Lecture Notes for *Introduction to Philosophical Problems*

J. Dmitri Gallow

Draft of January 23, 2019

Contents

What is it to consider your life sub specie aeternitatis*? How, according to Nagel, does considering your life* sub specie aeternitatis *help you to recognize its meaninglessness? Why does Nagel think that our lives are absurd?*

1. Reading Nagel, we should pay attention to the difference between the following two theses, both of which Nagel accepts:
   (a) Thesis #1: Our lives are meaningless.
      i. What does it mean for a life to be *meaningless*? It's not immediately clear, but, as a rough characterization: a life is meaningful if it *matters*, if it has a *valuable purpose*, or is directed towards *valuable ends*.
   (b) Thesis #2: Our lives are *absurd*.
      i. For Nagel, our lives are absurd because, even though we are in a position to recognize that our lives are meaningless (thesis #1), we act as though they are not. We treat our lives and our projects as though the were directed towards some valuable ends, even though we are capable of recognizing that they are not.

2. Though Nagel thinks our lives are meaningless, he thinks that many of the reasons people typically give for thinking so are bad ones. For instance:
   (a) Sometimes, to argue that our lives are meaningless, people will say "nothing you do will matter in a million years".
   (b) Nagel: even if this is true, it shouldn't give you any reason to think that our lives don't matter *now*. He reasons as follows:

      P1. Nothing we do will matter in a million years.

      P2. If nothing we do will matter in a million years, then (by the same token) nothing which will be the case in a million years matters now.

      C1. Nothing which will be the case in a million years matters. [from P1 and P2]

      C2. It does not matter that nothing we do will matter in a million years. [from P1 and C1]

   (c) Other times people will point out how small we are, or how short our lives are. Nagel also finds fault with these reasons for thinking that our lives are meaningless. If our lives are meaningless, would making us bigger endow them with meaning? If 70 years of our existence is meaningless, would an eternity of the life suddenly become meaningful?

3. Even though these reasons are not good reasons for thinking that our lives are meaningless, Nagel thinks that they can nevertheless help us to recognize that our lives are meaningless.
   (a) When we think about our lives from the perspective of a million years in the future, or from the perspective of a vast universe of which we are but a tiny and momentary speck, we 'step back' from, or 'bracket' all of our own values and commitments.
   (b) When we step back from and bracket our own values and commitments, Nagel says that we survey ourselves *sub specie aeternitatis* (or 'from the perspective of eternity').
   (c) From this perspective, we are capable of recognizing several facts, which Nagel mentions intermittently over the course of his article:
      i. Had we been constituted differently, we would have had different fundamental commitments—we would have valued different things and taken different considerations as reasons for and against various actions.
         A. For instance, as it is, we take the fact that an act will harm us as a reason to not do it (harm avoidance reasoning). But we could be so constituted that we took the fact that an act will harm us as a reason *to do it* (harm seeking reasoning).
      ii. We are incapable of saying anything to justify these fundamental commitments which does not simply presuppose those very commitments.

A. For instance, we may note that harm seeking reasoning will lead a species to go extinct very quickly. But this is only a reason to avoid the harm seeking reasoning according to the harm avoidance reasoning. The harm seeking reasoning would see this fact as a reason *in favor of* the harm seeking reasoning.

4. Nagel thinks that, once we are able to recognize that (1) our most fundamental commitments are contingent and that (2) we can only justify them circularly, by relying upon those very commitments, we are able to recognize that (3) our lives are meaningless.

   (a) It's not immediately obvious why Nagel thinks that recognzing (1) and (2) puts us in a position to recognize (3). But here is an attempted re-construction of his reasoning:

   P1. The only reasons we have for accepting our most fundamental commitments are circular reasons.

   P2. A circular reason is not a good reason.

   C1. We have no good reason to accept our most fundamental commitments. [from P1 and P2]

   P3. If we have no good reason to accept our most fundamental commitments, then we have no good reason to think that our lives are meaningful.

   C2. We have no good reason to think that our lives are meaningful. [from C1 and P3]

   P4. If we have no good reason to think that our lives are meaningful, then we should believe that they are not.

   C3. We should believe that our lives are not meaningful. [from C2 and P4]

5. What makes our lives *absurd*, according to Nagel, is not just that they are meaningless. What makes them *absurd* is that, even though we are in a position to recognize that they are meaningless, we go on taking them seriously, and treating them as though they did.

6. If Nagel is correct, if the manner in which we conduct our lives is absurd, what should we do about it?

   (a) Nagel: why think we should do anything about it? We should only do something about it if it is a *problem*. But we shouldn't think that it is. He reasons as follows:

   P1. Nothing matters.

   C1. So, it does not matter that nothing matters.

According to Wolf, what does it take for a life to be meaningful? On her account, is it possible for somebody to be mistaken *about whether their life is meaningful? If so, how? If not, why not?*

1. Wolf begins by distinguishing two questions: 1) 'What is the meaning of life?', and 2) 'What makes a life meaningful?'

   (a) The first question asks about the purpose of our existence. This question, according to Wolf, is easily answered: if we have a creator, then life has a purpose (specifically, whatever purpose the creator had in creating us); if there is no creator, then life has no purpose.

      i. Note a disagreement with Nagel. Nagel thought that our lives could be meaningless *even if* there is a creator with some aim in creating us.

   (b) Though she thinks the first question easily answered, the second question is more interesting. Her goal is to provide an answer.

2. Wolf's approach to answering this question is to consider some paradigm instances of *meaningless* lives, and try to discern that thing the lack of which makes them meaningless. Here are her examples:

   (a) *The Blob*: this person sits at home on the couch every day, drinking beer and watching situation comedies.

   (b) *The Idle Rich*: this person "flits about, fighting off boredom, moving from one amusement to another. She shops, she travels, she eats as expensive restaurants, she works out with her personal trainer."

   (c) *The Corporate Executive*: this person works twelve-hours a day, every day of the week, solely in order to accumulate a vast fortune.

   (d) *The Pig Farmer*: This person has only instrumental goals and no final goals: they want to grow corn in order to feed it to their pigs. They want to feed the pigs so they can sell the pigs at market. They want to sell the pigs so they can buy more land. They want to buy more land so they can grow more corn. They want to grow more corn so they can feed more pigs, and so on.

   (e) *The Scooped Scientist*: This scientist dedicates their life to developing a cure for cancer. The day before their discovery is to be revealed, another scientist announces that they have discovered the very same cure.

3. Reflecting on these cases, Wolf constructs an account of what makes a life meaningful.

   (a) Why are the lives of the Blob and the Idle Rich meaningless? Because their lives are not actively engaged in any projects. Theirs are lives of *passivity*. And such a life cannot be meaningful. So Wolf invites us to conclude that a meaningful life must one of *active engagement* with some project or projects. (Here, 'project' should be read broadly, so that it includes things like personal relationships.)

      i. As an aside: Wolf also thinks that cases like the blob and the idle rich show us that a purely *subjective* account of meaningfulness cannot be correct.

      ii. That is: consider the following proposal:

      Proposal A life is meaningful iff* the life seems meaningful to the person living it.

      iii. The Blob and the Idle Rich may very well *think* that their lives are meaningful. Their lives may *seem* meaningful to them. Even so, they are not.

   (b) Active engagement may be necessary for a meaningful life, but it is not sufficient. The corporate executive and the pig farmer are both actively engaged in their projects. But these projects have no real value. So Wolf invites us to conclude that a meaningful life must be actively engaged in projects *of positive value*.

      i. Wolf doesn't think this positive value need be *moral value*. A life actively engaged in the production of beautiful works of art can be meaningful. So Wolf wishes to include aesthetic value as well.

      ii. Note: Wolf thinks that the life must *actually have* positive value. It is not enough that the individual *thinks* that their life has positive value.

      iii. Consider the following alternative account of when a life is meaningful:

* 'iff' is shorthand for 'if and only if'.

An Alternative A person's life is meaningful iff it is actively engaged in a project that they believe to be of positive value.

iv. Wolf think that the case of the corporate executive shows that this alternative is false. The corporate executive may *think* that the accumulation of wealth for its own sake is valuable. But, since they are wrong, their life is still meaningless.

v. She additionally provides the following argument against the alternative account: sometimes, people realize that their lives up to that point have been meaningless—call this an 'awakening'. People who have awakenings like these often thought, before the awakening, that they were engaged in projects of positive value. But, if the alternative account were correct, then this would be impossible—since, if the alternative account were correct, what they would be realizing, when they realized that their lives up to that point were meaningless, was that they didn't *think* their projects were valuable. But of course they *did* think that their projects were valuable,. So, if the alternative account were correct, nobody could realize that they were blind to the meaninglessness of their life. That is:

P1. If the alternative account were correct, then it would be impossible to realize that your life has been meaningless.

P2. It is possible to realize that your life has been meaningless.

C1. The alternative account is not correct.

(c) The scooped scientist is actively engaged in a project of positive value, but their engagement ultimately leads nowhere. It would make sense, upon learning that their life's work was scooped, for the scientist to think that their life was meaningless. So Wolf thinks we should say that a life must not only be actively engaged in a project of positive value—the projects must also be to some degree *successful*.

4. In summary, Wolf's account says:

Wolf's Account A life is meaningful iff it is *actively* and at least somewhat *successfully* enaged in a project or projects of *positive value*.

5. If this is what it is to have a meaningful life, then should we *want* our lives to be meaningful?

(a) Reasons for doubt: meaningful lives are not necessarily *moral* lives, nor do they necessarily make us happier.

(b) Still, Wolf thinks we should want a meaningful life. Her reason: to do otherwise is to act as though you yourself were the only thing that mattered in the universe.

(c) You are but a speck in a vast and value-filled universe. Yours is but one of many perspectives from which your life may be perceived, and your perspective is no more important than any other. You should want live a life which accords with this fact. You should want to live a life actively engaged in projects which advance values beyond your own subjective pleasure and enjoyment. You should, that is, want to live a meaningful life.

Present two arguments from Epicurus and/or Lucretius for the conclusion that death is not a harm. For each of the two arguments you present, explain how Nagel responds to these arguments. (Does he reject one of the argument's premises? If so, what reason does he give for thinking that this premise is false?)

1. Perhaps there is life after the death of your body. Perhaps not. But this is not our topic for today. For today's class, when we say 'death', we are talking about the permanent end of your existence. (So that, if certain religions have it right, then you will *never* die.) Today, we are interested in the following question: Does your death harm you?

   (a) Everyone can agree that your death will harm *somebody*—in particular, all of your loved ones are harmed by your death. Our question isn't whether your death is bad *full stop*. Instead, the question is whether your death is bad *for you*.

Death is not a Harm

2. Epicurus (341–270 BC) thought that death was no harm. In the excerpt from his letter to Menoeceus (which you will read for next class), we can discern two arguments for this conclusion—call them '**No Harm without Awareness**' and '**No Harm without Existence**'.

   (a) **No Harm without Awareness.**

   P1. If you are not aware of something, then it cannot harm you.

   P2. No one is aware of their own death.

   C1. No one is harmed by their own death.

   i. Perhaps P1 is false. Perhaps we can be harmed by the *fear* of death, even if we are never aware of death itself.

   ii. Epicurus: We should only fear things that harm us. If death is not a harm, then there is no reason to fear it. And if we do not fear it, then we will not be harmed by the fear of death.

   (b) **No Harm without Existence.**

   P3. In order for someone to be harmed by something, they must exist at the same time as the thing that harms them.

   P4. Nobody exists at the same time as their death.

   C2. No one is harmed by their own death.

3. Lucretius (99–55 BC) was an Epicurean who wrote a didatic poem called *De rerum natura* (On the Nature of Things). In one section, titled *Folly of the Fear of Death*, we can find the following additional argument against death being a harm—call it the '**Prenatal-Posthumous Symmetry**' argument:

   (a) **Prenatal-Posthumous Symmetry**.

   P5. If people are harmed by their deaths, then so too are they harmed by the time before their conception.

   P6. No one is harmed by the time before their conception.

   C3. No one is harmed by their own death.

Death is a Harm

4. Nagel thinks these arguments are too quick. He wishes to respond to the three arguments above and show how death can be a harm.

(a) First, a clarification: if death is harmful, it is not because being dead is a bad state for you to be in. Rather, it is because *being alive* is a good state for you to be in, and death deprives you of this good state.

(b) Nagel argues for the conclusion that being dead is not a bad state to be in as follows: we don't view the time before our death as a misfortune for us, and people do not object to a *temporary* suspension of their existence.

5. In response to the argument **No Harm without Awareness**, Nagel rejects the titular premise, P1.

   (a) According to Nagel, people can be harmed by things that they aren't aware of.

   (b) For instance: Sarah believes that her husband loves her. However, he is only with her for her money, and has been cheating on her for years. Sarah goes to her grave without learning of the infidelity.

   (c) Sarah *believes* that things are going well for her romantically, but appearances are misleading. Things are not going well for her romantically. She is harmed by her husband's infidelity, even though she is never aware of it.

   (d) So P1 is false.

6. In response to the argument **No Harm without Existence**, Nagel again rejects the titular premise, P3.

   (a) According to Nagel, people can be harmed by things that happen when they are not alive.

   (b) For one example: if a dear friend breaks a deathbed promise to you, you have been harmed.

   (c) For another (?): suppose, before your birth, there is a change to your grandparent's will that has their vast fortune bequeathed to the Harvard business school, rather than their heirs. This change harms you, though you are not (yet) in existence.

   (d) So P3 is false.

7. In response to the argument **Prenatal-Posthumous Symmetry**, Nagel rejects the symmetry between the time prior to your conception and the time after your death, P5.

   (a) One asymmetry: Early death deprives us of further life. But late birth does not.

   (b) If you had been born earlier, you would not necessarily have had fewer of the goods of life (you could have lived just as long, though centuries earlier). However, if you were to die earlier, then you would necessarily have fewer of the goods of life.

   (c) So your death, by depriving you of the goods of life, is a harm. However, your not being born during the centuries preceding your conception is not a harm, since it did not deprive you of the goods of life.

   (d) So P5 is false.

What is the difference between the properties of 'goodness/badness' and 'rightness/wrongness', as we are using those terms in this class? What is the difference between consequentialist *and* non-consequentialist *ethical theories? What does the ethical theory of utilitarianism say?*

Ethics

1. *Descriptive* claims say something about the way that the world is. They do not make any judgment about whether the way the world is is the way that it ought to be, or whether the way that the world is is a good way for it to be. For instance:

   (a) The Braves won the World Series in 1995.
   (b) It is always sunny in Philadelphia.
   (c) Even the poorest of the poor would have more wealth if we abolished the capital gains tax.

   Note that (b)—and maybe (c), though it's controversial—is false. So a claim need not be true in order to be descriptive.

2. *Normative* claims say something about how the world *ought* to be, or something about which things are *good* or which actions are *right*. They don't merely describe the world—they additionally *evaluate* the world. For instance:

   (a) We should abolish the capital gains tax.
   (b) We shouldn't abolish the capital gains tax.
   (c) We have a moral obligation to not contribute to the suffering of animals.
   (d) Johann shouldn't think that we should abolish the capital gains tax.

   Either (a) or (b) is false, so a claim needn't be true in order to be a normative claim. Note also that a normative claim needn't be a claim about what's *moral*.

3. *Ethics* is the area of Philosophy which studies normative claims: it studies which normative claims are true or false, what it is that makes them true or false, and what normative claims even *mean*.

   (a) Descriptive claims may be relevant to determining which normative claims are true and which are false. But Ethics is not first and foremost concerned with which descriptive claims are true. Its primary goal is to determine which normative claims are true.

