Lecture Notes for *Central Problems of Philosophy*

J. Dmitri Gallow

Draft from Spring, 2026
Taught at the University of Southern California

Contents

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's a conflict in our views about *freedom* and *scientific inquiry*.

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

      i. E.g., if you are under hypnosis, you aren't free to do otherwise. And if there's a gun to your head, you aren't free to do otherwise.

   (b) We assume in scientific inquiry that *every* event has a cause—and, in particular, that our own decisions have causes. These causes trace back to events in the distant past, before we were born.

2. These thoughts seem to conflict with each other. Consider any arbitrary behavior of yours.

   P1. If you're not free to change something, then you're not free to change a necessary consequence of that thing.

   P2. Your behavior was a necessary consequence of the distant past and the laws of nature.

   ∴ C1. So, if you're not free to change the distant past or the laws of nature, then you're not free to change your behavior, either.

   P3. You are not free to change the distant past or the laws of nature.

   ∴ C2. So you are not free to change your behavior.

3. The *libertarian* rejects the premise P2. They deny that our actions are caused in the way that other physical processes are caused.

   - A challenge for the libertarian: their position looks to be in conflict with modern physics.
   - Another challenge: being *uncaused* doesn't seem to make actions free. Being uncaused makes actions *inexplicable*, not freely chosen.
     - E.g., Mother Teresa—for literally *no reason*, without *any cause*—throws a hand grenade into an orphanage. This doesn't look like a free or blameworthy act.

   ▷ A Libertarian response to this challenge: free acts *are* caused, but they're caused by *us*. There's a special kind of causation, *agent causation*, which is different from the kind of causation found in the physical world.

4. The *hard determinist* accepts the conclusion C2. They say that science has taught us that none of our actions are freely chosen.

   (a) A challenge for the hard determinist: we generally only praise or blame people when they could have done otherwise. If no one could have done otherwise, what becomes of our practices of praising and blaming?

   C2. You were not free to change your behavior.

   P4. If you were not free to change your behavior, then you are not to blame for that behavior.

   ∴ C3. So you are not to blame for your behavior.

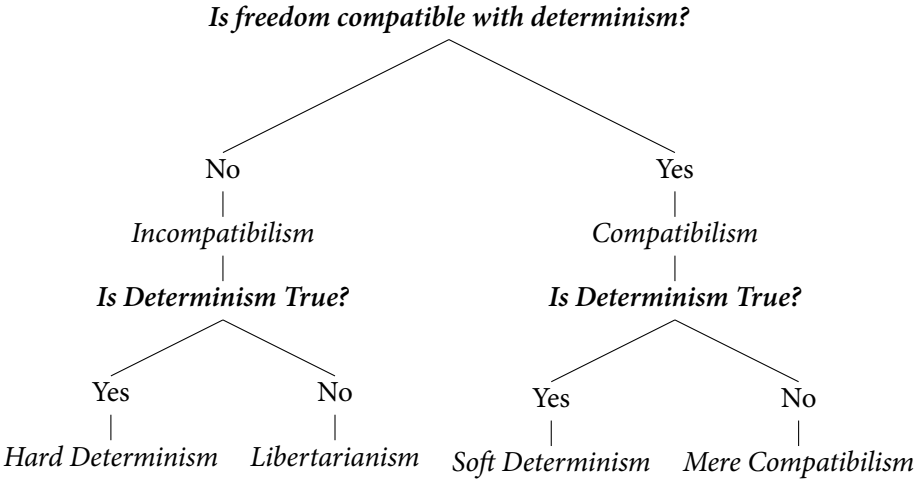
- ▷ The hard determinist could just say: no one is ever blameworthy, and no one is ever praiseworthy for anything they ever do. That's a hard attitude to maintain.
- ▷ Another response: find some *surrogate* notion—call it 'schreedom'—which plays the role we once thought freedom played: schreedom is needed for you to deserve praise or blame, *e.g.*, and schreedom is inconsistent with certain kinds of compulsion.

5. The *soft determinist* accepts P2, but denies P1. thinks that the hard determinist's 'schreedom' just is freedom. They think that you can be free to change your behavior, even if that behavior is a necessary consequence of something you're not free to change.

   (a) 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.

   - Proposal #1: An act is free if it is caused *by your beliefs and desires*.
   - ▷ Objection: The addict smokes cigarettes because of their desire to do so. But still, they are not free—their smoking is a *compulsion*.
   - Proposal #2: An act is free if it is caused by beliefs and desires which were freely chosen.
   - ▷ Objection: this proposal leads to an infinite regress.
   - Proposal #3: You are free to do otherwise if: you would have done otherwise, if you had tried.
     - — We will come back to this proposal later on.

6. A taxonomy:



What is 'moral responsibility'? 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. P1 says that, in order for a choice to be free, we must have been able to do something else. Call this the *Principle of Alternative Possibilities* (PAP_F).

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

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

   (b) In further support of the PAP_F, 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 did not *freely choose* to arrive in Rome. Why? Because there wasn't ever any *possibility* that he would *not* arrive in Rome.

3. Frankfurt affords us an argument against PAP_F. He gives 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.

4. In this story:

   (a) Jones killed Smith freely; yet

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

   So: PAP_F is false.

5. An argument that *Determinism* undermines moral responsibility (praise- and blameworthiness):

   P3. You deserve blame for some behavior only if it was possible for you to change your behavior.

   P4. If *Determinism* is true, then it is not possible for you to change any of your behavior.

   C2. If *Determinism* is true, then you not deserve blame for any of your behavior.

6. P3 says that, in order for you to be morally responsible for some behavior, you must have been able to do something else. Call this 'PAP_M' (the subscript is for *m*oral responsibility).

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

7. Frankfurt thinks that the case of Dr. Black *also* shows us that the PAP_M is false. For he thinks that

   (a) Jones deserves blame for killing Smith; yet

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

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

   You are not free to do change your behavior

   And we were blithely supposing that, if you're not free to *change* your behavior, then that behavior couldn't be chosen *freely*. But perhaps Frankfurt has given us reason to doubt this? Perhaps Jones was not free to *change* his behavior; but nonetheless, that behavior was chosen freely?

What is well-being? What is a philosophical theory of well-being attempting to do? What does the theory of hedonism *say about well-being? Present an argument from Epicurus for the conclusion that death is not a harm, and explain how hedonism could be used to justify one of its premises.*

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?

   ▷ 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.

   C. No one is harmed by their own death.

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

   ▷ 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 you to be harmed by something, you must exist at the same time as the thing that harms you.

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

   C. 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.

   C. No one is harmed by their own death.

Hedonism

4. Epicurus and Lucretius were *hedonists*, and their views about death were informed by their hedonism. Hedonism is a view about *well-being*. We'll discuss well-being, and then we'll discuss what hedonism says about well-being.

5. Distinguish two questions: 1) whether something is good; and 2) whether something is good *for you*. Things could be good without being good *for you*; and they could be good *for you* without being good.

   ▷ It was good that the Allies defeated the Axis powers in WWII, but it wasn't good *for Hitler*.

   ▷ The restoration of the Bourbon monarchy in France was bad, but it was not bad *for Charles X*.

6. *Well-being* is the term philosophers use for what's *good for you*. If something is good for you, then it raises your level of well-being. Bad for you, and it lowers your level of well-being.

   ▷ A *harm* is just something that lowers your level of well-being; a *benefit* is just something that raises your level of well-being.

   ▷ A punishment should lower your well-being; a reward should raise it.

   ▷ If you are selfish, then you care more about your own well-being than others' well-being. If you are selfless, you care more about others' well-being than your own.

   ▷ When you love someone, you want their well-being to be higher.

7. Distinguish two ways something could be good for you: it could be *instrumentally* good for you—meaning that that thing will help you achieve other things which are good for you. Or it could be *intrinsically* good for you. If something's intrinsically good for you, then it is good for you *in and of itself*.

   ▷ Often, money is instrumentally good for you. You can use it to make your life better. Even so, no one thinks that having money *just is* having things go well for you. It's possible to have money and have things go badly for you.

8. Philosophers who think about well-being and interested in the question of what things are *intrinsically* good for you. The Hedonist has a simple answer to this question: *pleasure* (or *happiness*).

   Hedonism The only thing which is intrinsically good for a person is pleasure. The only thing which is intrinsically bad for a person is pain.

   ▷ As the hedonist understands pleasure/happiness, *any* enjoyable experience is pleasurable. (Proving a theorem, listening to music, reading a crime novel—all these count as pleasurable.) And likewise, any experience which is distressing, upsetting, or unpleasant is *painful*. (Boredom, anger, heartbreak, and so on will all count as painful.)

9. According to the hedonist, only *experiences* are good or bad for you.

   ▷ Precisely because death involves no experiences at all, Epicurus thinks that it cannot be bad for you.

   ▷ If we accept hedonism, then we will straightaway accept P1 and P3. For the only things which can harm you are your own painful experiences, and you will always be *aware* of these, and they will never exist when you do not.

What is well-being? What is a philosophical theory of well-being attempting to do? What does the desire satisfaction *theory say about well-being? Present an argument from Epicurus that death is not a harm, and say how Thomas Nagel responds to that argument.*

Death is a Harm

1. Nagel thinks that the arguments from Epicurus and Lucretius are too quick. He wishes to respond to the arguments 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.

2. In response to the argument **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.

   C. No one is harmed by their own death.

   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.

3. In response to the argument **No Harm without Existence**,

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

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

   C. No one is harmed by their own death.

   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.

4. In response to the argument **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.
>
> ---
>
> C. No one is harmed by their own death.

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.

Desire Satisfaction

5. Nagel's objections could be further motivated with a rival theory of well-being:

Desire Satisfaction Theory The only thing intrinsically good for a person is getting what they want. The only thing intrinsically bad for a person is not getting what they want.

