, he is unhappy.

This argument appears to both count against OLT and in favor of DFT. For DFT contends that something is good for you only if it fulfills your desires. And, it seems implausible that having your desires fulfilled would alienate you for it is engaging precisely the thing you want engaged. Thus, we might summarize the entire argument against OLT as follows:

- P1: If a theory of well-being is alienating it fails as a theory of well-being.
- P2: A theory of well-being is alienating if and only if it is not consistent with the view that something is good for you only if you desire it.
- P3: OLT is a theory of well-being that is alienating.
- C: Therefore, OLT fails as a theory of well-being.

Objection 5.4.4 — P2 and P3 are false. The general objection to this argument is that it fails to be universal. That is, the objection would only apply to some forms of OLT. Various forms of OLT might not be alienating yet might be disconnected from an individual's explicit desires.

Consider the OLT that says pleasure is the only good (hedonism). According to Fletcher (p.60), an individual may not desire pleasure, receive pleasure, and while the pleasure is disconnected from his/her desires, it would not be alienating to them.

Or, consider the OLT that says that self-respect and happiness are the only goods. According to Fletcher (p.61), we could imagine Norman who is in a religious cult that brainwashes him into desiring to be worthless and not feel good about himself. According to this version of OLT, self-respect and happiness would not be alienating to Norman.

Exercise 5.5 Fletcher (p.60) claims that "hedonism is inconsistent with attitude-dependence but seems to be a clear case of a non-alienating theory". In order to support this, he uses the case of Miserable Maude. Review this case and consider whether it effectively supports his claim. Can you think of a better example?

5.5 Conclusion

We examined various forms of objective-list theories. These theories were divided into various types:

brute explanatory
monistic
pluralistic

6. Ethics and Religion

6.1 Introduction

For many individuals, religion plays a significant role in why and how they are moral. Religions often lay out a system of morality that individuals follow. Religious individuals often seek moral guidance from priests, rabbis, or other religious teachers. People seek moral guidance by emulating important religious figures and they often read and study religious texts for insights on how to live a moral life.

Because of the importance that religion has on the lives of millions of individuals, we will start by investigating whether religion can justify moral rules.

Exercise 6.1 How does religion shape your views about morality? You need not be religious to have an answer to this question.

6.2 Argument 1: Religious belief and moral motivation

Let's begin by considering four different positions about God's existence.

Definition 6.1 — atheism. The position that God does not exist.

Definition 6.2 — agnosticism. The position that one does not know whether God exists or does not exist.

Definition 6.3 — theism. The position that God exists.

Definition 6.4 — deism. The position that God exists but does not play an active role in the world, including not commanding us to do anything.

- P1: If one is an atheist, agnostic, or deist, then one has no motivation to behave morally.
 P2: If one is a theist, then one has a strong motivation to behave morally (fear of punishment, desire for reward).
 P3: It is beneficial for society that people are motivated to behave morally.
 C: Therefore, one ought to be a theist.

Note that this argument does not prove the existence of God. Instead, it aims to show that one ought to believe in God (regardless) for its practical benefits.

Objection 6.2.1 — P3 is false. Moral motivation can make society worse off. Theists may be more strongly motivated to be moral but this doesn't entail they are behaving more morally. In fact, their motivation to be moral may make society worse off. Suppose Tek is highly motivated to be moral. He believes that homosexuality is morally wrong and that one ought to kill anyone who is homosexual. Now suppose Liz who has no motivation to be moral. She has no opinion on the nature of homosexuality and is not motivated to kill anyone just because they are homosexual. Here, Liz's lack of motivation is more beneficial to society than Tek's strong motivation.

Exercise 6.2— Moral motivation for atheists, agnostics, deists. The above argument relies upon asserting that atheists, agnostics, and deists lack motivation to be moral. Is this the case? Does the motivation to be moral rest upon the existence of a God that plans to punish or reward us for behaving morally?

6.3 Argument 2: Divine Command Theory

The next theory we will examine asserts that moral rules are created not by human beings or nature, but by God.

Definition 6.5 — Divine Command Theory (DCT). Divine Command Theory (DCT) is the theory that what is morally right and wrong is determined by God's commands.

First, DCT is the theory that morality and religion cannot be separated since what is morally right and wrong is determined by God's commands. This means that morality is based on religion. Something is morally right if and only if God says that it is right, while something is morally wrong if and only if God forbids it.

Second, DCT is not offering an explanation for moral motivation. It is giving an explanation of why something is moral or immoral. According to DCT, something is immoral simply because God will's that act to be immoral.

DCT is an attractive moral theory for several reasons:

1. It allows for objective moral truths
2. It has a straightforward answer to the "why be moral" question. Answer: God will punish you if you are not and reward you if you obey.
3. It coincides with how many people understand the universe and their place in the universe.

The basic argument for DCT is that moral rules must come from somewhere and if laws are to have any binding they cannot simply be constructed by ("made up") by human beings. Thus, if there are objective moral rules, then the only being that could have created these is God.

- P1: Moral laws need to be created by someone or something.

P2: Moral laws cannot come from physical laws since physical laws tell us how to behave and moral laws tell us how to behave.
 P3: The best explanation then for moral laws is that they are made up by some being.
 P4: If moral laws are objective, then they cannot be created by human beings (otherwise they would be subjective).
 P5: Moral laws are objective.
 C: Therefore, the only possible (and therefore the best) explanation of the objectivity of moral laws is God.

Another way of thinking about this argument is as an inference to the best explanation.

P1: There are objective moral rules.
 P2: Moral rules are created.
 P3: The best explanation for the individual that created these moral rules is a supreme being (God).
 C: Therefore, what is morally right and wrong is determined by God's commands.

In articulating DCT, it is important to clarify that God's commands determine moral goodness. This is sometimes said to be a voluntaristic interpretation of God's relation to morality. Something is morally right or wrong because God commands it. In other words, an action is morally permissible because God makes it so. DCT adopts this voluntaristic interpretation of God's relation to morality. In contrast, were we to think that God commands us to do an action because it is morally good, then we have an intellectualist interpretation of God's relation to morality. On this understanding of God's relation to morality, God commands /prohibits certain actions because God recognizes that action as morally right /wrong. In other words, God does not make moral rightness or wrongness, God knows what is right and wrong and commands us accordingly.

DCT has an argument as to why it is preferable to the intellectualist interpretation of God's relation to morality.

P1: If God says that "x is wrong" because God recognizes that x is wrong, then the ultimate basis for morality is not rooted in God's will but in God's reasons.
 P2: These reasons are independent of God.
 C: Therefore, the basis of morality does not depend upon God's will or nature.

6.3.1 Objections to DCT

Objection 6.3.1 — God could make evil good and good evil. Suppose God says P is morally wrong. From DCT, it follows that what makes P morally wrong is that God willed it to be wrong. Since it is God's will rather than a set of reasons for why P is morally wrong, God could have chosen otherwise. And, if God could have chosen otherwise, God could have chosen evil acts to be morally good and morally good acts to be evil. But this is counterintuitive.

P1: If DCT is true, then God could have made actions such as rape, murder, genocide, etc. morally permissible.
 P2: If God chose these actions as morally permissible, then they would be morally permissible.
 P3: Such actions are not morally permissible even if God chose them to be morally permissible.
 C: Therefore, DCT is false.

Objection 6.3.2 — DCT is counter-intuitive on the assumption that God doesn't exist. P1 If God didn't exist, there would be no morals.
 P2 There are morals (and God doesn't exist) or there would be morals (even if God didn't exist).
 C Therefore, DCT is false

Objection 6.3.3 — We have a clearer idea of what is right and wrong than God's will. P1 If morality is determined by God's will, then our knowledge of good and bad should be identical with our

knowledge of God's will.

P2 Our knowledge of good and bad is not identical with our knowledge of God's will (we have a clearer idea of what is good and bad).

C Therefore, morality is not determined by God's will (DCT is false).

Exercise 6.3— Consider the following questions.

1. For those that believe in God: If it were discovered that God does not exist, would your views change on morality? If yes, give an example. If no, explain why.
2. For those that don't believe in God (or are agnostic): if it were discovered that God does exist, would your views change on morality? If yes, give an example. If no, explain why.

Exercise 6.4 Notice DCT involves articulating why something is morally right or wrong because God says it is morally right or wrong. However, people do not always appeal directly to God's will. Instead, they contend that an act is morally right or wrong because a religious text says that it is morally right or wrong. Take one of the objections above and apply that objection to someone who appeals to a religious text to justify that a particular action is morally right or wrong.

6.4 Argument 3: Moral wisdom through God

Recall that the intellectualist version of God's relation to morality is rejected by DCT because it would make the reasons that something is moral or immoral independent of God. However, one might contend that while God's will does not make an act moral /immoral, God knows what acts are moral/immoral. And, furthermore, God conveys this knowledge to us human beings. Such knowledge might be conveyed in a variety of different ways:

1. personal revelation: miracles, communication through prayer, divine insight, signs from God, etc.
2. scripture: holy texts written by authors or prophets who were divinely inspired
3. holy leaders: priests, rabbis, shamans, et al.

At least on the surface, there may be some plausibility to thinking that God's relation to morality is one of a conveyor (rather than the determiner) of moral truths:

1. historically, moral beliefs are often intertwined with religious texts
2. science does not seem capable of telling us what we ought to do and so it seems plausible to say that this knowledge must be divinely inspired or conveyed
3. our moral system is often guided by moral role models who are often religious or quasi-religious figures

However, there are several problems with this approach:

Objection 6.4.1 — God may not exist. If God does not exist, then it is impossible for God to impart wisdom.

Objection 6.4.2 — God may not convey any information about what God knows concerning morality.

Objection 6.4.3 — Inconsistency. Let's assume God does exist and let's assume that God's wisdom is imparted in some way. Even if we assume that this is the case, we do not know precisely how God is imparting God's wisdom.

1. It can sometimes be difficult to distinguish between a maniac or scam artist who says that God is speaking through him /her and an individual who is genuinely inspired by the Lord.
2. There are a wide variety of religious texts and figures whose messages conflict

Exercise 6.5 Let's assume that we cannot depend entirely on religion for morality. Even if this is the case, this does not mean that religion has no benefit for thinking about ethics or for discovering ethical truths. What, if anything, can religion provide in our thinking about ethical matters?

7. Natural Law Theory

In this lecture, we look at natural law theory (NLT).

7.1 The theory and its arguments

The core of natural law theory (NLT) is the following:

1. things have an essence or nature
2. an act is morally good if and only if it is in accordance with that nature (an act is right if it is natural or fulfills one's nature)
3. an act is morally evil if and only if it is against that nature (an act is morally wrong if it is unnatural or frustrates one's nature)

7.1.1 Argument 1: Objectivity

First, NLT can explain how morality is objective rather than subjective. An act is not moral simply because we desire it to be moral, but if it is in line with our nature.

7.1.2 Argument 2: Morality as applying to humans

Second, NLT can explain why morality seems to only apply to human beings. Non-human animals seem to lack the cognitive capacity to engage in self-control, reflection, and reasoning. These capacities seem necessary for free-will (autonomy). And, free will seems necessary for moral responsibility (how can I be held morally responsible for actions that were forced, viz., not freely chosen). Human beings are thus the only beings who have the sort of nature that allows for the capacity to be moral.

7.1.3 Argument 3: Origin of morality

Third, NLT can explain where morality comes from (its origin). If there were no humans, there would be no morality. And, if human beings were genetically modified to no longer be rational or free, then there would be no morality. Thus, NLT can explain these intuitive facts. Morality emerges because human beings have (or have acquired) a certain kind of nature.

7.1.4 Argument 4: Answers Hume's argument against morality

Fourth, NLT can show how Hume's Argument against morality is flawed. To see this clearly, we first need to formulate Hume's argument:

P1: We can only have knowledge of conceptual truths (e.g. two plus two equals four or all bachelors are unmarried men) or empirical truths (e.g. It rained in State College on 9 /24).

P2: Moral claims (e.g. you ought not to kill people) are neither conceptual truths or empirical truths.

C: Therefore, we cannot have knowledge of moral claims.

In support of P1, it seems that we can have knowledge of conceptual truths by reasoning alone (this is sometimes called a priori reasoning), by analysis of the meaning of words (e.g. all unmarried men are bachelors), by logical deduction (e.g. If P then Q, P therefore Q), or by the recognition that the denial of the claim would lead to contradiction (e.g. triangles have four sides). In contrast to conceptual truths where our understanding of what the sentence says (or mere formal analysis) entails a knowledge of the truth, our understanding of a moral claim does not entail that we know whether it is true or not. Furthermore, my denial of a purported moral truth (e.g. killing is not morally wrong) does not lead to contradiction. Thus, moral truths are not conceptual truths.

Further in support of P2, it seems that moral claims are not determined empirically. I can verify whether it is raining now or whether a diamond is hard through experimentation or observation. In contrast, a mere observation of killing does not tell me whether or not the activity is morally wrong or permissible. The difference between empirical and moral claims is rooted in the fact that empirical claims are descriptive (they aim to describe the way the world is) while moral claims are prescriptive (they aim to describe how the way the world should be). Thus, moral truths are not empirical truths.

Thus, Hume's argument concludes that if moral truths are neither conceptual truths nor empirical truths, and we only can have knowledge of the latter, then we cannot have knowledge about morality.

NLT has a response to Hume's argument. Namely, moral claims are empirical. They are empirical because, for NLT, to know what is morally good /bad, we only need to know two empirical truths:

1. our human nature
2. whether an act is in accordance or against our human nature

Exercise 7.1 Let's consider an application of NLT. Individuals will sometimes argue that a particular act is morally wrong because it is unnatural or it goes against human nature. For example:

P1: If an act is unnatural, then it is morally wrong.

P2: The use of contraception is unnatural.

C: Therefore, it is morally wrong.

Do you agree with the above argument or not? If not, which premise is problematic?

7.2 Conceptions of human nature

A key concept in NLT is the notion of human nature. This conception determines what is and is not morally permissible. Thus, much of NLT rests on precisely what is meant by “human nature”.

7.2.1 Human nature as animal nature

7.2.2 Human nature is what is innate

One conception of human nature is the set of characteristics that we are born with, viz., the set of biological characteristics that are not a product of upbringing, custom, education, and /or training.

Objection 7.2.1 — our knowledge of morality is greater than our knowledge of our nature. If what is innate determines what is morally permissible /impermissible, then our knowledge of what is morally good and bad should not extend beyond our knowledge of what is innate in human beings.

But, we seem to have greater knowledge about moral goodness and badness than we do about our own nature. For example, we take ourselves to know that rape, murder, etc. is wrong but might not be certain about whether human beings are born selfish, or good, or honest, or competitive, or trustful, or cooperative.

7.2.3 Human nature is what humans have in common

Another conception of human nature is our nature consists of all the characteristics that are shared by all human beings. Thus, suppose there were three human beings: A, B, C. Further, suppose there are three characteristics: x, y, z.

1. A: x, y
2. B: x, y
3. C: x, z

Given the above distribution of characteristics among humans, human nature would be defined by property x, which is the property that all human beings share.

Objection 7.2.2 — not clear if all humans share a common set of morally relevant properties. Human beings may share a set of common properties, e.g. we are all extended in space and time, we all give off heat, all of our bodies are material, etc. But not all of these shared properties seem morally relevant. The fact that we all give off heat does not seem to be relevantly related to whether murder is morally good or bad. Thus, NLT needs to show that there is evidence that we all share some set of characteristics that are relevantly related to morality.

Objection 7.2.3 — Even if there are morally relevant shared properties, this does not imply they give good moral guidance. Let's suppose there is some characteristic of our nature that is shared by all human beings. And, let's suppose that characteristic is significant enough to provide moral guidance. The mere fact that we have such a characteristic does not entail that the guidance it offers is good moral guidance.

For suppose that we all have the capacity to kill. According to NLT, this would mean that acting on this aspect of our nature would be a morally permissible act, and frustrating this part of our nature would be morally evil. But this is contrary to our intuitions about morals.

In short, NLT assumes that this shared part of our nature is not merely morally relevant but gives good rather than evil moral guidance.

Exercise 7.2 With respect to the last objection, NLT seems to assume that human beings are innately good. That is, if individuals were to more closely abide by their underlying nature (the common properties shared by all human beings), they would be behaving morally. However, if our underlying nature is evil, then acting on this part of our nature would be to act in a morally evil way. Do you think people are innately good or evil?

7.3 Natural design and human purposes

Let's try to formulate NLT in a different way:

Definition 7.1 — Natural Law Theory (NLT). Natural Law Theory (NLT) is the ethical theory that

1. The universe is an ordered system where everything has a specific purpose determined by God / Nature.
2. A thing's purpose not only specifies what it is for, but also how it should be.
3. We determine what a thing's natural purpose is through reason.

Let's examine of these claims. First, "The universe is an ordered system where everything has a specific purpose determined by God / Nature." There are two key components to this claim. First, everything in the universe has a purpose. That is, for any object in the universe, there is an answer to the question "what is it for?"

1. a fork — it is for eating
2. a home — it is for living in.
3. a television — it is for watching

The above have easy answers because these objects are made by humans to fulfill some need/speci c need/purpose. These are objects that were created to fulfill a function /purpose. Other objects are a bit more difficult:

1. the eye — for seeing
2. the heart — for pumping blood
3. the hand — for . . . typing, painting, etc.

Still others are more difficult:

1. a human being — ?
2. gravity — ?
3. the universe — ?

The second key component is that all of these purposes are determined by the designer of the universe, namely God or Nature. That is, things have purposes and these purposes are either due to creation (intelligent design) or evolution.

Second, a thing's purpose not only specifies what it is for, but also how it should be The idea here is that if the heart is for pumping blood, then a heart should pump blood. If it doesn't, then it isn't behaving as it should. We say that there is something wrong with the heart if it doesn't pump blood. To put this more abstractly,

- P1 If x is a natural object, then as a natural object, it has certain "natural" purposes a, b, c.
 P2 Failing to fulfill a natural purpose is to do something wrong (it goes against God's plan)
 P3 Assume that x is a natural object not fulfilling one of its purposes.
 C Therefore, x would be doing something wrong.

Let's apply this to human beings.

- P1 Human beings are natural objects and as natural objects, they have certain "natural" purposes.

P2 [insert some natural purpose here]

P3 Failure to fulfill a natural purpose is to do something wrong (it would go against God's plan)

C Therefore, if John is a human person and he doesn't fulfill one of his natural purposes (he goes against his nature), then he has done something wrong.

This sort of argument can be applied in all sorts of ways. Here is one:

P1 Human beings are natural objects and as natural objects, they have certain "natural" purposes. One of these is to socialize, to have friends, to forge relationships.

P2 Failure to fulfill socialize, to have friends, and to forge relationships goes against our nature and goes against God's plan.

P3 Failure to fulfill a natural purpose is to do something wrong (it would go against God's plan)

C Therefore, if John is a human person and he doesn't socialize, then he goes against his nature and has done something morally wrong.

The third claim of NLT is that "we determine what a thing's natural purpose is through reason." Thus far, the above theory has a strong religious tint as we have said that a thing's natural purpose is determined by God. This prompts the following question: how do we know what a thing's natural purpose is? A proponent of NLT might say that this requires we know God's will.

However, NLT contends that while all natural purposes are determined by God, we determine what our natural purposes are through the use of reason. The idea then is that God has given us all (even those that don't believe) the power to determine what is morally right and wrong by giving us the power of reason. We use reason to determine what natural purposes a thing has, and then reason that if a thing violates these purposes, it has done something wrong.

7.4 Objections to Natural Law Theory

There are several objections, religious and non-religious in nature, to NLT:

Objection 7.4.1 — What is natural is not always good.. There are many seemingly natural behaviors that are wrong and so NLT is mistaken in taking all natural behaviors to be moral:

1. People seem to be naturally selfish, but selfishness seems wrong.
2. It may be nature to kill or rape (they may increase our prospects of survival) but these are clearly wrong.
3. As people age, their body naturally breaks down, but we don't think of bodily deterioration and disease as naturally good.
4. Homosexuality sometimes gets classified as "unnatural" because it involves sexual activity that cannot lead to procreation, but it is not obvious that being homosexual (or engaging in homosexual sex) is wrong.

Objection 7.4.2 — What is unnatural is not always bad. NLT contends that if an action frustrates our natural purposes (procreation, reproduction, etc.), then it is morally wrong. But this is not clearly the case for consider the following:

1. Playing video games is not clearly a natural activity but it is not immoral
2. It might benefit me in terms of my survival to stay in a state of ready alertness, always listening, always watching my surroundings, always anticipating potential threats. This would mean that relaxing is unnatural and immoral. But this is counterintuitive.
- 3.

8. Psychological Egoism

8.1 Introduction

In this lecture, we consider the theory of psychological egoism (PE).

8.2 Psychological Egoism

Psychological egoism is the descriptive theory of human motivation that asserts that individuals always act in service of their own personal gain.

Definition 8.1 — psychological egoism (PE). Psychological egoism is the theory that contends that the ultimate motivation for any human individual's intentional (self-controlled) human action is that individual's self-interest (acquisition of personal gain and avoidance of personal loss).