4. The field of Ethics subdivides into three different subfields: ethical theory, applied ethics, and metaethics.

   (a) *Ethical Theory* attempts to provide a systematic theory to tell us which normative claims are true and which are false.
   (b) *Applied Ethics* attempts to answer particular ethical questions (with or without the aid of an ethical theory).
   (c) *Metaethics* attempts to answer questions about what we are talking about when we make ethical claims, or what (if anything) makes ethical claims true or false.

Kinds of Ethical Evaluation

5. We may evaluate:

   (a) *States of Affairs* as good or bad (or, in the comparative: better or worse)
   (b) *Actions* as right or wrong
      i. This is partly a stipulative use. As we'll use the terms in this class, it is a category error to say that an action is *good* or *bad*, or that a state of affairs is *right* or *wrong*.

6. It is possible for your evaluations of states of affairs and actions to come apart.

   (a) Suppose that I kill one person to save the lives of five. You could think that a) it's better for one to die and five to live than it is for five to die and one to live; but still b) it was wrong to kill the one person.

   Then again, our evaluations of states of affairs and actions *needn't* come apart in this way.

Consequentialism & Non-Consequentialism

7. An *ethical theory* is a theory which tells you, for some particular kind of normative claim, which normative claims of that kind are true and which are false.

   (a) One kind of ethical theory could tell us which states of affairs are better than which other states of affairs. Such an ethical theory is called an *axiological* theory—or just a 'theory of the good'.

      i. A sample axiology: *Welfarism* says that one state of affairs, S_1, is better than another, S_2, iff people have higher levels of *welfare* (or *well-being*) in S_1 than they do in S_2.

   (b) Another kind of ethical theory could tell us which actions are right and which actions are wrong. Such an ethical theory is called a *deontological* theory—or just a 'theory of the right', or a 'theory of right action'.

8. Theories of right action subdivide into two kinds:

   (a) A *consequentialist* theory of right action claims that whether an action is *right* or *wrong* is determined, in some way or other, by the goodness or badness of certain states of affairs.

      i. The consequentialist thinks that the right is determined (in some way or other) by the good.

      ii. A sample consequentialism: an act is right iff it has better consequences than any other available act.

   (b) A *non-consequentialist* theory of right action claims that whether an action is right is not just determined by the goodness or badness of states of affairs.

      i. The non-consequentialist thinks that the right is not determined by the good.

      ii. A sample non-consequentialism: an act is right iff it is does not violate the ten commandments (this is a form of 'divine command theory').

Utilitarianism

9. Utilitarianism is a consequentialist theory of right action. According to it, an act is right iff there's no alternative which promotes more net, aggregate happiness than it.

   Utilitarianism An act is right iff performing that act results in at least as much net, aggregate happiness as any other alternative act.

10. The Utilitarian accepts three other theses, from which **Utiliarianism** follows: firstly, a claim about *goodness*; secondly, a claim about *welfare*; and thirdly, a claim about *rightness*.

    Welfarism The goodness of a state of affairs is given the aggregate level of welfare in that state of affairs.

    Hedonism A person's level of welfare (or well-being) is their net level of happiness.

    Consequentialism An act is right iff performing that act results in a state of affairs which is at least as good as any other alternative act.

11. What does the utilitarian mean by 'net' and 'aggregate'?

    (a) *Net*: The hedonist thinks that what makes your life go well is happiness. What makes it go badly is unhappiness. Additional happiness does not necessarily make your life go better. If the additional happiness is accompanied with a massive amount of unhappiness, this could make things worse for you. When considering how good things are going for you, we should consider not only how much happiness you have, but also how much unhappiness. Taking their difference gives us your *net* level of happiness.

    (b) *Aggregate*: Your well-being isn't the only thing that matters, ethically. Everyone's well-being is equally important, according to the utilitarian. So when we ask whether an act is right or wrong, we need to consider not only its consequences for *your* net happiness. We should also consider its consequences for *everyone else's* net happiness. Aggregating everyone's net happiness gives us the goodness of the resultant state.

What does the ethical theory of utilitarianism say? Present an argument against utilitarianism. What does Kant's moral theory say? Illustrate Kant's moral theory by giving an example of an act the theory says is wrong, and explain why it says that act is wrong.

Objections to Utilitarianism

1. Recall: the *utilitarian* says that an act is right iff there's no alternative which promotes more net, aggregate happiness than it.

   Utilitarianism An act is right iff performing that act results in at least as much net, aggregate happiness as any other alternative act.

2. We could argue for utilitarianism by appealing to a claim about *well-being* (hedonism), a claim about *goodness* (welfarism), and a claim about *rightness* (consequentialism):

   P1. A state is good *for someone* to the extent that it makes them happy. (*Hedonism*)

   P2. A state is *good* to the extent that it is good *for* everyone. (*Welfarism*)

   P3. An act is right iff no alternative act brings about a better state. (*Consequentialism*)

   C1. An act is right iff no alternative act brings about more overall happiness. (*Utilitarianism*)

3. Consider the following three cases.

   (a) At the local hospital, there are five very sick patients who need organ transplants in order to survive. Tom—who has no relatives or job, and who the doctors know will not be missed—has come in to have his tonsils removed. When Tom is under anesthetic, the doctors painlessly kill him and remove his organs, distributing them to the five sick patients. The patients go on to lead lives which are each just as happy as the life Tom would have led, had the doctors only removed his tonsils.

   (b) Every day, Bill the bully beats up Vince the victim. When Sam learns of this, he intervenes, standing up to the bully and telling him to leave Vince alone. Bill (predictably) beats up Sam instead. Vince was quite used to being beat up by Bill, while Sam is new to the experience, so Sam is made much less happy by the beating than Vince would have been.

   (c) The outcome of the national election will be the same whether Daniel votes or not. Waiting in line makes Daniel unhappy. He'd be happier staying at home. So Daniel stays at home and doesn't vote. He lies to everyone about this, so that nobody else knows that Daniel didn't vote.

4. In the first case, utilitarianism says that the doctors acted rightly. In the second case, utilitarianism says that, by standing up to the Bully, Sam acted wrongly. In the third, utilitarianism says that, by staying home and not voting, Daniel acted rightly.

   (a) Insofar as we find these consequences objectionable, this gives us reason to worry about utilitarianism as a theory of right action.

Kant's Moral Theory

5. Consider again the case of Daniel: what kinds of things might we say to explain why Daniel acted wrongly (if we think he did)? A common refrain is the following: "What if everyone stayed home instead of voting?"

   (a) Note: Daniel can agree that it would be bad if *everyone* stayed home instead of voting. But he knows that that won't happen. And, since no one knows that he didn't vote, he knows that his not voting doesn't make it any more or less likely that others won't vote.

(b) When we point to the possibility of *everyone* not voting, we're not saying that this is likely to come about, nor that Daniel's action might play some role in bringing it about. But we still feel that this possibility can tell us something about how what Daniel has done is wrong.

6. Here's one way of developing this thought:

   Rule Utilitarianism The goodness of a system of rules is given by the amount of net, aggregate happiness that would result from everyone trying to follow those rules. An act is right iff it conforms to the *best* system of rules.

   (a) This is a kind of consequentialism. For, according to this theory, the right is determined by the good.

   (b) If we had a system of rules which permitted doctors to harvest the organs of their patients without their consent, nobody would go to the doctors. This would be worse than a state in which doctors required the consent of their patients in order to take organs. So the best system of rules will say that it's wrong to kill Tom and harvest his organs. (Think about what the theory would say about the other cases).

7. A worry about rule utilitarianism: it collapses back down to regular (act) utilitarianism.

   (a) Consider any system of rules which *always* forbids harvesting organs without the consent of the patients. Take that system of rules and emend it so that it includes the following opt-out clause: if Tom comes in to get his tonsils removed on September 17th, 2018, and there are five sick patients in need of organs, and you are certain that no one will find out, then remove Tom's organs and distribute them to the five.

   (b) If the doctors followed *this* system of rules, thing would be better off. And we can build in similar opt-out clauses for any act which promotes happiness in any particular case. So why doesn't rule utilitarianism just end up 'collapsing' back down to regular (act) utilitarianism?

8. Here's another way of developing the same thought (due to Immanuel Kant):

   Kant's Moral Theory Your act is morally right iff you can consistently will that the maxim on which you act can be universally followed.

   (a) A *maxim* is a general rule on the basis of which you act. Kant thinks that, whenever you act, you have some implicit maxim guiding the action.

   (b) If you can consistently will that everyone acts in accord with your maxim, then your act is morally permissible. If you cannot consistently will that everyone acts in accord with your maxim, then your act is not morally permissible.

   (c) Note an important difference between Kant's moral theory and rule utilitarianism. Both Kant and the rule utilitarian consider a possibility in which everyone follows a rule (or *maxim*). However, when considering that possibility, the rule utilitarian asks: 'how good is it?'. Kant, in contrast, asks: 'could you consistently will that possibility to be actual?'. On Kant's view, the rightness of the action isn't determined by the *goodness* of this possibility. So his theory is non-consequentialist.

9. An example:

   (a) You need money, but you know that you will be unable to repay a loan. Even so, you ask for money, promising to repay it. You do so on the basis of the maxim 'if I need money, then I will make a promise to repay a loan, even if I won't be able to'.

   (b) First, we 'generalize' your maxim, so that it applies not only to you, but to everyone else as well: 'if anyone needs money, then they will make a promise to repay a loan, even if they won't be able to'.

   (c) If this maxim were universally followed, then lenders would stop trusting that their loans will be repaid, and they will stop lending.

   (d) You cannot *consistently will* for this situation to be actual, for two reasons: 1) if there were no lenders, then nobody would be able to follow the maxim. So the situation in which everyone follows the maxim is contradictory. (This is a *contradiction in conception*). Also note that: 2) when you act, you will to obtain money; but, if your maxim were universally followed, you would not obtain money. For this reason, also, you cannot will that your maxim will be universally followed. (This is a *contradiction in will*.)

Does Mill think that we should have the freedom to discuss and defend any opinion, no matter how pernicious? If so, why? If not, why not?

1. Mill's thesis is that it is not permissible to prohibit the expression or defense of *any* opinion, however marginal, however pernicious.

2. Mill's go-to examples of pernicious opinions included:

   (a) that God does not exist;

   (b) that there is no afterlife; and

   (c) that Christian morality is false.

   When Mill was writing, these were regarded as the most pernicious ideas. In our age, we are more likely to hear calls to prohibit opinions like these:

   (a) that the Holocaust did not happen;

   (b) that some social groups (races, genders,...) have a higher moral status, are more intelligent, *etc*. than others; and

   (c) that being gay or trans is wrong, should be illegal, or should be socially stigmatized.

   Mill wants to take on the hard cases, so let's think about these cases when considering his arguments. Let's call all of these views 'heretical', and let's call their denial 'orthodoxy'.

3. There is a natural argument in favor of prohibiting heretical views: we know that they are false, and if they are false, then allowing them to be defended has the potential to lead to great harm.

   P1. We know that the heresy is false.

   P2. If we know that heresy is false, then we know that defending the heresy can lead to great harm.

   C1. We know that defending the heresy can lead to great harm.

   P3. If you know that something can lead to great harm, then it is permissible to prohibit it.

   C2. It is permissible to prohibit the defense of heresy.

4. There may be those who defend the freedom of opinion and discussion by denying that the heresies can harm ('Sticks and stones may break my bones...'), but Mill is not among them. Mill accepts that heresies can be extremely harmful. So he accepts both P1 and P2. What he *denies* is P3. Let's see why.

5. Mill begins by discussing the fallibility of human judgment.

   (a) When it comes to matters like politics, religion, and morality, individuals can be, and often have been, badly mistaken. Not only individuals, but entire communities can be, and often have been, badly mistaken. Not only communities, but entire ages, can be, and often have been, badly mistaken. Even when they were most badly mistaken, past societies were confident that they were right, in part because their opinions were so widely shared.

   (b) Mill: knowing the truth about politics, religion, and morality is hard. Even widespread consensus can be in error.

6. What is the relevance of human fallibility? Mill considers—***but does not endorse***—this argument:

   P4. Our opinions about religion, politics, and morality are fallible.

   C3. We should not prohibit the heresy.

   (a) Mill thinks that this argument is a bad one. After all, our opinions about *economics* and *justice* are equally fallible. But nobody would accept this as a reason to not pass or enforce laws against murder.

(b) In the past, society passed unjust laws—*e.g.*, laws against homosexuality—but this is no reason to think that we shouldn't outlaw rape or murder.

(c) Our judgment is fallible, but we still must rely upon it in order to decide how to act. So this argument is not compelling.

7. The point of bringing up the fallibility of human judgment is this: Mill thinks that we are fallible enough that, unless we are given an opportunity to consider the arguments against our opinions, we will not know that they are true.

   (a) On Mill's view, for society to have knowledge of an opinion about politics, religion, or morality, there must be some possibility of that opinion being refuted.

   (b) Compare: if you are to know that a scientific theory is true, then you must have performed an experiment which had the possibility of refuting the theory.

   (c) Mill: if an opinion cannot be gainsaid, then there is no possibility of that opinion being refuted.

8. The argument Mill favors is this one:

   P5. Orthodox opinions about politics, religion, or morality can be known to be correct only if it is possible for the orthodoxy to be refuted.

   P6. If heretical opinions are silenced, then it will not be possible for the orthodoxy to be refuted.

   C4. If heretical opinions are silenced, then we cannot know that the orthodoxy is correct.

   P7. It is permissible to silence heretical opinions only if we know that the orthodoxy is correct.

   C5. It is never permissible to silence heretical opinions.

9. Returning to the natural argument in favor of prohibiting the defense of heresy: Mill rejects P3,

   P3. If you know that something can lead to great harm, then it is permissible to prohibit it.

   (a) Note that Mill can accept a nearby variant of P3, P3*,

   P3*. If you know that something can lead to great harm—and you can retain this knowledge after prohibiting it—then it is permissible to prohibit it.

   However, with P3*, the original argument against prohibiting defense of the heresy fails. For, on Mill's view, as soon as you prohibit the defense of the heresy, you undermine your knowledge of its falsity, and thereby, you undermine your knowledge that its defense will lead to great harm.

Why does Stanley Fish think that 'there's no such thing as free speech'? How do his views about the distinction between speech and action support this opinion?

Speech versus action

1. At the start of chapter 3 of *On Liberty*, Mill makes an interesting claim. He says that "No one pretends that actions should be as free as opinions. On the contrary, even speech loses its immunity, when [it constitutes] a positive instigation to some mischievous act."

   (a) His example: society may punish somebody for yelling 'corn dealers are starvers of the poor' to an angry mob assembled outside of a corn dealer's home.

   (b) Mill denies that we may similarly punish somebody for writing an article saying that corn dealers are starvers of the poor.

   (c) But couldn't the publication of the article have precisely the same effect? What is the difference supposed to be?

   (d) Mill thinks that yelling 'corn dealers are starvers of the poor' to an assembled mob constitutes a kind of *action*; whereas writing an article does not.

2. Some exceptions to the absolute freedom of speech also appear in American first amendment jurisprudence. The freedom of speech is not taken to include state secrets, defamation, fighting words, obscenity, incitement, and speech which constitutes 'clear and present danger'.

3. Stanley Fish: any defender of free speech will have to draw a distinction between *speech*, on the one hand, and *action*, on the other. But this cannot be done. An action is something you do which has an effect upon the world. But everything you say or write has an effect upon the world. So there is no distinction between speech and action; whatever distinction we attempt to draw between the two will be artificial.