6. We can illustrate the difference between these theories with a thought experiment from Robert Nozick. Contrast two different possible worlds:

- ▷ Real World: all of your experiences are accurate. The things you see are really there, the people who you love exist, and they love you back. You achieve great things, and your achievements affect and inspire others.
- ▷ Fake World: since birth, you have been plugged into an *experience machine*, which simulates all of your experiences. While plugged into the machine, you undergo all of the same experiences you had in World 1. But none of these experiences are accurate. The things you see aren't really there. The people who you love do not really exist; they do not love you back. You only seem to achieve things, and these apparent achievements affect no one but yourself.

Nozick says that things are *better* for you in Real World than they are in Fake World. But if hedonism were correct, then things would be *equally* good for you in Real World and Fake World. So Nozick concludes that hedonism is not correct.

The desire satisfaction theorist, on the other hand, can say that Fake World is worse for you than Real World. Assume that you want for others to love you back, you want your achievements to affect others, and so on. Then, in Real World, you are getting less of what you want.

Why does Descartes worry about his beliefs? What is an a priori *belief, and what is an* a posteriori *belief? What is Descartes' 'method of doubt'? Which beliefs does the method of doubt lead Descartes to set aside, and why?*

The Meditations

1. Descartes is bothered by the fact that many things he once thought were true have turned out to be false. He decides to sit down in meditation and consider which of his opinions are truly trustworthy and free from doubt.
   - ▷ So he decides to reject outright any opinion which he can subject to any doubt. ("I shall proceed by setting aside all that in which the least doubt could be supposed to exist, just as if I had discovered that it was absolutely false.") This is his *method of doubt*.
   - ▷ His ultimate goal is to find some belief which cannot be subjected to doubt; and he hopes that, if such a belief can be found, it can serve as a *sure foundation* for all of science.
2. Rather than consider his opinions one by one, he instead decides that he will consider *their foundations*.
   - ▷ Descartes is assuming that, for each belief we hold—or at least, for any belief which we *justifiedly* hold—we hold is for some *reason*. Some of these beliefs are at the *foundation*; they are the beliefs upon which all others are built.
   - ▷ One foundational source of belief is *the senses*—things justified directly by experience.
   - ▷ Philosophers call a belief based in sense experience *a posteriori*. (Think: it is *after* or *posterior to* sense experience.) But it doesn't seem that all beliefs are *a posteriori*. There are also *mathematical* beliefs. These beliefs are called *a priori*, insofar as they can be justified *prior to* and *independent of* experience.
3. Descartes starts by raising a reason to doubt his *a posteriori* beliefs which are based upon sense experience.
   - ▷ He notes that he has often, in dreams, felt convinced of falsehoods. So, he thinks, he might well be dreaming *now*. He may not be sitting in his chair by the fire, but instead laying in bed asleep.
4. Even so, he thinks, this is no reason to doubt *a priori* beliefs like Arithmetic or Geometry. ("Whether I am awake or asleep, two and three together always form five.") But, he thinks that this, too, can be doubted:
   - ▷ "how do I know that I am not deceived every time that I add two and three, or count the sides of a square"?
5. He next imagines that there is an "evil genius not less powerful than deceitful [who] has employed his whole energies in deceiving me". In such a situation, Descartes could not trust any of his sense experience, nor any of his *a priori* beliefs about Arithmetic and Geometry.
   - ▷ *But*—Descartes notes that he could *not* be deceived in his belief that he exists! "I exist also if he deceives me; he can never cause me to be nothing so long as I think that I am something". Similarly, while we can be deceived about whether something is green, we cannot be deceived about whether it *appears* to be green.
   - ▷ From here, Descartes goes on to use his sure and certain knowledge that he exists to establish things about himself. For instance, since he can doubt that he exists, but not that his body exists, Descartes concludes that he must be distinct from his body. We will leave Descartes to it, and consider the kinds of arguments he turned up in his first two meditations.

Skeptical Arguments

6. 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*.

7. 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? Following Descartes, let's say that, at base, our evidence about the external world is our sensory experience.

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

8. **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.

9. 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.

10. The philosophical skeptic could proceed piecemeal, arguing against every one of your beliefs about the external world with arguments like these. Or, they could proceed like Descartes, and strike at the foundations—pointing to a *dream* scenario, or the case of an evil demon, or a brain-in-a-vat.

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. For each of the arguments we considered before, 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).

2. 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. Take any belief about the external world, p; then:

Possibility of Error Argument

P1. Your belief that p could be mistaken—you could be a brain-in-a-vat.

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

C. You don't know that p

Certainty Argument

P3. You're not absolutely certain that p—*after all, you're not certain that you're not a brain-in-a-vat*

P4. Knowledge must be absolutely certain.

C. You don't know that p

Indistinguishability Argument

P5. The case in which you are a brain-in-a-vat is indistinguishable from the case in which p

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

C. You don't know that p

Transmissibility Argument

P7. You don't know that you're not a brain-in-a-vat

P8. If you don't know that you're not a brain-in-a-vat, then you don't know that p.

C. You don't know that p.

3. 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. In general, ignorance is taken to *excuse* otherwise blameworthy bahavior. But, if the philosophical skeptic is correct, then we have a *universal* excuse.

4. 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**.

Fallibilism

5. 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.

6. Here's a version of fallibilism to consider:

   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.

   - ▷ E.g., in order to know that that's a zebra, your evidence must rule out the possibility that it's a horse, a giraffe, an elephant, and so on. But you don't have to rule out the possibility that it's a painted mule. That possibility is *irrelevant*.
   - ▷ Likewise, in order to know that Jack has a nut allergy, your evidence has to rule out possibilities in which he truthfully claims to not have a nut allergy. But you don't have to rule out possibilities in which he systematically lies about it; nor do you have to rule out possibilities in which you are a brain-in-a-vat. These possibilities are *irrelevant*.

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

   - ▷ ...you could be mistaken about whether he has a nut allergy;
   - ▷ ...you are not absolutely certain whether he has a nut allergy; and
   - ▷ ...your experience is indistinguishable from one in which you don't know that he has a nut allergy.
   - ▷ ...you don't know that Jack isn't lying, nor that you're not a brain-in-a-vat.

   It's worth focusing on this last response for a bit:

   - ▷ When it comes to your belief that Jack isn't lying, the possibility that he *is* lying is a relevant alternative possibility. Since your evidence doesn't rule this possibility out, you don't know that Jack isn't lying.
   - ▷ However, in order to know that Jack has a nut allergy, this possibility is not relevant. Since your evidence does rule this possibility out, you know that Jack has a nut allergy.

   So the relevant alternative fallibilist denies the following principle:

   Closure If you know that p, and there's no way for p to be true while q is false, then you know that q

Externalism about Evidence

7. The *externalist about evidence* denies **Evidence as Experience**.

   - ▷ 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*.
   - ▷ 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:

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

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.

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.

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'.

   (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*: 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*: 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.

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. 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 more unhappy 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.

3. 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

4. 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.

5. 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.

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

   (a) Emend the system of rules so that it includes an 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. So why doesn't rule utilitarianism just end up 'collapsing' back down to regular (act) utilitarianism?

7. Here's another way of developing the 'what if everyone did that?' 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.

8. 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*.)

Describe Peter Singer's 'shallow pond' example, and explain how it relates to his argument that we have a moral obligation to give substantial amounts of money to aid organizations.

The Shallow Pond

1. Peter Singer asks you to consider the following thought experiment:

   > *The Shallow Pond*: You are on your way to an important job interview, when you see a small child drowning in a shallow pond. You can wade into the pond and save the child from drowning, but it would ruin your expensive loafers.

   Singer thinks that you morally ought to rescue the drowning child, even though it would cost you the expensive loafers.

2. From reflection on this case, he extracts two general principles regarding our duties to aid—one weaker and one stronger:

   Weak Principle of Beneficence If it is in your power to prevent something bad from happening without sacrificing anything morally significant, you ought, morally, to do it.

   Strong Principle of Beneficence If it is in your power to prevent something bad from happening without sacrificing anything of comparable moral importance, then you ought, morally, to do it.

   Since your expensive loafers are not morally significant, the Weak Principle of Beneficence says that you ought to save the child. And since your expensive loafers are not *as* morally significant as the child's life, the Strong Principle of Beneficence says that you ought to save the child's life.

   ▷ The difference between the two principles has to do with *how much* you ought to sacrifice in order to prevent something bad from happening. Suppose that, in order to save the child, you would have to sacrifice your finger (which is morally significant, but not *as* morally significant as the child's life). Then, the weak principle doesn't say that you have to save the child, but the strong principle does.

Obligations to Give Aid

3. At the time Singer was writing, there was a famine in East Bengal (present day Bangladesh) affecting at least nine million people. Today, there are over 318 million people facing acute hunger,[1] with more than 800 million suffering from chronic undernutrition,[2] even though daily calories per person per day is in excess of 2500 in every region of the world.[3]

4. In the west, we think of charity as morally *optional*—something which it may be nice to do, but which is above and beyond, and not required. Peter Singer thinks that this attitude is mistaken. In our current situation, giving nothing is *wrong*.

5. Singer offers the following argument for this conclusion:

[1] `https://www.wfp.org/global-hunger-crisis`

[2] `https://www.fao.org/4/x0262e/x0262e05.htm`

[3] `https://ourworldindata.org/grapher/daily-per-capita-caloric-supply?country=OWID_NAM~OWID_AFR~OWID_EUR~OWID_ASI~OWID_SAM`

P1. Suffering and death from lack of food, shelter, and medical care are bad.

P2. If it is in your power to prevent something bad from happening without sacrificing anything morally significant, you ought, morally, to do it. (The Weak Principle of Beneficence)

P3. It is in your power to prevent suffering and death from lack of food, shelter, and medical care, by donating money to aid organizations.

C1. You ought, morally, to donate money to aid organizations.