There are several points to keep in mind with respect to PE:

1. PE is a descriptive theory not a prescriptive theory. It aims to describe what the ultimate motivation of human individuals is, not to prescribe what the motivation should be.
2. PE says individuals pursue their self-interest not that individuals are exclusively selfish. It is possible for the pursuit of what benefits you to also benefit others.
3. PE says that individuals always aim at increasing their personal gain, not that they are always successful in acquiring this personal gain.
4. PE says that not every action aims at increasing our self-interest. It is only our self-controlled or intentional actions that aim at increasing self-interest (e.g. it does not say our heartbeat aims to increase our self-interest).

One of the most important consequences of PE (and a reason why PE is controversial) is that PE implies that there are no cases of genuine altruism.

Definition 8.2 — genuine altruism. Genuine altruism refers to the motivation to help another individual independent of whether it benefits one's own self-interest.

According to PE, cases where individuals appear to be motivated to increase the gain of others are (in reality) cases where individuals are ultimately motivated by a concern for increasing one's self-interest.

Example 8.1 — Nice instructor. Suppose Tek is teaching a philosophy class. He is nice to his students. According to PE, he is nice to his students because Tek is trying to acquire some personal gain and would not be nice if it did not benefit him. For example, Tek is motivated by the fact that being nice will make his classes go smoother, or he wants positive reviews at the end of the semester, or because being nice makes him feel good about himself.

Example 8.2 — Nice student. Suppose Tek is a student in a philosophy class. He is nice to his instructor. According to PE, he is nice to his instructor because Tek is trying to acquire some personal gain and would not be nice if it did not benefit him. For example, Tek is motivated to be nice because he thinks that the instructor will grade his essay more charitably, or will grant him an extension if he needs one, or being nice makes him feel good about himself.

Exercise 8.1 Pretend you are a supporter of PE. Create a scenario where an individual is seemingly doing a good action for altruistic reasons but is actually motivated by personal gain.

8.2.1 Ethical consequences of psychological egoism

While psychological egoism is not an ethical theory, if PE is true, then there are significant consequences for any ethical theory. This is because if PE is true, then genuine altruism is impossible. And, furthermore, moral theories that require genuine altruism are mistaken.

The argument for why PE entails that moral theories that require moral altruism are false is straightforward.

P1: PE is true.

P2: If PE is true, then genuine altruism is impossible (cannot happen).

P3: We cannot be obligated to do something that is impossible (ought implies can).

C: Therefore, any moral theory that requires genuine altruism is false.

Exercise 8.2 What sorts of moral rules or propositions can you think of that require you to perform an act that might not benefit you personally?

8.3 Arguments for psychological egoism

In this section, two arguments are considered in support of PE:

8.3.1 Argument from desire

The first argument for PE is as follows:

P1: Individuals always act on their strongest desire.

P2: Our strongest desire is always and only what we think will benefit us most (what we think will increase our personal self-interest).

C: Therefore, individuals always act on what they think will benefit them most (PE is true).

P1 says that individuals may have a number of desires but what ultimately motivates them to act

is what they desire most deeply. That is, people always do what they want to do. That is, people always are motivated by what they desire the most.

P2 asserts that what people desire most is to increase their own personal self-interest. That is, while we may desire to help others or to improve the world, our deepest or most strongest desire is always directed at increasing our own personal gain (or minimizing personal loss).

Example 8.3 — Giving money to charity. In the case of giving money to charity, the psychological egoist says whenever we give money to charity we do so because it is our strongest desire to do so (P1) and this desire is linked to increasing our personal gain in some way, e.g. to get recognition or to promote a cause that will ultimately help us.

Objection 8.3.1 — P2 is false. Our deepest or strongest desires are not necessarily self-interested desires. My strongest desire is not always what I think will benefit me the most. Just because I always do what I want (act on my strongest desire) does not mean I'm trying to increase personal gain for me. Perhaps what I desire the most does not benefit me at all, e.g. sacrificing my life for my country or a friend

Exercise 8.3 A proponent of the desire argument for PE contends that our strongest desire is always what we think will benefit us most. However, the objection to this claim is that just because we have a desire for X does not mean that we think X benefits us. Do you think PE or the objector to PE is right; explain your reasoning.

8.3.2 Argument from expected results

The second argument for PE is as follows:

P1: If I do action X, then I expect X to make me better off than before I did X.

P2: If I expect action X to make me better off then I am trying to increase my self-interest.

C: Whenever you act, you are trying to increase your self-interest (PE is true).

There are two problems with the argument.

Objection 8.3.2 — P1 is false: expectations. Some individuals may perform actions without any expectation. In addition, some individuals may perform actions and anticipate that they will be worse off than before they did X. For example, consider an individual who has avoided arrest but turns themselves in for a crime.

1. Trivial acts of kindness: Suppose I hold the door open for someone. I may have done this intentionally but not intend to feel anything after performing this action.
2. Extra drink: Suppose I have been consuming alcohol and I know I've reached my limit. If I have another drink, I may expect to feel much worse off.

Objection 8.3.3 — P2 is false: expectation and motivation. Just because I expect to be better off in doing X does not mean that I did X to be better off. That is, one's expectation of benefit for doing X does not imply they were motivated to do X because of that benefit.

1. Trivial acts of kindness: Suppose I expect to feel good about myself for holding the door open for the person behind me. I may anticipate this as a result but I may also know that the good feeling is not worth the effort of holding the door. Nevertheless, I hold the door anyway.
2. Working with orphans: Suppose I do volunteer work involving orphaned children. This work makes me feel good about myself. While it makes me feel good about myself, I do not work with orphans because it makes me feel good. I would work with orphans even if it made me feel worse. The pleasurable feeling is a byproduct of the activity rather than a motivating reason.

Exercise 8.4 Articulate one of the two arguments in support of PE.

8.4 Arguments against psychological egoism

There are three main arguments against PE.

The first argument is simply an argument from cases of genuine altruism.

P1: If PE is true, then there are no cases of genuine altruism.

P2: There are cases of genuine altruism.

C: Therefore, PE is false.

P1 is accepted by PE and so the controversial premise is P2. In support of P2, individuals can point to cases of self-sacrifice. That is, cases where individuals perform acts that require giving up their own lives for the benefit of others.

Example 8.4 — Cases of self-sacrifice. Suppose Liz is the parent of Tek. Next, suppose that someone aims a gun at Tek, and Liz throws herself in front of Tek. Liz is shot and dies. In this case, Liz has sacrificed her own life for the benefit of Tek.

The second argument charges PE with being unfalsifiable because it is unfalsifiable. The general complaint here is that supporters of PE will always point to some secret, underlying, unverifiable motivation that is directed toward personal benefit. In other words, even if we humans are always motivated by our own personal gain, supporters of PE do not spell out how any being in any possible world could be motivated by something other than personal gain.

P1: PE is unfalsifiable in principle.

P2: We ought never to accept an unfalsifiable theory.

C: Therefore, PE is false.

The third argument charges PE with not providing an explanation for the fact that people are so deluded about their motivations. That is, suppose Tek says that he is not motivated by self-interest in performing some action. A supporter of PE will say that Tek is mistaken about his own motivations. He suffers from a kind of delusion about his own internal mental state. Given that so many people reason like Tek, it follows that millions (if not billions) of people suffer from a kind of delusion about their own motivations. If this is the case, then PE offers us a strong argument for the source or cause of this delusion. But they have none, and so we ought not accept PE.

P1: If PE is true, then millions (if not billions) of people suffer from a widespread delusion about their own motivations for their actions.

P2: Such a widespread delusion is implausible without an extremely strong explanation about the source of the delusion.

P3: PE offers no such explanation

C: Therefore, PE is false.

Exercise 8.5— Mock debate. We have now considered arguments in support and against PE. We will divide the class into two groups. The first group will be asked to defend PE. The second group will be asked to reject PE. Prepare a two-minute argument in support of or against PE. In addition, prepare a set of rebuttals to the other side.

9. Ethical Egoism

Psychological egoism is the theory that individuals always pursue their own individual happiness (self-interest). As such, it is descriptive theory about human motivation. As a descriptive theory, it does not say that we should or should not do, only what we do as a matter of fact. Let's suppose that PE is not clearly true or false.

We might nevertheless consider an ethical theory that says even if you don't always pursue your self-interest, you ought to pursue it. This is the theory of ethical egoism (EG). This theory contends that we ought to pursue our self-interest.¹

9.1 Ethical egoism explained

Let's first provide a definition of ethical egoism and then some clarifying points.

Definition 9.1 — Ethical Egoism (EG). Ethical egoism is the theory that each person ought to primarily pursue his /her interests. In other words, ethical egoism states that

1. there are objective moral facts and
2. if an individual acts to promote their self-interest, then they have done something morally good
3. if an individual acts to frustrate their self-interest, then they have done something morally wrong

Next, let's provide some points of clarification:

1. Not a theory of selfishness: A person is selfish if and only if he /she pursues his/her own interests without regard to the interests of others. EG does not say to be selfish in this way.

¹Ethical egoism is a claim about what is morally good or bad, whereas psychological egoism is a claim about human psychology.

Furthering your own interests and happiness often depends upon furthering the interests of others. If you are an egoist, the self-interest of other people matters but only insofar as it factors into promoting your self-interest, e.g. helping a friend so that your life is better.

2. Not a theory that promotes exclusive pursuit of short-term, base pleasure: EG does not say that you should pursue only your basest, immediate, and most sensual wants and desires. EG claims that an action is good if and only if it promotes your short-term and long-term happiness. Sometimes this is referred to as well-being or living the “good” life. This requires you to consider not only immediate pleasures but also “higher” pleasures, e.g. enjoying artworks, and long-term pleasures, e.g. having a long and healthy life.
3. Pursuit of what is in your actual self-interest: EG contends that you ought to pursue what promotes your personal good, not what you think promotes your personal good.

One notable point of clarification concerns a controversial aspect of EG. Ethical egoism is a controversial moral theory because many people believe that some acts of genuine altruism are morally good, i.e. some actions are good not because they are done for our own interests but because they benefit someone else, i.e. they promote another's interests. In other words, for some people, there are cases in life where their moral intuitions conflict with their self-interest.

Example 9.1 — Intuitions and self-interest. Some examples where our moral intuitions conflict with what is in our self-interest include:

1. cheating on an exam: intuitively this is morally wrong, but it may be in our self-interest to cheat if we know we won't be caught
2. stealing: intuitively stealing is wrong, but it may be in our self-interest to steal if we know we won't be caught
3. charity: intuitively this is morally good, but it may not be in our self-interest to give to the poor or to help others

EG is controversial because it says our moral intuitions are mistaken. We ought to ignore them because what is truly morally good is what increases our personal gain and what is morally wrong is what decreases our personal gain.

9.2 Arguments for Ethical Egoism

If psychological egoism is false, ethical egoism can still be true. This is because ethical egoism does not say that every action you do is to promote your own happiness. Rather, it states that if you fail to do something to promote your self-interest (e.g. altruistic act that is not in your self-interest), then you have done something morally wrong. If you do something to promote your own well-being, then you have done something morally right. In this section, arguments in support of EG are considered.

9.2.1 Argument from social welfare

P1: We best increase social welfare (the well-being of others) by always acting in a way that increases personal welfare.

P2: We ought to increase social welfare.

C: Therefore, we ought always act in a way that increases our personal welfare.²

P1 is supported by two related claims. First, we know our own wants and needs better than someone else. This is due to the fact that I know myself better than you know me and so I know what would make life better than you know what would make my life better. Second, trying to help others

²Another way of putting this argument is as follows: If we want to help others, the best way to do so is by pursuing our own interests (by being egoists). Altruists want to help others. Therefore, the best altruism is a kind of egoism. If the best altruism is a kind of egoism, then altruists should be egoists. Therefore, we should pursue our own self-interests exclusively.

often comes across as intrusive or paternalistic or a violation to my autonomy. And, intrusive or paternalistic behavior decreases the overall amount of well-being in society.

P2 is taken to be intuitively true.

Objection 9.2.1 — P1 is false. There are cases where we increase social good by helping others rather than ourselves.

1. Giving a person a starving person the rest of our sandwich when it would be easier for us to throw it away
2. Suppose someone breaks their leg. That person is an important doctor who helps children. You have no plans on having children ever. You could stop and help that person (or at least call an ambulance), but that would be an inconvenience to you.

9.2.2 Argument from reasons

Another argument for EG is as follows:

P1: We have moral obligations.

P2: If you have a moral obligation to do action X, then you have a good reason to do X

P3: If you have a good reason to do X, then action X must make you better off.

C: Therefore, if you have a moral obligation to do X, then X must make you better off.

With respect to P2, the idea here is that moral rules cannot be reasonless. There must be some sort of support or set of reasons that motivate you do the actions they prescribe.

With respect to P3, the idea here is that the primary reason you have for doing an action X is that doing X will make you better off than if you did not do X. Rhetorically, why else would you do X if it did not make you better off?

Example 9.2 Consider some random moral rules:

Suppose God commands you to do X. If you have a moral obligation to do X, then you have a reason, e.g. God will punish you if you don't do X, or God will reward you if you do X.

Suppose there is a moral rule to be nice. If you have a moral obligation to be nice, then you have a reason, e.g. It makes you feel better than if you were not nice or you thinking that being nice will somehow lead to more people being nice to you.

Objection 9.2.2 — P3 is false. Shafer-Landau (p.114) contends that we need to distinguish between two different propositions:

P3* If action X makes you better off, then there is a good reason for you to do X.

P3 If there is a good reason to do X, then X makes you better off.

P3* is true because if X will make you better off then you have a reason for doing it. Your reason is that you desire for your life to be better.

However, the argument from reasons involves P3* rather than P3 and P3 is false since there are a number of cases where we have good reasons to do X but X does not make us better off.

1. Imagine you witness someone break their leg. You have a reason for calling an ambulance (you want to help someone in need) but calling 911 won't necessarily make you better off.
2. Suppose you wish someone a good summer. You know you will never see them again. You have a reason for doing this (it is a nice thing to say and you want them to know you hope they do have a good summer) but wishing them a good summer won't make you any better off.

Exercise 9.1 Formulate one of the arguments for EG above.

Exercise 9.2 Some people argue that even heroic acts of kindness are done in order to benefit one's self. For example, people will say soldiers who run head first into gunfire to fight for their country or mothers who sacrifice their lives for their children are at bottom, motivated to do these acts because it will benefit them. What might you say to these arguments?

9.3 Arguments against ethical egoism

In this section, arguments against EG are considered.

9.3.1 Argument against common sense moral beliefs

P1: If EG is true, then it says our common sense moral beliefs (moral intuitions) are mistaken.

P2: Denying common sense requires powerful arguments.

P3: There is no compelling argument why our common sense moral beliefs (moral intuitions) are mistaken.

C: Therefore, EG is false.

Here is another way of putting the above argument: Assume ethical egoism is true. If ethical egoism is true, then individuals should do evil things (cheating, killing, etc.) as long as it benefits them. Any theory that says individuals should do evil things is false. Therefore, ethical egoism is false.

Note that P1 asserts that an action is morally good if and only if it promotes our personal gain. This means that there are cases where EG says that individuals who murder, rape, torture, steal, etc. are doing the morally good thing so long as it promotes their personal gain. However, there is the common sense belief that such actions are morally wrong.

P2 contends that individuals who deny common sense beliefs bear the burden of proof. For example, people generally believe that Big Foot does not exist. Thus, an individual who claims to have seen the Big Foot is saying something controversial (against common sense). It is up to them to prove (or support) their position with evidence. It is not up to those who accept the common sense belief to disprove the existence of the Big Foot.

P3 contends that the arguments in support of EG are less than persuasive. And so, C there is no reason to accept EG.

Another way of putting the above argument is that EG makes sinners saints and saints sinners. For consider that if EG is true, then it seems that the most capitalistic and self-interested of individuals would be the most moral and EG would justify their exploitation of the poor and disabled. But, the most capitalistic and self-interested of individuals are not the most moral. Therefore, EG is false.

Alternatively, if EG is true, then the most altruistic and self-abnegating individuals would be the least moral. The most altruistic and self-abnegating individuals are not the least moral. Therefore, EG is false.

9.3.2 Argument from lack of difference

Consider another argument against EG.

P1: If EG is true, then my self-interest is the only thing having intrinsic moral importance and your self-interest is unimportant. In other words, EG treats my self-interest as relevantly different than your self-interest.

P2: There is no relevant difference between my self-interest and the self-interest of others.

C: Therefore, EG is false.

In the case of P1, EG does not simply say that I'm allowed in certain cases to weigh my self-interest

as having more importance than yours. If we are both hungry and I have a sandwich, then the mere fact that I have the sandwich in my hands is enough a difference to justify that I eat the sandwich. Rather, P1 says that if EG is true, then there is a relevant difference between my self-interest and your self-interest such that I can treat my self-interest as having moral importance and yours as having none.

P2 contends that the only difference between our self-interests is that my self-interest belongs to me and yours to you. There is no argument for why this difference should make such a huge difference that I can treat my interest as having such importance and yours as having none.

Another way of thinking about this argument is that the critic of EG says that EG is arbitrary .

P1: The principle of equal treatment is the principle that we should treat people in the same way unless there is a compelling reason not to do so.

P2: Everyone accepts the principle of equal treatment.

P3: Any theory that rejects the principle of equal treatment is wrong.

P4: Ethical egoism rejects the principle of equal treatment as it says we should not treat ourselves in the same way we treat others. In fact, it says we should exclusively pursue our own interests.

C: Therefore, ethical egoism is false.

Exercise 9.3 First, try to formulate EG as a theory about moral right and wrong. Next, give an argument for EG in the clearest way you can. Third, give an argument against EG in the clearest way you can. Finally, even if EG isn't true, what about morality can you take away from it? What is its grain of truth?

10. Utilitarianism

10.1 Introduction

One intuition that people have concerning right and wrong action involves the consequences of an action. Many think that if one does something and that thing benefits many people, then one has done something morally good. In contrast, if someone does something and that action harms many people, then one has done something morally wrong.

10.2 What is Utilitarianism?

In this section, a consequentialist based theory of ethics is presented. The theory in particular that is considered is known as “utilitarianism”.

Definition 10.1 — Utilitarianism. Utilitarianism is the ethical theory that contends that an action is morally good if and only if it maximizes the overall amount of well-being in the world and an action is morally wrong if and only if it fails to maximize the overall amount of well-being in the world.

As a theory, utilitarianism makes several claims:

First, utilitarianism is a type of moral realism. Moral realism simply contends that there are moral facts, i.e. some state of affairs that makes a moral proposition true or false. This thesis contrasts with individuals who might adopt some sort of moral nihilism, viz., there is nothing that makes a moral proposition true or false..

Second, utilitarianism is a form of consequentialism in that actions are judged right or wrong not in themselves, but upon the consequences or state of affairs they produce. In other words, the consequentialist principle contends that no action is intrinsically right or wrong, actions are instead to be judged by their effects. In other words, actions cannot be judged as right or wrong in

terms of the nature of that action. Killing, saving someone's life, stealing, donating money to charity, being loyal to a friend, breaking a friend's trust all can be good or bad actions. What matters is what consequences such actions usher forth!

Third, utilitarianism accepts what might be called the well-being principle. The well-being principle asserts that the only relevant consequences are those that increase the well-being of individuals. In other words, the well-being principle contends that in evaluating whether or not an action is morally right or wrong, the only consequences that matter to such an evaluation are whether the action increases the overall well-being of individuals. Note that this theory thus depends upon a theory of well-being or prudential value (e.g. hedonism, objective list, desire-fulfillment theory, etc.).

In order to simplify things, we will consider a hedonist version of utilitarianism. Thus, we can restate the well-being principle as the hedonist principle. Namely, that in evaluating whether or not an action is morally right or wrong, the only consequences that matter to such an evaluation are whether the action increases the overall amount of pleasure relative to pain.²

Here it is important to note that "pleasure" refers not only to the quantity of pleasure but also the quality of pleasure. Let's distinguish between two different types of pleasure:³

pleasures of the flesh: short-lived feelings of elation, e.g. pleasure you get from eating a good meal
 pleasures of the intellect: more complex feelings that are dependent upon your "higher" faculties, e.g. the pleasure you get from being loyal to a friend or reading a good book.

At least one utilitarian (J.S. Mill) contends that the pleasures of the intellect are of greater quality than pleasures of the flesh. As these pleasures are of greater quality, they ought to count more in determining whether or not an action is right or wrong. Mill offers at least two (equivalent) arguments for why pleasures of the intellect are of higher quality than pleasures of the flesh. First, he contends that individuals who are acquainted with both of the pleasures would (more often than not) choose the pleasures of the intellect over the pleasures of the flesh. Second, Mill contends that except in cases of extreme depression human beings would never consent to be turned into a creature that was fully satisfied by the pleasures of the flesh.⁴⁵

Fourth, utilitarianism accepts what we can call the equality (impartiality) principle. The equality (impartiality) principle asserts that no individual's happiness is greater than any other. This principle follows from the claim that there is nothing that has intrinsic value. What makes something good or bad is the degree to which it produces pleasure (minus pain). As such, no individual thing capable of feeling pleasure or pain can, independently of appealing to the overall quality or quantity of pleasure produced by an action, claim that he or she somehow deserves to feel more pleasure than someone else.⁶

¹As Mill writes "pleasure, and freedom from pain, are the only things desirable as ends; and that all desirable things (which are as numerous in the utilitarian as in any other scheme) are desirable either for the pleasure inherent in themselves, or as means to the promotion of pleasure and the prevention of pain."