4. If there's no natural dividing line between speech and action—if the line is drawn by us arbitrarily—then the claim that speech should be free is empty.

   (a) What we choose to call 'speech', as opposed to action, are just the expressions which we want to protect.

5. Since there is no real distinction between speech and action, where we choose to draw this line will be based upon our own ideological prejudices.

   (a) If we say that state secrets, defamation, and fighting words are actions, but that racism slurs are speech, that's because we take the consequences of the first kinds of utterances more seriously than we take the consequences of the second.

6. On Mill's view, the freedom of opinion and discussion was a *right*. We were not to decide whether to allow somebody to speak by considering the consequences of them expressing their opinion and weighing it against the consequences of silencing them.

   (a) But, if we are forced to appeal to our ideological prejudices in distinguishing speech from action, Fish thinks that we must have a different understanding of what's involved in protecting free speech.

   (b) On this understanding, free speech is a value like any other. It is important to protect it within limits. But sometimes this value comes into conflict with other values—like, e.g., creating a society in which everyone is treated as equals. When it does, we must balance the value of free speech against these other values. And free speech is not *infinitely* valuable. There will always come a point at which the value of free speech gets trumped.

7. Fish presents the following argument against any form of free speech absolutism:

P1. Either free expression is valuable for its own sake, or else it is only valuable because it helps us achieve something else which is valuable.

P2. Free expression is not valuable for its own sake.

C1. Free expression is only valuable because it helps us achieve something else which is valuable.

P3. If free expression is valuable for the sake of some other end, then there are situations in which free express works against, rather than for, that other end.

C2. There are situations in which free expression works against, rather than for, the end towards which it is directed.

P4. If free expression works against the end towards which it is directed, then it should not be permitted.

C2. There are situations in which we should not permit the free expression of ideas.

Hate speech and the value of knowledge

8. Let's think through how to apply this argument against the value towards which Mill thought free expression was directed: the value of *knowledge*. Mari Matsuda gives some real-world examples of hate speech:

> *A black family enters a coffee shop in a small Texas town. A white man places a card on their table. The card reads "You have just been paid a visit by the Ku Klux Klan."*
>
> *A law student goes to her dorm and finds an anonymous message posted on the door, a caricature image of her race, with a red line slashed through it.*
>
> *A Hmong family in Eureka, California, was twice victimized by four-foot-high crosses burning on their lawn.*

9. Mill worried that, by silencing the voices of heretics, we would lose the ability to listen to what they had to say, and thereby lose our knowledge that what they say is false. And Mill worried not just about *government* silencing, but additionally silencing by society at large through boycotts and social stigma.

   (a) But if we should worry about boycotts and social stigma silencing voices, then surely we should also be worried about the hate speech above silencing the voices of the black family, the law student, and the Hmong family.

   (b) By silencing those voices, we lose whatever knowledge we might have gained from their opinions (in particular, their opinions about race in our country).

10. Matsuda has a suggestion: we should allow ourselves to proscribe speech which meets the following three criteria:

    1. It is a claim of racial inferiority
    2. It is directed against a historically oppressed group
    3. It is persecutory, hateful, and degrading

    (a) So, we should permit ourselves to proscribe signs which convey racial inferiority, including swastikas and Nazi regalia.

    (b) The third criteria exempts 'cool' social scientific research—even social scientific research which supports messages of racial inferiority.

11. Because the proposal allows 'cool' research, Matsuda believes that it allows us to retain our knowledge—it allows the heresy to be heard, albeit in a calm tone of voice. However, because it rules out hateful and degrading hate speech which could silence the targeted groups, it additionally allows us to learn from their perspective.

What is the distinction between a locutionary act*, an* illocutionary act*, and a* perlocutionary act*? Provide examples to illustrate the distinction. How could this distinction allow a free speech advocate to distinguish between speech and action in a non-ideological way?*

Locution, Illocution, and Perlocution

1. Let's take a detour through the philosophy of language. In the 1960's, J. L. Austin was thinking, not about ethics, but rather about what we do when we utter sentences. Consider the following utterances:

   (a) "I promise to repay you", said before you loan me money.

   (b) "I do", said while standing on an alter after being asked 'Do you take this woman...?'

   (c) "Nice family you got there; it'd be a shame if anything happened to them", said to someone associated with the mafia being interviewed by the police.

2. When I say (1a), one thing I do is *describe* the world as being a certain way. I describe myself as making a promise. But that's not *all* that I do. I additionally *make a promise*. That is, by uttering "I promise", I make my description true. The utterance itself *constituted* the promise.

   (a) Notice: I could hardly try to get out of paying you back by claiming that I was lying when I said "I promise". I *couldn't* have lied, or said something false. The claim "I promise" is automatically true when it's made in this way.

   Beyond just describing myself as making a promise, and beyond just making the promise, I also do something else: I *persuade* you to hand me the money—I cause you to hand the money over.

3. When I say (1b), I similarly *describe* the world. I affirm a certain proposition: namely, the proposition that I marry this woman. But that's not *all* that I do: I additionally (and *thereby*) marry her. My utterance constituted the act of marrying.

   (a) Notice: I could hardly protest in court that we were not married on the grounds that I was lying on the alter. I *couldn't* have lied, or said something false. The claim "I do" is automatically true when it's made in this context.

   I don't only describe myself as marrying, and I don't only marry. I also do something else: I *cause* the priest to continue the ceremony.

4. When I say (1c), I make a true claim. The family *is* nice. And it truly *would* be a shame if anything happened to them. But that's not all that I do. I additionally *threaten* the listener. My making the claim itself constituted a threat. Moreover, in addition to saying something true about how nice the family is, and in addition to threatening the listener, I do something more: I *cause* the listener to not squeal.

5. Austin distinguishes three kinds of *speech acts*—three kinds of things we do when we speak.

   Locutionary Act A *locutionary speech act* is an act of asserting or expressing a proposition. In saying "I promise", I perform the locutionary act of expressing the proposition that I make a promise.

   (a) Note: this same locutionary act could be performed by somebody else who said "Dmitri promised".

   Illocutionary Act An *illocutionary speech act* is any action your utterance constitutes beyond the expression of a proposition. In saying "I do", I performed the *illocutionary act* of marrying somebody.

   (a) Note: this same illocutionary act would *not* be performed by somebody else who said "Dmitri married".

   Perlocutionary Act A *perlocutionary speech act* is anything that you do with your utterance, where the act is not constituted by your utterance, but is instead *causally downstream* of your utterance. By saying "Shame if anything happened to them", I performed the *perlocutionary act* of shutting the witness up.

   (a) Note: this same perlocutionary act *need not* be performed by somebody else who says "Shame if anything happened to them" in some other context.

Speech versus action, again

6. Recall Fish's contention: there is no natural distinction between speech and action. Every time we speak, we impact the world in some way. There is only a distinction between utterances whose impact on the world we take to be consequential enough to be regulated and those whose impact on the world is not.

   (a) It was on this basis that Fish concluded the claim 'Speech should be free' was empty (*of course, if you thought it shouldn't be free, you wouldn't have called it 'speech'*).

   (b) It was also on this basis that Fish concluded that, however we distinguish speech from action, this will be informed by our own ideological prejudices.

7. Austin's distinctions between different kinds of speech acts gives free speech advocates the resources to respond to this challenge:

   (a) A free speech principle does not forbid punishing *illocutionary* acts and *perlocutionary* acts. What it forbids is the regulation of *locutionary* acts.

   (b) Call this the doctrine of *content neutrality*: we should allow people to express any proposition—though, in some contexts, we may forbid them from expressing a proposition when to do so would be to perform some bad illocutionary or perlocutionary act.

   (c) The state can forbid marrying more than one person (*e.g.*) without violating a principle of free speech—even though the act of marrying the second person is carried out in speech. What's targeted is not the locutionary act, but rather the illocutionary act.

      i. The law still allows the reporter to perform the locutionary act when they report on the bigamy.

   (d) Similarly, the state can forbid 'witness tampering' without violating a principle of free speech—even though the act of tampering with the witness is carried out in speech. What's targeted is not the locutionary act, but rather the *perlocutionary* act of causing the witness to withhold testimony.

      i. The law still allows people *in other contexts* to say that it would be a shame if something happened to the witness's family.

8. Austin's distinction allows Mill to explain the difference between yelling "Corn-dealers are starvers of the poor!" to an assembled crowd and writing the same sentence in an article.

   (a) The first utterance constitutes an *incitement to violence*; whereas the second does not.

9. Something to consider: the doctrine of content neutrality requires that we allow every propositions to be expressed in *some* context, but what if there are propositions which, *e.g.*, constitute incitements to violence—or incitements to hatred—no matter when they are expressed?

1. Distinguish between a *sentence* and a *proposition*. Consider the following three sentences:

   La nieve es blanca

   Schnee ist weiss

   Snow is white

   These are three different *sentences*. However, they all *mean the same thing*. They all *express the same proposition*.

2. Just as we can have two sentences express the same proposition, we can have one sentence express two different propositions (in two different contexts of utterance). Thus, if I say 'I am hungry', I say something different from what Sabrina says if she says 'I am hungry'.

3. Consider two utterances of the same sentence: "North Korea should allow its citizens to criticize the government"—one made by an American and one made by a North Korean.

4. According to cultural relativism, if we want to evaluate the truth of what these people say, we must first look to cultural norms. But there are (at least) three different ways that the cultural norms could affect the truth of what these people say.

   Truth Relativism Both the American and the North Korean express the very same proposition. In America, that proposition is true. In North Korea, that proposition is false.

   Meaning Relativism (speaker norms determine meaning) The American and the North Korean express two *different* propositions. The sentence "North Korea should allow its citizens to criticize the government" is like the sentence "I am hungry". What it means depends upon who says it.

   (a) The sentence "North Korea should allow its citizens to criticize the government" is incomplete. In order to express a complete proposition, we must supply a *moral standard*. The correct moral standard to supply is the moral standard of the speaker's culture.

   Since American culture values freedom of speech, what the American says is true. Since North Korean culture does not, what the North Korean says is false.

   Meaning Relativism (actor norms determine meaning) The sentence "North Korea should allow its citizens to criticize the government" is incomplete. In order to express a complete proposition, we must supply a *moral standard*. The correct moral standard to supply is the moral standard of the culture of the *actor* whose actions are being evaluated.

   So, both the American and the North Korean express the very same proposition: namely, that North Korea should allow criticism of the government according to the North Korean norms. Since this proposition is false, both the American and the North Korean speak falsely.

What is Cultural Relativism? What is the Cultural Differences Argument for Cultural Relativism? Why does Rachels think that this argument is bad?

1. Ruth Benedict:

> "*According to the Kwakiutl it did not matter whether a relative had died in bed of disease or by the hand of the enemy; in either case death was an affront to be wiped out by the death of another person. A chief's sister and her daughter [died]...[so his tribe] set out, and found seven men and two children asleep and killed them. 'They felt good when they arrived at Sebaa in the evening.'*
>
> "*The point which is of interest to us is that in our society those who on that occasion would feel good when they arrived at Sebaa that evening would be the definitely abnormal...[whereas] on the Northwest Coast those are favored and fortunate to whom that mood under those circumstances is congenial...*" (p. 480–81)
>
> "*We do not any longer make the mistake of deriving the morality of our locality and decade directly from the inevitable constitution of human nature. We do not elevate it to the dignity of a first principle. We recognize that morality differs in every society, and is a convenient term for socially approved habits. Mankind [sic] has always preferred to say, 'It is morally good,' rather than 'It is habitual,'...But historically the two phrases are synonymous.*" (p. 482)

2. According to Benedict, 'X is morally good' *just means* 'X is habitual', or 'X is culturally approved'. Since different cultures approve of different activities, there is no absolute, culture-independent, fact-of-the-matter of what is right and wrong.

   (a) Think about a claim like 'Suits are professional attire for men'. In some cultures throughout history, male professionals have worn suits, and in some cultures throughout history, they have worn other garments. There is no culture-independent fact-of-the-matter about whether suits are professional attire for men.

   (b) What we are *really* saying, when we say 'Suits are professional attire for men', is that suits are professional attire for men *in our culture*.

   (c) So the sentence 'Suits are professional attire' does not yet express a complete thought. Like the sentence 'It is raining', the sentence is incomplete. In order to know whether the sentence says something true or false, we must provide a *culture*. Relative to a given culture, then, we may say whether suits are professional attire for men in that culture. (Just as, relative to a given place and time, we can say whether it is raining at that place and at that time.)

   (d) Benedict thinks that sentences like 'The killing of innocents is wrong' are also culture-relative in precisely the same way. This sentence doesn't yet express a complete thought. In order for us to know whether the sentence says something true or false, we have to know which *culture* it is talking about.

3. Call this thesis *Cultural Relativism*.

   Cultural Relativism There is no absolute fact-of-the-matter about which actions are right and which are wrong. Moral claims are only true or false *relative to* a certain culture.

4. In support of this view, Benedict offers us examples of moral codes varying from society to society. With respect to the Kwakiutl, her argument goes something like this:

   P1. The Kwakiutl regard the killing of innocents as a morally permissible method of grieving.

   P2. Americans do not regard the killing of innocents as a morally permissible method of grieving.

   C1. There is no absolute, culture-independent, fact-of-the-matter about whether or not the killing of innocents is a morally permissible method of grieving.

5. More generally, we have what Rachels calls the *Cultural Differences Argument*:

   P4. Different cultures have different moral codes.

   C3. There is no absolute truth in morality. Moral claims are only true or false *relative to* a particular culture.

(a) Rachels: this isn't a good argument. The conclusion does not follow from this premise. To see this, note that the argument has the following form:

P5. Cultures disagree about X

C4. There is no absolute fact-of-the-matter about X. Claims about X are only true or false *relative to* a particular culture.

(b) And this is an *invalid* argument form. That is: it is possible for a premise like this to be true while the conclusion is false. To see this, consider if we let $X =$ the earth is round.

P5. Cultures have disagreed about whether the earth is round.

C4. There is no absolute fact-of-the-matter about whether the earth is round. Claims about whether the earth is round are only true or false *relative to* a particular culture.

(c) Rachels: just because different cultures disagree, that doesn't mean that neither culture is *wrong*.

(d) There is an absolute fact-of-the-matter about the shape of the earth, but it doesn't follow that we should expect everybody to *know* it. And, similarly, there is no reason to think that, if there were an absolute fact-of-the-matter about morality, everyone would know it.

(e) Additionally, we could explain cultural disagreement without even supposing that different cultures have different values. We can explain much of the disagreement by pointing to 1) disagreement about non-normative matters; and 2) different circumstances and environments.

i. As an example of (1): disagreement about whether killing cows is wrong could stem from a factual disagreement about whether we are reincarnated as cows—or as factual disagreement about whether cows suffer.

ii. As an example of (2): in a society in which caring for children can be prohibitively difficult, infanticide may be permissible; whereas, in a society in which caring for children is not so difficult, infanticide is not permissible.

6. Rachels: if we were to accept Cultural Relativism, then:

(a) We could no longer say that the customs of some cultures are morally inferior to others. For instance: we could not say that a culture which oppresses women and homosexuals is worse than a culture which does not.

i. We could say: *according to this culture's moral norms*, oppressing women and homosexuals is wrong. But we would also have to say that, *according to this other culture's norms*, oppressing women and homosexuals is permissible. And we could not say that the first culture's norms are better than the second's.

(b) We could decide whether actions are right or wrong just by consulting the standards of our own society.