- ▷ Singer is happy to take P1 for granted without argument; we will do so, too.
- ▷ It is an empirical question whether P3 is true. There are political scientists who think that aid organizations have in fact made matters worse in many parts of the world. Today, Peter Singer's organization *Giving What We Can* advises that people give to specific charities aimed at providing malaria nets, developing medicine to fight malaria, providing vitamin A to children, and providing vaccines to children.[4]
- ▷ How *much* should you donate? According to the Weak Principle, you should donate as much as you can, until donating more would require sacrificing something morally significant. Since coffee, clothes, and many other luxuries are not morally significant, it requires you to forgo all of these luxuries, and pass the savings on to those who need the money more.

  According to the Strong Principle, you should donate as much as you can, until donating more would require sacrificing something *as* morally significant as lack of food, shelter, and medical care. You should continue giving up to the point where giving more would leave you worse off than those who stand to benefit from your donations (what Singer calls 'the point of marginal utility').

6. An objection to P2:

   P4. If everyone in my position gave $6 to the Against Malaria Foundation, we would have enough nets for everyone who needs one.

   P5. We are only morally required to give enough to provide nets to everyone who needs one.

   P6. I am not morally required to give more than other people are morally required to give.

   C2. I am not morally required to give more than $6 to the Against Malaria Foundation.

   Since P2 says that I *do* have an obligation to give more than $6—since doing so would prevent suffering and death without sacrificing anything morally important—P2 is false.

   - ▷ Singer: this argument is invalid. What follows is only this: *if* everyone in my position gave $6 to the Against Malaria Foundation, *then* I would not be required to give more than $6. But since it's not true that everyone in my position is giving $6, I am required to give more than $6.
   - ▷ An analogy: you and another person walk by a shallow pond in which two children are drowning. Only one child can be saved at a time, and each time a child is saved, it will ruin a new article of clothing. You save one, but the other person stands by idly doing nothing. Given that the other person has done nothing, you ought, morally, to save the other child.

[4] `https://www.givingwhatwecan.org/best-charities-to-donate-to-2026`

What is Singer's argument that we have a moral obligation to provide substantial amounts of money to aid organizations? How does Arthur object to this argument? What aspects of morality does he think Singer's argument has neglected?

1. Arthur rejects Singer's Strong Principle of Beneficence.

   Strong Principle of Beneficence If it is in your power to prevent something bad from happening without sacrificing anything of comparable moral importance, then you ought, morally, to do it.

   He begins by pointing out that it has some very counterintuitive consequences:

   - ▷ You could prevent a stranger from dying by donating your kidney. Your second kidney is not as morally important as their life. So the principle says you ought to donate your kidney.
   - ▷ One farmer is industrious and grows more than enough plants to feed him and his family. A second farmer is lazy and does not grow enough for him and his family. The principle says that the first farmer ought, morally, to donate his surplus to the second farmer.

2. He thinks that Singer's principle has these consequences because it neglects two important aspects of common-sense morality: *moral rights* (note: not *legal* rights; but *moral* rights) and *desert*.

Moral Rights

3. Some paradigm moral rights: the right to life, the right to privacy, & the right to free expression.

4. Some of your rights concern *your own* action or inaction—they permit you to do or refrain from doing something. Other of your rights concern the action or inaction of *others*—they require others to do or refrain from doing something. For instance, some think that you have a right to free speech, which involves in part your right to speak and in part your right that *others* not punish you for speaking. We can call a right of yours which concerns your (in)actions a *privilege* right, and a right or yours which concerns other's (in)actions a *claim* right.

5. Rights also divide into *negative* and *positive*. A negative right concerns (your or other's) *inaction*, whereas a positive right concerns (your or other's) *actions*. An example of each is shown below (you might deny that these are moral rights, and you might deny that there are any moral rights in one of these categories; we're just illustrating the kinds of rights involved in each section).

	Negative	Positive
Privilege	Right to silence	Right to free speech
Claim	Right to privacy	Right to healthcare

6. Arthur also thinks that positive claim rights depend upon the existence of some kind of promise or contractual agreement. If I have made you no promises to provide you with healthcare, or signed no contract to provide you with healthcare, then you do not have a *right* that I provide you with healthcare.

   - ▷ So, in particular, people in need do not have a positive claim right that we provide them aid or rescue.

- ▷ Arthur thinks that we *can* have obligations or duties to give aid, but he doesn't think that these duties are grounded in anybody's *rights*.
- ▷ However, he thinks that we can invoke our own rights in order to justify not helping *distant* stranger, or when the cost to us of helping is substantial. You have a right to your own eye and kidney, and so you're under no obligation to donate them.

Desert

7. Arthur also think that Singer's principle ignores what people *deserve*—or, as philosophers refer to it: *desert*.
   - ▷ Desert can be *positive*, when what you deserve is something beneficial. Or it can be *negative*, when what you deserve is punishment.
8. In the two farmers case, Arthur thinks that the first farmer *deserves* their excess crop, whereas the second farmer does not, and that this is at least a factor to be taken into consideration when considering whether he ought to donate the surplus.
9. Arthur calls rights and desert *entitlements*, and he thinks that we should accept what common-sense morality says about how entitlements interact with the duty to give aid: if helping would require a substantial cost, you are entitled to not give. But if helping would only require a trivial cost, then you are required to help.

Moral Reform?

10. In describing these views about moral rights and desert, and how they relate to our duties to give aid, Arthur takes himself to be describing common-sense morality.
11. He suspects that Singer will respond that common-sense morality is in need of *reform*. What was common-sense morality in ages past is nowadays rejected as an inferior moral code. Similarly, Singer can suggest that our current moral code is inferior and in need of replacement.
12. Arthur gives two defenses of common-sense morality.
    - ▷ In the first place: he suggest that caring about desert is a way of caring about *fairness* and *justice*—requiring the productive farmer to give to the lazy one would be *unfair* and *unjust*. Likewise, caring about people's rights is a way of showing them *respect*.
    - ▷ In the second place: Arthur suggests that we should only accept a moral code that could be supported by almost everyone. "Rules that would work only for angels are not the ones it is rational to support for humans". Since he thinks most people cannot live by Singer's proposed moral code, this is a reason to reject it.

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.

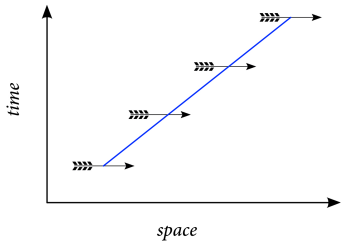
3. Zeno of Elea (5th century BC) thought that motion was impossible. He provided several arguments for this surprising conclusion. Let's consider just one of them:

   (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 at a time is not to be *instantaneously* in motion at that time. Rather, all 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 at Tuesday, and so on.

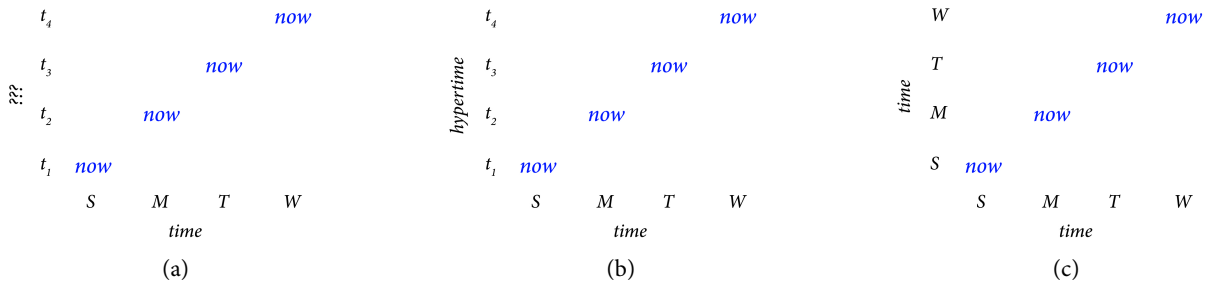


Figure 1

(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.

(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. 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.)

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

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, Tenet): 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 2024 to the year 2034. 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, the time separating Tim's departure from his arrival is five minutes.

   P2. If Tim's journey is possible, 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 (or God, or the infinite past) has no explanation.In normal universes, only the initial causes lack explanations. 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 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 can change the past.

   P8. Nobody can change the past.

   C3. Time travel is not possible.

   (a) 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.

   (b) 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 100 · 3rd March, 2026

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 an objection to Mackie's argument, and explain how Mackie responds 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 he knows about all preventable evil.

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

   P3. If God exists and is omnibenevolent, then he wants to prevent all preventable evil.

   C1. If God exists and is omniscient, omnipotent, and omnibenevolent, he will 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 considers several objections.

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. So, if God were omnibenevolent, he would want to permit some evil as a necessary means to there being good.

   ▷ 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.

      ▷ 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 he cannot lift 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.

   ▷ 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, he 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.

- ▷ 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.
- ▷ 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.

   - ▷ 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.
     - ▷ 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.
     - ▷ 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.

   - ▷ 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?
   - ▷ 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.
   - ▷ 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.
   - ▷ 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 100 · 4th March 2026

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.
   - ▷ If free will is compatible with determinism, then God could have created us so that we always choose what's best.
   - ▷ 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?
   - ▷ 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.
   - ▷ 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 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*.
   - ▷ 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.
     - ▷ 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*.
     - ▷ Perhaps: 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.

- ▷ If free will is compatible with determinism, then God could have created us so that we always choose what's best.
  - ▷ *Reply*: Free will is not compatible with determinism. If we are to be truly free, then God cannot decide *for us*.
- ▷ Why would *shreedom* be worse than *freedom*?
  - ▷ *Reply*: Because only freedom allows for genuine and consequential responsibility.
- ▷ Why would God allow *natural evil*?
  - ▷ *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.
    - ▷ *Retort*: But why couldn't God simply *tell us* how to seriously harm our fellow humans?
    - ▷ *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.
  - ▷ *Reply 2*: The presence of natural evil allows for the possibility of *second order* goods like courage, perseverance, benevolence, *etc*.
- ▷ God could allow free choice without allowing every *consequence* of those choices.
  - ▷ *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.

   - ▷ *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.
     - ▷ 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.
   - ▷ *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. 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.