²Consider the following quotation from John Stuart Mill's Utilitarianism: "The creed which accepts as the foundation of morals. Utility, or the Greatest Happiness Principle, holds that actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness. By happiness is intended pleasure, and the absence of pain; by unhappiness, pain, and the privation of pleasure."

³(1) a short-lived experience or feeling of elation, e.g. the pleasure you derive from eating a good meal (2) a condition of being, typically lasting over a significant stretch of time, e.g. "a happy life"

⁴Mill writes: "Few human creatures would consent to be changed into any of the lower animals, for a promise of the fullest allowance of a beast's pleasures; no intelligent human being would consent to be a fool, no instructed person would be an ignorant, no person of feeling and conscience would be selfish and base, even though they should be persuaded that the fool, the dunce, or the rascal is better satisfied with his lot than they are with theirs."

⁵Elsewhere Mill writes "It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied." Mill saw this feature of the theory to be a part of the Epicurean theory for he writes "But there is no known Epicurean theory of life which does not assign to the pleasures of the intellect, of the feelings and imagination, and of the moral sentiments, a much higher value as pleasures than to those of mere sensation."

⁶As Mill writes, "But it is by no means an indispensable condition to the acceptance of the utilitarian standard; for that standard is not the agent's own greatest happiness, but the greatest amount of happiness altogether." [...] I must again repeat, what the assailants of utilitarianism seldom have the justice to acknowledge, that the happiness which forms the utilitarian

Thus, utilitarianism is the theory that contends that when deciding whether an action is morally good as opposed to morally wrong we need only consider whether that action maximizes everyone's happiness or decreases it. When deciding what to do in a given situation, we ought to perform the action that maximizes happiness (pleasure relative to pain). Thus, the utilitarian asks us not simply to maximize short-term pleasures but also long-term pleasures. That is, it encourages us not to be gluttons and focus on satisfying our basest desires but also to work toward accomplishing things that will make our overall condition more pleasurable, e.g. reaching goals.

Fifth, utilitarianism does not imply that we need to increase the overall amount of well-being for the greatest number of people. It asserts that we need to increase the overall amount of well-being in the world.

Example 10.1 If Tek has 10,000 dollars and there are 100 people. He could give each person 100 dollars. This might increase the well-being of each individual. However, he could instead give 1,000 dollars to the neediest (or poorest) 10 people. If giving 1,000 dollars to the neediest 10 people would bring more well-being into the world than giving 100 dollars to 1,000 people, then utilitarianism requires Tek to do the former rather than the latter.

Sixth, suppose we adopt a hedonistic version of utilitarianism. The theory does not say that we have to increase the overall amount of pleasure in the world. Instead, it says that we are required to increase the overall amount of pleasure relative to pain.

Example 10.2 Suppose Liz gains great joy from torturing Tek. This increases her overall amount of pleasure. The mere fact that it brings her pleasure does not mean she should torture Tek since this pleasure needs to be balanced against the amount of pain Tek receives from being tortured.

Seventh, utilitarianism contends that both the short-term and long-term consequences matter when evaluating whether or not an action is morally good or bad.

Exercise 10.1 Suppose that my friends and I steal a TV from a local electronics store. In weighing the consequences, I note that the day of the theft, I receive a great deal of pleasure from my action (with zero pain). I am up all night with my friends watching our new TV and the store owner receives a good deal of pain (with no pleasure) as a result of my action. Let's assume my action increases the overall amount of pleasure in the world. If I only analyzed these short-term consequences, I might reason that my act is a morally good act.

However, let's suppose that a few months later the store owner's insurance premium becomes more costly because of the theft. And this causes him to close his business. This causes him a great deal of pain (with no pleasure). Later still, a detective finally discovers that I am responsible for the theft and I am incarcerated. This causes me a great deal of pain.

Thus, when we take into account both the short-term and long-term consequences, my act may have decreased the overall amount of well-being in the world. It is thus seen, from the utilitarian point of view, as a morally wrong act.

Eighth, utilitarians generally contend that an act is morally good if and only if its actual consequences maximize well-being not the expected consequences

Example 10.3 Let's assume that Tek is driving and Tek intentionally runs Vic over with Tek's car. The expected result of the action is that Tek will actually decrease the overall amount of well-being in the world. And so, Tek has done something morally wrong.

However, let's suppose that Vic was on his way to torture a group of individuals he has been

standard of what is right in conduct, is not the agent's own happiness, but that of all concerned. I must again repeat, what the assailants of utilitarianism seldom have the justice to acknowledge, that the happiness which forms the utilitarian standard of what is right in conduct, is not the agent's own happiness, but that of all concerned. As between his own happiness and that of others, utilitarianism requires him to be as strictly impartial as a disinterested and benevolent spectator."

torturing on a regular basis. In stopping Vic, Tek has actually increased the overall amount of well-being in the world.

Ninth, utilitarians contend that an act is morally good if and only if it maximizes well-being not merely increases well-being

Let's suppose that Tek has a decision between three actions:

Action A (save people in a burning building) which results in a net increase of 100 units of happiness.

Action B (call the fire department to alert them to a burning building) which results in a net increase of 10 units of happiness.

Action C (lock the doors so people cannot escape the burning building) which results in a net decrease of 100 units of happiness.

Utilitarians contend that action A and only action A is morally good. It is the only act that is morally obligated. Action C is clearly morally wrong (both intuitively and from the standpoint of the utilitarian). However, utilitarians also contend that action B is morally wrong because while it does increase the overall amount of happiness in the world it does not maximize the overall amount of happiness in the world.

Exercise 10.2— Applications. Consider the one of the following three topics:

1. euthanasia of an individual in extreme, untreatable pain
2. recreational use of marijuana
3. the eating of nonhuman animals

According to the utilitarian, are any of these acts ever morally permissible? If so, under what conditions? If not, why not?

10.3 Three core arguments against Utilitarianism

In this section, we consider a number of objections to Consequentialism (Utilitarianism). In the next, we respond to these objections in an effort to further defend the theory.

Let's begin with the following objections to the consequence principle.

10.3.1 Argument from injustice

The utilitarian thesis allows for the ends to justify the means. This strikes us as counterintuitive when the means involve a violation of basic human rights, fairness, and justice. An argument for this objection would go as follows:

P1: If utilitarianism is true, then sometimes unjust acts would be morally permissible.⁷

P2: Unjust acts are not morally permissible.⁸

C: Therefore, utilitarianism is not true.

With respect to P1, since there are cases where acting unjustly may maximize well-being, utilitarians would say that these acts are morally obligated. That is, there are cases where violating someone's right to life, privacy, freedom of speech, bodily-integrity, or freedom from unjust punishment are not merely morally permissible but morally obligated.

Example 10.4 — false conviction. Suppose there has been a string of murders in a city. Next,

⁷ That is, provided it would increase the overall amount of well-being in the world (good consequences), it would be acceptable to violate someone's basic human and legal rights.

⁸ That is, it is not acceptable to violate someone's basic human and legal rights.

suppose that Tek (a police detective) is responsible for these murders. The city is in a complete panic: schools are closed, people are afraid to go out at night, etc. Now suppose that Tek feels bad about these murders, decides to dedicate his life to increasing well-being in every way he can. However, he decides that first he must pin the murders on an individual. He reasons as follows:

P1: I ought to maximize well-being

P2: If I pin the murders on someone else, then the well-being of the inhabitants of the city will increase by 100 units (I plan to do a lot of good in the free world).

P3: If I confess to the murders, then the well-being of the inhabitants of the city will increase by 50 units (I won't be able to do all the good I plan to do if I am in prison)

P4: If someone is not held responsible, then the well-being of the inhabitants of the city will decrease by 50 units (the citizens will live in fear).

C: Therefore, I ought to pin the murder on someone else.

Example 10.5 — organ harvesting. Suppose that there are 1000 important people (people who greatly increase the well-being of the world) who need organ transplants. Now suppose there are 1000 people who do not greatly increase the well-being of the world and who would not be missed if they disappeared. According to utilitarianism, it would be morally permissible to capture these individuals and harvest their organs for those who will increase the overall amount of well-being.

With respect to P2, we intuitively think that individuals have basic moral and legal rights. That is, that all individuals are born with certain rights that only can be forfeited by the individual (not simply taken away without due process).

In short, utilitarianism is counter to our intuitions about what qualifies as good action as it requires us to do what produces the best overall outcome in terms of happiness even if the means by which we produce this happiness is horrific.

10.3.2 Argument from supererogation

Consider that Utilitarian accepts the consequence principle. If this is the case, then it isn't that your intentions to do good that matter nor is it the intrinsic nature of the act. Rather, it is the whether or not your act maximizes the overall amount of well-being in the world. For some, this makes utilitarianism too demanding. To see this argument more clearly, let's define the notion of supererogation.

Definition 10.2 — supererogation. Supererogation refers to action that is above and beyond your moral duty (e.g. morally heroic action). An act is said to be supererogatory if and only if that act is morally above and beyond your moral duty.

Example 10.6 Some acts that are supererogatory include

sacrificing your life for a total stranger

donating the majority of your money to the poor

spending all of your free time doing volunteer work at a local children's hospital

Critics of the utilitarian contend that utilitarianism incorrectly makes supererogatory acts morally obligatory. In other words,

P1: If utilitarianism is true, then supererogatory acts are morally obligated acts.

P2: Supererogatory acts are not morally obligated but those above and beyond our moral duty.

C: Therefore, utilitarianism is not true.

To see that P1 is true, suppose that you have two choices:

1. Running into a burning building and saving an important doctor on the top floor but dying in

the process

2. Running into a burning building and saving a doctor of lesser importance but living in the process.

Intuitively action (1) is supererogatory while action (2) is simply morally good albeit less heroic. According to utilitarianism, your obligation is to sacrifice your life for the important doctor and not to save the doctor of lesser importance. That is, utilitarianism says (1) is not supererogatory but is instead your moral duty.

P2 is said to be intuitively true.

Thus, utilitarianism is too demanding. Given the current state of the world and the impoverished state so many people live in, treating everyone's happiness equally would currently demand that we give away nearly all of our money and do everything we can to help others (of course, we should not give away so much that we decrease happiness). On the basis of this, people reject utilitarianism because (i) they think that this asks too much of us and it (ii) confuses acts that are supererogatory (one's that go above and beyond our moral requirements) rather than obligatory.

10.3.3 Argument from partiality

Utilitarianism relies on the principle of equality (impartiality). This principle contends that the well-being of each individual is of equal value. In other words, no individual's well-being is intrinsically worth more than any other individual's well-being. This principle, at first glance, is an attractive part of the theory since it would contend that billionaires using their wealth to increase their own personal well-being a fraction of a degree is morally wrong since they could significantly increase the overall well-being in the world by say building a school or feeding the poor.

However, individuals often have two moral intuitions:

1. there are certain circumstances where we are morally obligated to increase the well-being of people we know or care for even if it we could increase the overall well-being in the world to a greater degree
2. there are certain circumstances where it is morally permissible to increase the well-being of people we know or care for even if it we could increase the overall well-being in the world to a greater degree

Example 10.7 — moral obligation to increase well-being of loved ones. Suppose you have a dependent (e.g. baby, dependent parent). This dependent is your responsibility. Intuitively, it seems that you have a moral obligation to increase the well-being of this dependent even if you could increase more well-being by neglecting your dependent and helping others.

Example 10.8 — morally permissible to increase well-being of loved ones. Suppose my child needs reading glasses. I could pay for those reading glasses or I could donate the money to purchase reading glasses for 10 children in a far away land. Many think that it is morally permissible for me to buy my child reading glasses even if I could increase more well-being by buying 10 children in a far away land reading glasses.

To summarize this argument against utilitarianism:

- P1: If utilitarianism is true, then the equality /impartiality principle is true.
- P2: If the equality/impartiality principle is true, then we are not allowed to be preferential toward family members, loved ones, or friends.
- P3: There are cases where it seems that it would be morally permissible to be preferential toward family members, loved ones, or friends.
- C: Therefore, utilitarianism is not true.

10.4 Other arguments against Utilitarianism

10.4.1 Argument from the past

Utilitarianism Ignores the Past and is thus incompatible with how past actions bind future behavior (e.g. keeping promises, loyalty in relationships).

P1: If utilitarianism is true, then we should solely focus on the consequences that will result from an action and ignore any factors that led to that action.

P2: There are times when factors that lead up to an action matter for how we evaluate the goodness or wrongness of an action, e.g. making/breaking promises, evil actions, etc.

C: Therefore, utilitarianism is not true.

Consider the following example. After several cancelled meetings with a friend, you swear on your life that you will meet me her at 5PM at a restaurant you both like, but when you weigh the pros and cons, you find that it would very slightly increase the overall happiness of the world if you did not show up. Utilitarianism says that this very slight increase in happiness allows you to break the solemn oath you made.

10.4.2 Argument from distribution

Utilitarianism commands individuals to increase the overall amount of happiness in the world. In other words, it says good action corresponds to those actions that lead to an increase in the net amount of happiness. This is problematic as it does not account for the distribution of happiness.

Utilitarianism asks us to increase the total amount of happiness independent of how that happiness is distributed. This is compatible with two counterintuitive results. First, it is compatible with increasing the total amount of happiness by increasing the total number of people yet decreasing how happy each individual is. Second, it is compatible with a society where some individuals are insanely happy but others are living in misery.

10.4.3 From intentions

Utilitarianism does not consider the intentions of the actor. Consider three individuals: Liz, Jon, and Vic. Suppose that Jon is attempting to hurt Vic while Liz is attempting to make him happy. However, in their respective efforts, Jon ends up improving Vic's life (making him happy) while Liz ends up ruining his life (making his life worse). According to the utilitarian theory, it is Jon, not Liz, who has done something good.

10.4.4 Argument from experience machine

Utilitarianism is a form of hedonism and hedonism is false: [experience machine] As an illustration of this objection, consider the following example: ⁹

Suppose there were an experience machine that would give you any experience you desired. Superduper neuropsychologists could stimulate your brain so that you would think and feel you were writing a great novel, or making a friend, or reading an interesting book. All the time you would be floating in a tank, with electrodes attached to your brain. Should you plug into this machine for life, preprogramming your life's experiences? If you are worried about missing out on desirable experiences, we can suppose that business enterprises have researched thoroughly the lives of many others. You can pick and choose from their large library or smorgasbord of such experiences, selecting your life's experiences for, say, the next two years. After two years have passed, you will have ten minutes or ten hours out of the tank, to select the experiences of your next two years. Of course, while in the tank you won't know that you're there; you'll think it's all actually happening. Others can also plug in to have the experiences that they want, so there's no need to stay

⁹Based on Robert Nozick, "The Experience Machine," pp. 42-5 from *Anarchy, State and Utopia*. New York: Basic Books, 1974.

unplugged to serve them. (Ignore problems such as who will service the machines if everybody plugs in.) Would you plug in? What else can matter to us, other than how our lives feel from the inside? Nor should you refrain because of the few moments of distress between the moment you've decided and the moment you're plugged. What's a few moments of distress compared to a lifetime of bliss (if that's what you choose), and why feel any distress at all if your decision is the best one?

P1: The utilitarian accepts the happiness principle and so reduces all morality to pleasure-seeking.

P2: However, even if the type of pleasure we seek is not base in nature, there are things we think are valuable that are not pleasure, e.g. honesty, bravery, etc.

C: Therefore, utilitarianism is false.

10.5 Defending Utilitarianism

Given the arguments against utilitarianism, utilitarianism might seem as though it does not offer us a guide for picking out the ideal course of action, but let's consider a defense of this theory.

10.5.1 Response to the argument from injustice

Recall the argument from injustice:

P1: If utilitarianism is true, then sometimes unjust acts would be morally permissible.

P2: Unjust acts are not morally permissible.

C: Therefore, utilitarianism is not true.

First, utilitarians might deny P1. While we might recoil at the horrible acts that utilitarianism seems to require, the utilitarian might require that is because we are looking at things in a short-sided way. If we were to consider all of the consequences of our actions, we would see that these unjust acts are not as morally repugnant as they first seem.

Case 1: Wrongfully convicting someone makes the person who was convicted unhappy, but what about all of the good that came of this, e.g. makes the family of an unknown killer happy, eases hysteria, maybe causes less crime, etc.

Case 2: If 1000 people are killed for their organs and these organs are gifted to 1000 people who will greatly increase the well-being of the world, we might focus on the 1000 deaths. However, suppose that the 1000 people who increase the world's well-being save 10,000 lives. Here we are sacrificing the few for the many.

The point is that even though utilitarianism does not square with our intuitions about what is right and wrong, it may be that we are selectively focusing only on the negative consequences and not paying enough attention to the positive consequences.

Second, utilitarians might deny P2. They might contend that the rights of individuals are not absolute. Consider that we often put individuals to death, restrict their freedom of speech when their speech is hateful, or seize their property when it is in the interests of the public good. The idea here is that there are cases when the public good will override the rights of the individual. And, in these cases, while the act is unjust, it is nevertheless the morally right thing to do.

10.5.2 Response to impartiality

Recall the argument from impartiality

P1: If utilitarianism is true, then the impartiality principle is true.

P2: If the impartiality principle is true, then we are not allowed to be preferential toward family members, loved ones, or friends.

P3: There are cases where it seems that it would be morally permissible to be preferential toward family members, loved ones, or friends.

C: Therefore, utilitarianism is not true.

Utilitarians might simply reject P2. For consider that the individuals that we have the greatest impact in terms of improving or harming their well-being are those that are closest to us. For example, if Tek lies to his wife, she not only be harmed by the lie itself but further damage comes from the fact that Tek is her husband. On the other hand, if a good friend pays you a compliment, it may mean more than if a stranger pays you a compliment.

In short, while utilitarians may contend that you may not value the well-being of loved ones (or yourself) over that of others, this does not imply that we ought not to be preferential. We ought to be preferential since our actions can have the greatest impact on those that are closest to us.

10.5.3 Response to the argument from supererogation

One argument that we considered was that utilitarianism is too demanding. This is the case because it has no place for supererogation in it for all supererogatory acts are morally obligatory.

There are two responses to this argument.

First, the utilitarian can contend that its critics have no successful response to the following argument:

P1: If you have a choice between doing better or doing worse, you should always do better.

P2: Doing better is best defined as increasing well-being in the world.

C: Therefore, we ought to always maximize well-being.

Second, another response for the utilitarian is to bite the bullet. That is, they simply need to admit that utilitarianism is a very demanding theory but it is nevertheless true. In other words, utilitarianism does go against certain moral intuitions we have about what is morally permissible and impermissible, but this is because our intuitions are wrong. Yes, utilitarianism is demanding in terms of what it asks us to do, but this is because some of our moral beliefs are wrong and we should change them to fit the utilitarian theory. How can the utilitarian motivate this claim? There are several options:

1. point to cases where our intuitions have been wrong before. For example, remember that intuitions about what is right and wrong have been wrong in the past (e.g. slavery) and there are probably other beliefs we have that are also wrong.
2. point out that when we want to do something less than maximize well-being, this is often for self-interested gains. For example, I don't give most of my money away to needy people because I want to increase my well-being.

10.5.4 A Modification to Utilitarianism: Act vs. Rule

The version of utilitarianism we have been considering is act utilitarianism. This species of utilitarianism says that particular acts are morally good or bad according to whether that act increases or decreases happiness. One problem with the act utilitarian approach is that it seems to allow for some exceptionally bad actions in certain situations, e.g. murdering someone and selling their organs. This kind of outcome is horrifying even if we focus on ALL of the positive consequences.

One way to potentially fix this problem is not to consider the positive /negative consequences of individual acts but the positive /negative consequences of rules of action. Rule utilitarianism says that a particular act is good or bad provided it is done according to certain rules that generally maximize happiness. In other words, rule utilitarianism says an action is morally good if it is done in accordance with a rule that yields (overall) happy consequences and morally wrong if it is done

in accordance with a rule that does not yield a favorably happy outcome.

Whereas act utilitarianism might allow for an occasional act where an individual is killed and their organs are harvested for certain individuals that may produce great happiness in the world, the rule utilitarian would say that a governmental policy that rounds up “undesirables” and harvests their organs for the “desirables” would be morally wrong as it would produce an extreme amount of unhappiness in the world (rampant fear, outrage, etc.).

The above response can be generalized to a variety of other objections to utilitarianism. For consider the fact that utilitarianism will allow for lying and someone individuals may object to this.

P1: If utilitarianism is true, then it is morally permissible to lie to your friend.

P2: Ethical theories should not allow lying.

P3: Therefore, utilitarianism is not true.