(c) It becomes difficult to understand the idea of *moral progress*.

i. We may say: according to the moral codes of the antebellum south, slavery was permissible.

ii. And we may say: according to our current moral codes, slavery is not permissible.

iii. But we can't say: our current moral codes are *better* than the moral codes of the antebellum south.

7. Rachels: these consequences of Cultural Relativism are absurd. They give us sufficient reason to reject it.

(a) Of course, that doesn't mean that we shouldn't be *tolerant* of other cultures. We can think that cultural norms which are tolerant of other cultural norms are *better* than those which are not. Intolerance doesn't follow from rejecting relativism.

8. Rachels: Cultural Relativism still has some important insights. For instance,

(a) It correctly reminds us that many of our own practices are simply matters of convention—*e.g.*, whether we honor our dead by burning, burying, or eating them. This is not inevitable or based on any absolute, culture-independent standard.

i. But: just because *some* of our practices are like this, it doesn't follow that *all* of them are.

(b) By emphasizing the extent to which our ethical views can depend upon the prejudices of our own culture, it calls our attention to the fallibility of our own ethical judgments, and encourages us to keep an open mind.

Explain Harman's moral relativism. Describe one of Rachels' criticisms of relativism and say how Harman responds to this criticism.

1. Harman draws an analogy between Einstein's special theory of relativity (STR) and moral relativism.

   (a) According to the STR, claims like 'the stick is one meter long' are incomplete. In order to get a proposition—in order to get something which is capable of being true or false—we must specify a frame of reference, R.

      i. Relative to a frame of reference in which the stick is at rest, it is true that the stick is one meter long. But, relative to a reference frame in which the stick is moving, it could be false that the stick is one meter long. (Similarly, relative to one reference frame, the stick could have one mass; and, relative to another reference frame, it could have another. And, relative to one reference frame, my dropping the stick could be simultaneous with your clapping your hands; while, relative to another frame of reference, they are not simultaneous.)

      ii. No frame of reference is privileged. So there's no sense in which we can say that the stick is *really, absolutely* one meter long.

   (b) Similarly, Harman says: according to moral relativism, claims like 'abortion is permissible in cases of rape' are incomplete. In order to get a proposition—something which is capable of being true or false—we must specify a moral framework.

      i. Relative to one moral framework, this claim is true. Relative to another, the claim is false.

      ii. No moral framework is privileged. So there's no sense in which we can say that abortion is *really, absolutely* permissible in cases of rape.

2. Harman isn't making a claim about what people *intend* their sentences to mean. He isn't saying that, when I say 'Abortion is permissible in cases of rape', I am *trying to say* that abortion is permissible in cases of rape according to my moral framework.

   (a) Compare:

      i. Einstein isn't making a claim about *meaning*. The STR doesn't say that, when people say 'The stick is one meter long', they are *trying to say* that it is one meter long in *their* reference frame.

      ii. What the STR says is that, when it comes to an object's length, there simply is no reference-frame-independent fact-of-the-matter about length, mass, or simultaneity.

      iii. But people often talk about objects' lengths and masses. They often talk about which events happened at the same time as others, even when those events were far apart.

      iv. We face a choice, then: we could either say that their talk about length, mass, and simultaneity is all in error—it is like talk about witches. Or we could say that their talk about length, mass, and simultaneity should be understood as talking about length, mass, and simultaneity *in their own reference frame*.

   (b) Similarly, then:

      i. The moral relativist doesn't say that, when people say 'Abortion is permissible in cases of rape', they are *trying to say* something about their own moral framework.

      ii. However, when it comes to claims about morality, there simply isn't any moral framework-independent fact-of-the-matter.

      iii. We face a choice: we could either say that people's talk about rightness and wrongness, goodness and badness, *etc.* is all in error—that it is like talk about witches. Or we could say that this talk should be understood as talk about rightness and wrongness, goodness and badness, *etc. in their own moral framework*.

3. Note an important difference between the cultural relativist position we considered last time and Harman's form of moral relativism:

   (a) When we considered a moral claim like 'Female genital mutilation in Djibouti is wrong', the cultural relativist wanted to say that this claim was false. According to them, the correct moral standard to use when evaluating this claim is the moral standard of Djibouti culture. That is, the cultural relativist says that, for the purposes of determining whether the claim is true or false, we should evaluate the claim relative to the moral framework of the person whose actions are being discussed. Then, the claim becomes:

FGM in Djibouti is wrong according to Djibouti norms

which is false. However, Harman thinks that the correct moral standards to be using are the moral standards of *the speaker*. Then, the claim becomes:

FGM in Djibouti is wrong according to my norms

which is true.

4. Harman argues for moral relativism by pointing to the prevalence of intractable moral disagreement. Last class, Rachels pointed out that the existence of disagreement doesn't necessarily mean that this disagreement is *faultless*. Harman responds: 'No, it doesn't. But I am not claiming that the disagreement *necessarily entails* that there is no moral-framework-independent fact-of-the-matter. Rather, I am making an *inference to the best explanation*.' Here is Harman's argument, applied to the case of disagreement about moral vegetarianism:

P1. Disagreements between moral vegetarians and nonvegetarians can survive full discussion and full information.

C1. If there is a non-relative truth concerning the moral importance of animals, then it cannot be discovered. [from P1]

P2. As a general methodological principle, it is better to not suppose that there are undiscoverable truths.

C2. It is better to not suppose that there is a truth about the moral importance of animals. [from C1 and P2]

5. Rachels: if moral relativism is true, then we cannot say that the norms of some cultures are morally inferior to others. For instance, we could not say that a culture which oppresses women and homosexuals is worse than one which does not. This is implausible, so relativism is false.

   (a) Harman: nothing in moral relativism says that we cannot criticize others.

   (b) Just because people have different moral frameworks doesn't mean that they can't *disagree*.

      i. Compare: you and I have different plans: I want to go to the movies, while you want to go to the beach. We can disagree about where to go, and I can criticize your choice, even if there's no plan-independent fact about where to go. In order to settle where to go, we will have to *bargain*. And this process of bargaining can involve criticism.

   (c) Similarly, if you and I have different moral frameworks, we can disagree about how to treat women. In order to settle this disagreement, we will have to *bargain*. And this bargaining can involve criticism.

6. Rachels: if moral relativism is true, then we could decide whether actions are right or wrong just by consulting our own moral framework. This is implausible, so relativism is false.

   (a) Harman: moral frameworks are like legal frameworks. Existing legislation and case law need not settle which acts are legal and illegal; there could be disagreements between laws; some laws may be unconstitutional. Sorting all this out will take time and thought.

   (b) Similarly, your current moral views need not settle which acts are right and wrong. Some of your moral views could be contradictory. Some may be inconsistent with more fundamental values. Sorting all this out will take time and thought.

   (c) So, we shouldn't confuse your moral framework with your moral beliefs. Rather, your moral framework is the set of moral beliefs you *would* have, if you *were* to be aware of all the facts, and if you *were* to rationally revise your views so as to make them consistent.

Explain the 'conflict' in our ordinary ways of thinking about freedom, moral responsibility, and scientific inquiry. List and explain the possible ways of resolving this conflict.

1. Sider: there is a conflict between our ordinary way of thinking about moral responsibility and our ordinary ways of thinking about scientific inquiry. The conflict:

   (a) We tend to think that people are *morally responsible* for an action only if they chose that action *freely*.

      i. Say that you are *morally responsible* for an action when it is appropriate to *blame* or *praise* you for the performance of that action.

      ii. Then, suppose that you were hypnotized and given the order to donate money to charity. We do not think that it is appropriate to praise you for this action. Why? Because you didn't donate the money *freely*—you did it while under the control of the hypnotist.

      iii. Alternatively, suppose that a gun is held to your head and you are commanded to lie on your taxes. Again, we do not think that it is appropriate to blame you for this action. Why? Because you didn't lie on the taxes *freely*—you did it under duress.

   (b) We tend to think that an act is unfree when it is caused by things external to us that we are not free to control.

      i. For instance, the hypnotist's orders and the gun to your head both caused you to act as you did, but these causes are external to you and not under your control.

   (c) We tend to suppose in scientific inquiry that *every* event has a cause—and, in particular, that our own decisions have causes.

2. These thoughts conflict with each other. Take some arbitrary act of yours, A.

   P1. Your doing A was a necessary consequence of the conditions prevailing shortly after the big bang, together with the laws of nature. (Determinism)

   P2. You are not free to change the conditions prevailing shortly after the big bang and the laws of nature.

   P3. For any X, if you are not free to change X, then you are not free to change any necessary consequence of X.

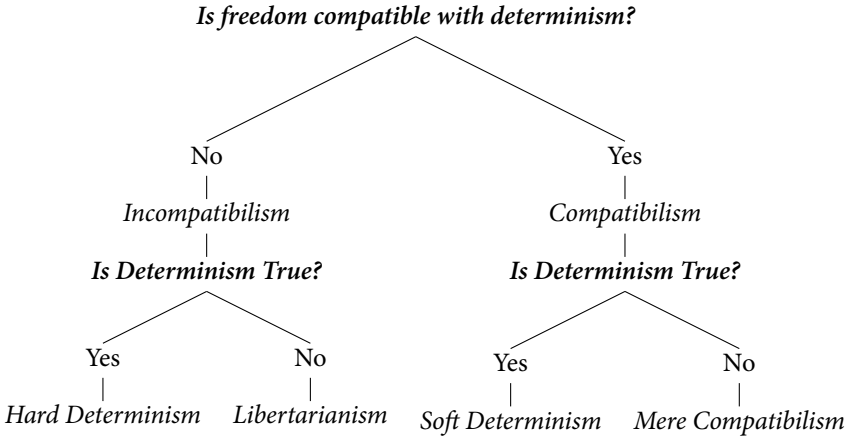
   C1. You are not free to not do anything other than A. [from P1, P2, and P3]

   P4. You are only morally responsible for an action if were free to do otherwise.

   C2. You are not morally responsible for doing A. [from C1 ad P4]

3. The conclusion C1 tells us that no act of yours (or anybody else's) is every chosen freely. The conclusion C2 tells us that you are not morally responsible for any of your actions (and neither is anybody else).

4. A taxonomy of responses to the first argument (for C1):



5. The *libertarian* rejects the premise P1 (Determinism). They deny that our actions are caused in the way that other physical processes are caused.

   (a) A challenge for the libertarian: their position looks to be in conflict with modern physics. Moreover, the libertarian says that psychology and neuroscience will never find psychological/neurological causes of our decisions. But this seems premature—how could we discover this just from the armchair?

   (b) Another challenge for the libertarian: being *uncaused* doesn't seem to make actions free. Being uncaused makes actions *inexplicable*, not freely chosen.

      i. Suppose Mother Teresa—for *no reason*, without *any cause*—pulls the pin from a hand grenade and throws it into an orphanage. This behavior is not caused by any intentions or desires of Mother Teresa (she is just as horrified to watch her arms behave in this way as anyone else). It wasn't caused by any external compulsion. It was just *uncaused*.

      ii. In this case, it doesn't look as though Mother Teresa is to *blame* for throwing the grenade. Moreover, it doesn't seem that she was acting *freely* in throwing the grenade.

      iii. Lesson: freedom seems to *require* causation.

   (c) A Libertarian response to the 2nd challenge: for an act to be free is for it to be caused (in the right way) by the *agent themselves*. This kind of causation, *agent causation*, is distinct from the kind of causation found in physics. And it is the kind required for freedom. Though our acts are caused by *us*, they are not caused by factors outside of our control. So P1 is still false.

6. The *hard determinist* accepts the conclusion C1. They say that science has taught us that none of our actions are freely chosen.

   (a) A challenge for the hard determinist: what should we do about our practices of praise and blame?

      i. Presumably, we shouldn't let people off the hook for terrible actions, even though they were determined to take those actions. Similarly, we don't want to stop praising people when they do praiseworthy things. But how do we avoid this, if we are hard determinists?

   (b) The hard determinist could just bite the bullet. They could say: no one is ever blameworthy, and no one is ever praiseworthy for anything they ever do. That's a hard attitude to maintain.

   (c) Another response: find some *surrogate* notion—call it 'schreedom'—which plays the role we once thought freedom played: schreedom is necessary for moral responsibility, *e.g.*, and schreedom is inconsistent with certain kinds of compulsion.

7. The *soft determinist* thinks that the hard determinist's 'schreedom' just is freedom. They deny P3. They think that you are free to not take some action, A, even though you are not free to change things which necessitate that you take the action A.

   (a) A challenge for the soft determinist: how could P3 possibly be false?

   (b) The soft determinist says you are free to do otherwise when your actions are caused *in the right way*. If so, then they owe us an explanation of what *the right way* is.

      i. Proposal #1: An act is free if it is caused *by your beliefs and desires*.

      ii. Objection:

         A. The addict smokes cigarettes because of their desire to do so. But still, they are not free—their smoking is a *compulsion*.

         B. Phineas Gage's actions were caused by his beliefs and desires, but because of the spike in his brain, those desires do not appear to be *his own*. For this reason, his choices do not appear to be free.

      iii. Proposal #2: An act is free if it is caused by beliefs and desires which were freely chose.

      iv. Objection: this proposal is circular. We might just as well have said: 'An act is freely chosen iff it is freely chosen.'

      v. Proposal #3: A *first-order desire* is a desire which does not have another desire as its object (*e.g.*, I want cigarettes, or I want ice cream). A *second-order desire* is a desire which has a first-order desire as its object (*e.g.*, I don't want to want cigarettes, or I want to want ice cream). Then, say that an action is freely chosen if it is caused by a first-order desire which is itself second-order desired.

If we are determined to act as we do, can we be held morally responsible for our actions?

What does the 'principle of alternative possibilities' (PAP) say about moral responsibility? Why does Frankfurt think that what the PAP says about moral responsibility is false?

1. An argument that *Determinism* undermines freedom (*i.e.*, an argument for incompatibilism):

   P1. Your act is freely chosen only if you could have done otherwise.

   P2. If *Determinism* is true, then you could not have done otherwise.

   C1. If *Determinism* is true, then your act is not freely chosen.

2. An argument that *Determinism* undermines moral responsibility:

   P3. You are morally responsible only if you could have done otherwise.

   P4. If *Determinism* is true, then you could not have done otherwise.

   C2. If *Determinism* is true, then you are not morally responsible.

3. P1 and P3 say that, in order for us to be free/morally responsible for what we've done, we must have been able to do something else. Call this the *Principle of Alternative Possibilities* (PAP). The PAP comes in two flavors, depending upon whether we are talking about freedom, on the one hand, or moral responsibility, on the other.

   PAP$_F$ In order for you to freely choose to do something, it must be possible for you to not do that thing.

   PAP$_M$ In order for you to be morally responsible for something you've done, it must be possible for you to not do that thing.

   (a) In support of the PAP: in paradigm cases of coercion, it appears to be precisely because the coercion deprives you of alternatives that the coercion undermines your freedom.

      i. The hypnotist deprives you of the ability to do otherwise; the gun to your head deprives you of the ability to do otherwise (without incurring great personal harm), *etc.* In these cases, you are not free; and you are not blameworthy or praiseworthy.

   (b) In further support of the PAP$_M$, consider the following case (due to Peter van Inwagen):

      All Roads Lead to Rome A man, Ryder, directs his horse along a series of forking roads. At each fork, he directs the horse either to the left or the right. The man arrives in Rome—but this is unsurprising, since every single road leads to Rome. No matter how Ryder directed his horse, he would have arrived in Rome.

      Ryder could not have failed to arrive in Rome. In that case, Ryder is not *praiseworthy* for having arrived in Rome. It's natural to think that the *reason* he's not praiseworthy for arriving in Rome is just that there wasn't ever any *possibility* that he would arrive in Rome.