Introduce the 'puzzle of the statue and the clay'. Explain which assumptions lie behind the puzzle. And explain one way of rejecting (one of) these assumptions.

The Puzzle of the Statue and the Clay

1. Michelangelo took a lump of clay and sculpted it into a statue of David. After the statue is sculpted, there is still a lump of clay (now statue-shaped). And there is also a statue, composed out of the clay. Let's call the lump of clay which exists post-sculpting 'Lump'. And let's call the statue 'David'.

   - ▷ Here's an argument that David and Lump must be distinct: Lump existed outside of Michelangelo's studio, before Michelangelo started sculpting, but David did not. One and the same thing cannot have two different properties. So David and Lump must be distinct.
   - ▷ This argument appeals to a principle of metaphysics, known as *Leibniz's Law*:

     Leibniz's Law For any x and y: if x has a property which y lacks, then $x \neq y$.
   - ▷ In this principle, a *property* is just anything which is true of an object. If we can truly say 'Lump is made of clay', then Lump has the property of being made of clay. Since 'Ishmael is made of clay' is false, Ishmael lacks this property. (So, Leibniz's Law says: Ishmael and the clay must be distinct.)
   - ▷ Since we can truly say 'Lump was once outside of Michelangelo's studio', Lump has the property of having once been outside of Michelangelo's studio. But we cannot truly say 'David was once outside of Michelangelo's studio'. So David lacks this property. By Leibniz's Law, David and Lump must be distinct.
   - ▷ If Lump and David are distinct, then it looks like it's possible for two things to be in the same place at the same time, and to be made up of all the same matter!

2. This conclusion can seem absurd. Can we resist it? Let's think of the argument as a collection of inconsistent claims. (As we learned to think about arguments before the midterm.) If we take Leibniz's Law for granted, then these claims are inconsistent.

   Creation Michelangelo really did create David—that is, David exists but did not exist *before* the sculpting.

   Survival Lump existed before the sculpting, and Michelangelo did not *destroy* Lump by forming it into a statue.

   Composition If David exists after the sculpting, then David is made up of the clay, and exists wherever the clay does. And if Lump exists after the sculpting, then Lump is also made up of that clay, and exists wherever the clay does.

   Absurdity It is not possible for distinct objects to share the same matter and spatial location at one time.

   After all, by *Creation*, David did not exist before Michelangelo sculpted it. By *Survival*, Lump existed before Michelangelo sculpted David. By Leibniz's Law, David $\neq$ Lump. By *Survival* and *Creation*, both Lump and David exist after David is sculpted. So, after the sculpting, there are two objects in existence at the same time, Lump and David. By *Composition*, Lump and David share the same matter and location at one time. But this contradicts *Absurdity*.

The Just-Matter Theory

3. Perhaps we should reject *Creation*. Perhaps Michelangelo didn't *create* David. Rather, he simply *rearranged* Lump to make it David-shaped, but this rearrangement didn't make any new objects.

    ▷ According to the *Just-Matter Theory*, the only things that exist are hunks of matter. Rearranging hunks of matter doesn't create any new objects. If matter is neither created nor destroyed, then nothing ever comes into or goes out of existence.

    ▷ Sider: the *Just-Matter Theory* has unacceptable consequences. In particular: it says that Socrates exists today, since the matter that composed him is still around (though dispersed).

    ▷ Actually, it's not like there was *just one* hunk of matter which constituted Socrates. The matter in everyone's body is constantly changing, and is entirely replaced every 5 years. So the *Just-Matter Theory* says that you are not the same person you were five years ago—though that thing (the person you were five years ago) still exists. This looks like an unacceptable consequence of the view.

The Takeover Theory

4. Perhaps we should reject *Survival*. Perhaps, when Michelangelo created David, he thereby *destroyed* Lump.

    ▷ Some background: each hunk of matter will belong to many different *kinds*. For instance, outside the studio, the matter making up Lump and David belonged just to the kind *lump of clay*. Inside the studio, it belonged to *two* kinds: *lump of clay* and *statue*.

    ▷ According to the *Takeover Theory*, a hunk of matter can only constitute one object of a given kind at a time. When a hunk of matter belongs to many different kinds, only one of them will be *dominant*. And there will only be a single thing that matter constitutes, corresponding to that dominant kind. For instance, when the clay belongs to both the kinds *lump of clay* and *statue*, the kind *statue* is dominant. So, the matter which is David-shaped constitutes a statue, but does *not* constitute a lump of clay.

        ▷ The dominant kind will determine which changes an object can and cannot survive. Since statues must be statue-shaped, David cannot survive being squashed. If David is squashed, the *statue* will stop being the dominant kind, and *lump of clay* will again be the only dominant kind—so the statue will go away and a lump of clay will take its place.

    ▷ Sider: the takeover theory seems athropocentric, since the *kinds* that it's using in its theory are the kinds into which humans typically sort hunks of matter. Imagine that Martians sort matter into *inclays* and *outclays*. An *inclay* is clay located indoors, whereas *outclay* is clay located outdoors. A Martian takeover theorist would say that taking clay outside destroys one thing and creates another. The Earth takeover theorist has no explanation of why reality cares about *our* kinds and not *Martian* kinds.

Introduce the 'puzzle of the statue and the clay'. Explain the solution to the puzzle Sider calls 'Cohabitation', and raise an objection to Cohabitation. Explain how Sider's Space-Time Theory *allows him to respond to this objection.*

The Puzzle of the Statue and the Clay

1. Recall: assuming Leibniz's Law, the following claims are inconsistent.

   Creation Michelangelo really did create David—that is, David exists but did not exist *before* being sculpted.

   Survival Lump existed before Michelangelo sculpted it, and Michelangelo did not *destroy* Lump by forming it into a statue.

   Composition If David exists after being sculpted, then David is made up of the clay, and exists wherever the clay does. And if Lump exists after David is sculpted, then Lump is also made up of that clay, and exists wherever the clay does.

   Absurdity It is not possible for distinct objects to share the same matter and spatial location at one time.

   Last time, we considered a way of rejecting *Creation*—the Just-Matter Theory. And we considered a way of rejecting *Survival*—the Takeover Theory. And we saw reasons to worry about each of these reactions. Let's press on.

Nihilism

2. Perhaps we should reject *both Creation and Survival*. Perhaps we should say that *neither* David *nor* Lump exist at any point in time.

   ▷ The *Nihilist* says that there are no such things as statues or lumps of clay. According to them, there are only *particles*—things without any smaller parts. Some of these particles are arranged statue-wise. But there is no statue.

   ▷ Sider: this isn't as absurd as it might seem. If there were no statue, and only particles arranged statue-wise, surely everything would look exactly the same as it currently does. So we can't easily dismiss nihilism

   — Would it look that way *to us*? If we don't exist, then surely *nothing* looks any particular way to us. Notice that the nihilist denies the one thing Descartes thought was beyond doubt: that you exist.

   ▷ Sider: we should, however, reject nihilism. For it could be that *everything* has smaller parts. That is, it could be that atoms are made of protons and electrons, which are made of quarks, which are made of tiddlywinks, which are made of mini-kebabs... and so on and so forth, without end. (Matter which is is always further divisible in this way—matter in which every part of the matter has a still smaller part—is called *gunk*.) The nihilist says that, if everything were made of gunk, then *nothing* would exist—since, if, everything were made of gunk, then everything would have smaller parts. And the nihilist denies the existence of anything with smaller parts. But surely, it's at least *possible* for there to be gunk. And surely, in such a possibility, *something* exists.

Cohabitation

3. Perhaps we should reject *Absurdity*. Perhaps we should say that David and Lump are just two different objects, residing in exactly the same place at the same time, and sharing exactly the same matter at that time. Call this theory *Cohabitation*.

4. Our question this week is: what is there in Michelangelo's studio? We've considered three answers so far: 1) *only Lump* (the Just Matter Theory); 2) *only David* (The Takeover Theory); and 3) *neither David nor Lump* (Nihilism). Cohabitation gives the answer *both Lump and David*.

5. Sider considers two objections to *Cohabitation*

   ▷ *Objection #1*: Suppose Michelangelo smushes the statue in his hand. According to *Cohabitation*, this destroys David, but it does not destroy Lump. So David must be much more *fragile* than Lump is. But how could this be? David and Lump are made up of the very same matter—David should therefore be just as durable as Lump. Since *Cohabitation* implies that how durable an object is depends upon more than the matter it is made up of, *Cohabitation* must be false.

   (a) *Objection #2*: a whole is nothing more than the sum of its parts. So any two wholes which have all the same parts must be identical. There's an area of metaphysics known as *mereology* which studies the nature of parthood. The following is an axiom of the classical theory of mereology:

   Extensionality If all the smaller parts of x are smaller parts of y, and all the smaller parts of y are smaller parts of x, then $x = y$.

   David and Lump have all the same smaller parts. So by *Extensionality*, they must be identical. So *Cohabitation* must be wrong.

6. Sider thinks that we can respond to both of these objections if we *also* accept the *Space Time Theory* we learnt about earlier in the course.

   ▷ According to the Space Time Theory, both David and Lump are four dimensional worms. Since Lump existed before David was sculpted, and continues to exist after David is smooshed, David is a *temporal part* of Lump. So Lump and David *do not* share all the same smaller parts. Lump has temporal parts which David does not have. So the second objection made a false assumption about the parts of Lump and David. It neglected their temporal parts.

   ▷ According to the Space Time Theory, the relationship between David and Lump is like the relationship between Santa Monica Blvd and California State Route 2. Santa Monica Blvd is the *part* of State Route 2 which exists inside of Los Angeles. But State Route 2 extends much further east than Santa Monica Blvd does. Nobody asks why Santa Monica Blvd is more fragile than State Route 2—this is just a consequence of us using 'Santa Monica Blvd' for the segments of State Route 2 which lie inside Los Angeles. And similarly, David going out of existence when its smooshed is just a consequence of us using 'David' for the temporal parts of the clay which are state-shaped.