However, a utilitarian can argue that P1 is false because a careful consideration of all of the consequences would show that certain acts of lying actually increase the overall amount of well-being in the world. That is, in the short-term you may think that lying will help you get away with x, y, or z, but it seems as though you will ultimately get caught and this will have very bad consequences, e.g. no one will trust you, your friends will dislike you, etc.

Exercise 10.3 Come up with your own example of a case where an isolated act might be morally good (according to the act utilitarian) but morally bad (according to the rule utilitarian)

10.5.5 Defense 3, Utilitarianism and Common Sense

One objection to utilitarianism is that it goes against common sense by being too demanding.

First, the utilitarian might say that rather than running counter to common sense, it actually explains common sense. If you consider the moral principles you hold (e.g. don't lie to friends), ask yourself the following question: “why do we accept these principles?”

One answer to this question is that because holding to such a principle, on the whole, makes people happy more often than not (rule utilitarianism here!). And so, the utilitarian can argue as follows:

P1: We would only accept an ethical theory that can explain the moral principles we currently hold.

P2: Utilitarianism best explains our current moral principles.

C: Therefore, we should accept utilitarianism.

Table 10.1: Utilitarianism and Common Sense

10.6 Utilitarianism: Applications

Exercise 10.4 1. What is utilitarianism?

2. For the utilitarian, is euthanasia ever morally acceptable?

3. For the utilitarian, is the use of marijuana ever morally acceptable?

4. According to the utilitarian, must animals and humans be treated the same way?

5. What is “speciesism”?

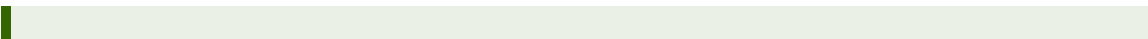
6. What is the difference between act and rule utilitarianism?

7. What is act utilitarianism?

8. What is rule utilitarianism?

9. Give one argument for utilitarianism

10. Give one argument against utilitarianism?



11. Deontology

11.1 Introduction

The utilitarian theory contends that what makes a moral act good or bad depend upon whether or not that act maximizes well-being. While we may disagree with the theory in several places, intuitively it seems that the moral goodness or moral badness is somehow connected to the consequences of an act. If an individual saves someone from a burning building, we regard them as morally heroic. If an individual preys upon the weak and exploits the less fortunate, then we regard them as morally evil.

These moral intuitions seem to focus on the consequences but perhaps these do not matter at all. Perhaps, what truly matters when we morally evaluate an action is not the consequences of the action but the intentions of the agent who performs the act. That is, provided an individual had good intentions or a good will when acting, then their action is morally good. In contrast, if the agent failed to have a good will when act, then their action is morally bad.

In what follows, we will explore the ethical theory of deontology (specifically Kant's deontological theory), the theory that insists that an act is morally good if and only if it is done for the right reason (with the right intentions) and this reason is to do one's duty.

Exercise 11.1 Create a scenario where we say an action was morally good even though the act had bad consequences.

11.2 The good will

There are two problems with respect to the claim that what matters in the moral evaluation of an act are the individual's intentions and not the consequences of the act. These problems are as follows:

- is it true that a good will is the only thing that matters?
- what does it mean for an individual to have a good will ?

11.2.1 Problem 1: whether the good will is the only thing that matters

The first claim we will look at is the claim that the good will is the only thing that is absolutely good. It is the only thing that matters in the moral evaluation of an action.

Kant argues that it is the intentions (or the reason, principle underlying why we did the act). In Section I, he argues that the only thing that can be considered absolutely good is the good will. Why is this the case?

On Kant's account, everything else can be corrupted (or used for evil in some way).

Example 11.1 — intelligence. We might contend that being smart is a good thing. Intelligent people are capable of helping others in all sorts of different ways, e.g. curing the sick, but intelligence is not absolutely good since we can imagine all sorts of ways that intelligence might be put to evil use.

Example 11.2 — any other admirable character trait. Next, consider certain admirable character traits like courage or perseverance. These too can be put to evil use. Lastly, consider gifts of fortune e.g. being rich, being healthy, having a political power. Again, we can think say that these may be good for a person to have, but they are not good in every way. Being politically powerful can make you arrogant (corrupt your character) and you might use your wealth to do evil things.

Kant claims, the good will—that is, the will to do good because it is good—is absolutely good in every way. What makes the good will absolutely good? Kant says that the good will is not good because of what it can accomplish or because it allows individuals to increase happiness or make the world a better place. He writes that “[a] good will is not good because of what it effects or accomplishes” and that “usefulness or fruitlessness can neither add anything to this worth nor take anything away from it”.

One reason that the good will is the only thing that is absolutely good in every way is because it is the only thing that we can control.

Example 11.3 To illustrate this idea, consider an individual who is born with limited intellectual gifts, rather weak character traits, and into highly unfortunate circumstances. Finally, suppose this individual has a good will. Now in virtue of having a good will, this individual aims to do good because it is good in every circumstance. Unfortunately, due to this individual's circumstances, s/he always fails. Kant contends that this shouldn't matter in the moral evaluation of the individual's actions. Kant contends that even if “this will should wholly lack the capacity to carry out its purpose [...] then, like a jewel, it would still shine by itself, as something that has its full worth in itself.”

To put Kant's point in different words, Kant is claiming that the goodness of the good will is intrinsic rather than extrinsic (or instrumental).

Definition 11.1 — extrinsic (instrumental) good. Something is extrinsically (instrumentally) good if and only if the goodness of that thing is a means of obtaining something else.

Definition 11.2 — intrinsic good. Something is intrinsically good if and only if that thing is not a means of obtaining something else, if that thing is good in and of itself, if that thing is worth pursuing for its sake alone.^a

^aIntuitively, we can distinguish between extrinsic and intrinsic goods by asking a series of “why” questions. For example, having money might be something regarded as good, but it is usually not something people pursue just to have it. That is, if we ask someone “why do you want money?” They are likely to give us answer like “to buy things, e.g. houses, cars, vacations, etc.” Money is thus an instrumental good. It is a means to some other end. In other words, money is an instrumental good for something else. But again, we can ask the individual “why they want possessions”. At some point, individuals are likely to no longer have an answer to our “why” questions. For example, suppose an individual were to

say they desire possessions because those possessions make them happy. If we ask them why they want to be happy, they may not have an answer for us. They might simply say that they simply want to be happy and they don't want to be happy for some other reason or for something else. Here "happiness" is regarded as intrinsically valuable

11.2.2 Problem 2: what does it mean to have a good will

Let's suppose that if Tek does action X, then X is good if and only if Tek had "good intentions" or a "good will" in doing X. However, what does it mean to have "good intentions" or a "good will"?

Exercise 11.2 Can you think of a case where an individual (i) believes themselves to have good intentions but (ii) their act is not morally good? That is, where they think they have a good will but are mistaken about what it means to have good intentions?

To get a better understanding of the absolutely goodness of the good will (and what this implies), we need the following points:

1. It is the role of reason is to purify the will (to make it a good will)
2. Action done from duty is conceptually independent from action done on the basis of inclination/impulse.
3. Evaluating the moral worth of an act involves evaluating its maxim
4. The moral worth of an action done from duty comes from a maxim that is independent from any particular human desire.
5. Duty is the necessity of an action from respect for law.
6. The categorical imperative is the general form of law that is worthy of our respect

11.2.2.1 Point 1: The role of reason is to purify the will

The ethical theory that we will present depends upon a view of rational agents. That is, rational agents (of which human beings are included) consist of three different powers:

1. The power to reason (our reason)
2. The power to quickly decide (our instinct)
3. The power to choose (our will)

Let's further suppose that each power has its own special function. In particular, instinct and reason must have their own specific functions. ¹

First, note that the function of reason cannot be the preservation of life or happiness since this would be better served by instinct. Thus, Kant argues that the faculty of reason must have some other end or function. ²

Second, one hypothesis then is that if reason does not serve the aim of furthering our goals or life, then its function must be related to the influence the will. In particular, the goal of reason is to perfect the will, to make it good. ³

11.2.2.2 Point 2: Action from duty is conceptually distinct from action from inclination

We can conceptually distinguish between two different types of actions:

1. actions done from duty

¹"In the natural constitution of an organized being, that is, one constituted purposively for life, we assume as a principle that there will be found in it no instrument for some end other than what is almost most appropriate to that end and best adapted to it."

²The argument for this can be put forward as follows: No instrument in nature exists for one end that is better suited for another end. Reason would not be best-suited for human happiness since instinct could secure happiness more quickly and directly with less interference. Therefore, reason must have another end or function.

³Kant writes "the true vocation of reason must be to produce a will that is good, not perhaps as a means to other purposes, but good in itself for which reason was absolutely necessary."

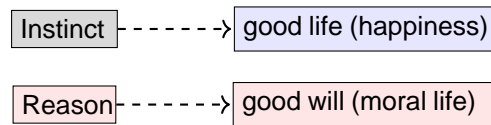


Figure 11.1: Role of reason is to produce a good will

2. actions done from impulse/inclination

To see this distinction more clearly, let's consider the following cases:

1. acting against duty from inclination.
Example: Doing something you know is wrong because it benefits you.
2. acting in conformity with duty but from inclination.
Example: Doing something you know is right because it benefits you.
3. acting from duty in conformity with inclination.
Example: Doing something that benefits you but not because it benefits you but because you know it is right.
4. acting from duty against inclination.
Example: Doing something that does not benefit you because you know it is right.

Kant contends that while we can conceptually distinguish acting from duty from acting from inclination we often cannot practically distinguish one from the other. The reason for this is particularly evident in cases (2) and (3).

Example 11.4 — Case 2 and Case 3. Suppose, for example, we have a duty not to lie (or to be honest). With respect to case (2), an individual might tell the truth and so is acting in conformity with the duty. But, the reason s/he tells the truth might simply be because it is beneficial to him or her (e.g. s/he might tell the truth to a friend because s/he thinks that the friend might ultimately find out s/he is lying and enact revenge).

With respect to case (3), an individual might tell the truth and while this might benefit the individual, s/he only tells the truth because telling the truth is what one ought to do.

Now notice that there is no real way to distinguish between cases (2) and (3) because we would need to know the individual's inner motivation for doing the action (something Kant thinks we can't have access to). Nevertheless, cases (2) and (3) are different.

Exercise 11.3 Can you think of an example of each of the above cases?

11.2.2.3 Point 3: Evaluating the moral worth of an act involves evaluating its maxim

If what is relevant to evaluating the moral worth of an act are the intentions or the will of an agent who does an act, then we want to know two things.

1. We want to know what they are trying to do
2. We want to know why they are trying to do that act

The answers to these two ingredients are known as an action's maxim.

Definition 11.3 — maxim. The maxim of an action is (i) a formulation of what you aim to do and (ii) why you aim to do it

Example 11.5 — Lie to get a loan. Suppose I am going to lie to get a loan. My maxim would be: (i) I plan to say I will pay the loan back but not pay the back and (ii) I will lie because if I told the truth I would not get the loan

The idea here is that when we evaluate whether an action from duty is morally good or bad, we do not look at the outcome of that action from duty, but instead look at (i) what the action aimed to do and (ii) why it was performed.

11.2.2.4 Point 4: The maxim must be independent of any particular desire

The moral worth of an action done from duty comes not from the action's purpose (the goal of action) but from a maxim (rule, principle) that is independent from any particular human desire.

The idea here is that in order for a maxim to be morally good, it must be impartial. That is, it cannot take into account the special interests of you or your neighbor. The reason for this is that if the maxim were formulated with the interests of a specific person, it would be a maxim whose goodness would depend upon its consequences rather than its intent.

11.2.2.5 Point 5: When we act from duty, that act is necessitated by respect of a law

The idea here is that given some principle or law, if we respect that law, then we necessarily act in a way with respect to that law. For example, consider a moral law like "you should never lie; always be honesty". If I respect that moral law, then I will necessarily act in accordance with it. That is, I will not lie.

Example 11.6 — Parents. Suppose you respect your parents and they tell you to do X. The point here is that if you respect your parents (or some moral rule) then you would do X, not because it is of benefit to you or because you calculate the consequences but simply because you think that X is worthy of your respect.

11.2.2.6 Point 6: The moral law is worth our respect

Let's take stock. We have shown that the good will is a will that is purified or perfected by reason. It is a will that acts out of duty rather than inclination. It does not act from a particular desire but instead acts out of a respect for certain types of laws (see Figure 11.2).

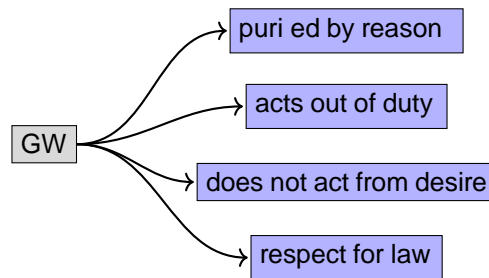


Figure 11.2: The good will is purified by reason, acts out of duty or respect for law (not desire).

The question this is what sort of rule, maxim, principle, or law—when it is represented to us—elicits our respect and impels us to act in accordance with that law? For Kant, the type of law that could command the respect of reason is the moral law which is articulated as a categorical imperative. Before we articulate this moral law, we need to make a few distinctions.

11.2.2.7 Clarification: imperatives, hypothetical and categorical

Note that the rule that the good will respects is what is known as a categorical imperative. Let's get clear on this by first defining an imperative.

Definition 11.4 — imperative. An imperative is a command expressed using terms like “ought”, “should”, etc.

Imperatives command in two different ways: hypothetically or categorically. Kant says that hypothetical imperatives “represent the practical necessity of a possible action as a means to achieving something else that one wills (or that it is at least possible for one to will).”

Definition 11.5 — hypothetical imperative. A hypothetical imperative is a rule that specifies what we ought to do if we want to reach some end, satisfy some desire, or accomplish some goal.

Hypothetical imperatives are typically expressed as “if you want x, you ought to do y.” For example, if you want an A on the exam, you ought to study is a hypothetical imperative. Hypothetical imperatives only apply to those individuals who have the relevant end, desire, or goal they want to accomplish. And so, if I don't desire to get an A on the exam (maybe I am happy with a D), the hypothetical imperative if you want an A on the exam, you ought to study does not apply to me. In short, desires make hypothetical imperatives possible.

Kant says that categorical imperatives are those commands which represent “an action as objectively necessary of itself, without reference to another end.”

Definition 11.6 — categorical imperative. A categorical imperative is a rule that specifies what we ought to do something (period), i.e., not on the condition that we have some desire we want to satisfy or goal we want to accomplish.

Whereas hypothetical imperatives are typically expressed as “if you want x, you ought to do y”, categorical imperatives are expressed as “you ought to do y”. Whereas hypothetical imperatives only apply if I have the relevant inclination, categorical imperatives are said to apply to all rational beings and are independent of our inclinations or goals. Whereas desires make hypothetical imperatives possible, it is thought that the faculty of reason makes categorical imperatives possible.

Both hypothetical and categorical imperatives represent some action as good. Hypothetical imperatives represent an action as good insofar as it is a means to something else (extrinsic good). Categorical imperatives, by contrast, represent an action as good in itself (intrinsic good).

11.2.3 Point 6: Returned, the moral law as a categorical imperative

For Kant, the law worthy of our respect is the moral law. Kant provided several formulations of this ultimate moral law that he took to be expressed as a categorical imperative:

Act only according to that maxim by which you can at the same time will that it should become a universal law.⁴

Another version of this same categorical imperative reads as follows:

Act as if the maxim of your action were to become by your will a universal law of nature.

Here Kant formulates the rule (or law) that a truly good will follows. This rule is known as the categorical imperative. It states that an act is morally good if and only if the maxim can be consistently willed after being universalized (or made into a law of nature). It is this type of rule

⁴Kant writes this concerning the type of law: “But what kind of law can that be, the representation of which must determine the will, even without regard for the effect expected from it, in order for the will to be called good absolutely and without limitation? Since I have deprived the will of every impulse that could arise for it from obeying some law, nothing is left but the conformity of action as such with the universal law, which alone is to serve the will as its principle, that is, I ought never to act except in such a way that I could also will that my maxim should become a universal law.”

that is worthy of our respect . When an individual acts according to such an imperative, then the individual has good intentions or a good will .

11.2.3.1 Illustration

Let's give an example of the good will and its following of the categorical imperative.

Example 11.7 — Lying. Suppose John lies to Sally. He tells her he needs a loan and will pay her back but he knows that he will not pay her back. To determine if this act is morally wrong, we only need to follow the following process:

Precisely formulate the maxim: I (John) aim to acquire money from Mary by lying to her that I will pay it back so that she gives me money.

Universalize that maxim: Imagine a world where everyone knows and acts on your maxim. Whenever someone wants money and cannot pay it back, they will say they will pay it back even when they will not pay it back.

Consistency check: Can you aim be accomplished if the maxim is universalized? If yes, then the action is morally permissible. If no, then the action is morally impermissible. In the case, the answer is the act is morally impermissible for if everyone lied when they wanted money but did not (or could not) pay it back, then no one would lend money. And so the aim of the action cannot be accomplished.

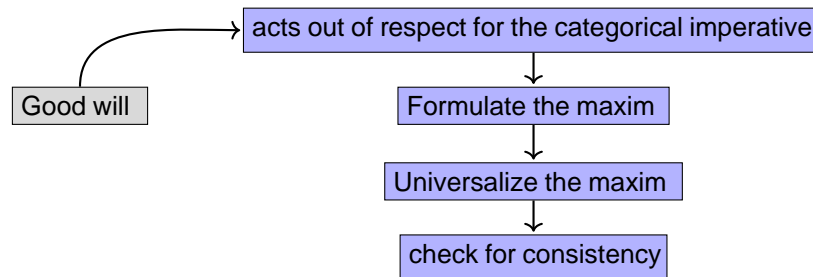


Figure 11.3: The good will acts out of respect for the categorical imperative.

Exercise 11.4 Suppose Tek kills John because he wants John's money. Follow the step-by-step approach laid out above. Is the act of killing morally permissible.

11.2.4 The good will: a summary and a theory

Recall that after establishing that the good will is the only thing that is absolutely good, the next issue we had to tackle was to articulate the nature of the good will. An individual may contend that they have a good will but it isn't clear what that involves.

In our effort to articulate what it would mean to have a good will, we have articulated a deontological theory of ethics .

Definition 11.7 — deontology. Deontology is the ethical theory that contends that an act is morally good if and only if the maxim of that action is done out of respect for (or duty to) the moral law.

The theory contends that a will is good if and only if a universalized version of the maxim upon which it is based can be consistently willed. In other words, an act is morally good if and only if it conforms to (or is done out of duty to) a particular type of rule, namely the categorical imperative.

11.2.5 A problem with the categorical imperative

One issue with the categorical imperative is that it says an act is morally permissible if and only if you can consistently will that it become a universal law. But this seems to exclude the possibility of exceptions. For consider the case of lying. The deontological theory might contend that it is never permissible to lie, we might be willing to make exceptions in some scenarios. For example, consider a case where lying would prevent the death of someone. If this is the case, then P2 would be false since the maxim would be it is acceptable for me to lie when it is to save someone's life. But, it seems that it is possible to will that this maxim be universalized. And so, there is no absolute moral rule not to lie.

11.3 The principle of humanity

Kant formulates a second version of the categorical imperative. Much like the first version (the universal version), the second version specifies a condition on what it would mean for an individual to have a good will. Namely, an individual with a good will would act out of a sense of duty to (or respect of) the categorical imperative rather than being motivated by personal desires and /or inclinations.

The second version of the categorical imperative is which we will call the "Humanity Formula of the Categorical Imperative". Kant contends that rational beings (human beings included) are something of absolute worth, having intrinsic value, and that they are deserving of our respect. For this reason, Kant contends that they should not be treated merely as means to some end, but as ends-in-themselves.

Note [The Categorical Imperative - Humanity Version] [T]he human being and in general every rational being exists as an end in itself, not merely as a means to be used by this or that will at its discretion; instead he must in all his actions, whether directed to himself or also to other rational beings, always be regarded at the same time as an end

The humanity version of the categorical imperative makes a number of claims:

- there is a distinction between rational beings and non-rational beings
- our actions must always be directed toward treating rational beings (including ourselves) as an end
- we can treat rational beings as a means to an end but never merely as a means

First, Kant contends that there is a distinction between rational beings and non-rational beings. This distinction might be clarified as the distinction between things and persons.

Definition 11.8 — thing. a thing is a being without reason

Microwaves, televisions, insects, and some (perhaps all) animals are all things. For Kant, they are things because they lack the power to reason and self-legislate (autonomy). Their decision making (if they have the power to decide at all) is governed not by the capacity to evaluate choices and select from those choices in a rational way but by instinct, habit, or inclination.

Definition 11.9 — person. a person is a being with reason and autonomy

In contrast, persons are individuals with the power to reason. They can use this faculty to evaluate options, determine the best way to achieve goals, engage in complex discursive tasks. In addition, persons are autonomous. This means that they are able to self-legislate: they are not a slave to their passions, are capable of resisting some temptations, and are able to make decisions based on

consciously-selected rules rather than mere urges and instincts.