4. Frankfurt argues against PAP$_M$. He gives what he takes to be a *counterexample* to this claim. That is, he gives a case in which he thinks that someone could not have done otherwise, yet they are still morally responsible for what they've done. Here's the case:

   Dr. Black Dr. Black is a brilliant but evil neuroscientist. He has designed a device which, when implanted in a person's brain, is capable of both monitoring and intervening upon the person's brain, effectively controlling that person like a puppet. Dr. Black wants Jones to kill Smith. However, Black would rather not get his hands dirty if he doesn't have to. So he implants his device in Jones' brain, and forms the following plan: he will stand back and watch Jones. If Jones forms the intention to kill Smith, then Black will stand down and allow him to do so. If, however, Jones does not form the intention to kill Smith, then Black will use his device to have Jones kill Smith. As it happens, Jones forms the intention to kill Smith on his own, Black never has to intervene, and Jones kills Smith.

5. Frankfurt says: in this story:

   (a) Jones is morally responsible for killing Smith; yet

   (b) It was not possible for Jones to not kill Smith.

   So: PAP_M is false.

6. We could use this conclusion to argue against PAP_F. Consider the following argument:

   P5. You are morally responsible for an action only if that act was freely chosen.

   P6. Jones is morally responsible for killing Smith.

   C3. Jones' killing of Smith was freely chosen.

   P7. It was not possible for Jones to not kill Smith.

   C4. PAP_F is false.

   (a) Even if we reject P5, we may think that C3 is plausible *on its own*, and we may think that, for this reason, we should reject the PAP_F.

7. Rejecting the PAP, in both of its flavors, means rejecting the arguments for incompatibilism we considered in points 1 and 2 above.

8. Recall that, last class, we considered an argument for this conclusion:

   > You are not free to do anything other than A.

   And we were blithely supposing that, if you're not free to do anything *other* than A, then A could not have been performed freely. But Frankfurt may have also given us reason to doubt *this*. For perhaps we should say:

   (a) Jones was not free to do anything other than kill Smith; yet

   (b) Jones killed Smith freely.

Describe the view which Sider calls 'the Space-Time Theory'.

1. Today we're going to be considering the common-sensical idea that time flows forward. More carefully,

   The Moving Present The present moment ('now') moves from the past into the future. Only the present exists.

   (a) This thesis is incredibly intuitive. However, today we're going to encounter some reasons to think that it is nevertheless false.

2. To think more carefully about **The Moving Present**, let's begin to think about motion in general.

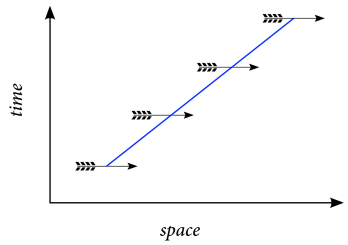
3. Zeno of Elea (5th century BC) thought that motion was impossible. He provided several arguments for this surprising conclusion. Let's consider just one of them:

   (a) Consider something that you think moves: an arrow in flight, for instance. If the arrow moves, then it must move at some time. Time consists of nothing more than a collection of instants. So, if the arrow moves, then it must move at some instant. But, at any instant, the arrow is motionless. So the arrow does not move at any instant. And if the arrow does not move at any instant, then the arrow *never* moves.

4. This argument has a false conclusion—that much is certain. Things do move. I suggest: we should reject the premise that, if an arrow is ever in motion, then it is in motion *at some instant*. To be in motion is not to be *instantaneously* in motion. Rather, what it is for the arrow to move is for it to be at one place at one time and to be at another place at another time. This is sometimes called the 'at-at' theory of motion.

   At-At Theory of Motion What it is for something to move is just for it to be *at* one place earlier and for it to be *at* a different place later. What it is for something to be in motion *at an instant* is just for it to be ***at*** different places ***at*** nearby earlier and later times.

5. We can visualize what the at-at theory says with the aid of a *space-time* diagram. We can visualize the motion of the arrow with by showing that, at earlier times, it's *at* one location; while, at later times, it's *at* a different location:



6. But now let us return to **The Moving Present**. If we understand all motion in terms of the at-at theory, then what do we *mean* when we say that the present moment *moves* from the past into the future?

   (a) Presumably: At one time, *now* is at Sunday. At another time, *now* is at Monday. At a later time, *now* is Tuesday, and so on.

   (b) But hold on—we've mentioned *two* times here. There's the sequence of times *Sunday*, *Monday*, *Tuesday*, and so on. But then there are the times relative to which *now* is at each of these times. What is this second sequence of times? How should be label the vertical axis in figure 1a?

   (c) One answer: this vertical axis is *another* temporal dimension (figure 1b). It is *hypertime*. Just as ordinary objects move relative to time, the present moment moves relative to hypertime.

      i. Sider: if time flows, and hypertime itself is a kind of time, then hypertime itself must flow. But that means that we need *hyper*hypertime. And, of course, if hyperhypertime is a kind of time, then it, too, must flow, so we'll need *hyper*hyperhypertime, and so on, without limit. Better to do without hypertime.

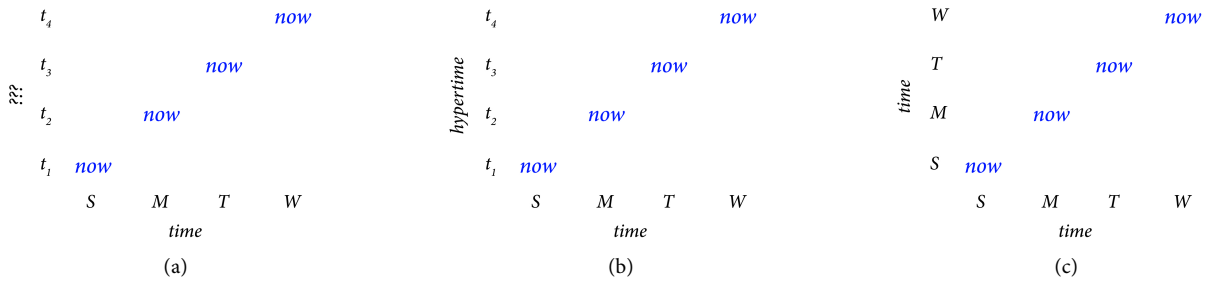


Figure 1

(d) Another answer: the vertical axis is just *time* itself (figure 1c).

i. This doesn't run into the trouble with hypertime, but neither does it capture the sense in which time is moving. True, every time is present relative to itself. But we could say the same about *space* and *here*. Every place is *here* relative to itself. But space does not flow from left to right, or in any direction.

ii. The off-diagonal locations in figure 1c also don't seem to make much sense. What would it even *be* for now to be at Sunday on Monday? That feels like nonsense. So it looks like all we've really said is that times are *now* when they occur. Likewise, places are *here* where they are located.

7. Sider thinks we should accept these consequences. When we think hard about it, we see that time is just like space. Space doesn't flow in any interesting sense. And neither does time. He calls this the **Space-Time Theory**.

Space-Time Theory Time is just another dimension, like space. Just as all *places* are equally real, so too are all *times* equally real. Just as there is no objective 'here', so too is there no objective 'now'. Just as there are *spatial* parts (my left hand is a part of me), so too are there *temporal* parts (my childhood is a part of my life).

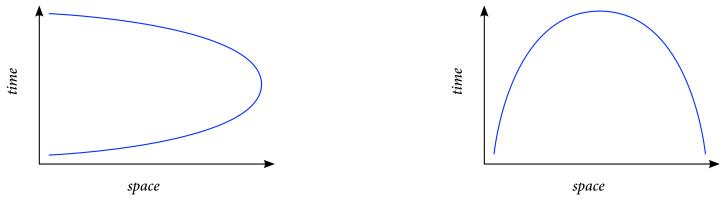
8. Two objections to the Space-Time Theory, and Sider's replies:

(a) Objection: if the Space-Time Theory is true, then things do not really ever *change*. With four-dimensional eyes, we could survey all of space and time laid out before us, and it would be static and unchanging. On the Space-Time Theory, the only sense in which things change is the sense in which a painting changes color (in this place, it's red, in this other place, it's yellow.)

i. Sider: True, but this shouldn't bother us. Really, there *is* no deep difference between change across space and change across time.

(b) Objection: if the Space-Time Theory is true, then things should be able to move back and forth through time, the same way they move back and forth through *space*. Since things can't move back and forth through time, the Space-Time Theory is false.

i. Sider: things *can* move back and forth through time! Consider the spacetime diagrams below. On the left, something moving back and forth in space. What it is for something to move back and forth through space is for it to be at one place at one time, a different place later, and then back to the same place later still. Similarly,



what it is for something to move back and forth in *time* is just for it to be at one *time* in *space*, a different *time* further on in space, and then back to the same *time* further on still. That's just what we see on the right. But this happens anytime a spatially-separated thing comes together (*e.g.*, when the two halves of a locket are reconnected).

Describe one argument against the possibility of time travel, and explain how Lewis defends the possibility of time travel from this argument.

1. We are familiar with stories of time travel from science fiction. Time travel could happen in a variety of ways.

   (a) Way #1 (Terminator): at some time in the future (past) you pop out of existence. At another time in the past (future) you pop into existence, looking just as you did when you popped out of existence.

   (b) Way #2 (Primer): at some time in the future (past), you enter the time machine. At some time in the past (future), you exit. You exist in the time machine throughout the intervening time.

   (c) Way #3 (Gödel, Interstellar?): you head off in a certain direction, always heading into your own local future. However, because spacetime is curved in odd ways, you end up returning to the same point in spacetime from which you departed.

2. According to the Space-Time theory, we can think of each of these forms of time travel in terms of space-time diagrams. Non-time travelers are just 4-dimensional space-time worms. Time travelers are more interesting space-time worms—they are split in two (way #1), or else they age in uncommon ways (way #2), or else they loop back on themselves (way #3).

3. But some have alleged that time travel is impossible—it leads to contradictions. If so, this spells trouble for the Space-Time theory. David Lewis wishes to defend the possibility of time travel from these objectors.

4. Objection #1: Suppose that Tim is a time traveler. Tim travels from 2018 to the year 2028. If Tim is a time traveler in any interesting sense, then this journey takes him less than ten years. Say it took him five minutes. So there are five minutes separating Tim's departure and his arrival. But wait—if we say this, we've contradicted ourselves, since we *also* said that there are ten years separating Tim's departure and his arrival!

   P1. If Tim's journey is possible, then the time separating Tim's departure from his arrival is five minutes.

   P2. If Tim's journey is possible, then the time separating Tim's departure from his arrival is ten years.

   P3. The time separating Tim's departure from his arrival cannot be both five minutes and ten years.

   C1. Tim's journey is not possible.

   (a) Lewis: the argument equivocates between two senses of *time*. We must distinguish *personal time* from *external time*. Personal time is (roughly) time measured by the time traveler's wristwatch. *External* time is real time (roughly, time as measured by the non-time traveler's wristwatch). Premise 1 is only true when 'time' means 'personal time'. Premise 2 is only true when 'time' means 'external time'. And if 'time' means something different in premises 1 and 2, then the argument is invalid.

   (b) The difference between personal time and external time is like the difference between the distance along a train track and distance as the crow flies. The stations are 20 miles apart as the crow flies, though the train must travel 40 miles to get from one station to the other. There's no contradiction here.

5. Objection #2: Tim travels back in time to visit his favorite author, J.D. Salinger. While in the past, he drops his copy of *Catcher in the Rye*. Salinger finds the novel with his name on it, and publishes it. It is very popular and a young Tim reads it, falls in love with it, and travels back in time to see it's author, accidentally leaving his copy behind. But wait—who wrote the book? This is a book with no author! So there's no explanation of why it has the plot that has.

   P4. If Tim's journey is possible, then there could be no explanation for why *Catcher in the Rye* has the plot it does.

   P5. There must be an explanation for why books have the plots they do.

   C2. Tim's journey is not possible.

(a) Lewis: P5 is false. Some things lack any explanation. For instance, the big bang has no explanation—or, if you think that the big bang was caused by God, then God's existence and nature lacks an explanation—or, if you think that there is no big bang, but rather an infinite past, then the existence of this universe in the first place lacks an explanation. In normal universes, these are the only things which have no explanation. In possible universes with time travelers, there are further things without explanations.

6. Objection #3: Tim goes back in time and decides to kill Hitler as a baby. Unbeknownst to Tim, he is descended from Hitler (his parents kept this secret from him). Tim can kill baby Hitler. He's got everything it takes—he's got a gun, a clear shot, there's no one there to stop him, and little baby Hitler cannot defend itself. But, at the same time, Tim *couldn't* kill baby Hitler. If he kills baby Hitler, then Tim won't be born, so he won't travel back in time, and he won't kill baby Hitler. Now our story has contradicted itself.

   In fact, it doesn't really matter whether Tim is descended from Hitler or not. Suppose he isn't. Then, if Tim kills Hitler, then Hitler won't rise to power in Germany, and a young Tim won't learn of Hitler in the history books. So Tim won't go back in time to kill baby Hitler. Contradiction.

   In fact, it doesn't even matter whether Tim knows about Hitler when he travels back in time. Suppose he just wants to kill a random German baby in 1889, and the one he's aiming his gun at happens to be Hitler. If Tim were to kill the baby, then Hitler wouldn't rise to power in Germany in the 1930s. But Hitler *does* rise to power in Germany in the 1930s. Contradiction.

   P7. If time travel is possible, then the time traveler could change the past.

   P8. You cannot change the past.

   C3. Time travel is not possible.

   (a) Note that this is only a problem if we assume that there is only one static timeline, and therefore, only one version of 1889.

      i. Suppose that time *branches*, so that, with 4-dimensional eyes, we could see two histories: one in which Tim never arrives in 1889, Hitler is born and rises to power; and another, in which Tim arrives, kills Hitler, and another demagogue takes his place in the Nazi party in the 1930s. Here, there is no contradiction.

      ii. Similarly, suppose that the timeline itself *changes*. As we saw last class, change requires the passage of time, so, if the entire timeline is to change, then there must be some *other* time relative to which it changes—that is, there must be hypertime. Okay, so suppose there's hypertime. Then, at hypertime t_1, *now* is at 1888 and, in the future, Hitler rises to power. Then, at hypertime t_2, *now* is at 1889, Tim has popped into existence to kill baby Hitler, and in the future, another demagogue rises to power.

   (b) Lewis: this argument equivocates in its use of 'could'. When we say that somebody 'can' or 'could' do something, we mean that it is possible for them to do the thing, *holding some facts fixed*. Unlike a chimpanzee, I can speak Finnish; even though, since I haven't taken any lessons, I can't speak Finnish. I don't contradict myself in saying these two things, since the 'can' means different things each time it's used.

   (c) The first premise is only true when we don't hold fixed facts about the future. Tim *could* kill baby Hitler, in the sense that this is consistent with most of the facts before 1889. But in a *different* sense, Tim *couldn't* kill baby Hitler, since this is inconsistent with facts *after* 1889.

7. Note: if we think that time travel is possible, then we have another argument, in addition to Frankfurt's, that we can be free even if we couldn't have done otherwise.

   (a) Tim stands before baby Hitler, gun in hand. If you think determinism is incompatible with free will, then give Tim all the libertarian agent causal powers you wish. So Tim freely refrains from killing baby Hitler. But he couldn't have done otherwise.

   (b) Note also: events which happened before Tim was born (namely, WWII) *necessitate* that Tim doesn't kill baby Hitler. So, if you think time travel is possible, then, even if you're a libertarian, you should accept that an act can be free even if it is necessitated by events which pre-date your birth.