7. A worry about Sider's defense of *Cohabitation*: suppose that God creates a statue of Goliath *ex nihilo*, and then, after displaying it for five minutes, destroys the entire thing, leaving no matter behind at all.

   ▷ During these five minutes, there is a lump of clay—call it 'Lumpl'. And there is also a statue—call it 'Goliath'.

   ▷ Goliath could not have survived being smooshed. But Lumpl could have survived being smooshed. So Lumpl has a property which Goliath lacks. So, by *Leibniz's Law*, Lumpl $\neq$ Goliath.

   ▷ But Lumpl and Goliath are precisely the same spacetime worm, so they have all the same spatial and temporal parts.

What is the inference rule modus ponens*? Justify the use of* modus ponens *by using truth-tables. What is the inference rule known as* affirming the consequent*? Explain why this inference rule is bad using truth-tables.*

1. Consider the following two ways of reasoning:

If the Butler did it, then the Gardener didn't.	If the Butler did it, then the Gardener didn't.
The Butler did it.	The Gardener didn't do it.
The Gardener didn't do it.	The Butler did it.

   (a) It seems as though the first way of reasoning is good. If both premises are true, then the conclusion *must* be true.

   (b) On the other hand, it seems like the second way of reasoning is bad. Maybe the Gardener didn't do it, but it doesn't *necessarily follow* that he didn't. If we know that there was only one murderer, then the first premise seems true. And if we know that the Gardener was at home at the time of the murder, then it looks like we know that the second premise is true. But it could still be the maid. So the conclusion doesn't *necessarily follow*.

2. *Logic* is the field of philosophy which studies arguments, or ways of reasoning, like these. And logic can help us understand why one of these inferences is good and one is bad. For logic teaches us that the first argument has a *valid* form, and the second one does not. And, if an argument has a valid form, then, if the premises are true, the conclusion will be true as well.

3. First, let's talk about the notion of an *argument form*. The two arguments above have the following forms:

If p, then q.	If p, then q.
p.	q.
q.	p.

   (a) Here, I've used 'p' and 'q' as placeholders for any arbitrary sentence. Notice that, if we let $p =$ 'the Butler did it', and let $q =$ 'the Gardener didn't do it', then we get back our original arguments. (Of course, there are many other arguments which have these forms; we get those other arguments by plugging in different sentences for p and q.)

   (b) The inference rule *modus ponens* licenses the first way of reasoning. An *inference rule* tells you which kinds of inferences are good. The rule of *modus ponens* says: if you know that 'if p, then q' is true, and you know that 'p' is true, then you may *infer* that 'q' is true, as well.

   (c) The inference rule *affirming the consequent* licenses the second way of reasoning. This rule says: if you know that 'if p, then q' is true, and you know that 'q' is true, then you may *infer* that 'p' is true, as well.

4. Next, the notion of *formal validity*. An argument form is *valid* iff any argument which has that form and whose premises are true has a true conclusion.

   Valid Argument Form An argument form is *valid* iff every argument with that form which has true premises has a true conclusion. Or, equivalently: an argument form is *valid* iff there is no argument with that form which has true premises and a false conclusion.

(a) An argument with a valid form is therefore *guaranteed* to not lead you from truth to falsehood. If you know that the premises of such an argument are true, then there is no way that its conclusion could be false.

5. We may *prove* that *modus ponens* is a valid argument form. And we may *prove* that *affirming the consequent* is not. To do so, we'll first need to make a stipulation about how to understand the meaning of 'if p, then q' (claims like this are known as *conditionals*). The stipulation is this: a conditional of the form 'if p, then q' is false when (and only when) 'p' is true and 'q' is false. Using a 'truth table' like the following, we can specify the conditions under which 'if p, then q' is true and the conditions under which it is false.

p	q	if p, then q
true	*true*	*true*
true	*false*	*false*
false	*true*	*true*
false	*false*	*true*

(a) The rows of this table represent all of the possible truth-values of 'p' and 'q'.

(b) You may object that this table doesn't get the meaning of 'if..., then...' in English quite right. If you have that objection, I think you are exactly right. English 'if..., then...' is more complicated than the table suggests. But it's not bad as a first pass. Suppose I tell you 'if it's Monday, I'm teaching'. What would it take for you to prove me wrong? You'd have to show that it's Monday and I'm not teaching. And that's exactly when the table above would say 'false'.[5]

6. Then, in order to see whether an *argument form* is invalid—that is, to see whether it is possible that, when we plug in sentences for 'p' and 'q', we will get true premises and a false conclusion—we may simply look at a *truth-table* that contains the premises and the conclusion. For instance, on the left we have the argument form of *modus ponens*; and, on the right, we have the argument form of *affirming the consequent*.

p	q	if p, then q	p	q	
true	*true*	***true***	***true***	***true***	✓
true	*false*	*false*	*true*	*false*	
false	*true*	*true*	*false*	*true*	
false	*false*	*true*	*false*	*false*	

p	q	if p, then q	q	p	
true	*true*	***true***	***true***	***true***	✓
true	*false*	*false*	*false*	*true*	
false	*true*	***true***	***true***	***false***	×
false	*false*	*true*	*false*	*false*	

(a) Look at *modus ponens* on the left: notice that the only way for both premises to be true is for both 'p' and 'q' to be true. (If 'q' is false while 'p' is true, then 'if p, then q' is false.) So, if both premises are true, then we must be in the first row of the truth-table. 'q' is true in the first row of the truth-table. So, if both premises are true, then the conclusion is true as well. So the argument form is valid.

(b) Look at *affirming the consequent* on the right: notice that the premises of the argument could be true if we are in either the first row or the third row of the table. But, in the third row, the conclusion is false. So there are choices of 'p' and 'q' we could plug in which would give the argument true premises but a false conclusion. So the argument form is not valid.

[5] In any case, in order to show that *modus ponens* is a valid argument form, we don't need to concern ourselves with the last two rows at all. And, in order to show that *affirming the consequent* is not valid, we need only assume that it is possible for 'if p, then q' to be true while 'p' is false and 'q' is true. Exercise for the ambitious student: show these claims to be true.

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 and that defending it can lead to great harm

   P2. If you know that something can lead to great harm, then it is permissible to prohibit it.

   C1. It is permissible to prohibit the defense of heresy.

4. There are 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 P1. What he *denies* is P2. 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:

P3. Our opinions about religion, politics, and morality are fallible.

C2. 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:

   P4. Orthodox opinions about politics, religion, or morality can be known to be correct only if it is possible for the orthodoxy to be refuted.

   P5. If heretical opinions are silenced, then it will not be possible for the orthodoxy to be refuted.

   C3. If heretical opinions are silenced, then we cannot know that the orthodoxy is correct.

   P6. It is permissible to silence heretical opinions only if we know that the orthodoxy is correct.

   C4. It is never permissible to silence heretical opinions.

9. Returning to the natural argument in favor of prohibiting the defense of heresy: Mill rejects P2,

   P2. 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 P2, P2*,

      P2*. 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 P2*, 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? How does Fish argue against free speech absolutism?

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. 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.

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

   P1. Free speech is not *instrinsically* valuable—it is only *instrumentally* valuable.

   P2. If free speech is *instrumentally* valuable, then it can work against, rather than for, the end to which it is directed.

   C1. Free speech can work against, rather than for, the end to which it is directed.

   P3. If free speech works against the end towards which it is directed, then it should not be permitted.

   C2. In some cases, we should restrict free speech.

6. On Fish's 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.

PHIL 100 · 3rd April, 2026

Explain J.L. Austin's three kinds of speech acts. *Explain how this theory could be used to formulate a free speech principle. And explain how you could respond to Fish's arguments, if you accepted this free speech principle.*

Speech Act Theory

1. Let's take a detour through the philosophy of language. In the 1960's, J. L. Austin was thinking, not about free speech, 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.

   Saying When you *say* something with your words, your utterance expresses a proposition, or describes the world as being a certain way.

   Doing When you *do* something with your words, your utterance *constitutes* an action.

   Causing When you *cause* something with your words, your utterance has certain *effects* on the world.

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 the things you *do* and *cause* with your words. What it forbids is the regulation of what you *say* with your words.

   (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 *do* something bad or *cause* something bad.

   (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 *saying*, but rather the *doing*

      i. The law still allows the reporter to *say* 'they did [take that person...]' 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 thing *said*, but rather *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 or how they are expressed?

PHIL 100 · 6th April, 2026

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 Los Angeles', 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.

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*. I 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.

   (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

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.

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.

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.

PHIL 100 · 8th April, 2026

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
P2.	The second F was G
	⋮
PN.	The Nth F was G
C.	All Fs are G

P1.	The sun rose on the first day
P2.	The sun rose on the second day
	⋮
PN.	The sun rose on the Nth day
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
UN.	The future will resemble the past in the future

P1.	The sun always rose in the past
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. But this belief is exactly the one we were trying to justify. We have reasoned in a circle. 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. 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.

PHIL 100 · 10th April, 2026

What is enumerative induction? Describe either Hempel's 'paradox of the ravens' or Goodman's 'new riddle of induction', and explain how they pose a challenge to the idea that our beliefs about the future are formed through enumerative induction.

1. Recall, Hume thought that we formed our beliefs about cause and effect through *enumerative induction*—we see lots of *C*s that cause *E*s, and we conclude that all *C*s cause *E*s.

   Enumerative Induction In enumerative induction, you begin with the premises that some collection of *F* things each have a certain property, and you conclude that *all F* things have that property.

	The first raven is black		The first emerald is green		The first *F* is *G*
	The second raven is black		The second emerald is green		The second *F* is *G*
	⋮		⋮		⋮
	The *n*th raven is black		The *n*th emerald is green		The *n*th *F* is *G*
∴	All ravens are black		All emeralds are green	∴	All *F*s are *G*

   Hume's skeptical challenges got off the ground in part because he was assuming that we formed our beliefs about unobserved matters through enumerative induction.