Second, our actions must always be directed toward treating rational beings (including ourselves) as an end. What this means is treating persons as individuals capable of rational self-determination. Individuals are free and the humanity principle requires that we respect this fact about other rational agents.

Third, the humanity version of the categorical imperative says that all things can be used merely as a means to an end but persons can never be used merely as a means to an end

What this means is that you can use things wholly to achieve some goal without consideration of what that thing is or what that thing wants or what is good for that thing.

Example 11.8 — microwave. I can use a microwave to heat my food without regard to whether the microwave wants to be used in this way. I can use it simply to achieve some end.

What this means for persons is they cannot be used purely instrumentally. That is, you cannot use someone to achieve some goal or end without taking into consideration that the fact that the person is rational and autonomous. In other words, the person cannot be used without their rational consent.

Example 11.9 — piggy-back ride. Suppose I want a piggy-back ride. Now imagine two scenarios:

A: If I were to jump on your back and force you to give me a ride, this would be to use you purely as a means to an end am using you as a tool or thing and not taking into consideration your status as a person.

B: If I were to ask to jump on your back and you agreed on the condition that I would give you a ride later, then I would be using you as a means to an end but not purely as a means to an end would be taking into consideration you as a rational, self-determining being and noting that this is something you willingly accept.

Why can't human beings be merely treated as a means? For one, if human beings could be used merely as a means to an end, then their value is conditional. That is, they would only be valuable for such-and-such purpose or to use to fulfill such-and-such desire. In other words, the value of human beings is not found in what consequences or outcomes they can achieve (or be used to achieve). Instead, rational beings have intrinsic value. They are valuable in virtue of being the type of being that they are.

Exercise 11.5 Briefly consider the following topics and consider what the humanity version of the categorical imperative would say about each:

1. slavery
2. brainwashing
3. a parent forcing their child to eat their vegetables
4. the government taking away the guns of its citizens
5. a political group lying to voters to get them to vote for the "right" candidate
6. experimenting on human beings without their consent
7. abortion
8. eating animals

11.3.1 Problems with the the humanity version

We will consider three problems associated with the humanity version of the categorical imperative.

11.3.1.1 Argument from non-autonomy

The humanity version of the categorical imperative implies that we are rational and autonomous. This implies that our decisions are free. If it were discovered that we are not free, then the imperative would not apply. There are some powerful arguments to imply that human beings lack autonomy. And so, this ethical theory rests on a deeper, metaphysical issue concerning the possibility of free will.

Let's consider one such argument against autonomy:

P1: Our choices are either determined by prior state of the world and the laws of nature or they are not.

P2: If they are determined, then we are not free since the prior state of the world and the laws of nature determined by action.

P3: If they are not determined, then they are random since they are not determined by the prior state of the world and the laws of nature.

C: Therefore, we are not free and if we are not free, we are not autonomous.

Exercise 11.6 Critically evaluate the argument from non-autonomy. State the argument in your own words and then try to criticize one of the premises.

11.3.1.2 Argument from moral luck

The humanity version of the categorical imperative implies that there is no such thing as moral luck.

Definition 11.10 — moral luck. Moral luck refers to various cases where the moral goodness or badness of an action is determined by factors outside of our control.

It asserts this because the only thing that matters is our intentions and /or whether or not we treat persons as ends-in-themselves. But, we do, as a matter of fact, judge people based upon factors outside of their control.

Example 11.10 — distracted driving. Consider the following two cases:

1. Case A: Suppose a variety of cases where I am driving distracted (eating, texting, intoxicated, tired, etc.) but were no one is hurt
2. Case B: Suppose a variety of cases where I am driving distracted (eating, texting, intoxicated, tired, etc.) but were I run over an individual and kill them.

According to Kant, cases A and B should be treated the same. The only difference between the two cases is that in case B I was unlucky. But, as a matter of fact, we tend to think case B is morally worse than case A (and this is further reflected in the punishments for case B over case A).

In short, the argument against the humanity version is as follows:

P1: If the deontological theory of ethics is true (along with the human version of the categorical imperative), then there is no such thing as moral luck (aspects of the action do not matter morally)

P2: There is moral luck.

C: Therefore, the deontological theory of ethics is false.

Exercise 11.7 Consider the case of distracted driving outlined above. Now consider the following questions:

1. Is there such a thing as moral luck?
2. Suppose we accepted that moral luck does not exist. How ought we to view individuals

who engage in distracted driving? Should we view individuals who engage in it with no consequence more harshly? Should we view individuals who engage in it but with serious consequence less harshly?

11.3.1.3 Argument from the scope of moral community

Ethical theories have scope concerning the objects that have moral value or importance. The humanity version of the categorical imperative contends that the objects that have moral worth are persons and persons are those beings that have the power to reason and are autonomous.

The problem with this approach then is that it overly restricts the scope of the moral community. Here are a list of beings that have absolutely no intrinsic moral worth (some are debatable):

1. fetuses
2. babies
3. human beings with serious cognitive disorders
4. animals
5. natural objects like rivers, lakes, mountains, forests, etc.

We have no duty to the objects above. They have worth merely as a means to some end.

Thus, the following argument can be raised against the humanity version of the categorical imperative:

- P1: If the humanity version of the categorical imperative is true, then fetuses, babies, human beings with serious cognitive disorders, etc. have no intrinsic moral value (they not persons)
P2: If they have no intrinsic moral value, then they are things.
P3: If they are things, then they can be used as things and so they can be destroyed, tortured, their parts may be harvested, etc. That is, if they are things, then we have no moral duties to these objects.
P4: Those objects are not things and do have intrinsic moral worth.
C: Therefore, the humanity version of the categorical imperative is false.

Exercise 11.8 Consider the above argument against the categorical imperative. Is there any way you can object to it? Is there any way we might say we have duties to those objects even if they are not persons?

12. Contractarianism

In this unit, we will examine social contract theory (also known as contractarianism). This is the view that moral right and wrong is determined by a social contract.

12.1 The theory

12.1.1 Contractarianism

In discussing contractarianism, we will first give a definition of the theory and then proceed to clarify it.

Definition 12.1 — contractarianism. The ethical theory that contends that an act is morally good if and only if it does not violate the social contracts that free and rational contractors would make. In contrast, it is morally wrong if and only if an action violates these contracts.

Several points of clarification are required.

First, the definition of contractarianism requires the notion of a social contract. Let's call any case where we limit our pursuit of certain self-interest on the condition that others do so as well, a social contract.

Definition 12.2 — social contract. An explicit or implicit contractual agreement with others to limit the pursuit of certain self-interests on the condition others do so as well.

Second, it is important to note that the social contract need not be explicit. That is, it is not necessary that individuals explicitly agree to give up their self-interest provided others do the same.

Example 12.1 — Killing. It might benefit me to kill you and take your items. This leaves you terrified. But, it is terrifying to think that you might kill me to take my items. This leaves me terrified.

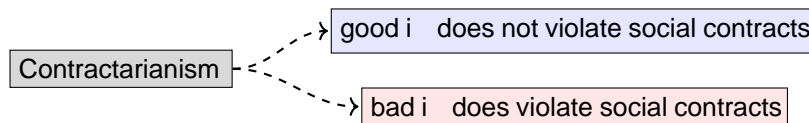


Figure 12.1: Contractarianism states that an act is morally good if and only if it does not violate the free and rational contracts (agreements) made between individuals and morally wrong if and only if it does violate these contracts.

As a result, each of us lie awake at night waiting for the other attack. This is less than ideal. We thus make an implicit agreement not to kill each other.

Example 12.2 — Walking on the right side. In the United States, we drive and walk on the right and pass on the left. It might be better to simply walk on the left or even drive on the left. But this is dangerous. And others might do the same. We want to reduce danger and inconvenience. We thus make an implicit or explicit agreement to walk and drive on the right.

Third, contractarianism makes morality a social construction. In other words, there is only moral right and wrong (nothing morally binding) once we enter into a social contract. Morality thus becomes real once we enter into a contract. This is not to say that morality is subjective or unreal. Instead, it is to say that morality is conditional.

Example 12.3 Suppose Tek makes an explicit agreement with John to return a book he borrowed. Just because the promise depends upon Tek making the explicit agreement does not mean that the promise is subjective. The agreement is real (objective) even though it depends upon a certain act by Tek.

Fourth, free, egoistic, and rational individuals would enter into social contracts. That is, individuals would actually agree to limit the pursuit of their self-interest on the conditions that (i) others would do the same and (ii) the contracts agreed to would be enforceable. The reason that individuals would not exclusively pursue their self-interest and agree to cooperate is because: there are situations where if everyone rationally pursues their own self-interest then the result is not as beneficial for you or for them compared to entering into a social contract.

Objection 12.1.1 This cannot be the case. Why would it ever be rational for an egoistic person to sacrifice the pursuit of their own interest for mutual benefit?

In order to support the point that it is rational to enter into the social contract, let's consider two different scenarios.

12.1.2 The student's dilemma and the prisoner's dilemma

Before considering the scenarios, let's make two assumptions:

Egoistic assumption: We will assume that human beings are, in general, self-interested. That is, when human beings act, they almost always aim to increase their well-being. In making this assumption, we are not asserting that people are, in fact, egoistic. Instead, the assumption is entirely tactical. That is, the goal will be to show how an ethical system can emerge out of this rather pessimistic view about human beings.

Rationalistic assumption: Human beings behave (or ought to behave) rationally and so give an option between A and B, and knowledge of the relevant facts, they will choose the better of the two options.

With these assumptions, let's consider two different cases.

12.1.2.1 The student's grade dilemma

Suppose I pair you up with another student in this class. You both will be anonymous to each other. I will never reveal who you are to the other student. In fact, I won't even know which student is paired with which student.

I've grown tired of teaching and have decided to determine grades in the following way:

1. You will pick your grade and the other student's grade
2. You can choose between two grades: A and F.
3. The first grade you will choose is yours, the second grade you will choose is the other (paired) student's grade. Thus, AF is you choosing A for your grade and F for their grade.
4. The other student do the same.
5. Your choice along with the other student's choice will determine your actual grades (see [Table 12.1](#)).

| | | |
|----|----|----|
| | AA | AF |
| AA | BB | FA |
| AF | AF | CC |

Table 12.1: If you choose AA and they choose AA, you will receive a B and the other student will receive a B (viz., BB). If you choose AA and they choose AF, the result is FA. If you choose AF and they choose AA, the result is AF. If you choose FA and they choose FA, the result is CC.

Exercise 12.1 Which grade would you choose for yourself and which would you choose for your fellow student?

The most rational and egoistic choice for you to make is AF and the most rational and egoistic choice for the other student to make is AF. The reason is that (i) you do not know what the other student will choose and (ii) if you choose AF

1. the best outcome for you in choosing AF is better than the best outcome of AA
2. the worst outcome of AF for you is still better than the worst outcome of AA

However, note that the rational and most self-interested choice (a grade of C) is worse than the choice you would sacrifice the pursuit of your self-interest and cooperated. That is, while the best outcome possible is the pursuit of your self-interest (receiving an A), it is not possible given the behavior of other individuals.

Thus, the best outcome would be if you and the other student collectively choose AA, but this would require you giving up on some of your self-interest and cooperating with your fellow student. ¹

12.1.2.2 The prisoner's dilemma

Let's consider a variation on the above dilemma. Suppose that Al and Bob are partners in crime and both are arrested, put in separate cells, and interrogated independently. The interrogator makes the following offer to both Al and Bob:

1. if you remain silent and your partner snitches, you will receive 6 years in prison and your partner will receive 0 years in prison.
2. if you remain silent and your partner remains silent, you and your partner will receive 2 years in prison
3. if you snitch and your partner snitches, you will both receive 4 years in prison
4. if you snitch and your partner remains silent, you will receive 0 years in prison and your partner will receive 6 years in prison.

¹What you would need to ensure this result, however is (i) the capacity to cooperate with each other and (ii) once the decision is made to both pick AA, the capacity to enforce this agreement (so the other student doesn't backstab you).

| | Bob silent | Bob snitches |
|-------------|------------|--------------|
| Al silent | 2,2 | 6,0 |
| Al snitches | 0,6 | 4,4 |

Again, the most and self-interested rational choice for both Al and Bob is to snitch. The reason is that (i) neither knows what the other will do and (ii) if they snitch

1. the best outcome in each snitching is better than the best outcome of not snitching
2. the worst outcome of snitching is still better than the worst outcome of not snitching

However, note that the rational and most self-interested choice (4 year prison sentence) is worse than the choice you would sacrifice some of your pursuit of your self-interest and cooperate (2 year sentence). That is, while the best outcome possible is the pursuit of your self-interest (receiving an A), it is not possible given the behavior of other individuals. And so, the best outcome would be if you and the other prisoner collectively remained silent, but this would require you giving up on some of your self-interest and cooperating with your fellow prisoner.

12.1.3 Contractarianism again

We have now argued that contractarianism asserts that given that we live with others, it is in our interest to enter into a social contract provided other individuals do the same. And, it asserts that an act is morally good if and only if the act does not violate the social contracts that free and rational individuals would make.

Another point then is that those social contracts that free and rational individuals would enter into are only those that they would benefit from by sacrificing their self-interest. That is, legitimate social contracts are those that my giving up my self-interest and cooperating with others would make me better off than if we both exclusively pursued our self-interests. Thus, a social contract is only legitimate if entering into the contract would make you better off than if you were not in the contract.

Example 12.4 — slavery. Let's suppose that you are forced into slavery. And life is miserable. You would be better off pursuing your own self-interest and others pursuing theirs than living a life of slavery. Slavery is thus not legitimate.

Sixth, it is important to note that in order to escape the prisoner's dilemma, we need two things: (i) the willingness of others to sacrifice their self-interest and (ii) some sort of external constraint that would make what is in our collective interest the same as what is in our immediate self interest. The former is necessary for cooperation, the latter is necessary for ensuring fulfillment of the contract once the contract is initialized.

12.2 Arguments for social contract theory

In this section, we consider several arguments in favor of social contract theory.

12.2.1 Morality as a social phenomenon

The first argument contends that only social contract theory can explain our moral intuition that while duties to others exist, there are no duties to one's self.

P1: Moral duties to others exist, there are no duties to one's self.

P2: Social contract theory can explain why this is the case since morality is the product of a social phenomenon (a social contract).

C: Therefore, social contract theory is preferable to other moral theories.

With respect to P1, note that while people contend it is morally wrong to harm others in various

ways, it is at best controversial whether it is morally wrong to harm one's self. That is, consider the following acts that we might say are morally acceptable (although not practical) provided they do not harm others:

1. using drugs
2. spending all of your money
3. engaging in sloth
4. harming one's self
5. engaging in extremely dangerous behavior

Divine command theory, Natural law theory, utilitarianism, etc. would (with some exceptions) regard such behaviors as morally wrong, but social contract theory can explain why they are permissible. Namely, they are permissible because morality emerges when we engage in an agreement with others to limit our self interest provided they do the same. No such agreement is ever made with one's self.

12.2.2 Proof of immoral acts even to the egoist

Social contract theory allows for the proof of every moral rule. In particular, it can show (even the most skeptical of individuals) why a particular act is morally unacceptable.

P1: A moral theory is viable if and only if it can persuade egoistic individuals that a particular action is morally wrong.

P2: Social contract theory is the only theory that can explain even to egoistic individuals why particular acts are morally wrong.

C: Therefore, social contract theory is the only viable theory.

With respect to P1, we might say that egoistic individuals are some of the most reluctant to accept certain moral theories. To every moral obligation that seemingly requires them to put aside their self-interest, they can ask why should I? That is, they can constantly press ethical theorists as to why they should be moral in the sense required?

However, with respect to P2, recall that social contract theory argues for its position from the most minimal of assumptions; namely, it assumes that even if we are egoists, then provided we are rational we have a self-interest to enter into the social contract. And the reason we have a reason to enter into the social contract is because doing so benefits us.

12.2.3 Explains moral exceptions

Contractarianism contends that morality depends upon the willing cooperation of others for it is only with their cooperation that your decision to cooperate benefits you. If the other prisoner won't agree to remain silent and aims to snitch, then you ought to snitch. If everyone agrees to take an equal share of the pot but then immediately takes more, then you are not obligated to take the agreed upon equal share.

Given that morality is conditional upon the cooperation of others, contractarianism can best explain why moral exceptions are permissible.

P1: Any ethical theory ought to be able to explain why it is permissible to break certain moral rules.

P2: Contractarianism can best explain why it is permissible to break such rules.

C: Therefore, contractarianism is preferable to other moral theories.

12.3 Arguments against contractarianism

In this section, arguments against contractarianism are considered.

12.3.1 Argument from free riders

First, let's put forward the argument.

P1: If contractarianism are true, then I ought to only give up the pursuit of my self-interest when it bene ts me.

P2: There are cases where I can be a free rider and reap the bene ts of others giving up the pursuit of their self-interest without giving up my own self-interest.

P3: It is not permissible to be a free rider.

C: Therefore, contractarianism is false.

Next, let's talk through each premise.

With respect to P1, if we assume contractarianism is the case, then it is only necessary to sacri ce the pursuit of one's self-interest given that we are compensated by others doing the same.

However, with respect to P2, there are cases where we can be compensated without sacri cing anything (where we can be a free rider).

De nition 12.3 — free rider. A free rider is an individual who gains bene ts without making any sacri ces (giving up anything).

Example 12.5 — not paying taxes. We all agree to relinquish some of our wealth in the form of taxation for certain public goods (e.g. roads, re, police, etc.). If I could get away with not paying taxes, but reap the bene ts of the public goods associated with other people paying taxes. I'm a free rider.

Example 12.6 — not littering. While it may bene t me to throw my trash away whenever and wherever I want, we all sacri ce this freedom and agree not to litter so we can live in a clean community. But, I may decide not to sacri ce this freedom yet I can still reap the public bene t of others doing so. I'm a free rider.

P3 simply asserts that being a free rider is against our moral intuitions.

Exercise 12.2 Can you think of an example of a free rider on a social contract? That is, where they reap the bene t of others giving up the pursuit of their self-interest for a collective bene t.

One objection to the above argument is to deny P2.

Objection 12.3.1 — P2 is false. Dangers. One objection is to say P2 is false since it is never in your bene t to be a free rider. That is, being a free rider is extremely dangerous and one runs the high probability of being harmed as a result of being a free rider.

12.3.2 Argument from non-consent

First, let's put forward the argument.

P1: If contractarianism are true, then I only have an obligation to follow those social contracts that I have either explicitly or tactitly consented to.

P2: There are numerous social contracts that I have not given my explicit or tacit consent.

IC: Therefore, if contractarianism is true, then I need not follow any social contract that I have not given my explicit or tacit consent.

P3: I ought to follow certain social contracts regardless of whether I have consented.

C: Therefore, contractarianism is false.

With respect to P1, if contractarianism is true, then it looks like I am bound to those social contracts that I have either willingly agreed to or have tacitly accepted. Our tacit acceptance of a contract can be seen in our willingness to take advantage of the mutual goods that are the result of others giving up the pursuit of their exclusive self-interest.

Example 12.7 If I drive on the roads, or call the police when there is a crime, or call the fire department when there is a fire, I have tacitly accepted the tax system that allows for such services.

But note P2. There seems to be a variety of social contracts that I have not explicitly accepted and some I might explicitly reject and so not reap any benefits from.

Example 12.8 — jaywalking and marijuana prohibition. Suppose Tek jaywalks and uses marijuana. He rejects any societal agreement to give up his rights to do either.

However, note that it seems that we ought to obey some rules (social contracts) regardless of whether or not we have consented.

Example 12.9 There are perhaps a number of obligations we have to certain social contracts regardless of whether or not we have consented to them. Some might include:

1. obligation to help those less fortunate
2. obligation to be polite (civil) to our neighbors
3. obligation not to treat people poorly

Objection 12.3.2 — P1 is false. The contractarian can reply that P1 is false. They can argue that contractarianism does not require that we either explicitly or tacitly consent to a social contract in order to be morally bound by that contract. Rather, it asserts that we are bound by those social contracts that a free and rational person would accept in order to gain the benefits of mutual cooperation.

12.3.3 Argument from the scope of the moral community

First, recall that contractarianism contends that the only reason that we would ever enter into a social contract (sacrifice the pursuit of our own self-interest and cooperate) is because if the outcome of not entering the social contract is worse than everyone individually pursuing their own self-interest.

Second, what this implies then is that the individuals that we enter into social contracts with are our equals: those individuals whom if they pursued their interests would interfere with the pursuit of our interests.

Example 12.10 I may want to kill you and take your belongings. But I know that you are capable of killing me and taking my belongings. To avoid the terror of constantly thinking we might be killed, we enter into a social contract.

Third, it is important to point out that the contract exists not simply between individual persons, but also exists between single persons and collections of individuals. That is, a rich individual might believe s/he is better than a poor individual and thus seek to exploit that person. However, the poor individual may form a community with other poor individuals and thus a social contract may be formed. The rich person agrees not to use his/her wealth to exploit the poor and the poor people agree not to seize the wealth of the rich person.

Nevertheless, a serious problem remains.