Why would an all-powerful, all-knowing, and all-good god not prevent evil and suffering?

PHIL 0080 · October 22nd, 2018

What is the problem of evil*, and why does Mackie think that it means that we must deny the existence of a certain kind of god? Present one objection to Mackie's argument, and explain how Mackie replies to this objection.*

1. The *problem of evil* is a problem for a certain kind of theist. The problem is raised for those who think that there is a god—call them 'God', though it needn't be the Christian god—with the following three properties:
   (a) *Omniscience*: God is all knowing—they know all truths.
   (b) *Omnipotence*: God is all powerful—they are able to do anything.
   (c) *Omnibenevolence*: God is all good—they want things to be best.
2. The problem of evil is that these three properties appear to be incompatible with an empirical observation: there is needless and preventable suffering and evil. More carefully:

   P1. If God exists and is omniscient, then they know about all preventable evil.

   P2. If God exists and is omnipotent, then they could prevent all preventable evil, if they wanted to.

   P3. If God exists and is omnibenevolent, then they would want to prevent all preventable evil.

   C1. If God exists and is omniscient, omnipotent, and omnibenevolent, then they would prevent all preventable evil.

   P4. There is preventable evil.

   C2. God is either not omniscient, not omnipotent, not omnibenevolent, or non-existent.

3. J. L. Mackie thinks this argument is good. In order to defend it, he considiers several objections, all of which deny P3.
4. First objection: P3 is false. In order for *any* good to exist, there must be some evil in the world. That's because, in order for something to exist, its opposite must exist as well. We cannot understand what it is for something to be big, *e.g.*, if there is not also something which is small. Similarly, we cannot understand what it is for something to be good if there is not also something which is evil. So, if God were omnibenevolent, he would want to permit some evil as a necessary means to there being good.
   (a) Mackie's first reply to the first objection: to say that God cannot create goodness without also creating evil is to deny that God is omnipotent.
      i. A reply to Mackie's first reply to the first objection: To say that God is omnipotent is not to say that God is capable of doing literally *anything*. God cannot make a square triangle, or make 2+2=5, or create a rock so heavy that even they cannot lift it. If something is impossible, God cannot do it. What it is for God to be omnipotent is for God to be able to do anything *possible*. Since it's not possible for there to be good without evil, saying that God could not create good without evil is not putting a limit on God's power.
   (b) Mackie's second reply to the first objection: even if *some* evil is necessary in order for good to exist, it is not necessary for *so much* evil to exists. If God were omnipotent, they would want to create a world with as little evil in it as possible.
5. Second objection: P3 is false. Evil is necessary for good, not because there cannot be *any* goodness without evil, but rather because there are *certain* goods which cannot exist without evil.
   (a) Call pain and suffering *first-order evils*, and call pleasure and happiness *first-order goods*. In addition to first-order goods, there are *second-order goods*. Second order goods are goods which cannot exist without first-order evils. For example: heroism, compassion, and benevolence. Without the threat of first-order evil, heroism is impossible. Without suffering, no one can show or be show compassion.

(b) Mackie's reply to the second objection: very well. Let us grant that first-order evils are needed for second-order goods. Even so, there are additionally *second-order evils*. A second-order evil is an evil which cannot exist without a first-order goods or evils. For instance, envy, cowardice, and cruelty are second-order evils. Even if first-order evils are required for second-order goods, second-order evils are not. And second-order evils are preventable. So the problem of evil rises again with respect to these second-order evils.

6. Third objection (the *free will* defense): P3 is false. An omnibenevolent god would want things to be best, but *human freedom* is required in order for things to be best. Giving humans freedom means not preventing the evils that they themselves bring about. So an omnibenevolent god would not want to prevent all preventable evil. They would want to allow whatever evil resulted from the free choices of humans.

   (a) Mackie's reply to the third objection: God could have made humans so that they always freely choose good. Then, we would have the good of freedom without the bad of humans choosing evils.

      i. Note an important assumption of Mackie's reply to the third objection: he assumes that it is *possible* for God to make a free agent while simultaneously determining them to choose well. The incompatibilist will deny this. So Mackie's reply presupposes compatibilism.

      ii. Note also that the third objection assumes that it is *impossible* for God to make a free agent while guaranteeing that they choose well. So the third objection assumes incompatibilism. It also assumes that God *did* give us freedom. So it assumes libertarianism.

7. We can strengthen Mackie's reply to the free will defense in four ways.

   (a) Firstly: Even if freedom is incompatible with determinism, *shreedom*—that property which the compatibilist takes to be freedom—is not. Granting that freedom is incompatible with determinism, why should we think that freedom is more valuable than shreedom? And, even granting that freedom is more valuable than shreedom, why should we think that it is *so much more* valuable than shreedom so as to make all the world's preventable evil worthwhile?

   (b) Secondly: even the libertarian can accept that people's freedom can be compromised by, *e.g.*, insanity. This is why we don't punish people who commit horrible crimes due to mental illness. But consider someone who, *non compos mentis*, commits murder. We cannot use the free will defense to explain why God did not prevent this preventable evil, since this preventable evil was not brought about by a *free* agent.

   (c) Thirdly, and relatedly: not all preventable evils are brought about by a free agent. There is so-called *natural evil*—evil which is brought about by nature. For instance, the tsunami which, in 2004, brought about the death of 30,000 people in Sri Lanka. This tsunami was not the consequence of anyone's free choice. So we cannot use the free will defense to explain why God did not prevent this preventable evil, since this preventable evil was not brought about by a free agent.

   (d) Fourthly: God could allow free *choice* without allowing every *consequence* of those free choices. Suppose Jones freely chooses to murder Smith, and pulls the trigger. If Jones' free choice is good, but Smith's death is bad, then God could bring about the best consequences by allowing Jones to make his choice, but then causing the gun to jam, so that Smith lives. So we cannot use the free will defense to explain why God not only allows people to make wrong choices, but also allows those wrong choices to have their bad consequences.

Does the value of freedom explain why an all-powerful, all-knowing, and all-good god would not prevent evil and suffering?

PHIL 0080 · October 22nd, 2018

Discuss one objection to the free will defense. What else, besides freedom, does Swinburne think is needed in order to explain why God allows evil? How does this additional piece of the story allow him to respond to the objection you discussed?

1. Recall, last class, we raised four objections to the free will defense.
   (a) If free will is compatible with determinism, then God could have created us so that we always choose what's best.
   (b) Even if free will is compatible with determinism, why couldn't God have just given us SHREEDOM—why would us being *shree* but not *free* be so much worse than us being free?
   (c) There are many *natural evils*—earthquakes, disease, and famine, *e.g.*. Since these evils are not brought about through any free choice, the free will defense does not explain why God would allow such evils.
   (d) God could allow free *choice* without allowing every *consequence* of those choices. God could intervene so as to prevent wrong choices from having bad consequences.
2. Swinburne: free will on its own does not explain why God allows evil. More is needed. The missing part of the story is *responsibility*. Humans can have freedom, but only because they are given a narrow range of choices whose consequences are unimportant.
   (a) A thought experiment (due to Robert Nozick): there is an *experience machine*. If you decide to get into the machine, your brain will be stimulated to so that you experience whatever pleasurable experiences you wish for the remainder of your life. Hedonists about well-being like J. S. Mill think that things would be better *for you* if you plugged into the machine. Nozick thinks that things would be worse for you if you plugged in. Insofar as we agree, it seems that we value something more than pleasure—and also, that we value more than free will. We can choose whatever we want inside of the machine, but there are no real *stakes*; nothing we do in the machine *matters*.
   (b) A thought experiment (due to screenwriter Andrew Niccol): Truman Burbank is raised in a very large television studio. Everyone around him is an actor being controlled by a benevolent director, Christof. Christof orchestrates the people around Truman so that Truman lives a happy and pleasurable life. Truman is allowed to choose whatever he wishes, though, if he chooses incorrectly, Christof intervenes to prevent his choices from having negative consequences.
      i. The near universal reaction viewers have when witnessing this reality play out before them is that there is something wrong with Truman's world—that living inside of the television studio is worse *for Truman*.
      ii. One way of making sense of this reaction: while Truman has freedom, he is not *responsible* for anything. His choices do not *matter*. There is nothing of any significance *at stake* when Truman chooses.

   Swinburne: a world in which we have, not only *freedom*, but also *responsibility* for the wellbeing of others—a world in which we have the genuine freedom to help and harm other people—is a better world than one in which we have a limited freedom, or a world in which our choices have no serious consequences.
3. By appealing not only to *free choice*, but (what Swinburne calls) *free and responsible choice*, we can respond to the objections we considered last class.
   (a) If free will is compatible with determinism, then God could have created us so that we always choose what's best.
      i. *Reply*: Free will is not compatible with determinism. If we are to be truly free, then God cannot decide *for us*.
   (b) Why would *shreedom* be worse than *freedom*?
      i. *Reply*: Because only freedom allows for genuine and consequential responsibility.
   (c) Why would God allow *natural evil*?
      i. *Reply 1*: If we are to have real responsibility over our own lives and the lives of our fellow humans, we must genuinely have available the choice to seriously harm them. In order to have this choice, we must *know how* to do this. Natural evil exists in order to teach humans how to do this.

A. *Retort*: But why couldn't God simply *tell us* how to seriously harm our fellow humans?

B. *Reply*: Knowing that you are being constantly watched by God would make it more difficult for to choose evil—it would deprive you of the ability to make a free and responsible choice. The most morally serious choices you make are the ones you make in private.

ii. *Reply 2*: The presence of natural evil allows for the possibility of *second order* goods like courage, perseverance, benevolence, *etc*.

(d) God could allow free choice without allowing every *consequence* of those choices.

i. *Reply*: this is the reality which Truman lives. Even though he has free choice, he is not *responsible* for the consequences of those choices. But this is worse than Truman being provided free *and responsible* choice.

4. An objection to Swinburne: even if we agree that a world in which we have free and responsible choice is better than a world in which we are deprived of any consequential choices—that a world in which Jones can, if he chooses, kill Smith is better than a world in which Jones can only choose between apples and bananas—even so, Smith must *consent* to being exposed to the possibility of being killed by Jones. It is wrong for doctors, *e.g.*, to subject their patients to great danger without their consent, so it is similarly wrong for God to expose Smith to this great danger without his consent.

(a) *Reply 1*: Parents have certain rights over their children because they created them. They can make choices for them, without their consent, precisely because they brought them into existence. Similarly, since God brought humans into existence, he has the right to make some choices for them.

i. Of course, not *any* choice is a permissible one to make. God could not rightly decide to subject Smith to the possibility of an eternity of pain brought about by Jones. But our lives on earth are short, and so there is necessarily a limit to the amount of danger to which God exposes Smith.

(b) *Reply 2*: When God is making this choice, Smith *does not exist*. So God cannot get Smith's consent. If something is impossible, it's not morally required. So God is not morally required to get Smith's consent before putting him into a world in which people have free and responsible choice.

5. Swinburne concludes by noting that you may think that it is permissible to expose Smith to the danger of death at the hands of Jones only if Smith is in some way *compensated* for this danger.

(a) If we believe that there is an *afterlife*—as many theists do—then may appeal to this afterlife to justify God's decision to expose Smith to these dangers.

(b) Perhaps, however, this afterlife brings in additional troubles. If the afterlife is not all roses—if, as many theists hold—there is not only *compensation* in the form of eternal bliss, but also *punishment*, in the form of eternal torment, then there is a worry that the creation of *this* evil is inconsistent with God's omnibenevolence. And it's not easy to see how this evil could be explained in the way Swinburne suggests to explain away worldly evil. (See David Lewis's *Divine Evil* for more.)

Why does Al-Ghazali worry about his religious beliefs? What kind of knowledge does he seek? Does he think that the evidence of his senses is enough to provide him with this kind of knowledge? Why or why not? Does he think that 'self-evident truths' provide him with this kind of knowledge? Why or why not?

1. In everyday conversations about whether someone knows something, it is commonplace to argue that they *don't* know that thing by pointing out that they are relying upon the testimony of an unreliable source. For instance:

   A: "The Democrats are going to retake the House."

   B: "Why do you think that?"

   A: "I don't think it; I *know* it. All the polls have them ahead."

   B: "Okay, maybe they will retake the House, but you don't *know* they will. The polls have been wrong before."

   In the dialogue, A makes a claim to knowledge. B attempts to undermine A's claim to knowledge by pointing out that the *source* of A's belief are not entirely trustworthy. They aren't trustworthy because those sources have been wrong in the past. Here's a commonsensical claim:

   No Knowledge from An Untrustworthy Source If the source of your belief that p is not completely trustworthy, then you don't know that p.

2. In everyday conversations about whether someone knows something, it is also commonplace to argue that they *don't* know that thing by pointing to a possibility, consistent with their evidence, in which that thing is false.

   (a) Some terminology: Say that your evidence *rules out* a possibility if and only if your evidence is incompatible with that possibility. For instance: if I have the evidence that Sam was at home at the time of the burglary, then my evidence *rules out* that Sam committed the burglary.

   (b) Here's another common-sensical claim:

   Knowledge Precludes the Possibility of Error If your evidence doesn't rule out a possibility in which p is false, then you don't know that p.

3. Appealing to **No Knowledge from An Untrustworthy Source**, Al-Ghazali begins to worry about his religious beliefs.

   (a) Al-Ghazali notices that there is a wide divergence of opinion about religious questions, all derived from a common source:

   > *I saw that the children of Christians always grew up embracing Christianity, and the children of Jews always grew up adhering to Judaism, and the children of Muslims always grew up following the religion of Islam.*

   (b) The source of his religious beliefs—the testimony of parents—is not completely trustworthy. So Al-Ghazali does not think that this source is enough for him to truly *know* that his religious beliefs are true.

4. Al-Ghazali wishes to find a source for his religious beliefs which is completely trustworthy and free from the possibility of error. What he seeks is *sure and certain knowledge*:

   > *sure and certain knowledge is that in which the thing known is made so manifest that no doubt clings to it, nor is it accompanied by the possibility of error and deception, nor can the mind even suppose such a possibility.*

   (a) Upon surveying all of his beliefs, Al-Ghazali decides that none are candidates for being sure and certain knowledge except for two:

      i. The testimony of his senses (sight, smell, sound, *etc.*); and

      ii. The testimony of his reason—or 'self-evident truths' (*e.g.*, his belief that 2+2=4, that no proposition is both true and false, that three is less than ten, that the sum of the square of the legs of a right triangle equals the square of its hypotenuse).

5. Unfortunately, Al-Ghazali does not believe that he is able to have sure and certain knowledge of the testimony of his senses.

   (a) His senses have deceived him in the past. He once thought stars were very small. But then his reason, through geometrical proof, taught him that the stars were much bigger than the entire earth. He thought that shadows were motionless, but then his reason teaches him that it is in fact moving, just imperceptibly.

6. Let's think through Al-Ghazali's reasoning here more carefully. He appeals to the fact that his senses have deceived in the past, and uses this to argue that they are *never* a source of knowledge. In so doing, he appeals to **No Knowledge From An Untrustworthy Source**. So take any arbitrary belief which has the testimony of the senses as its source—for instance, the belief that I am wearing a blue jacket.

   P1. Your only evidence that Dmitri is wearing a blue jacket is the testimony of your senses.

   P2. The testimony of your senses is not completely trustworthy—it has misled you before.

   P3. If the source of your belief that p is not completely trustworthy, then you don't know that p.

   C1. You do not know that Dmitri is wearing a blue jacket.