2. Today, we're going to discuss two paradoxes which both raise the question of *what exactly* enumerative induction is meant to be. And you could understand them as challenges to Hume's psychological claim—that we form our beliefs about cause and effect through enumerative induction.

The Paradox of the Ravens

3. The first paradox comes from Carl Hempel. Hempel was trying to give a *formal theory* of inductive inference to match the formal theory of deductive inference we spoke about earlier in the course.

   (a) Hempel assumed that we would want a formal theory of enumerative induction to satisfy the following condition:

   Equivalence Condition If some evidence gives us reason to believe a proposition, then it also gives us reason to believe anything which is logically equivalent to that proposition.

   If one proposition is logically equivalent to another, then it's not possible for one to be true while the other is false. So they are true in exactly the same circumstances. So surely, any reason to think that the one is true is a reason to think that the other is true also.

   (b) Unfortunately, Hempel realized that, when you combine enumerative induction with the equivalence condition, some odd things follow. Suppose, for instance, that you go out and gather the evidence of a white piece of chalk and a green blade of grass. It doesn't seem that this gives you any reason at all to conclude that all ravens are black. But notice that 'All ravens are black' is logically equivalent to 'All non-black things are non-ravens'. (Why? Well, suppose that 'All ravens are black' is false. Then, there must be some raven which isn't black. But that's just a non-black thing which isn't a non-raven. Or

suppose that 'All non-black things are non-ravens' is false. Then, there's some non-black thing which isn't a non-raven. So there's a non-black raven, and 'All ravens are black' is false.)

(c) So then, it looks like enumerative induction should allow us to reason as follows:

The first non-black thing is a non-raven

The second non-black thing is a non-raven

⋮

The nth non-black thing is a non-raven

$\therefore$ All non-black things are non-ravens

$\therefore$ All ravens are black

The white piece of chalk, the green blade of grass, and the blue sweater all give us reason to think that all non-black things are non-ravens. And this is logically equivalent to 'all ravens are black'. So they also give us reason to think that all ravens are black.

(d) But this is *absurd*, right? We can't do 'indoor ornithology' in this way. So what's gone wrong? Perhaps enumerative induction isn't quite right as a description of the way we arrive at generalizations?

The New Riddle of Induction

4. Nelson Goodman argued that we cannot give a *formal* theory of inductive inference. He suggested that, in order to know whether seeing a bunch of G Fs gives us a reason to think that All Fs are Gs, we have to know something about what 'F' and 'G' *mean*.

   (a) Say that a thing is grue if and only if it has been observed before 2027 and is green, or else it has *not* been observed before 2027 and is blue.

   (b) Then, there is no *formal* difference between these two enumerative inductions:

The first emerald is green	The first emerald is grue
The second emerald is green	The second emerald is grue
⋮	⋮
The nth emerald is green	The nth emerald is grue
$\therefore$ All emeralds are green	$\therefore$ All emeralds are grue

   Yet it seems clear that we should expect the first emerald observed in 2027 to be *green* and not *blue*.

5. What is the difference between 'green' and 'grue'? Why do we generalize about one but not the other?

   (a) Perhaps: 'grue' is defined relative to a time?

   (b) But Goodman points out that, we could likewise define 'green' in a way that makes reference to a time. Say that a thing is bleen iff it is first observed before 2027 and is blue or else it is not first observed before 2027 and is green.

   (c) Imagine meeting someone who only knows about grue and bleen. How would you explain to them what green and blue are? You'd have to say:

i. Something is *green* iff it is first observed before 2027 and is grue, or else it is not first observed before 2027 and is bleen.
ii. Something is *blue* iff it is first observed before 2027 and is bleen, or else it is not first observed before 2027 and is grue.

So whether *green* or *grue* is defined with reference to a time depends upon which language you start out speaking.

What is moral realism? What is expressivism? What is 'the error theory'? What is moral skepticism?

1. Some normative sentences: You shouldn't believe that the earth is flat. Animal suffering is bad. Torturing babies for fun is wrong. We should lower the capital gains tax. Climate change is an inter-generational injustice. Happiness is good for you.

2. Metaethics asks: are normative claims objectively true or false? Or are they merely subjective, like matters of taste? Can we know them to be true or false?

Bad Metaethics: Facts vs Opinions

3. Unfortunately, you've probably been exposed to a bad conceptual scheme, which can make metaethics more difficult to understand. This is the grade school distinction between *facts* and *opinions*.

   (a) Grade school handouts teach "A **fact** is something that can be proven true", whereas "An **opinion** is someone's feelings about a particular topic".

   (b) These handouts end up teaching you *four, importantly different* distinctions, and implying (incorrectly) that those distinctions are one and the same.

   (c) Firstly, there is a distinction between propositions which are *true* and those which you *believe* to be true. But this distinction doesn't have anything to do with what you can prove. (The number of blades of grass on planet earth is either odd or even, but I can't prove either one.) And (hopefully!) some of your beliefs are true, so some propositions can be *both* fact *and* belief.

   (d) Secondly, there is a distinction between propositions which are *known* and those which you *believe* to be true. Again, this is a false dichotomy—barring skepticism, many of your beliefs are known!

   (e) Thirdly, there is a distinction between propositions which are *subjective* and those which are *objective*. This is a tricky distinction, and there's many ways we could try to draw it.

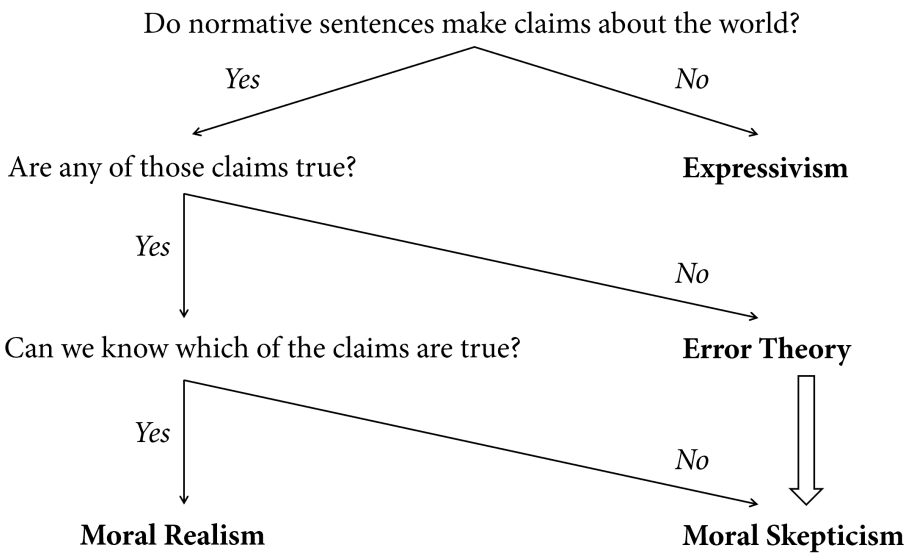
      i. For instance, we might say something is *subjective* if people can genuinely disagree about it without either of them making a mistake. (John thinks that chocolate is tastier than vanilla, and Anna thinks that vanilla is tastier than chocolate; they are disagreeing, but neither has made any kind of mistake.) And something is objective if, whenever two people genuinely disagree, one of them has made a mistake.

      ii. Or we might say that something is *subjective* if it is *about* your own experience. ('Chocolate is tastier *to me* would be subjective in the second sense but not subjective in the first sense.)

   This distinction is importantly different from the first two. If *subjective* means *about your experience*, then it could be a *fact* that something tastes good to you; and you could *believe* or *know* this fact.

   (f) Finally, there is the distinction between *normative* and *non-normative* propositions. A proposition is normative if it is broadly prescriptive or evaluative—if it says something about what's good or bad, about what you ought to do, about what is rational or irrational, about what's just or unjust, and so on. And a proposition is non-normative if it's not like this.

      i. Metaethics is all about how this *final* distinction relates to the previous ones: are normative questions subjective or objective? Are they the kinds of propositions which can be true or false? Can we know them to be true?



Four Metaethical Positions

4. We can break metaethical questions down into three kinds:

   (a) What are these sentences saying about the world? (Semantic)

   (b) Are any of these sentences objectively true? If so, what makes them objectively true? (Metaphysical)

   (c) Can we know which of these sentences are true? (Epistemic)

5. Four metaethical positions:

   (a) The *moral realist* interprets normative claims straightfowardly and says that many of them are both *true* and *knowable*. According to the moral realist, when we say 'animal suffering is bad', we are ascribing a property of *badness* to the state of affairs of animals suffering. And our sentence is true iff the suffering of animals in fact has this property. Finally, the moral realist thinks that we can come to *know* things about moral claims.

   (b) The *expressivist* denies the moral realist's semantic claim. They think that moral claims aren't ascribing any properties at all. Instead, they understand the sentence 'animal suffering is bad' as meaning something like 'boo animal suffering!'.

   (c) The *error theorist* accepts the moral realist's semantic claim, but they deny their metaphysical claim. They say that the sentence 'animal suffering is bad' is ascribing the property of *badness* to the suffering of animals. And they say that the claim is true iff the suffering of animals possesses this property. However, unlike the moral realist, the error theorist says that there are no normative properties, so all of these claims are *false*.

   (d) The *moral skeptic* denies the moral realist's epistemological claim. They say that we cannot come to know things about goodness/badness/rightness/wrongness. Note that a moral skeptic could agree with the moral realist's semantic and metaphysical claims. Just like you think that there's an unknowable fact of the matter about whether the total number of atoms in the universe is odd or even, the moral skeptic could think that there's a fact-of-the-matter about whether animal suffering is bad or good. But they do not think that we can come to know this fact one way or another.

What is 'the error theory'? Give two of J. L. Mackie's arguments for the error theory.

1. Recall the four metaethical positions we discussed last time:

   (a) The *moral realist* says that moral claims are saying things about the world (and not just expressing attitudes), that many of those things are true, and that we can know those things to be true.

   (b) The *expressivist* thinks that moral claims are not saying things about the world. Instead, like 'boo!', they are just expressing attitudes.