P1: If contractarianism are true, then our moral obligations are limited to our equals: individuals that we cannot exploit and whom it is in our interest to cooperate.

P2: This would imply that the moral sphere is extremely narrow, e.g. it would apply to animals,

young people, the elderly, sick people, individuals with serious disabilities, the environment, etc.

P3: The scope of the moral community is not that limited. We have moral obligations to all sorts of beings beyond our equals.

C: Therefore, contractarianism is false.

Exercise 12.3 What do you think of the above argument? Is there any way of rejecting P2? Can you think of an example of P3?

13. Pluralistic Absolutism

13.1 Ethical monism

Thus far, we have considered a variety of theories that are referred to as ethical monisms.

Definition 13.1 — ethical monism. Ethical monism is the metaethical theory that says that the basis of morality relies upon a single rule that is both fundamental and absolute rule.

A rule R is fundamental if and only if there is no R' that is more basic than R , viz., there is no R' which we could use to derive R ¹

A rule R is absolute if and only if there are no exceptions to R , viz., one cannot break the rule

Every ethical theory we have considered is an ethical monism:

1. Utilitarianism is an ethical monism as it asserts there is a single, absolute, and fundamental rule: you ought to maximize well-being.
2. Divine Command theory is an ethical monism as it asserts that there is a single rule that is absolute and fundamental you ought to obey God's commands.
3. etc.

13.2 Ways to object to ethical monisms

Since (i) all of the theories that we have thus far considered are ethical monisms and (ii) all of these theories have been shown to have some serious problems, we have three different options:

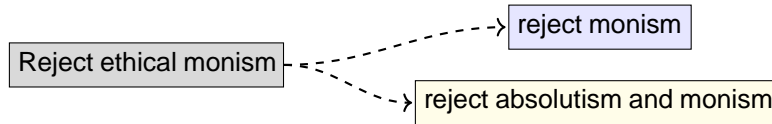
1. develop a new ethical theory that is a new type of ethical monism
2. show how one of the existing ethical theories that we have already considered is correct by providing new arguments or responding to all of the available objections

¹In contrast to a fundamental rule is a derived rule. You ought to hold the door open for people might be considered derived from a more fundamental rule like You ought to be nice to people

3. abandon ethical monism for a new type of theory.

We will consider two different ways that we might reject ethical monism:

1. accept absolutism but reject monism (pluralistic absolutism) ²
2. reject absolutism and reject monism (Ross's view or ethical particularism) ³



13.3 Pluralistic absolutism

Suppose we hold onto the idea that there are absolute moral rules but give up on claim of ethical monism that says there is a single, fundamental moral rule. Let's call this view pluralistic absolutism.

Definition 13.2 — pluralistic absolutism. The metaethical theory that asserts that there are at least two absolute moral rules.

13.3.1 Argument from disaster prevention

The first argument against pluralistic absolutism is that there are simply no absolute moral rules.

P1: If there are absolute moral rules, then there are no exceptions to following these rules.

P2: For any absolute moral rule, there are scenarios (even if they are extremely implausible) where it is necessary to make an exception to the absolute moral rule to prevent a disaster.

C: Therefore, there are no absolute moral rules.

P1 is definitional. That is, a rule is absolute if and only if there is no exception to that rule.

The controversial premise is P2. The supporter of P2 will argue that give them any absolute moral rule that you please, and it will be possible to come up with a scenario where it would be necessary to make an exception to that moral rule.

Example 13.1 — torturing. Suppose it is proposed that you ought to never torture is an absolute moral rule. The supporter of P2 will argue that there are exceptions to this rule. For example, suppose there is an individual who has planted several nuclear devices in several cities. Those devices will detonate unless we obtain information from the bomb maker who is currently in custody. Intuitively, if torturing the bomb maker would force the bomb maker to offer up information about the locations of the bomb, individuals might be willing to make an exception so as to prevent a disaster.

Exercise 13.1 Consider the following candidates for absolute moral rules. Can you think of an exception?

1. stealing from a poor person
2. stealing from a hungry person
3. abortion (from the pro-life perspective)
4. killing an innocent person

²Reject the idea that there is a single ethical rule that is fundamental and absolute. For example, argue that there are at least two ethical rules that are fundamental and absolute.

³Deny the claim that there are absolute moral rules. For example, you could assert that there are exceptions to every moral rule.

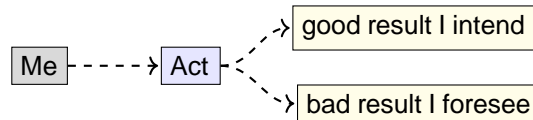
5. killing a child

13.3.2 Response: the doctrine of double effect

Defenders of pluralistic absolutism will try to respond to the argument from disaster prevention. One way they attempt to respond to this argument is by introducing a position called the “doctrine of double effect”:

Definition 13.3 — the doctrine of double effect (DDE). If the intent (goal) of your act is morally good, then it is sometimes permissible to perform that act even if you can foresee that the act has a bad side-effect.

The general idea of DDE is that my act can be classified as morally good even though I know it will harm others (have a bad effect). According to DDE, the act is good provided that the bad effect is merely a side-effect of the effect I am intending.



Let's look at two examples of the doctrine of double effect.

Example 13.2 — abortion. Suppose Liz is pregnant. The fetus is threatening her life. If she continues with her pregnancy, she and the fetus will die. However, Liz is pro-life (from a moral perspective) and so she does not want to kill the fetus. Liz reasons as follows: “I do not want to kill the fetus but the fetus and I will die if I bring the fetus to term. What I will do instead is have an operation that separates the fetus from me. The goal is to save my life (a morally good act) but one side-effect of this morally good act is that the fetus will likely die. According to the doctrine of double effect, my intent is to save my life; it is not to kill the fetus, although I recognize that the effect of saving my life is the death of the fetus.

With this example in place, there are several key points of clarification required.

First, in order for DDE to apply, the goal (the intention) of the actor must be morally good. Liz's goal is good in that she aims to save her own life. If Liz's goal was to abort the fetus because she wanted to upset her mother, then her goal would not be morally good and so DDE would not apply.

Second, the DDE specifies cases where acts that have bad side-effects are morally permitted. Since Liz is pro-life, she sees the abortion as morally wrong (a bad effect).

Third, the morally good act you intend must be achieved independently of the bad side-effect. That is, the bad side-effect cannot be the means used to achieve the morally good result. If Liz had the doctor's kill the fetus, then DDE would not apply. If Liz had the doctors remove the fetus from her body and the fetus died as a result, the DDE would apply.

Example 13.3 — death to cure pain. Suppose Tek is sad because his mom died. I tell Tek I can cure his pain. Now consider the following two scenarios:

1. DDE does not apply: The goal of my act is to cure his pain, but I do it by killing him. Killing him was the means to cure his pain.
2. DDE does apply: The goal of my act is to cure his pain. I give him some homemade painkillers I made. I give him too many and he dies. Killing him is simply a side-effect of trying to help him.

Fourth, there cannot be other measures that would achieve the same good result with a less severe

side effect. In the scenario involving Liz, the removal of the fetus from Liz's body is proposed as the only way to save Liz's life. If, however, she could take a simple drug that would be safer and preserve the fetus, the DDE would not apply.

Example 13.4 — drugs. Suppose my goal is to save lives and I am faced with the following two options:

1. save ten people by killing one person
2. save ten people by killing two people

The doctrine of double effect would not apply to the second action since there is another action that accomplishes the same morally good goal with less evil as a side-effect.⁴

Exercise 13.2 Come up with your own example where an individual uses the doctrine of double effect to argue that an action is morally permissible. The structure of the example is this: action X that has (a) a good result and (b) a negative side effect, but where person Y intends (a) even though knows (b) will result.

With the doctrine of double effect now clarified, we can now explain how this concept can be used against the argument disaster prevention. That is, the doctrine of double effect can be used to argue for the claim that there are absolute moral rules. Let's recall the argument from disaster prevention:

- P1: If there are absolute moral rules, then there are no exceptions to following these rules.
 P2: For any absolute moral rule, there are scenarios (even if they are extremely implausible) where it is necessary to make an exception to the absolute moral rule to prevent a disaster.
 C: Therefore, there are no absolute moral rules.

Advocates of the doctrine of double effect can assert that there are certain absolute moral laws and these are that there are certain acts (e.g. killing, torture, rape, theft, etc.) that you can never intend as your goal. That is, while it may permit various acts of killing, torture, etc., this does not change the fact that there are absolute moral rules governing what is morally acceptable to intend (to take as your goal).

Thus, provided that DDE is legitimate, then there are absolute moral rules and so pluralistic absolutism is not rejected.

13.3.3 The argument from moral conflict

Putting aside the argument from disaster prevention and whether DDE is a legitimate principle, there is another problem with any form of pluralistic absolutism:

- P1: If there are at least two absolute moral rules, then they will conflict in some way.
 P2: If they can conflict, then the moral system would (i) morally require you to do X and (ii) morally forbid you to do X.
 P3: No moral theory should require you to do X and forbid you to do X.
 C: Therefore, pluralistic absolutism is false.

Example 13.5 — promise keeping and no harm to others. Suppose Tek helps Vic. In payment, Vic promises to do whatever Tek wants. Tek tells him that he wants Vic to go on a rampage, attacking random people. He is thus required and forbidden to do that act.

Example 13.6 — telling the truth and no harm to others. Jane asks Liz if she looks good. Liz likes Jane and doesn't want to hurt her feelings but also doesn't want to lie. Liz knows that no matter

⁴Another example. Suppose Tek has a serious headache. He needs some medication to manage the pain. My intention is to give him a drug to relieve him of pain. I have a drug with some minor side effects and a drug with some major side effects. Both are equally effective. If I were to give him the drug with the major side effects, even if my intent was to reduce his pain, this would not be permissible because there is an equivalent act with less serious side-effects.

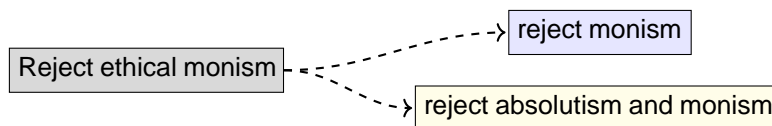
how kindly or delicately she tells Jane that she does not look good, it will hurt her feelings. Liz is both required not to hurt her feelings but also to tell the truth.

Objection 1 (P1 is false) The absolutist can respond by contending that P1 is false. They can argue that only certain forms of pluralistic absolutism are viable and these are versions that forbid certain actions (e.g. never kill, never lie, etc.) but never require action (e.g. keep promises, tell the truth).

14. Prima facie duties and ethical particularism

In the previous lecture, we noted two different ways that we might reject ethical monism:

1. accept absolutism but reject monism (pluralistic absolutism) ¹
2. reject absolutism and reject monism (Ross's view or ethical particularism) ²



In that prior lecture, we considered the possibility of accepting absolutism but rejecting monism. This was the theory of pluralistic absolutism. In this lecture, we consider the possibility of rejecting both monism and absolutism.

14.1 Ross theory of prima facie duties

Let's begin by defining a theory that rejects both monism and absolutism.

Definition 14.1 — Ross's theory of prima facie duties. The theory that there are at least two prima facie duties.

¹Reject the idea that there is a single ethical rule that is fundamental and absolute. For example, argue that there are at least two ethical rules that are fundamental and absolute.

²Deny the claim that there are absolute moral rules. For example, you could assert that there are exceptions to every moral rule.

Definition 14.2 — prima facie duty. An individual has a prima facie duty to do action X when there is at least one permanent, fundamental, strong, and non-absolute reason to do X.

A couple points of clarification:

In calling prima facie duties fundamental what is meant is that no single duty is not based on (or derivable from) other duties.

In calling them “non-absolute”, what is meant is that there are cases where some of these duties may have exceptions.

In calling the reasons permanent, what is meant is that there is always a reason for doing X even if that reason does not require you to do X because it is overruled by more powerful reasons. In other words, if we have a prima facie duty to do action X, then X is always morally important even if we have other, more important prima facie duties.

It is sort of misleading to call the duties themselves “prima facie duties”. What is really the case is that there are some moral obligations that there are fundamental, non-absolute, and good reasons for being the case.

Example 14.1 Consider the following duties that might be considered prima facie duties (from Ross):

1. Fidelity: promise keeping
2. Reparations: repairing harm that one has caused
3. Gratitude: acknowledging the gifts others have given us
4. Justice: the punishment of vice and rewarding of virtue
5. Benevolence: increasing intelligence, virtue, or pleasure in others
6. Self-improvement: increasing one's own intelligence or character
7. Non-malevolence: decreasing harm towards others

In each case, there are exceptions. In the case of fidelity, while there is always a good reason to keep one's promises, there are exceptions, e.g. if telling the truth would cause serious harm.

14.1.1 Arguments for Ross's theory

14.1.1.1 The argument from exceptions

P1: There is the moral intuition that there are cases when it is morally permissible to break moral rules (that is, there are always exceptions to moral rules).

P2: Ross's theory of prima facie moral duties fits with this intuition while various forms of absolutistic monism and pluralistic monism do not.

C: Therefore, there is reason to accept Ross's theory of prima facie moral duties

14.1.1.2 The argument from moral regret

To understand the argument from moral regret, it is necessary to define moral regret.

Definition 14.3 — moral regret. Moral regret occurs when an individual feels that they have done something morally wrong.

P1: We often feel moral regret even when we do the right thing.

P2: Various forms of absolutism cannot explain this phenomenon.

P3: Ross's theory of prima facie moral duties can explain this phenomenon.

C: Therefore, there is a strong reason to prefer Ross's theory of prima facie moral duties over other theories.

The critical premises are P2 and P3.

In the case of P2, various forms of absolutism cannot explain why it is rational to feel moral regret when we do the right thing. If one is faced with a decision between X and Y where X is the morally right thing to do, then Y has no moral value in comparison to X. For a consequentialist, if one acts in a way to maximize well-being, then one has done the right thing and therefore has no reason to feel moral regret. One should not feel moral regret for not having done some act that increases (but does not maximize) well being.

In the case of P3, when we face moral conflicts (e.g. promise keeping and doing no harm), we are faced with prima facie reasons to do two different acts X and Y. Now suppose X is the morally right thing to do. In facing these reasons for doing X over Y, we recognize not that that Y is of no value in comparison to X, but simply that X has more value in comparison to Y (the reasons are stronger). For this reason it is naturally to feel moral regret.

Example 14.2 — truth-telling and non-maleficence. You ask me if you look nice. I can tell the truth and hurt your feelings or lie and not hurt your feelings. There are prima facie reasons to tell the truth and prima facie reasons not hurt your feelings. However, let's say that there are stronger reasons for me to tell the truth so I act accordingly. It is rational, however, for me to feel moral regret even though I did the right thing. The reason is because I recognize that there is a strong reason for me to also not hurt your feelings.

14.1.2 Arguments against Ross's theory

There are two primary problems with Ross's theory. The first is that there is no argument for which duties we have prima facie reasons to accept. The other, and the one we consider here, is that there is no way to determine when it is permissible to not act on our prima facie duties. For example, suppose I have a prima facie reason not to harm others (non-maleficence) but I also have a prima facie reason not to lie. There is no explanation or argument for which one of these two reasons ought to have more moral weight in cases where they conflict.

P1: If Ross's prima facie theory is true, then it is possible for there to be exceptions to our prima facie moral duties (especially in cases where the duties conflict).

P2: There is no explanation for what one ought to do in cases where they do conflict.

P3: A moral theory ought to provide a determination of how one ought to live, especially involving cases where our moral intuitions conflict.

C: Therefore, Ross's prima facie theory is false.

Ross can object to P3.

Objection 14.1.1 — P3 is false. Ross might argue that P3 is false. While the theory can provide guidance on how one ought to live in a number of situations, no theory can be expected to provide a decision procedure (algorithmic step-by-step solution to every moral problem). In fact, decision procedures are ought of place not only in ethics but also in the sciences. For suppose a scientific theory T. The theory T makes predictions about the world. Now suppose that one of the predictions it makes turns out to be false. In this case, scientists are left with two options: (i) reject the theory or (ii) reject the observation (data). Science has proceeded successfully doing both (i) and (ii).

14.2 Ethical particularism

A more extreme rejection of monism and absolutism also involves a rejection of prima facie duties.

Definition 14.4 — ethical particularism. Ethical particularism is the theory that ethical monism is false, absolutism is false, and we have no prima facie moral duties. An act X is morally good or bad is always determined by the particulars of the specific situation (context, the details of the

case) involving X and not by any duty.

A couple points of clarification

First, contrast ethical particularism against absolutism and Ross's prima facie theory of duties:

| | absolutism | prima facie theory | particularism |
|--|------------|--------------------|---------------|
| some features are always morally important | x | x | |
| some features are always morally decisive | x | | |

Table 14.1: Contrast between absolutism, Ross's prima facie theory, and particularism

Second, ethical particularists will contend that there are no prima facie duties because there are situations where we do not even have a prima facie reason to do action X.

Example 14.3 — promising and kidnappers. Suppose Tek is kidnapped. He tells the kidnappers that he promises not to reveal their identity if they release him. According to Ross, he has a prima facie reason to keep his promise even though there may other prima facie reasons that override this promise. According to particularists, there is not even a prima facie reason to keep your promise. You have no reason at all to keep a promise with kidnappers.

Third, ethical particularists will often argue for their position by showing that there are always some scenarios where certain moral rules (e.g. being honest, promise keeping, non-maleficence) have no importance at all. That is, being honest, promise keeping, non-maleficence has no moral importance.

Fourth, ethical particularists will contend that our moral knowledge (knowledge of what is morally right or wrong) will always depend upon a knowledge of the context (the situation in which one plans on acting). This means that one has a better chance of doing morally good if one increases one's knowledge of the context in which an act is being committed.

14.2.1 Arguments against ethical particularism

Only one argument against ethical particularism is considered here.

P1: If ethical particularism is true, then there is nothing that possesses permanent moral importance (no rules that are decision, nor prima facie duties that are always morally important).

P2: If nothing possesses permanent moral importance, then ethical particularism provides us no guidance concerning what we ought to do in any particular ethical situation.

P3: If an ethical theory is useful at all, then it ought to offer some ethical guidance.

C: Therefore, ethical particularism is useless.

15. Virtue Ethics

We might distinguish between two different questions:

what ought I to do? (which acts are morally obligated, forbidden, permissible?)
What type of person should I be? (what character traits should I strive for?)

We can imagine an individual who does the right things but whose character is morally flawed.

Example 15.1 Imagine Tek who does the right thing but only because he fears being arrested if he does wrong. If the laws were different, Tek would steal, kill, and defraud. Here, Tek may be said to act morally but he has a low moral character.

Virtue ethics shifts the focus from a discussion of whether such and such act is morally obligated to what type of character traits we ought to seek to establish in ourselves and others.

One of the reasons for this shift in focus is that it might provide a better, more reliable way for determining good and bad moral action. It may be the case that it is impossible to construct a moral theory that can take into account every situation. The moral particulars of a situation vary too much to know, in advance, whether an act is morally obligated or forbidden. At best we can develop certain ideal habits of actions or character traits that will give us a foundation for thinking through the specifics of that situation.

15.1 Defining moral good

Let's define virtue ethics as follows.

Definition 15.1 — virtue ethics (VE). Virtue ethics refers to a family of ethical theories that contends that an action X is morally good if and only if X is an action a virtuous person would do (assuming they are acting freely and in line with their general behavior).

A couple points of clarification.

First, VE does not provide absolute moral rules that anyone can apply to any moral situation. Instead, the theory contends that an action is morally good if and only if a virtuous person (a moral exemplar) would do that act.

Second, the specification of what a virtuous person would do involves that virtuous person acting in accordance with their character. That is, the theory relies upon how a virtuous person would normally behave or behave when they are acting in line with their general character. It does not refer to how a virtuous person would behave in unusual scenarios where they are “not themselves”.

Example 15.2 — Tek on drugs. Suppose Tek is a virtuous person. But now suppose that Liz spikes his drink with an extremely hallucinogenic substance. After consuming the substance, Tek begins to behave wildly and aggressively toward random individuals. Tek is not behaving in line with his general character and so his behavior does not determine what is morally appropriate.

Third, VE is a form of ethical pluralism as there are a number of virtuous individuals and these individuals may behave differently in similar circumstances. Thus, there is not a single right and wrong way that an individual should act in any given situation.

Example 15.3 — Tek and Liz and helping Jon. Suppose Tek and Liz are both virtuous people. Now suppose Jon who is homeless and is in need of help. Tek decides to help Jon by giving him food, new clothes, and a place to stay for the night. Tek provides Jon some immediate comfort for the harshness of homelessness. In contrast, Liz tells Jon to meet her tomorrow at a restaurant. She knows the owner of the restaurant and arranges for Jon to wash dishes at the restaurant for pay. While the utilitarian would say that the act that is morally good is the one that maximizes well-being, the virtue ethicist will say that both Tek and Liz (since they are virtuous people) have done morally good and we can use them as a guide on how to navigate similar scenarios.