7. There is another argument that the testimony of the senses do not provide knowledge—this argument appeals to the second commonsensical claim about knowledge: that knowledge precludes the possibility of error.

   P4. Your only evidence that Dmitri is wearing a blue jacket is the testimony of your senses.

   P5. The testimony of your senses does not rule out the possibility that you are dreaming, and that Dmitri is not wearing a blue jacket.

   P6. If your evidence does not rule out a possibility in which p is false, then you don't know that p.

   C2. You don't know that Dmitri is wearing a blue jacket.

8. What of the testimony of reason? Here, too, Al-Ghazali finds reasons for doubt. For, just as, in dreams, we seem to perceive things which are not present, so too, in dream, do things *seem* to make sense when in fact they are *nonsense*. Al-Ghazali does not discuss this, but even in waking life, we also frequently encounter '*brain farts*'—occasions when something *seems* self-evident to us, even though it is false. So we may run similar arguments with respect to the testimony of reason:

   P7. Your only evidence that the area of a right triangle is 1/2 times its base times its height is the testimony of your reason.

   P8. The testimony of reason is not completely trustworthy—it has misled in dreams and brainfarts.

   P9. If the source of your belief that p is not completely trustworthy, then you don't know that p.

   C3. You do not know that the area of a right triangle is 1/2 times its base times its height.

   And, likewise:

   P10. Your only evidence that the area of a right triangle is 1/2 times its base times its height is the testimony of your reason.

   P11. The testimony of reason does not rule out the possibility that you are dreaming, confused, or suffering a brainfart, and the area of a right triangle is something other than 1/2 bh.

   P12. If your evidence does not rule out a possibility in which p is false, then you don't know that p.

   C4. You don't know that the area of a right triangle is 1/2 times its base times its height.

What is the position of the philosophical skeptic? Present one argument for philosophical skepticism. Explain how the fallibilist *and* the externalist about evidence *would reply to this argument.*

1. According to one view, knowledge is *justified, true belief*. Call this the '**K=JTB**' view.

   K=JTB You know that p iff:

   (a) p is true

   (b) You believe that p

   (c) Your belief that p is sufficiently well *justified*.

2. What is it for a belief to be *justified*? Let us say this: a belief of yours is *justified* iff it is supported by *your evidence*.

   (a) What is your evidence? We saw last class that we hold most of our beliefs about the world on the basis of our *sensory experience*. So let us say that, when it comes to our beliefs about the external world, at least, this is our evidence: our evidence about the external world is our sensory experience.

   Evidence as Experience Our sensory experience is our total evidence about the external world.

   (b) What is it for our evidence to *support* a belief of yours? Let us say: your evidence *supports* your belief that p iff your evidence makes p likely. Since we've said that a belief of yours is justified iff it is supported by your evidence, we arrive at the view that a belief of yours is justified iff your evidence makes it likely.

3. **K=JTB** says that, in order for your belief that p to constitute *knowledge*, the belief must be, not only justified *to some extent*, but justified *sufficiently well*. How well justified must the belief be?

   (a) The philosophical skeptic says: in order for a belief to count as *knowledge*, it must be *maximally justified*—that is, it must be *certain*.

   (b) That is: in order for a belief to constitute *knowledge*, your evidence must eliminate *every* possibility of error. In order to know that p, your evidence supporting p must be *infallible*. Call this view *infallibilism*.

   Infallibilism In order to know that p, your evidence must rule out every possibility in which p is false.

4. With **Evidence as Experience** and **Infallibilism** in hand, the philosophical skeptic argues against most of our everyday knowledge. Take any mundane belief about the external world—for instance, that your friend Jack has a nut allergy. Feldman identifies four skeptical arguments:

   Possibility of Error Argument

   P1. Your belief that Jack has a nut allergy could be mistaken.

   P2. If a belief could be mistaken, then it is not knowledge.

   C. You don't know that Jack has a nut allergy.

   Certainty Argument

   P3. You're not absolutely certain that Jack has a nut allergy.

   P4. Knowledge must be absolutely certain.

   C. You don't know that Jack has a nut allergy.

   Indistinguishability Argument

   P5. The case in which Jack has lied about his nut allergy is indistinguishable from one in which he has not.

   P6. Cases of knowledge cannot be indistinguishable from cases of non-knowledge.

   C. You don't know that Jack has a nut allergy.

   Transmissibility Argument

   P7. You don't know that Jack hasn't lied about his nut allergy.

   P8. If you don't know that Jack hasn't lied about his nut allergy, then you don't know that Jack has a nut allergy.

   C. You don't know that Jack has a nut allergy.

5. The philosophical skeptic could proceed piecemeal, arguing against every one of your beliefs about the external world with arguments like these. Or, they could attempt to establish, in one fell swoop, that *none* of your beliefs about the external world are knowledge.

   (a) For the arguments above, the skeptic was pointing to a *bad case*—one in which your belief about Jack's nut allergy was false, but your experiences were exactly the same as they are in the *good case* (the case in which Jack isn't lying).

   (b) However, they could instead point to a *very bad case*: one in which *all* of your beliefs about the external world are false. Here's one: you are a brain-in-a-vat, being stimulated by electrical signals to re-create experiences qualitatively identical to the ones you are currently having. Then, they could run versions of the four arguments above, not just for your belief that Jack has a nut allergy, but for *all* of your beliefs about the external world.

6. Why not simply accept the conclusion? One reason is that knowledge plays an important normative role in our lives. For instance: if you didn't know that Jack had a nut allergy, we will excuse you for offering him brownies with nuts in it. If you *did* know that Jack had a nut allergy, we will blame you for offering him the brownie. In general, ignorance is taken to *excuse* otherwise blameworthy bahavior. But, if the philosophical skeptic is correct, then we have a *universal* excuse.

7. Here are two ways of resisting the skeptic's conclusions: on the one hand, we could deny **Infallibilism**. On the other hand, we could deny **Evidence as Experience**.

   (a) The *fallibilist* says that, in order to know that p (for any proposition p), your evidence needn't eliminate *all* possibilities of error. In particular, you needn't eliminate the *bad case* from the arguments above.

   Fallibilism You can know that p even when your evidence does not rule out every possibility in which p is false.

   The fallibilist will reject the second premise of each of the arguments above. You can know that Jack has a nut allergy even though...

   i. ...you could be mistaken about whether he has a nut allergy;
   ii. ...you are not absolutely certain whether he has a nut allergy; and
   iii. ...your experience is indistinguishable from one in which you don't know that he has a nut allergy.

   The fallibilist *could* reject the first premise of *the Transmissibility Argument*. They could insist that you *do* know that Jack wasn't lying (and that, *e.g.*, you are *not* a brain-in-a-vat). However, they need not. They might try something fancier than that. Here's a version of fallibilism that denies the second premise of that argument:

   Relevant Alternatives Fallibilism For any belief, p, there is a set of *relevant* alternative possibilities in which that p is false. In order to have enough justification to count as knowing that p, your evidence must rule out—not *every* possibility in which p is false—but every *relevant* possibility in which p is false.

   i. When it comes to your belief that Jack isn't lying, the possibility that he *is* lying is a relevant alternative possibility. In order to know that Jack isn't lying, your evidence must rule out the possibility that he is lying. Since your evidence doesn't rule this possibility out, you don't know that Jack isn't lying.
   ii. However, in order to know that Jack has a nut allergy, the relevant possibility you have to rule out is just the one where he doesn't have a nut allergy, and tells you that he doesn't. Since your evidence does rule this possibility out, you know that Jack has a nut allergy.

   (b) The *externalist about evidence* denies **Evidence as Experience**.

   i. They claim that, in the good case, you don't just have the evidence that you heard Jack say that he has a nut allergy. You *also* have the evidence that *Jack has a nut allergy*.
   ii. Of course, in the bad case, you don't have this evidence (since, in the bad case, it's not *true* that Jack has a nut allergy). So your evidence is different in the good case and the bad case, even though you have precisely the same experiences in the good case and the bad case.

   The externalist about evidence will deny P1, P3, P6, and P7 of the arguments above. For, so long as you are in the good case, and Jack *isn't* lying:

   i. your belief could not be mistaken (your evidence rules out the possibility of error)
   ii. you *are* absolutely certain that Jack has a nut allergy (your evidence makes it certain that he does)
   iii. cases of knowledge *can* be indistinguishable from cases of non-knowledge;
   iv. you *do* know that Jack hasn't lied.

According to Hume, how do we form beliefs about things we haven't directly observed like the future? According to Hume, can we know that this method of forming beliefs won't lead us into error? Why or why not?

1. David Hume draws a distinction between two different kinds of beliefs: beliefs about *relations of ideas* and beliefs about *matters of fact*.

2. Beliefs about *relations of ideas* are beliefs—like 'Singles are unmarried' and 'No woman is taller than herself'—which we can know to be true just by consulting our ideas of 'singles', 'unmarried', and 'taller than'.

   (a) These beliefs are knowable by reason alone.

   (b) We cannot even *conceive* of these beliefs being false. We cannot *conceive* of a woman who is taller than herself, or a triangle whose area is greater than 1/2 its base times its height.

3. If a belief is not a relation of ideas, then Hume calls it a *matter of fact*.

   (a) These are beliefs—like 'Trump is president', 'I am in Pittsburgh', and 'The Earth travels around the Sun about once every 365.25 days'—which we can only know by having some sense experience.

   (b) In contrast to beliefs about relations of ideas, the beliefs are not knowable by reason alone.

   (c) In contrast to beliefs about relations of ideas, we can conceive of these beliefs being false. We can conceive of Trump not being president, we can conceive not being in Pittsburgh, and we can conceive of the Earth's trip around the sun taking longer.

4. Some of our beliefs about matters of fact are based directly on sense experience—like, for instance, my belief that I have hands. However, many of our beliefs are not.

   (a) I see smoke in the distance and conclude that there must be fire there. But I haven't seen the fire. I hit 'send' and believe that you will read my email, but I haven't seen you reading the email. I watch the weather channel, and believe that there will be a storm. But I haven't seen the storm.

   (b) In all these cases, we have beliefs about matters of fact which were not *directly* observed.

5. So, Hume inquires: what *is* the basis of these beliefs about unobserved matters of fact?

   (a) His answer: these beliefs must be based upon relations of *cause* and *effect*. believe that fire caused the smoke, and I did directly observe the smoke. I believe that hitting 'send' causes emails to arrive in your inbox, which causes you to read them. And I did directly observe myself hitting 'send'. I believe that the weather report and the weather have a common cause, and I did directly observe the weather report.

6. Hume is not yet satisfied. He pushes on: what, then, is the basis of our beliefs about *cause and effect*?

   (a) Hume thinks that there are two options: either these are beliefs about *relations of ideas* or they are beliefs about *matters of fact*.

   (b) If these were beliefs about relations of ideas, then we could not conceive of them being false.

   (c) But, Hume contends, we *can* conceive of our beliefs about cause and effect being false. We can conceive of fire not causing smoke. We can conceive of me hitting 'send' but an email not arriving in your inbox. We can conceive of the weather report being caused by something that doesn't cause the weather.

   (d) So: our beliefs about cause and effect cannot be beliefs about relations of ideas.

   (e) So: they must be beliefs about matters of fact. So they must be knowable through sense experience.

7. Hume presses on: which experiences justify us in believing, *e.g.*, that fire causes smoke?

   (a) His answer: our previous observations of fire being followed by smoke. We observe fire causing smoke on several occasions and conclude that fire *always* causes smoke. That is, our inference a kind of *enumerative induction*.

P1.	The first F was G
P2.	The second F was G
	⋮
PN.	The Nth F was G
C1.	All Fs are G

P1.	The first fire caused smoke
P2.	The second fire caused smoke
	⋮
PN.	The Nth fire caused smoke
C.	All fire causes smoke

(On the left, the general schema of enumerative induction; on the right, the schema applied to fire causing smoke.)

8. But wait—what justifies us in accepting *this* inference? It is conceivable that these premises are true but the conclusion is false. So the inference cannot be determined by reason alone.

   (a) Another way of making Hume's point: enumerative induction is a *deductively invalid* argument form. If we plug in 'is a swan' for 'F' and 'is white' for 'G', then before 1500, the premises of the argument form were all true, and the conclusion was false.

9. If this inference is not justified on the basis of reason alone, then it must be justified on the basis of some other beliefs that we have. What beliefs are those?

   (a) Hume: we don't *only* believe that the first N fires caused smoke. We *also* believe that nature operates in a uniform manner—there are laws governing the behavior of things like fires, and they apply in the same way in all places and all times.

   (b) If we add *this* additional premise to the reasoning above, then we will have a deductively valid argument form.

P1.	The first fire caused smoke
	⋮
PN.	The Nth fire caused smoke
UN.	Nature operates uniformly
C.	All fire causes smoke

10. Hume is still unsatisfied—what justifies us in believing that nature operates uniformly?

    (a) It is either a belief about relations of ideas or matters of fact.

    (b) If it were a belief about relations of ideas, then we would not be able to conceive of the belief being false.

    (c) But we can conceive of nature operating in a non-uniform manner. So this belief is not about relations of ideas. It must be a belief about matters of fact.

11. If our belief that nature operates uniformly is a belief about matters of fact, then we must know it through sense experience. But which sense experience could justify it?

    (a) The most natural answer: in our past experience, nature has operated uniformly. Thus far, when we go looking for regularities in nature, for the most part we find them; and, for the most part, those regularities continue into the future.

    (b) So this must be the basis of our belief that nature operates uniformly: nature has always operated uniformly in the past.

P1.	Nature has always operated uniformly in the past
C.	Nature always operates uniformly

12. But—and this is Hume's first punchline—we can easily conceive of this premise being false while the conclusion is true. We can easily conceive of nature operating uniformly up until the year 2020, and operating in a haphazard, nonuniform manner ever after. So *this* inference cannot be determined by reason, either.

    (a) Another way of making Hume's point: this inference is of the invalid form 'X has always operated uniformly in the past; therefore, X will always operate uniformly in the future.' Think about the chicken who supposes that, because the farmer feeds them every day, they will continue to feed them. The farmer operates uniformly throughout most of their life, until one day he does not feed them, but slaughters them. If we replace 'X' with 'The farmer', the premise is true but the conclusion false. So this argument form is not deductively valid.

Does Hume think that we have any good reason to think that induction will lead to truth? Why or why not? Illustrate Hume's position by discussing 'counterinduction'.

1. Recall: Hume argued that we form beliefs about unobserved matters of fact—*e.g.*, the future—by relying upon our beliefs about relations of cause and effect; and that we form beliefs about relations of cause and effect by relying upon *enumerative induction*. On the left, the general form of an enumerative inductive inference. On the right, an example.

P1. The first F was G	P1. The sun rose on the first day
P2. The second F was G	P2. The sun rose on the second day
⋮	⋮
PN. The Nth F was G	PN. The sun rose on the Nth day
C. All Fs are G	C. The sun will rise tomorrow.

   Hume then argued that this method of forming beliefs—unlike deductively valid inference—is not guaranteed to lead us to truth. It is possible for the premises of an enumerative inductive inference to be true while the conclusion is false. It is conceivable that, tomorrow, the sun does not rise.

2. Let's agree with Hume. Inductive inferences are not guaranteed to lead to truth—but what of it? Hume has shown us only that we don't have *conclusive* reason to accept the conclusion of an enumerative inductive inference. But we may still have *excellent reason* to accept the conclusion.

3. Hume wishes to call this into question, too. Hume doesn't think that we have *any* good reason to think that the sun will rise tomorrow morning.