   (c) The *error theorist* thinks that moral claims are saying things about the world, but that the things they are saying are simply false. Saying 'animal suffering is bad' is like saying 'witches corrupt the youth.' There are no witches, and there is no badness.

   (d) The *moral skeptic* thinks that we cannot know any moral claims. They might think this because they think no moral claims are true. Or they might think that they are true, but unknowable.

2. J. L. Mackie defends the error theory. He thinks that all moral claims are false. He also defends moral skepticism. He thinks that—even if there *were* normative properties, we could not come to know things about them.

3. He gives two main arguments for the error theory. Firstly,

<u>The Argument from Disagreement</u>

P1. Different people and cultures disagree about normative matters.

P2. The best explanation of this disagreement is that there are no objective moral facts.

C1. There are no objective moral facts.

   (a) Mackie recognizes that P1 on its own does not support his conclusion. People disagree about questions like whether the earth is round, whether Neil Armstrong walked on the moon, and so on. The mere fact of disagreement doesn't mean that there's no fact of the matter about whether the earth is round, or whether Armstrong walked on the moon.

   (b) Mackie thinks that moral disagreement is different from empirical disagreement. He thinks that people disagree about morality *because* they participate in different ways of life; he doesn't think that people participate in different ways of life *because* of their different moral values. For instance: people approve of monogamy *because* they were raised in a monogamous society.

   (c) A potential response to the argument: we can explain moral 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 from a disagreement about whether cows suffer.

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

4. Mackie's second argument for the error theory is that, if there *were* normative facts and properties, they would be very *strange* (or *queer*) kinds of facts and properties. Moreover, any way of coming to *know* things about these facts and properties would be a very strange faculty.

   (a) The feature of normative properties that Mackie finds queer is that they have a kind of objective authority over us—that they are *objectively prescriptive*. As he puts it: rightness would have to have a kind of *to-be-doneness*, and wrongness a kind of *to-be-avoidedness*.

   The Argument from Queerness (Metaphysical)

   P3. If there were normative properties, they would be objectively prescriptive.

   P4. Nothing is objectively prescriptive.

   C2. There are no normative properties.

   He also notes that, if there were moral properties, then we would only be able to come to *know* things about them with a very strange (queer) kind of perceptual capacity.

   The Argument from Queerness (Epistemological)

   P3. If there were normative properties, we would only be able to know things about them with a very strange kind of perceptual capacity.

   P4. We do not have any strange perceptual capacities.

   C2. If there are normative properties, we cannot know anything about them.

5. After giving these arguments against moral realism, Mackie gives an alternative picture of what's going on with moral discourse. on which what we take to be objective moral properties of external states of affairs are in fact just subjective reactions in ourselves that we incorrectly project out onto the external world.

   (a) We have internal attitudes and responses to states-of-affairs and to other people's behavior. For instance, if we see someone being cruel, we are horrified, we condemn the cruel person, and we demand that they change their behavior.

   (b) On Mackie's view, we (incorrectly) *project* these internal, subjective attitudes and reactions out onto the world. We take our subjective responses to correspond to some objective property of the world.

   (c) Compare with the case of *color*: a *subjectivist* about color thinks that objects do not have objective colors. Instead, color is a feature of our subjective experience that we (incorrectly) project out onto the world. Mackie says the same thing about things like *goodness* and *badness*.

Give two of Enoch's tests for objectivity, and explain why these tests indicate that morality is objective. What is moral realism? Give one of Mackie's arguments against moral realism and explain how Enoch responds to this argument.

1. Mackie thinks that all normative claims are false, but he doesn't encourage us to stop making or deliberating about normative claims. He writes:

   > Morality is not to be discovered but to be made: we have to decide what moral views to adopt, what moral stands to take...the object of the exercise [is] to decide what to do, what to support and what to condemn, what principles of conduct to accept and foster as guiding or controlling our own choices and perhaps those of other people as well.

   (a) Mackie therefore views normativity the way some people view *fashion* or *taste*. They might think that there's no objective truth about what's fashionable or tasty, but nonetheless think that we should try to develop a sense of fashion or taste. The point of developing this sense isn't to discover an objective truth, but rather to decide what to wear or what to eat. According to Mackie, it's the same with normativity.

2. In his essay, Enoch wants to emphasize just how *alien* this way of understanding normativity really is. He has two goals: 1) firstly, to persuade you that you are *already* an objectivist about morality—you already think that there are objective moral truths to be investigated and discovered—and 2) secondly, to convince you that people like Mackie haven't given you any good reason to give up on this view.

3. Enoch gives three 'tests' for whether we treat a subject matter as objective or subjective.

   (a) **The 'Spinach Test'**:

      i. A child says "I'm glad I don't like spinach. Because, if I did like spinach, then I'd eat it, and spinach is yucky." We find this funny, because we don't think that *yuckiness* is objective. There's no real reason to not want to eat spinach, if you aren't disgusted by it. But we don't find a corresponding speech about morality funny: "I'm glad I wasn't born in the antebellum south. Because, if I were, I'd be a racist, and racism is wrong". Or consider: "I'm glad I wasn't raised in the cult; because, if I were, I'd have lots of irrational beliefs". These speeches aren't funny. According to Enoch, this reveals that we don't think about normativity as subjective.

   (b) **The Disagreement/Deliberation Test**:

      i. When we disagree about whether chocolate or vanilla is better, it feels like you are simply stating your own preferences, and trying to get others to align their preferences with yours. But when we disagree about whether abortion is permissible, or animal suffering is bad, or whether it's rational to believe that Neil Armstrong walked on the moon, it doesn't feel like that.

      ii. When you deliberate about whether to get chocolate or vanilla, you look *inwards* at yourself, and try to discern your own preferences. But when you deliberate about whether abortion is permissible, or animal suffering is bad, or whether it's rational to believe that Neil Armstrong walked on the moon, you look *outward*. It doesn't feel like you're trying to discover what your response is, or decide upon a response. The feeling of normative deliberation patterns with factual deliberation, and not with deliberation about subjective preference.

   (c) **The 'What If' Test**: on Mackie's view, If we think about what would happen, were our subjective attitudes to change, we don't tend to think that the normative facts change with them.

i. Consider the following claims:
   A. If everyone wore and liked the look of tophats, then tophats would be fashionable.
   B. If everyone believed the earth was flat, the earth would be flat.
   C. If everyone liked and approved of torturing innocent people for entertainment, then torturing innocent people for entertainment would be moral.
   D. If everyone decided what to believe by consulting Tarot cards, then consulting Tarot cards would be a rational way to decide what to believe.

ii. (A) seems true, but (B) seems false. When we ask about what things would be like, were our attitudes to change, it seems like subjective matters like fashion *would* change, but objective matters like the shape of the earth would *not* change.

iii. But Enoch suggests that here, too, the normative patterns with the objective, and not the subjective. Both (C) and (D) seem false.

4. On Mackie's view, normativity is *response-dependent*. Since we are just projecting our internal subjective responses out onto the world, normativity depends upon your internal subjective responses. But Enoch's three tests suggest that we treat normativity as though it were *response-independent*.

   (a) Relatedly, on a subjectivist view, it is difficult to make sense of the idea of moral *progress* or *regress*. Attitudes can change over time, but in order for these changes to constitute *progress*/*regress*, they must be getting closer to/further from some truth or ideal which is independent of us and our responses.

5. After suggesting that we already think of normative matters as objective, Enoch goes on to respond to Mackie's arguments that normativity is subjective.

   (a) In response to Mackie's argument from Disagreement, Enoch suggests an alternative explanation of disagreement: people engage in *motivated reasoning*. They tend to endorse the normative claims that advantage them. This can explain why there's so much disagreement without undermining the objectivity of normativity. Compare: the rich tend to think that lowering the capital gains tax would make even the poorest of the poor richer, whereas the poor tend to disagree. This isn't a normative question, though.

      i. Enoch also raises a separate objection to Mackie's argument from Disagreement: it is *self-undermining*. After all, there is disagreement about whether normativity is subjective or objective. If disagreement is enough to undermine the objectivity of normativity, why isn't it also enough to undermine the objectivity of metaethical positions like subjectivism?

   (b) In response to Mackie's epistemic argument, Enoch points out that there are many areas where we seem to be able to acquire knowledge without direct perception. For instance, we know things about mathematics, even though we do not directly perceive mathematical reality. Enoch draws a distinction between the *a priori* and the *a posteriori*. Knowledge is *a priori* if the knowledge does not depend upon sensory perception—this includes knowledge about mathematics and philosophy. And knowledge is *a posteriori* if it depends upon sensory perception—this includes knowledge about physics and biology.

      i. Again, Enoch raises a separate worry about Mackie's epistemic argument: it is *self-undermining*. After all, we don't come to know things about metaethics *via* perception. So we could raise a parallel epistemic argument against Mackie's own subjectivism: even if it were true, we could never come to know that it's true.

PHIL 100 · 20 April 2026

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 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.

   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, it was that it was far more *falsifiable*.

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 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.

   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.

   (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 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*):

If T, then E	If T, then E
E	It is not the case that E
T	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*.

   (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.

PHIL 100 · 22 April, 2026

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.

   (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.

   (a) Consider, for instance, Newton's theory of universal gravitation (TUG).

   (b) If we want to derive a *testable prediction* about the orbit of the planets 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.

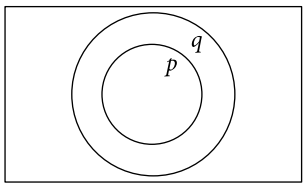
What is Susan Haack's 'problem of deduction'? Explain the problem by talking about the justification of modus ponens *and the rule Haack calls* modus morons. *What is rule externalism, and how would a rule externalist respond to Haack's problem?*

The Problem of Deduction

1. Recall, Hume raised a problem with our method of forming beliefs about the future. The problem is that it looks as though the only reasons we have for holding these beliefs are circular.