Fourth, VE also tells us what not to do. We ought not to do what the virtuous person would not do. Perhaps another way of putting this is that we ought not to engage in vice, or behave in a way that follows the moral lead of the vicious individual.

Fifth, VE often assert that moral knowledge is a type of practical knowledge. Practical knowledge can be understood as a kind of knowing how rather than knowing that.

Example 15.4 — knowing that and knowing how. One individual may have a lot of knowledge about how car engines work but no practical knowledge concerning how to fix or create a car engine. You might know how the physics of skateboarding or bike-riding but not know how to skateboard or ride a bike.

We often learn how to do something (practical knowledge) through experience and practice. An individual learns how to swing a tennis racket not simply by studying the mechanics of tennis swings but repeatedly swinging the racket and refining this skill. VE sometimes assert that moral knowledge is a type of knowing how. One becomes virtuous or learns how to behave in various situations not simply by studying various ethical rules but through experience, by learning from experts (moral exemplars), through experimentation, and by repeated exposure to (and reflection on) morally-significant situations.

Sixth, VE often takes virtuous behavior to be something that you learn not only through training but also by having the right sort of models. This makes your moral situation one depend upon conditions that are outside of your control.

If you grow up around individuals who are not virtuous and never have exposure to moral exemplars, then you are likely to not be virtuous yourself. You don't have a virtuous person to learn from.

If you grow up around virtuous individuals and take the time to learn from them, then you

are likely to be virtuous yourself.

Seventh, virtuous individuals are said to model excellent behaviors. The character traits that they manifest in habitual activities and that determine what is morally good are virtues.

Definition 15.2 — virtue. A virtue is a character trait that is manifested in habitual activity and that is indicative of excellent action.

Some virtues include:

Cardinal virtues: justice, fortitude, courage, bravery, temperance, prudence

Theological virtues: faith, hope, Love (charity)

Capital virtues: humility, liberality, brotherly love, meekness, chastity, temperance, diligence

Other virtues: courage / bravery, temperance, liberality, amiability, sincerity, wit, modesty, magnanimity

Let's look at one virtue in detail: bravery.

Definition 15.3 Bravery is the property of standing firm against frightening things because we know that running away is shameful and standing firm would be honorable.

Bravery is thus an action that is response to two different pieces of knowledge

The brave person knows that to run away from the thing that causes fear would be shameful (a psychological feeling of disappointment). Running away would cause them to feel that they failed to do what would have made them excellent (not a fear of failing to do one's duty but failing to do what would make them an excellent person)

The brave person knows that standing firm in the face of fear would be doing something excellent (a psychological motivation to do the excellent thing)

Example 15.5 — Imposters. There are a number of people who appear brave but lack the virtue of bravery.

Fearless individuals: If you have no fear of the consequences, then you cannot be brave.

Overly hopeful people: If you think that everything will work out, then you may again lack the requisite fear to be brave.

Ignorant people: If you cannot appreciate the danger of a situation, then you may lack fear

Exercise 15.1 One of the key claims of VE is that morality is something that you learn and develop through training, experimentation, and exposure to role models. Is virtuous behavior something you are born with or learn?

15.2 Arguments for virtue ethics

15.2.1 The argument from the right amount of precision

P1: Each area of knowledge requires its own degree of precision

P2: An area of study ought not to formulate laws, principles, and proposition that are more precise than the area of knowledge allows.

P3: Requiring an ethical theory to provide a precise rule that allows for determining whether an act is morally right or wrong for every situation is to provide a rule that is more precise than the area of morality allows.

P4: Virtue ethics provides the right level of precision as it allows for a high degree of sensitivity and variability concerning what one should and should not do.

C: Therefore, we have a reason for accepting virtue ethics over other disciplines

P1 asserts that every area of study involves varying degrees of precision. This seems true since we don't think that knowledge about Shakespearean novels needs to be proven with the same degree of precision and rigor as mathematical truths. In other words, mathematics may require a high degree of precision, physics a high degree but not as precise as most physical predictions involve a standard of error, biology even less precise, and morality may be even less so.

P2 says that it would be a mistake to formulate the key propositions, laws, or truths of a discipline in a way that is more precise than the phenomena allows. Clouds may have fuzzy boundaries. I would be inaccurate if I were to give a highly precise determination of a cloud when the cloud does not have precise boundaries.

P3 contends that ethical theories that propose absolute rules that apply in every situation are engaging in this kind of over precision. They incorrectly presume that there is a moral law that can be applied in every situation regardless of the complexity or the particulars of the situation.

P4 asserts that VE allows for a great deal of flexibility when it comes to moral decision making. This is because VE offers general guidelines as to what one should or should not do and it gives this guidance in the form of how various moral exemplars would behave in situations. These moral exemplars, however, are attentive to the nuances of the specific moral situation and respond in different ways.

15.3 Arguments against virtue ethics

15.3.1 Virtue theory is circular

The first argument against VE is that it is circular. VE contends that what makes an action a virtue is that it is done by a virtuous person rather than saying that a person is virtuous because they do virtuous actions. This raises the question of how we are able to identify who is and is not virtuous.

P1: If VE is false, then an action X is morally good if and only if X is an action a virtuous person would do.

P2: If VE is true, then virtuous people determine what is virtuous.

P3: But we don't know who is and isn't virtuous unless we already know what virtue is.

C: Therefore, VE is circular and so is false.

In order to see the circularity, consider [Figure 15.1](#).

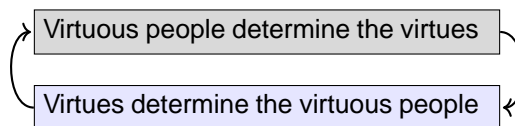


Figure 15.1: Illustration of the circularity of virtue theory.

Another way of thinking about the circularity of virtue theory is to consider disagreements about determining who is and is not virtuous.

Exercise 15.2 Create a list of five people you think are virtuous (moral exemplars). Next, consider whether any of the the following people below are virtuous:

1. Kanye West, 50 Cent, Meek Mill (rappers)
2. Donald Trump, Kim Jong Un, Vladimir Putin (political leaders)
3. Martin Luther King, Jim Brown, Frederick Douglas (civil rights activists)

4. LeBron James, Stephen Curry (football players)
5. Naomi Osaka or Serena Williams (tennis players)

15.3.2 Virtue ethics leads to contradiction

P1: If virtue theory is true, then an action X is morally good if and only if X is an action a virtuous person would do.

P2: There are cases where one virtuous person would do action X and one virtuous individual would do action not-X.

IC: Therefore, if VE is true, then X and not-X are morally good. Contradiction.

C: Therefore, virtue theory is false.

The key premise is P2. Let's consider an example.

Example 15.6 — Colin Kaepernick and his virtuous critics. Colin Kaepernick was a professional football player (playing quarterback for the San Francisco 49ers from 2011-2016). In a preseason game in 2016, Kaepernick sat during “The Star-Spangled Banner”. When asked why he didn't stand (which is the convention), he stated that he wasn't willing to stand and show pride for a flag that oppresses people of color. In subsequent games, Kaepernick knelt during the national anthem rather than standing. Let's suppose that Kaepernick is a virtuous person. A number of individuals have been critical of Kaepernick's decision to kneel during the national anthem (many might also be regarded as virtuous). Former football player and civil rights activist Jim Brown said that while he would like to be “be in his [Kaepernick's] corner [...] I'm an American. I don't desecrate my flag and my national anthem. I'm going to work within those situations. But this is my country, and I'll work out the problems, but I'll do it in an intelligent manner.”¹ Supreme Court Justice Ruth Bader Ginsburg criticized Kaepernick's decision to kneel as “dumb and disrespectful”, stating that “I think it's a terrible thing to do, but I wouldn't lock a person up for doing it. I would point out how ridiculous it seems to me to do such an act.”²

Exercise 15.3— How ought we to deal with conflicts. VE says that we ought to follow the lead of virtuous individuals, but if Colin Kaepernick, Jim Brown, and Justice Ginsburg are all virtuous then we have no clear direction on how to act when it comes to whether it is morally appropriate for football players to kneel during the national anthem. Using VE as your guide, how might you try to solve this problem?

¹<https://www.usatoday.com/story/sports/nfl/2017/08/25/jim-brown-colin-kaepernick-i-dont-desecrate-my-flag-and-my-national-anthem/602349001/>

²https://www.washingtonpost.com/news/early-lead/wp/2016/10/12/colin-kaepernick-responds-to-justice-ruth-bader-ginsburgs-criticism-of-anthem-protests/?utm_term=.58476b4800cc

16. Care Ethics

16.1 Care ethics

Care ethics is sometimes said to not be capable of being characterized systematically. It is instead said to be composed of a motley of insights and intuitions that are sensitive to contextual considerations. However, some care ethicists provide more systematic accounts or, at the very least, some central features of the theory. Here we adopt the latter approach.

First, care ethics takes the caring relation to be central to moral obligation. While there are a number of definitions of “care”, the one we will use defines it as follows.

Definition 16.1 — care. A functional state whereby one person A (the caring-for individual) meets the needs, develops the capacities, helps to maintain the capacities, and/or works to alleviate harm of person B (the cared-for individual) via a face-to-face interaction so that they can flourish.

This caring relation is due to the fact that we, as human beings, spend large tracts of our lives dependent upon others and much our interaction with others involves experiences where others depend upon us.

Example 16.1 — children and the elderly. Consider, for example, that as children you depend upon your parents (or some adult) to survive. In addition, as we age, we sometimes depend upon our children or other adults for continued survival.

Not only do we spend a large portion of our lives depending upon others and this dependency is necessary for our survival, but it is also crucial for human progress and production.

Example 16.2 — food production, etc.. Consider that, perhaps, you are not self-taught, you don't grow your own food, you don't write or enforce the laws, you don't fix your own car, you don't make your own roads, or build your own bridges, you don't make your own phone, etc.

Second, care ethics contends that in trying to determine what is morally good and bad for us to

do in any situation emotion plays an important role. This is neither to say that every emotion is important for evaluating moral situations or that one can unreflectively read off certain emotions to moral truths. Rather, care ethics insists that emotions like sympathy, empathy, sensitivity, and responsiveness are moral emotions that when they are cultivated and reflected upon can

1. help us to act on what our reason recommends as the right moral act
2. help us to better determine what is morally right and wrong

Virginia Held argues that even anger can help us to determine the right moral act, writing “anger may be a component of the moral indignation that should be felt when people are treated unjustly or inhumanely, and it may contribute to (rather than interfere with) an appropriate interpretation of the moral wrong.”¹

Exercise 16.1 One of the core claims of care ethics is that emotions not only can help to act on our reason but it can also inform us as what is morally right and wrong. Consider the following emotions and try and determine how these emotions would help you to behave more morally: sympathy, empathy, sensitivity, disgust, sadness, fear, pride.

Third, the traditionally dominant ethical theories often place a high premium on constructing abstract, universal, and impartial moral rules. When it comes to one’s actual relations with those one is partial toward, these dominant ethical theories allow for such partiality only when it is permitted by the universal moral rules.

Example 16.3 Consider the objection of partiality to utilitarianism. This objection stated that since you are required to maximize well-being in the world, you are not permitted to help those closest to you if another act would increase the well-being of someone you don’t know to a greater degree. The utilitarian, however, can respond that utilitarianism is not inconsistent with being partial to those nearest to you especially given the fact that your actions often have a greater impact on their well-being than a stranger. Thus, partiality is permitted provided it adheres to the universal moral law laid out by utilitarianism.

In addition, these ethical theories often construct the central terms of the debate between two extremes: (i) being naturally compelled to seek out one’s interests exclusively (egoism) and (ii) being rationally compelled to obey some universal moral principle (utilitarianism, deontology, divine command theory, etc.). The debate thus is phrased in terms of a concern of our individual interests and a consideration of humanity in general. Or, more basically, a consideration of acting for one’s self and acting solely for others

In contrast, the ethics of care focuses neither on the importance of our own individual interests exclusively nor humanity in general, but instead contends that when we care for someone, (i) we engage in an act of engrossment and (ii) undergo a motivational displacement.

Definition 16.2 — engrossment. An act of care is an act of engrossment if and only if the caring-for individual attends to the cared-for individual on their own terms, viz., attend to them without some selfish intent, without projecting one’s own interests onto the cared-for individual, and without judging or evaluating them

Example 16.4 — engrossment. Suppose Tek is recovering from a knee surgery. He has difficulty walking. His friend Liz comes over to help him. Tek wants to try to do most things himself since he thinks this will help his recovery. Liz respects this choice and tries to help him out in a way that is considerate of this fact.

Example 16.5 — non-engrossment. Suppose Tek needs help with her Math homework. Liz could help her but instead thinks Tek could use help with his English homework instead. This would not

¹Held, Virginia. *The Ethics of Care: Personal, Political, and Global*.

be a case of engrossment. Liz does not attend to Tek on his own terms.

Definition 16.3 — motivational displacement. We undergo a motivational displacement when we take on the interests of others as though they were our own (when our own interests become intertwined with the interests of the cared-for)

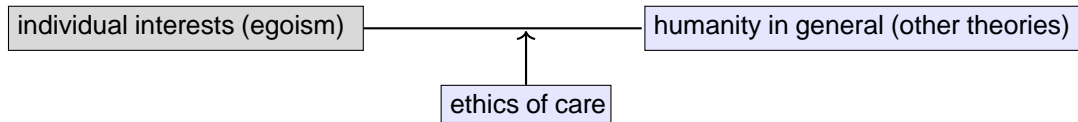


Figure 16.1: The ethics of care points out that we ought neither to understand moral problems exclusively in egoistic terms or in terms of humanity as a whole. Instead, we ought to consider moral problems with specific individuals that we care for.

In adopting this middle ground, the ethics of care recognizes the impossibility of caring for every single individual. We cannot always see everyone's interests as our own nor can we act in a way that develops these interests. Instead, it recognizes the moral importance of our existing relations between family and friends and uses these relations to draw broader conclusions about how we should act more generally.

Fourth, care ethics adopts a relational view of persons. That is, in traditionally dominant moral theories, persons are viewed as independent, rational, self-interested, and autonomous. On this approach, we are individuals first and then enter into relationships with others. In contrast, care ethics understands persons to begin as dependent upon others and then retain a kind of interdependence with others throughout one's life.

16.2 Arguments for care ethics

16.2.1 Argument from partiality

One argument in favor of feminist ethics is that it better accommodates our intuitions behind partiality than theories that adopt a more universalistic approach to ethical theory.

P1: We have moral obligations to specific individuals (partial obligation)

P2: The dominant theories of moral obligation cannot make sense of partial obligation.

P3: Care ethics can provide an account of our partial obligation in its theory of care.

C: Therefore, care ethics is preferable to the traditionally dominant ethical theories.

With respect to P1, the assertion here is that we have no abstract or generic moral obligations. Instead, we find ourselves in dependency relations where we have obligations to specific individuals because they depend upon us (e.g. children, friends, coworkers, parents, etc.).

With respect to P2, the majority of ethical theories specify generic ethical rules that do not take into account the specific dependencies others have on us or we have on others.

With respect to P3, care ethics does provide an account of these dependencies through its notion of care.

16.3 Arguments against care ethics

16.3.1 The scope of the moral community

One of the main problems for care ethics is that in making the primary moral concept that of care it seems to create an overly narrow moral community. That is, if we have moral obligations only to those that we care-for, then the status of those who we do not care-for is either (i) undefined or (ii) we have no moral obligation to these individuals.

P1: If care ethics is true, then we cannot care for everyone.

P2: If we cannot care for everyone and the scope of the moral community is determined by those that we care-for, then it seems that (i) we have no moral obligations to those we don't care for or (ii) the notion of care is insufficient to explain all of our moral obligations.

P3: We do have moral obligations to those that we do not care-for.

C: Therefore, care ethics is either false or incomplete.

A care ethicist might reject P2 and contend that we do not simply have ethical obligations to those we care for. The care ethicist might weaken P2 and contend that we have moral obligations to individuals if (i) there is the potential to have a relationship with them and (ii) that relation has the potential to grow into a relationship where we would care for them.

Exercise 16.2 How might a care ethics approach respond to the above argument? Which premise would it reject?

16.4 Further Readings

This handout relied upon a number of works, including:

Held, Virginia. 2006. *The Ethics of Care: Personal, Political, and Global*. Oxford: Oxford University Press.

Shafer-Landau, Russ. 2018. *The Fundamentals of Ethics*, 4th ed. New York: Oxford University Press.

Engster, Daniel. 2007. *The Heart of Justice*. Oxford: Oxford University Press.

17. Ethical Relativism

17.1 Introduction: moral nihilism and moral relativism

This handout introduces the topic of moral (ethical) relativism.

Definition 17.1 — moral skepticism. Moral skepticism refers to the general position that objective moral standards, principles, or truths are problematic either because (i) they do not exist or (ii) there is no way of knowing them.

First, moral skepticism comes in two forms. The first is the form that denies the *existence* of objective moral standards, principles, or truths. The second is the form that raises doubt about our capacity to *know* these objective moral standards, principles, or truths

Second, defining moral skepticism involves the notion of **objective moral standards, principles, or truths**.

Definition 17.2 — objective moral standard. An objective moral standard is a rule, principle, or maxim that states what one should or should not do that applies to everyone at all times regardless of whether they believe, know, or disagree with that standard.

Third, there are two principal varieties of moral skepticism of the sort that assert that object moral standards, principles, or truths do not exist. These are (i) moral nihilism and (ii) moral relativism (we won't consider forms of moral skepticism that have to do with knowledge).

Definition 17.3 — moral nihilism. Moral nihilism asserts that there are no objective moral truths because there are no moral truths at all.

Moral nihilism are moral skeptics about objective moral standards not because they assert that moral standards fail to be objective (apply to everyone independent of whether or not they agree

with them) but because no moral standards simply do not exist at all.

Definition 17.4 — moral relativism. Moral relativism asserts that there are moral truths but none of these moral truths are objective; instead, all moral truths are relative.

First, in the case of moral relativism, there are moral truths but these moral truths are **not objective** in the sense that simply apply to everyone at all times. Instead, moral relativists contends that a moral judgment is true or false only **relative to** some standard of evaluation.

Second, it is important to note that this is not a claim about the context-sensitivity of language (or a theory about what people *mean* when they utter various moral propositions). The claim is about how we ought to evaluate the truth and falsity of moral propositions. If John says that “it is raining”, he can be understood (from the context), to assert something more precise. Namely, John can be understood to assert “it is not raining at this minute where I am”. Fleshing out John’s shorter sentence using the context would be a theory about what John *means* or the meaning of certain sentences.

Ethical relativism is not a theory of meaning. Instead, it is a theory about how to evaluate the truth and falsity of moral sentences. In the case of the relativist, the relativist contends that if John says “murder is wrong”, John is not asserting that “murder is wrong for me” or “murder is wrong in my community”. He is simply asserting that *murder is wrong*. Relativism asserts, however, that this proposition cannot be determined as true or false universally or objectively. It can only be determined ot be true or false by considering something *X* relative to the moral sentence.

Third, there are two principal forms of moral relativism: (i) cultural relativism (CR) and (ii) individual relativism, also known as ethical subjectivism (ES). See [Figure 17.1](#).

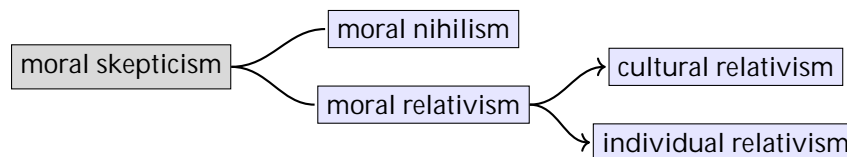


Figure 17.1: Types of moral skepticism

In what follows, we consider both (i) cultural relativism (CR) and individual relativism in more detail.

17.2 Cultural relativism

The first form of moral relativism we will consider is **cultural relativism**. One intuition people have about morality is that it varies from culture to culture. That is, what is morally permissible in culture *A* may not be morally permissible in culture *B* and what is forbidden in culture *A* might be permissible in culture *B*.

This intuition gives rise to the view that there are true moral propositions and standards but they are not true or false objectively but only *relative* to specific cultures or societies.

Definition 17.5 — ethical cultural relativism (ECR). Cultural relativism is the theory that contends that moral standards, principles, or propositions are true or false relative to specific cultures or societies. An act is morally obligated if and only if it is obligated by the underlying principles, values, or ideals of a culture, morally impermissible if and only if it is forbidden by the culture,

and permissible if and only if it is not forbidden.

Several points of clarification. **First**, ethical cultural relativism is distinct from what might be called **descriptive cultural relativism about morals**.

Definition 17.6 — Descriptive (sociological) cultural relativism (DCR). DCR is the descriptive claim that *what* people believe to be morally right or wrong varies from culture to culture.

DCR is just the account of what different cultures believe is morally permissible and impermissible. It **says nothing about whether these beliefs are true or false**. DCR **does not** say that there is no cultural overlap. It says instead that there are some differences about what people think concerning morality. Thus, it is possible that all cultures have some values in common.