   (a) Recall: in order to reach the conclusion that the sun will rise tomorrow, Hume believes that we must assume that nature operates *uniformly*, that the future will resemble the past.

P1. The sun rose on the first day
⋮
PN. The sun rose on the Nth day
UN. The future will resemble the past
C. The sun will rise tomorrow.

   Your past observations of the sun rising give you reason to think that the sun will rise tomorrow only if you have reason to think that nature operates uniformly. That is: you should think that *this* inference is good only if you have some good reason to accept UN.

   (b) But, Hume contends, you *don't* have any good reason to think that nature operates uniformly. The only reason you could give for accepting UN is that nature has always operated uniformly *in the past*. *But this is just another inductive inference!* The inference on the left is no different from the inference on the right.

P1. The future has resembled the past in the past	P1. The sun always rose in the past
UN. The future will resemble the past in the future	C. The sun will rise in the future

   (c) In both cases, we only have reason to accept the conclusion on the basis of the premises if we suppose that nature will always operate uniformly,

P1. The future has resembled the past in the past	P1. The sun always rose in the past
UN. The future will resemble the past	UN. The future will resemble the past
UN. The future will resemble the past in the future	C. The sun will rise in the future

(d) But this belief is exactly the one we were trying to justify. We have reasoned in a circle. Our premise UN *on its own* entails our conclusion. We've simply taken for granted the thing we were trying to show. Our reasoning in this case is no different from our reasoning here:

P1. Whatever Billy says is true
P2. Billy says that whatever he says is true

C. Whatever Billy says is true

This argument does not give us a good reason to think that Billy speaks truly. Why not? Because it only gives us *circular* reasons. And circular reasons are not good reasons.

4. Spelling it out more carefully, *this* is why Hume thinks you don't have any good reason to think that the sun will rise tomorrow.

P1. The only reason you have to think that nature operates uniformly is a circular reason.
P2. Circular reasons are not good reasons.

C1. You have no good reason to think that nature operates uniformly. (*from P1 and P2*)
P3. If you have no good reason to think that nature operates uniformly, then you have no good reason to think that the sun will rise tomorrow.

C2. You have no good reason to think that the sun will rise tomorrow. (*from C1 and P3*)

5. Consider a competitor to induction, called *counterinduction*.

(a) When the counterinductivist sees a regularity in nature, they predict that the regularity will *not* continue into the future. The counterinductivist reasons as follows (on the left, the general pattern; on the right, an example):

P1. The first F was G
⋮
PN. The Nth F was G

C. All other Fs are not G

P1. The sun rose on the first day
⋮
PN. The sun rose on the Nth day

C. The sun will never rise again.

(b) Asked to justify these inferences, the counterinductivist explains that they are relying upon a principle of the *non-uniformity of nature*: that the future will not resemble the past.

P1. The sun rose on the first day
⋮
PN. The sun rose on the Nth day
NU. The future will not resemble the past

C. The sun will never rise again.

(c) When you ask them why they accept the principle of the non-uniformity of nature, they explain to you that the principle has never held true in the past—and, since the future won't resemble the past, this means that the principle *will* hold true in the future.

P1. The future has always resembled the past in the past
NU. The future will not resemble the past

NU. The future will not resemble the past in the future

(d) It looks as though the counterinductivist's reasons for thinking that the sun won't rise tomorrow aren't very good. And the *reason* they don't look very good is that those reasons are circular.

(e) But it also seems as though our own reasons for thinking that the sun *will* rise tomorrow are circular in precisely the same way.

(f) So it looks as though we don't have any good reason to think that the sun will rise tomorrow.

What does Karl Popper think distinguishes science from pseudo-science? Illustrate Popper's views by discussing Einstein's theory of gravitation and Marx's theory of history. Does Karl Popper think that good scientific theories are verified by evidence? Why or why not? Explain how Popper's views afford him a response to Hume's Problem of Induction.

Popper on Demarcation

1. Many think that there is an important difference between *sciences*, like the fields on the left, and *pseudo-sciences*, like those on the right.

Science	Pseudo-Science
Physics	Astrology
Medicine	Homeopathy
Chemistry	Phrenology
Biology	Intelligent Design

   The *problem of demarcation* is the problem of saying what it is that distinguishes the intellectual activities on the left from those on the right. It is the problem of saying what distinguishes science from pseudo-science.

2. Here is a popular answer to the problem of demarcation (call it *the verificationist's* answer):

   (a) The fields on the left have theories which are *well verified by evidence*. Their theories were formed using the *inductive method*. Their theories offer *good explanations* a wide variety of phenomena.

   (b) The fields on the right are not well verified by evidence. Their theories were not formed using the inductive method. They cannot explain a wide variety of phenomena.

3. Popper thinks that the verificationist's answer is wrong. He explains why by making reference to three theories which were popular in the Vienna of his youth: Einstein's theory of gravitation, Marx's theory of history, and Freud's' theory of the unconscious.

   (a) Both Marx's' and Freud's theories have *tons* of evidence verifying them. As soon as you adopt Marx's or Freud's theory, you begin seeing evidence for them *everywhere*. Every headline is further evidence of class struggle; every psychological pathology further evidence of *repression*.

   (b) Similarly, both Marx's and Freud's theories could *explain* tons of phenomena. Marx's theory offered an explanation of every major historical development. And Freud's theory offered an explanation of otherwise puzzling psychological phenomena.

      i. So, according to the verificationist, both Marx and Freud should count as paradigm instances of science.

   (c) Popper, however, began to think that their abundance of verifying evidence and the ease with which they offer explanations is in fact a *weakness* of those theories, and not a strength.

   (d) He contrasts the theories of Marx and Freud with the theory of Einstein. Einstein's theory predicted the phenomenon of *gravitational lensing* (light bending around massive bodies like the sun). The theory said that, during a solar eclipse, the apparent position of the stars in the sky would be different than they are at night. In 1919, Eddington traveled to South America during a solar eclipse, and saw that the theory's prediction was in fact true—the apparent position of the stars *was* different.

   (e) In Popper's eyes, what made Einstein's theory superior to the theories of Marx and Freud wasn't that Einstein's theory had more *verification*. Rather, what made the theory superior was that it was far more *falsifiable*.

      i. While both Marx and Freud's theories could easily accomodate and explain any evidence whatsoever, if the apparent position of the stars had not shifted, Einstein's theory would not have been able to accommodate or explain this observation.

      ii. Unlike the theories of Marx and Freud, Einstein's theory could have been easily *refuted*. If light had not bent around the sun as predicted, then the theory would have been falsified.

4. Popper thinks that *this* is the feature which distinguishes science from pseudo-science: it is falsifiable. It is capable of being refuted by evidence. It *sticks its neck out*. Pseudo-science, in contrast, is not falsifiable. It fails to stick its neck out. It only gathers evidence in its favor. It never bothers to make risky predictions which could potentially refute it.

Popper on Induction

5. In fact, Popper goes further: on his view, verification is not a goal of science at all. In fact, the only goal of science is to *refute* and *falsify* theories. While the verificationists think that verifying evidence can give you positive reason to *believe* a theory is true, Popper thinks that it can never do this.

6. This claim goes beyond what Popper has said before. Let's clearly distinguish the following two claims:

   Refutability is Scientific Science is refutable. In order for an inquiry to count as scientific, there must be some evidence you could receive which would lead you to reject the theory you currently accept. If you are not prepared to reject your theory in the light of any evidence, then your inquiry is not scientific.

   Verification is Unscientific Science does not attempt to *verify* its theories, but only *refute* them. In order for your inquiry to count as scientific, you must not attempt to verify your theory. If you attempt to verify your theory, then your inquiry is not scientific.

   (a) Thus, Popper claims that science does not *establish* scientific theories. A scientific field will not conclude that its theories are *true*. A scientist will not *believe* that Einstein's theory of gravitation is true.

   (b) Instead, a scientist will only believe that Newton's theory of gravitation is *false*. For Newton's theory of gravitation has been *refuted*.

   (c) If a theory sustains repeated attempts at falsification, then Popper says that that theory has been *corroborated*. But this does *not* mean that we have any reason to think that the theory is true. Nor does it mean that we have any reason to think that the theory is *likely*. (On Popper's view, the probability of any scientific theory is always *zero*.) Rather, it means only that we tried several times to falsify the theory, and we failed.

7. Popper believes that his second claim, **Verification is Unscientific**, affords us a response to Hume's problem of induction. Compare the following two patterns of reasoning (or *argument forms*):

   $$\frac{\text{If } T\text{, then } E \qquad E}{T} \qquad\qquad \frac{\text{If } T\text{, then } E \qquad \text{It is not the case that } E}{\text{It is not the case that } T}$$

   (a) The argument form on the left takes a theory to be true because its predictions turn out to be true. This is how you would reason to *verify* a scientific theory. This argument form is *deductively invalid*. It is the argument form we earlier called *affirming the consequent*.

   (b) The argument form on the right takes a theory to be *false* because its predictions turn out to be *false*. This is how you would reason to *refute* a scientific theory. This argument form is *deductively valid*. Consider the following truth-table (I've assumed that 'it is not the case that p' is true when 'p' is false and false when 'p' is true):

T	E	if T, then E	it is not the case that E	it is not the case that T	
true	*true*	*true*	*false*	*false*	
true	*false*	*false*	*true*	*false*	
false	*true*	*true*	*false*	*true*	
false	*false*	***true***	***true***	***true***	✓

   (c) So: if science is only in the business of refuting theories, and not at all in the business of verifying them, the science never has to engage in *inductive* inference at all. And Hume's problem of induction does not arise.

8. Thus, Popper's 'solution' to Hume's problem of induction is to agree with Hume that we have no good reason to believe that observed regularities will continue into the future. However, this is not a problem for science, because good science doesn't attempt to conclude that they will. Science does not seek to verify its theories; so science does not have any need of inductive inferences.

Briefly, what is Hume's problem of induction? What is Popper's 'solution' to Hume's problem of induction? Why does Putnam think that this solution does not work?

1. Recall, Popper's position is that science does not use induction. For science is not in the business of *verifying* theories; rather, science is only in the business of *falsifying* theories. While the logic of *verification* is inductive, the logic of *falsification* is not—it is *deductively valid*.

Verification	*Falsification*
If T, then E	If T, then E
E	Not E
T	Not T

   (a) 'T' is a scientific theory. 'E' is some *evidence* which the theory predicts. It is an *testable prediction* of the theory.

   (b) So, if science is only in the business of falsifying theories, then science need not use induction.

   (c) For this reason, Popper claims to have solved Hume's problem of induction. The solution is this: Hume was right; we have no good reason to think that induction will lead to truth. But, that's fine, since good science doesn't actually *use* induction.

   (d) Let us begin by focusing on *this* claim that Popper makes:

   Deductivism Good science only reasons *deductively*. Good science does not reason inductively.

2. Popper's position relies upon the idea that scientific theories make definite *predictions*. It relies upon the idea that we can say: *if* this theory T is true, *then* we will observe this evidence, E. Putnam wishes to call this assumption into question. Putnam denies that any scientific theory *implies* any particular test implication *all by itself*.

   (a) Consider, for instance, Newton's theory of universal gravitation (TUG). This theory tells us that that all massive bodies exert a gravitational force upon each other, that force is proportional to the mass of the two bodies, and inversely proportional to the square of the distance between them. And it tells us that bodies will accelerate in the direction of the forces acting upon them, with a magnitude equal to the magnitude of the force acting upon them divided by the body's mass. That is, the theory provides us with the equations,

$$a = F_{\text{total}}/m$$
$$F^{1,2}_{\text{grav}} = G \cdot \frac{m_1 \cdot m_2}{d^2_{1,2}}$$

   (b) Suppose that we wish to use this theory to derive a testable prediction about the orbits of the planets. In order to do so, we must assume that the sun and the planets in our solar system are the only relevant masses—that all other masses are small enough or far enough away that their gravitational masses are negligible. We must also assume that the planets are only subject to the gravitational forces they exert upon each other. That is, we must assume that there are no electromagnetic forces (e.g.) which make a substantial difference with respect to the orbits of the planets.

   More generally, any motion of the planets could be squared with Newton's theory by simply postulating additional forces. So the TUG, on its own, does not entail anything about the motion of the planets.

   (c) If we want to derive a *testable prediction* from a theory like TUG, then we will have to make additional assumptions—Putnam calls these *auxilliary statements*—like:

   AS_1 All bodies except the sun and the seven planets are far enough away that their masses are negligible.

   AS_2 All forces other than mutually induced gravitational forces are negligible.

   AS_3 The sun and the seven planets exist in a vacuum.

   These auxiliary statements are not a *part* of the theory; however, no prediction can be obtained without making auxiliary assumptions such as these.

3. If a theory on its own does not yield testable predictions, but instead only does so with the aid of additional auxiliary assumptions, then the logic of verification and falsification we began with should be emended to reflect this fact. *Really*, this is the logical form of verification and falsification:

*Verification**	*Falsification**
If T and A, then E	If T and A, then E
E	Not E
T	Not T

   (a) But—and this is Putnam's first point: the argument form of *Falsification** is *not deductively valid*. (Let $T =$ 'Trump loses the popular vote', $A =$ 'The winner of the election wins the popular vote', and $E =$ 'Trump loses'. Then, the premises of the argument are true, but the conclusion is false.)

   (b) The most we could conclude from the premises of *Falsification** is this: *either* not T *or* not A. So, if we ever reject a theory on the basis of a failed test prediction, we are reasoning *inductively*. So **Deductivism** is false. So Popper's solution to the problem of induction fails.

4. Popper is aware that science needs auxiliary assumptions in order to text its theories. He says:

   > *Some genuinely testable theories, when found to be false, are still upheld by their admirers—for instance by introducing ad hoc some auxiliary assumption. Such a procedure is always possible, but it rescues the theory from refutation only at the price of destroying, or at least lowering, its scientific status.*

   (a) This does not answer the objection. If 'such a procedure' is always possible, then it appears as though we can never reason *deductively* from evidence to the rejection of the theory. So Popper's 'solution' to the problem of induction fails.

   (b) However, we can understand this claim as Popper articulating his claim that good science is refutable. He is saying that, even if it is *possible* to hang onto your theory after its testable predictions are not borne out, doing so is *unscientific*.

   Falsificationism Good scientists reject theories when their testable predictions are discovered to be false. Good scientists do not reject the auxiliary assumptions in an attempt to save their theory from refutation.

5. So, Popper's **Falsificationism** says that the first inference pattern is scientific; whereas the second is unscientific.

Scientific	*Unscientific*
If T and A, then E	If T and A, then E
Not E	Not E
Not T	Not A

6. Putnam's second point is that Popper is wrong about even *this*.

   (a) In the early 19th century, Leverrier used Newton's TUG to make predictions about the orbits of the planets in our solar system. While doing so, he assumed that all objects other than the 7 planets and the sun were small enough or distant enough that their gravitational influence was negligible. However, TUG, together with this assumption, makes false predictions about the orbit of Uranus.

   (b) Popper says that the scientific reaction is to reject the TUG.

   (c) In fact, this is not what Leverrier did. Instead, he postulated that there is an *eighth* planet, which Leverrier named 'Neptune'. By making additional assumptions about Neptune's mass and its orbit, Leverrier was able to show that TUG makes the correct predictions about Uranus's orbit. Later, the planet Neptune was discovered, and Leverrier's assumptions about it were found to be correct.

7. Putnam's point isn't just that Leverrier was *correct* about the existence of Neptune. His point is that Leverrier's reasoning here was both *good* and *scientific*. So **Falsificationism** is false.