2. Susan Haack notes that this is a problem for *deduction* as well.

3. Recall the justification of the argument form *modus ponens* (if p, then q; p; therefore, q) we considered earlier in the course. If 'if p, then q' is true, then every relevant possibility in which 'p' is true must be one in which 'q' is true, too. We can represent this with an Euler diagram If we know 'if p, then q' is true, then we know



that the 'p' circle is inside the 'q' circle. And if we know that 'p' is true, then we know that we're in the p circle. So we have to be in the 'q' circle.

4. But wait!—this reasoning *used modus ponens*, which is the very rule we were trying to justify:

If we're in the p circle, then we're in the q circle
We're in the p circle

We're in the q circle

So the justification of deduction we gave earlier in the course is circular in exactly the same way that our justification of induction was circular.

5. We can give a similar justification of the deductively invalid argument form *affirming the consequent*:

If we're in the p circle, then we're in the q circle
We're in the q circle

We're in the p circle

This inference *uses* the rule of affirming the consequent to *justify* the rule of affirming the consequent—just as the counterinductivist used counterinduction to justify counterinduction.

What the Tortoise Said to Achilles

6. Lewis Carroll shares a story about a dialogue between the tortoise and Achilles. The Tortoise references the first proposition proved in Euclid's *Elements*:

   A. Things equal to the same are equal to each other
   B. The two constructed sides of the triangle are equal to the original side

   ---

   Z. The two constructed sides of the triangle are equal to each other

   (a) The Tortoise is happy to accept the premises, but he does not think that the conclusion follows. He asks why logic should compel him to accept Z, just because he accepts A and B. Achilles responds: because *if* A and B are true, then Z must be true, too. The Tortoise is willing to accept this additional premise, but he still doesn't think that Z follows.

   (b) Achilles responds that, if A, B, and C are all true, then Z must be true, too. The Tortoise is again happy to accept this premise, but doesn't think that Z follows.

   (c) The Tortoise is willing to accept all of the following,

   D. If A and B and C are true, Z must be true
   E. If A and B and C and D are true, Z must be true
   ⋮

   But not willing to concede that Z follows.

7. We can draw two lessons from this story:

   (a) First, we must draw a distinction between *premises* and *rules of inference*. A rule of inference is a way of reasoning. A premise is something *from which* you reason. The Tortoise's trick is treating each proposed rule of inference as if it were just another premise. We shouldn't treat a rule of inference as though it were among the reasons you have for reaching the conclusion.

   (b) Second, if we're speaking with someone who refuses to use the same rules of inference as us, we may not be able to give them reasons to accept a conclusion that they will recognize as reasons.

8. Return to the problem of induction and the problem of deduction. With the Tortoise's lesson learnt, let's distinguish bewteen two different ways for an argument to be *circular*:

   (a) A *premise-circular* argument contains its conclusion as one of its premises.

   (b) A *rule-circular* argument is an argument *for* using a particular rule of inference which itself *uses* that very rule of inference.

   (c) If rules of inference are not to be counted among the *reasons* upon which our conclusion is based, then perhaps there is nothing wrong with reasoning *rule circularly*.

   Rule Externalism If you infer a conclusion from known premises using inference rule R, then you come to know the conclusion iff R is a good rule of inference. (You needn't independently know that R is a good rule of inference.)

   (d) The rule externalist says that our justification of *modus ponens* can put us in a position to know that *modus ponens* is a valid way of reasoning, even if we didn't already know that it was a valid way of reasoning. All that matters is that *modus ponens* in fact *is* a valid way of reasoning. They similarly say that our justification of inductive can put us in a position to know that induction is a good way of reasoning, so long as induction in fact *is* a valid way of reasoning.

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, distinguish the following two theses, both of which Nagel accepts:
   (a) Thesis #1: Our lives are meaningless.
   (b) Thesis #2: Our lives are *absurd*.
      ▷ 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.

2. Though Nagel thinks our lives are meaningless and absurd, he thinks that many of the reasons people typically give for thinking so are bad ones. For instance:
   (a) To argue that our lives are meaningless, some 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 now. [from P1 and P2]

      C2. It does not matter now that we will not 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? How does prolonging a meaningless existence make it 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:
      ▷ 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.
         ▷ 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).
      ▷ We are incapable of saying anything to justify these fundamental commitments which does not simply presuppose those very commitments.

- ▷ 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 recognizing (1) and (2) puts us in a position to recognize (3). But here is an attempted reconstruction of his reasoning:

   P1. In order for a life to be meaningful, you must be in a position to answer doubts about the values of your projects and commitments.

   P2. We are unable to answer doubts about the values of our projects and commitments.

   C. Our lives are meaningless.

   (a) In support of P1, Nagel gives the following example: suppose we learn that we are being raised by some alien civilization for consumption. These aliens may be engaged in some worthwhile project, and by feeding the aliens, we may be participating in this project. But Nagel insists that this isn't enough to make our lives meaningful:

   > ...although we might acknowledge that this culinary role would make our lives meaningful to *them*, it is not clear how it would make them meaningful to *us*.

   According to Nagel, what's needed to make our lives meaningful for *us* is an ability to answer doubts about the value of our lives and (non-circularly) justify our most fundamental commitments. Since he thinks we are unable to do this, our lives cannot be meaningful for us (though we may often *think* or *feel* that our lives are meaningful).

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 were meaningful.

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. It does not matter that nothing matters.

PHIL 100 · 28 April, 2026

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? Explain why she thinks it's important to say that we can be wrong about whether our lives are meaningful.*

Meaning

1. Distinguish two questions: 1) What is the meaning of life?; and 2) What makes a life meaningful?

   - ▷ The first question asks about the purpose of our existence. According to Wolf, this question is easily answered. If we have a creator, then life has a purpose (namely, whatever purpose the creator had in creating us). If there is no creator, then life has no purpose.
   - ▷ But she thinks the second question is more interesting; and she has an answer to give.

2. Some paradigm examples of meaningless lives (according to Wolf):

   - ▷ *The couch potato*: sits at home on the couch every day drinking beer and watching situation comedies.
   - ▷ *The Idle Rich*: "flits about, fighting off boredom, moving from one amusement to another. She shops, she travels, she eats at expensive restaurants, she works out with her personal trainer."
   - ▷ *The Corporate Executive*: works twelve hours a day, every day of the week, solely in order to accumulate a vast fortune.
   - ▷ *The Pig Farmer*: has only instrumental goals, and no final goals. They want to grow corn in order to feed pigs in order to sell pigs in order to buy more land in order to grow more corn in order to feed more pigs, in order to sell the pigs, in order to buy more land...
   - ▷ *The Scooped Scientist*: dedicates their life to developing a cure for cancer. The day before their discovery is revealed, another scientist announces that they have discovered the very same cure.

3. Reflecting on these cases, Wolf argues against some proposals about what makes a life meaningful:

   Proposal Someone's life is meaningful iff their life seems meaningful to them

   - ▷ Wolf: the couch potato and the idle rich may well *think* that their life is meaningful; their lives may *seem* meaningful to them. Even so, they are not. What's missing is that their lives are purely *passive*. They are not actively engaged in any projects, nor do the projects they're engaged in have any value.

   Proposal #2 Someone's life is meaningful iff they are actively engaged in projects they take to be valuable.

   - ▷ Wolf: The corporate executive shows that this is false. Even if the corporate executive takes the accumulation of wealth to be valuable, that isn't enough to make their lives meaningful.

4. Neither of the first two proposals allow for someone to *realize* that their life so far has been meaningless. Suppose the corporate executive spends their twenties and thirties accumulating wealth, only to have a midlife crisis and conclude that their life to date has been meaningless. During their twenties and thirties, they took accumulating wealth to be meaningful, and their life seemed meaningful to them. So proposals #1 and #2 say that their life in their twenties and thirties *was* meaningful. So, when they have their midlife crisis and conclude that their life wasn't meaningful, proposals #1 and #2 say that they are *wrong*.

P1. If proposal #1 (#2) is correct, then it is impossible to discover that your life was meaningless.

P2. It is possible to discover that your life was meaningless.

C1. Proposal #1 (#2) is incorrect.

5. According to Wolf, what's missing from proposal #2 is something *objective*. You not only need to be actively engaged in projects you *take* to be valuable. Those projects also have to *be* valuable.

   Wolf thinks that the case of the scooped scientist gives us a reason to think that your engagement with the valuable projects also have to be at least somewhat *successful*. So she offers the following proposal about what makes a life meaningful:

   Wolf's Account Someone's life is meaningful iff they are *actively* and at least somewhat *successfully* engaged in objectively valuable projects.

6. 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.

Meaning as a Component of Well-being

7. In the reading, Wolf's main topic is *self-interest*—or, to use different terms for the same thing, *welfare* or *well-being*.

   Recall earlier in the semester, we encountered these two theories of well-being:

   Hedonism The only thing intrinsically good for a person is pleasure; the only thing intrinsically bad for them is pain.

   Desire Satisfaction The only thing intrinsically good for a person is getting what they want; the only thing intrinsically bad for them is not getting what they want.

   - ▷ Wolf: both of these theories are wrong. *Meaning* is intrinsically good for a person, whether or not a meaningful life makes them happy, and whether or not they want a meaningful life.
   - ▷ A hedonist can grant that meaning is good for a person. But they'll say it's only *instrumentally* good for that person. What's intrinsically good for them is the *feeling of fulfillment*. Wolf thinks that this is a mistake. Suppose the scooped scientist never discovers that they have been scooped. They have the feeling of fulfillment, but not the meaning. Contrast them with the scientist whose cure for cancer is not scooped. Wolf thinks that things are going better for the unscooped scientist than the scooped one.
   - ▷ A desire satisfaction theorist can grant that meaning is good for a person. But they'll say that it is only because we usually want meaning in our lives. Wolf also thinks this is a mistake. If this were so, then our desire for meaning would be *ungrounded* and *arbitrary*. But Wolf thinks we desire meaning in our lives *for a reason*—we want to have a meaningful life precisely because we recognize that it is good for us to lead meaningful lives.