Example 17.1 In Herodotus's *Histories*, DCR is illustrated when the King Darius notes the different beliefs of the Greeks and Callatians. The Greeks took it to be morally impermissible to eat the dead bodies of relatives and presumably morally acceptable to cremate the dead. In contrast, the Callatians took eating dead relatives to be morally permissible but cremations to be morally impermissible.

In contrast, **ethical cultural relativism** is the theory that, as a matter of fact, the moral rightness / wrongness of an act is not something that you can specify independently of the culture in which the act is performed. It says that there is a fact of the matter as to whether an action is morally good, bad, or permissible but this fact cannot be specified without reference to a specific culture.

Example 17.2 — An Extreme Example. The Jivaro are an Indian clan. They are known to be an intensely warlike group, tremendously protective of their freedom and unwilling to subordinate themselves to other authorities. In this fierce society, headhunting and shrinking the heads is a process much revered and honored. This practice is morally permissible. In contrast, in the U.S., chopping off someone's head then trying to preserve it is not only a criminal offense, but held by almost everyone to be immoral. According to ethical cultural relativism, there is no objective, universal rule that says headhunting is wrong. Rather, it is morally permissible (even admirable) for the Jivaro, while it is morally wrong from those in the U.S.

Second, ECR is not to be confused with the claim that *everyone ought to be tolerant of other cultures* is objectively true. If you accept ECR, you deny that there is some objective standard by which various moral statements can be evaluated. And so the truth or falsity of *everyone ought to be tolerant of other cultures* will vary depending upon the culture under which it is evaluating.

everyone ought to be tolerant of other cultures may be true relative to culture C_1
everyone ought to be tolerant of other cultures may be false relative to culture C_2

Third, ECR does not exclude the possibility that every culture treats some moral proposition as true. That is, it is possible that cultures may behave extremely different from each other but be committed to the same ethical principles. The apparent difference is not due to a disagreement about the truth or falsity of moral principles but to either (i) situational differences or (ii) differences in beliefs about the world.

Situational differences: two cultures *A* and *B* might appear different because *A* allows infanticide while *B* does not. On the surface, it looks like they have different ethical rules about infanticide. However, it may be the case that they have *the same moral views*, it is just that infanticide is permitted in culture *A* because living conditions are extremely harsh and individuals would starve if infanticide were not permitted. Thus, two cultures can have the same ethical rules but their implementation of these rules differ due to situational differences. Differences in beliefs about the world: suppose *A* and *B* belong to two different cultures.

Suppose also that both believe you should never kill a being that has a soul. However, suppose that *A* thinks only full-grown people have souls, while *B* thinks that children also have soul. If this is the case, it is morally permissible for *A* to kill children whereas it is not for *B*. Thus, two cultures *A* and *B* can *share the same moral beliefs* but the implementation may differ due to their having different beliefs about the world.

Exercise 17.1 In most Western cultures, suicide is regarded as immoral. Some cultures or subcultures at various periods of time have taken a different attitude to the moral status of suicide, e.g. the Japanese tradition of hara-kiri or the practice of suicide bombing. Do these differences in attitudes toward suicide reflect a fundamental difference in basic moral principles regarding the intrinsic value of human life?^a

^aPartially taken from Judith A. Boss's *Ethics for Life: A Text with Readings*. 3rd ed. p.107.

17.2.1 Argument for cultural relativism from descriptive cultural relativism

Here is an argument for ethical cultural relativism.

P1: If different cultures have different moral codes, then right and wrong are only matters of opinion, and opinions vary from culture to culture.

P2: Different societies/cultures have different moral codes.

C: Therefore, there is no objective moral "truth," only cultural relativism.

First, notice that **P2** is simply **descriptive cultural relativism**, namely that different societies believe different things about which actions are morally good as opposed to evil. The key premise then is **P1**, which there are several reasons to deny.

Objection 17.2.1 — P1 is not argued for. The fact that two cultures disagree about what is morally right or wrong *does not imply that there is no fact to the matter*. For example, two cultures might disagree about whether the world is flat or spherical. From this disagreement, it does not follow from their disagreement that the world is not either flat or spherical. Even if different cultures have different moral codes, this does not imply that *they ought to have different moral codes*. What is the case does not determine what ought to be the case. For example, suppose you decided to skip class on Friday, this does not mean you ought to have skipped class (there could have been a big test!).

17.2.2 Argument against cultural relativism: cultural nihilism

Cultural relativism depends upon the existence of cultures. That is, there are only moral truths if there are cultures relative to which moral propositions can be true or false. If, however, there are no cultures or the existence of a definable cultural norms, values, or ideal fails to exist, then **ECR** is false.

Thus, we can raise two related arguments against cultural relativism. The first argues against the existence of cultures.

P1: If cultural relativism is true, then cultures, societies, communities must exist.

P2: The only things that exist are individual material beings and/or collections of these beings that have definite limits.

P3: A culture is not a material being (it is a sociological construct) and it does not have definite limits.

C: Therefore, cultural relativism is false.

This argument is somewhat problematic in that while the argument might be rationally persuasive, people generally think there are things like society, cultures, communities, etc.

The second argument against cultural relativism takes a more nuanced approach in that it accepts

the possibility of there being cultures but rejects the idea that there is any sort of definable cultural moral standard that any individual can appeal to.

P1: If cultural relativism is true, then there must be some definable moral standard(s) within a culture (if there were not, then moral propositions could not be true or false).

P2: Cultures are too variegated and complex to find a single definable moral standard, ideal, etc.

C: Therefore, cultural relativism is false.

The key premise in this second argument is P2. One way that it might be defended is to point out that not everyone in a culture or a community think the same about every moral issue. Since there is not universal agreement about moral topics within a culture, this raises the question of *what* determines the moral standards of a culture.

Exercise 17.2 What do you think determines the moral standards of a community or culture? The political leaders? The majority of the people? A certain subset of the people? Are certain individuals excluded (e.g. criminals)?

17.3 Individual ethical relativism (ethical subjectivism)

The second form of moral relativism we will consider is **individual ethical relativism** (ethical subjectivism).

Definition 17.7 — individual (ethical) relativism (IR). Individual (ethical) relativism is the theory that contends that moral standards, principles, or propositions are true or false relative to specific subjects (persons). An act is morally obligated if and only if it the subject approves of it, morally impermissible if and only if the subject disapproves of it, and permissible if and only if it is not disapproved of.

17.3.1 Arguments for individual (and cultural) relativism: from moral facts and knowledge

One of the more powerful arguments in support of both ethical cultural relativism and individual relativism is that the theories can give straightforward, non-spooky accounts of both the nature of moral facts and the possibility of moral knowledge.

P1: There are no objective moral facts.

P2: But there are moral facts nevertheless.

P3: Moral relativism can provide a straightforward, non-spooky explanation for the nature of moral facts and the possibility of moral knowledge (in contrast to other theories).

C: Therefore, we ought to prefer moral relativism over other ethical theories.

We will examine **P1** in more detail in the next handout. **P2** can be taken for granted as uncontroversial. The key premise then is **P3**. **First**, the premise implies that other theories give spooky or mysterious accounts of the object that makes various moral propositions true or false. **Second**, the premise asserts moral relativism provides an extremely straightforward account of what makes various moral propositions true:

With respect to **moral facts**, moral relativism can straightforwardly explain what makes a moral proposition true or false. In the case of ethical cultural relativism, it is the values, ideals, or standards set by the culture. In the case of individual relativism, it is simply the judgment or belief of the individual.

With respect to **moral knowledge**, moral relativism can straightforwardly explain how we learn what the moral facts are. In the case of ethical cultural relativism, we simply learn the values, ideals, or standards of a culture. In the case of individual relativism, we simply

reflect upon our own judgments or beliefs or beliefs (or the judgments or beliefs of another individual).

17.3.2 Arguments against relativism: moral disagreement is impossible

Consider that when I say “you have done something wrong”, if IR is true, then you understand me to be evaluating the truth of this sentence using my personal standard of moral evaluation. You understand me to take that sentence as true *for me*. You also, however, take it as *false for you*. So, there is no disagreement about the moral facts. But this is an implausible consequence given how much disagreement there is about moral matters and so the IR owes a further explanation. Thus, we can formulate the following argument:

P1: If IR is true, then if I say action X is morally wrong and you say action X is morally acceptable, then we are not disagreeing. In fact, moral disagreement is impossible.

P2: People do disagree about moral matters.

C: Therefore, **IR** is false.

18. Moral Nihilism

18.1 Moral nihilism

A moral fact is a fact that would make a moral proposition (e.g. “x is morally wrong” or “x is morally obligated”) true or false. There are three intuitions about moral facts:

- there are objective moral facts
- moral facts are relative to either an individual or a culture
- there are no moral facts

In this handout, the third intuition is examined. The theory that asserts that moral facts are neither objective nor relative but instead are not real is known as **moral nihilism**.

Definition 18.1 — moral nihilism. Moral nihilism is the metaethical moral theory that there are no moral facts and thus there is no moral knowledge.

There are two types of moral nihilism: (i) moral error theory (or simply “error theory”) and (ii) moral expressivism (or simply “expressivism”). See [Figure 18.1](#).

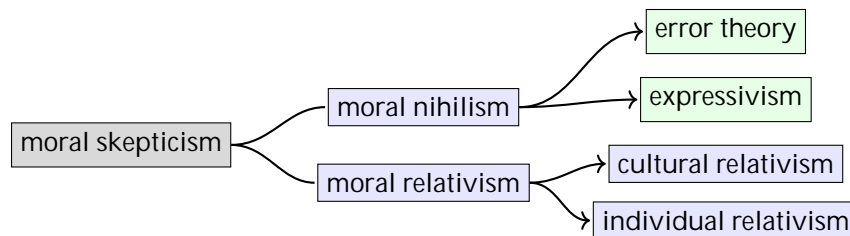


Figure 18.1: Varieties of moral skepticism, including the two types of moral nihilism

18.2 Error theory

Error theory is comprised of the following core claims:

Moral facts claim: the world consists of only scientific facts; moral terms refer to non-scientific (non-physical) entities and thus there is nothing in the world that corresponds to moral goodness, moral badness, moral permissibility, etc.

Moral propositions claim: all moral propositions are false since there are no moral facts, there is nothing in the world that would make moral propositions true

Language claim: that is, when we utter a moral proposition like “murder is wrong” we are trying to say something that is true or false, viz., we assume there is something in the world (a truthmaker) that would make our proposition either true or false.

It is important to keep in mind several points concerning error theory.

First, error theorists are not asserting that it is morally permissible to do whatever you want. That is, they are not asserting that we ought to flip traditional moral rules on their head and engage in immoral behavior. Since there are no moral facts, they are not saying that murder or any action for that matter is morally good or bad, and so the theory is not making any sort of moral recommendation about you should or should not continue living your life.

Second, error theorists are not contending that when we sincerely assert a moral proposition, we are unclear about what we are asserting. That is, if someone utters “murder is wrong”, they *mean* to say that there is some moral property (or fact) in the world that makes the act of murder morally impermissible.

Third, error theory rejects both objectivist theories of ethics and relativistic theories. As a form of moral nihilism, error theory denies that moral claims are true or false relative to this or that culture or this or that individual. Error theorists contends that moral claims are not true even when we use some sort of relativist standard of evaluation.

Fourth, error theory contends that all of our moral claims are **false**. This is because we are saying that reality (or the world) has certain properties that it does not (in actuality) have. Consider, for example, if I asserted that “there is a unicorn in my backyard”, this assertion would be false since *unicorns do not exist*.

18.2.1 Argument for error theory

The general intuition that supports error theory is that there is no reason (or evidence) to believe that there is something that would make our moral propositions true or false. That is, there are no transparent objects or properties that we could point to or experience that would correspond to the meaning of moral terms.

Using this intuition, we might formulate the following argument in favor of error theory:

P1: If a proposition *P* is true, then the terms of *P* refer to real things and properties.

P2: The terms of moral propositions do not refer to real things and properties.

IC: Therefore, all moral propositions are false.

P3: When an individual utters a moral proposition *P*, they intend to refer to a real moral thing or real moral properties.

C: Therefore, when individuals utter moral propositions, they utter false propositions.

P1 makes the claim that if a proposition *P* is true, then its terms must refer to real objects or properties. If I uttered the proposition “unicorns are dancing on my bed”, since there are no unicorns, this proposition would be false. Furthermore, if I uttered a proposition that does not refer to any real property, the proposition would be false. For example, if I asserted that “100 is the largest integer”, since there is no object that corresponds to the property *is the largest integer*, then what I have asserted

would be false.

While there is some controversy associated with P1, the key premise for error theorists is P2. This premise asserts that the various terms of moral propositions fail to refer to real objects or real properties. That is, expressions like *morally good*, *morally bad*, *morally neutral* are like unicorns and greatest numbers; they refer to no real object.

P2 can be supported in a variety of different ways. One way that it might be supported is that the objects of moral terms are not empirical or observational. That is, many would argue that while you can see trees, taste cheese, or smell the moisture in the air, you cannot perceive moral properties like moral goodness or badness. Without an empirical correlate, moral terms seem as though they fail to refer to anything real.

IC draws the intermediate conclusion that all moral propositions are false. This follows from P1 and P2.

P3 asserts the key language claim of the error theorists. This claim is that when individual do utter moral propositions, they mean (or intend) to say something about real moral objects or properties. But, P2 asserts that these do not exist, and so C concludes that all moral propositions are false and whenever we use moral language to utter moral propositions, we say something false we err).

18.2.2 Against error theory: error theory does not meet the burden of proof

Perhaps the most pressing argument against error theory and all forms of moral nihilism in general are that in denying moral facts and thereby denying the truth of any moral propositions, nihilists take on an extremely heavy burden of proof. That is, since many individuals hold at least one moral proposition to be true, and thereby made true by some moral fact, moral nihilists must provide not merely good reasons for their theory but excellent reasons for their position.

P1: If error theory is true, then there are no true moral propositions.

P2: The error theorist bears an extremely high burden of proof.

P3: The error theorist has not met their burden of proof.

C: therefore, error theory is false.

P1 is a straightforward consequence of moral error theory.

P2 involves two parts to consider.

First, P2 asserts that the burden of proof is assigned to the moral nihilist and they bear a high burden of proof. When one bears the burden of proof one has the responsibility of proving what one asserts to be true. The burden of proof is typically assigned (although there are some exceptions) to the individual who:

puts forward a highly controversial or unusual claim, or
who rejects a widely accepted view

Example 18.1 — Bigfoot. Suppose I assert I saw Bigfoot. This is a highly unusual or controversial claim and so I would bear the responsibility of proving this claim. You would not bear the burden of proof in proving me wrong.

Exercise 18.1 Quickly look through the following and determine who bears the burden of proof:

Tek asserts that he saw a UFO but Jon denies that Tek saw a UFO.

Tek asserts that the earth is flat but Jon asserts it is spherical.

Tek asserts that he knows who will win the next World Cup while Jon says that he doesn't know who will win.

Second, P2 asserts that error theory is highly controversial. It is highly controversial because the belief that there are at least some moral facts is generally supported by our common sense intuitions.

Example 18.2 — genocide. When the majority of individuals are asked about the moral permissibility of a genocide of a group of individuals that they take to be valuable, they firmly deny that such an act is morally permissible. In fact, while we might theorize in general in a way that is relativistic or nihilist, individuals often will make positive moral claims when given specific cases to evaluate.

Finally, P3 asserts that error theorists have not met their burden of proof. We won't examine this premise here in much detail, but we can say that (i) we haven't provide enough arguments to meet this burden of proof here and (ii) pointing out some flaws in the argument in support of error theory might be sufficient to show that error theorists have not met their burden of proof.

18.3 Expressivism

Expressivism is comprised of the following core claims:

Moral facts claim: same as error theory

Moral propositions claim: same as error theory

Language claim: moral language is deceptive in that when we utter a moral proposition like "murder is wrong" we are not actually uttering a moral proposition. Rather, if Jon asserts that "murder is morally wrong", Jon is not describing the world and is not asserting that there are moral feature to the world. Instead, John is either (i) expressing his psychological attitude toward emotively or (ii) trying to command others not to murder.

A few points of clarification.

First, the distinguishing point between error theory and expressivism concerns **moral language**. For error theorists, **moral language aims to describe some fact about the world** but moral language always fails as there are no moral features in the world. By contrast, for emotivists **moral language does not describe some fact** (either about reality or a description of one's psychological attitudes) but instead functions to (i) express some attitude or (ii) persuade others of some action.

Let's look at two applications of the emotivism with respect to moral sentences.

1. Homosexuality is wrong.
2. Homosexuality is morally permissible.

According to emotivism, what (1) and (2) really say is something like this:

- 1a Don't be homosexual! (1 is really a command that aims to change your mind)
- 1b Homosexuality, gross! (1 is really an exclamation that expresses an attitude)
- 2a Be homosexual, if you want! (2 is really a command)
- 2b Homosexuality, whatever! (2 is really an exclamation that expresses an attitude)

In short, moral language that appears to express propositions is an illusion. In reality, these apparent moral propositions express *exclamations* or **imperatives**.

Second, apparent moral propositions do not actually express propositions (content that can be true or false), expressivists contend that our moral claims and assertions are **neither true nor false**. This is because exclamations and commands do not express the sort of content that can be true or false. This differs from error theorists who contend that all of our moral propositions are **false**.

18.3.1 Arguments for emotivism

One argument for emotivism is that it is a better skeptical alternative than forms of moral relativism.

- P1: Some form of moral skepticism about objective moral facts is true.
- P2: Various forms of moral relativism make moral disagreement impossible.
- P3: Emotivism does not make moral disagreement impossible.

C: Therefore, emotivism is the best form of moral skepticism.

P1 asserts that there are no objective moral facts. That is, moral skepticism is true.

P2 rehearses the objections raised against cultural relativism and individual relativism discussed earlier. To review, if Tek asserts “murder is wrong”, this is true relative to Tek’s moral standards. If Jon asserts “murder is not wrong”, this is true relative to Jon’s moral standards. While Tek and Jon appear to disagree about the moral status of murder, they are **not in disagreement since both statements are true**. Thus, moral disagreement is impossible.

In contrast, P3 can account for moral disagreement even though this type of disagreement is not about whether a particular act is morally permissible or impermissible. Instead, the disagreement is a disagreement in emotional attitudes. To see this more clearly, let’s consider the following example:

Jon: You betrayed my trust. What you did was morally wrong!

Vic: Yes, I did betray your trust. But, you are wrong, betraying someone’s trust is not morally wrong!

Here is how the emotivist sees this situation:

Jon: You betrayed my trust. Don’t betray my trust! Ugh!

Vic: Yes, I did betray your trust. But, betray your trust, not betray your trust, whatever!

Jon and Vic do not disagree in the sense they assert propositions that say contradict each other, but they disagree in another sense. They disagree in the sense that people have two different attitudes toward betraying someone’s trust. In other words, with respect to P3, for emotivists, moral disagreement is possible except that Jon and Tek are not disagreeing about the facts but instead have a clash of emotions or what types of acts they want performed or not performed.

18.3.2 Argument against emotivism: cannot explain the denial of past moral beliefs

One argument against emotivism stems from the attitude that we do not take our moral propositions to express either exclamations or commands but instead take them to assert facts. This is not only evident when we analyze our moral propositions here and now, but also when we reflect upon the **denial** of moral propositions we once accepted. In other words, emotivism cannot adequately explain the denial of our prior moral beliefs.

P1: If emotivism is true, then whenever we change our moral beliefs then we are simply changing our attitudes toward certain acts.

P2: When we change our moral beliefs, this is not merely a result of changing our attitudes toward certain acts but a **denial** of the truth of our prior moral beliefs (thus emotivism cannot explain our change in attitude toward past moral beliefs).

C: Therefore, emotivism is false.

With respect to P1, suppose that an individual changes their mind on the morality of abortion. At time t_1 , Liz asserts that abortion is morally wrong. Later, at t_2 she changes her mind. She contends that abortion is morally permissible. For the emotivist, at t_1 Liz had one attitude (or emotion) toward abortion but at t_2 she had a different attitude. For the emotivist, this change in attitude is nothing more than a change in one’s emotional attitude toward that act.

P2 asserts that when we change our minds about the status of moral propositions, this is not simply a change in our attitudes, emotional states, or preferences toward particular act. Rather, when we deny a moral proposition that we previously held to be true, we regard ourselves as having been mistaken.

We may like certain foods, movies, activities, art, and music at certain points in our life. Later, we may have different tastes. We acknowledge that this is nothing more than a shift in tastes. We were **not wrong** to like the foods we no longer enjoy; we were not mistaken in liking the movie or art we

once liked; and, our old hobbies were not necessarily flawed in some critical way.

The emotivist contends that changing speech about moral propositions reflects this same attitude. But, according to P2, this is not the case. When Liz asserts that abortion is morally permissible at t_2 she will also routinely take (and sometimes assert) that her prior position was false. That is, unlike the case in one's shifting preferences in art, music, and food, a shift in moral claims comes with a negative evaluation of the prior position.

That is, when Liz asserts that abortion is morally permissible at t_2 , she also asserts that *she was wrong to believe that abortion was morally wrong at t_1 .*